



IGFBP2 Is a Potential Master Regulator Driving the Dysregulated Gene Network Responsible for Short Survival in Glioblastoma Multiforme

Manasa Kalya^{1,2}, Alexander Kel^{2,3*}, Darius Wlochowitz¹, Edgar Wingender² and Tim Beißbarth¹

¹ Department of Medical Bioinformatics, University Medical Center Göttingen, Göttingen, Germany, ² geneXplain GmbH, Wolfenbüttel, Germany, ³ Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia

OPEN ACCESS

Edited by:

Alessandro Laganà,
Icahn School of Medicine at Mount
Sinai, United States

Reviewed by:

Balaji Banoth,
St. Jude Children's Research
Hospital, United States
Daniela Albrecht-Eckardt,
BioControl Jena GmbH, Germany

*Correspondence:

Alexander Kel
alexander.kel@geneXplain.com

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 20 February 2021

Accepted: 06 April 2021

Published: 15 June 2021

Citation:

Kalya M, Kel A, Wlochowitz D,
Wingender E and Beißbarth T (2021)
IGFBP2 Is a Potential Master
Regulator Driving the Dysregulated
Gene Network Responsible for Short
Survival in Glioblastoma Multiforme.
Front. Genet. 12:670240.
doi: 10.3389/fgene.2021.670240

Only 2% of glioblastoma multiforme (GBM) patients respond to standard therapy and survive beyond 36 months (long-term survivors, LTS), while the majority survive less than 12 months (short-term survivors, STS). To understand the mechanism leading to poor survival, we analyzed publicly available datasets of 113 STS and 58 LTS. This analysis revealed 198 differentially expressed genes (DEGs) that characterize aggressive tumor growth and may be responsible for the poor prognosis. These genes belong largely to the Gene Ontology (GO) categories “epithelial-to-mesenchymal transition” and “response to hypoxia.” In this article, we applied an upstream analysis approach that involves state-of-the-art promoter analysis and network analysis of the dysregulated genes potentially responsible for short survival in GBM. Binding sites for transcription factors (TFs) associated with GBM pathology like NANOG, NF- κ B, REST, FRA-1, PPARG, and seven others were found enriched in the promoters of the dysregulated genes. We reconstructed the gene regulatory network with several positive feedback loops controlled by five master regulators [insulin-like growth factor binding protein 2 (IGFBP2), vascular endothelial growth factor A (VEGFA), VEGF165, platelet-derived growth factor A (PDGFA), adipocyte enhancer-binding protein (AEBP1), and oncostatin M (OSMR)], which can be proposed as biomarkers and as therapeutic targets for enhancing GBM prognosis. A critical analysis of this gene regulatory network gives insights into the mechanism of gene regulation by IGFBP2 via several TFs including the key molecule of GBM tumor invasiveness and progression, FRA-1. All the observations were validated in independent cohorts, and their impact on overall survival has been investigated.

Keywords: glioblastoma, master regulators, upstream analysis, IGFBP2, FRA-1, FOSL1, short term survivors, transcription factors

INTRODUCTION

Glioblastoma multiforme (GBM) is the most common, highly malignant primary brain tumor (Wen and Kesari, 2008). Despite huge developments in treatment strategies, GBM poses unique treatment challenges due to tumor recurrence (34%) and drug resistance leading to poor survival rates of less than 15 months even after advanced chemoradiotherapy (Krex et al., 2007). As few as

2% of patients respond to standard therapy and survive beyond 36 months (Krex et al., 2007; Das et al., 2011), clinically called long-term survivors (LTS). Another group termed short-term survivors (STS) are those who survive less than 12 months (Shinawi et al., 2013). The factors that determine the long survival are not well understood.

Though several factors like age, gender, Karnofsky Performance Score, the extent of tumor resection, radiotherapy, and chemotherapy are associated with survival and treatment response (Scott et al., 1999; Lee et al., 2008; Sonoda et al., 2009; Zhang et al., 2012), it is evident from recent research that certain molecular signatures can be connected with treatment response and thereby survival. Promoter methylation of the gene MGMT, mutations in the genes IDH1/2, and loss of heterozygosity in chromosome 1p/19q have been confirmed to be highly informative (Krex et al., 2007; Das et al., 2011; Zhang et al., 2012; Han et al., 2014; Reifenberger et al., 2014; Franceschi et al., 2015; Chen et al., 2016). Furthermore, CHI3L1, FBLN4, EMP3, IGFBP2, IGFBP3, LGALS3, MAOB, PDPN, SERPING1, and TIMP1 gene expression has repeatedly been reported to be decreased in LTS patients (De Vega et al., 2009; Bi and Beroukhi, 2014; Han et al., 2014; Franceschi et al., 2015). A better characterization of these extreme survival groups at the molecular level will likely shed important light on the biological aspects that drive their malignancy and survival.

With the advent of gene expression profiling and remarkable developments in high-throughput technologies, it is possible to gain deeper molecular insights into disease biology. Databases like Gene Expression Omnibus—GEO (Barrett et al., 2013), Array Express (Athar et al., 2019), and The Cancer Genome Atlas—TCGA (Grossman et al., 2016) serve as open platforms for retrieval of high-quality multi-omics data to search for new markers in cancer research. The analysis of differentially expressed genes (DEGs) is already an important and established *in silico* strategy to identify potential drivers of cellular state transitions. For a more refined analysis, annotation of DEGs, using *a priori* known biological categories from the Gene Ontology (GO; Ashburner et al., 2000) and pathway databases, e.g., TRANSPATH® (Krull et al., 2003), KEGG (Kanehisa et al., 2020), PANTHER (Thomas et al., 2003), and Reactome (Jassal et al., 2020), has proven to be an effective hypothesis-driven approach in cancer research. Moreover, with the advent of state-of-the-art promoter analysis, it is now possible to establish gene regulatory networks computationally that can be used to understand the causes of gene dysregulation and for identification of causal master regulators driving them. In this regard, we applied the Genome Enhancer¹, a multi-omics analysis tool that makes use of the open-source programming environment BioUML (Kolpakov et al., 2019) and incorporates an automated pipeline for the previously published “upstream analysis” (Koschmann et al., 2015; Boyarskikh et al., 2018) and the “walking pathways” (Kel et al., 2019) approach. There are two major steps that constitute this strategy: (1) analysis of the promoters of DEGs to identify relevant transcription factors (TFs): this is done with the help of the TRANSFAC®

¹<https://genexplain.com/genome-enhancer/>

database (Matys et al., 2006) and the binding site identification algorithms, MATCH™ (Kel et al., 2003, 2006) and CMA (Waleev et al., 2006); (2) reconstruction of signaling pathways that activate these TFs and identification of master regulators on the top of such pathways: for this, the signaling pathway database TRANSPATH® (Krull et al., 2003) has been employed in conjunction with special graph search algorithms that identify positive feedback loops (Kel et al., 2019).

In this study, we applied the upstream analysis to publicly available datasets of GBM from the GEO database to understand the gene-regulatory networks contributing to short survival in GBM. This regulatory network revealed a set of 12 TFs binding to the regulatory regions of the genes of interest and five master regulators regulating them, namely, (a) vascular endothelial growth factor A (VEGFA), a mediator of angiogenesis (Xu et al., 2013) and a promoter of stem-like cells in GBM; (b) PDGF, a highly amplified gene and key player of tumorigenesis (Martinho and Reis, 2011); (c) oncostatin M (OSMR), which orchestrates feed-forward signaling with EGFR and STAT3 to regulate tumor growth (Jahani-As et al., 2016); (d) adipocyte enhancer-binding protein (AEBP1), which plays a key role in pathogenesis through NF-κB activation (Majdalawieh et al., 2020); and (e) IGFBP2.

Insulin-like growth factor binding protein 2, a well-established molecule of interest in GBM (Yao et al., 2016), was found to be more highly expressed in STS and to have an impact on overall survival. IGFBP2 expression is said to be higher in all four (classical, mesenchymal, proneural, and neural) GBM subtypes (Lindström, 2019). It also drives gene programs for immunosuppression in the mesenchymal subtype and is suggested as an immunotherapeutic target (Liu et al., 2019). In non-mesenchymal subtypes (classical, proneural, and neural), it modulates cell proliferation (Phillips et al., 2016; Cai et al., 2018). It has also been found to be a marker of tumor aggressiveness and a prognostic marker for survival (Lindström, 2019). However, the molecular mechanism by which IGFBP2 affects disease progression and patient prognosis is not fully understood.

This work focuses on understanding gene regulatory networks that drive short survival in GBM and their master regulators, which we suggest as biomarkers and therapeutic targets. Later, we critically discuss the role of IGFBP2 in the gene regulatory network.

RESULTS

Identification of Differentially Expressed Genes

Identifying DEGs gives us insight into the biological semantics of a cellular state and helps to identify promising biomarkers of various disease states. The differential gene expression analysis between STS and LTS groups of GBM, from the batch-corrected GSE dataset, was performed using linear models for microarray data (LIMMA) (Ritchie et al., 2015) with FDR cutoff of 5%. The analysis revealed 957 genes that are significantly differentially expressed (DEGs) (adjusted p -value < 0.05). Furthermore, the analysis revealed 115 significantly upregulated ($\log_2FC > 0.5$) and 83 significantly downregulated [$\log_2FC < (-0.5)$] genes.

The top five upregulated and downregulated genes and their corresponding log₂FC are shown in **Table 1** and the full list is given in **Supplementary Table 1-A**.

Functional Annotation of Differentially Expressed Genes

Functional annotation was performed to investigate the biological roles of these DEGs. As shown in **Supplementary Figure 1A**, the top GO biological processes are extracellular structure and matrix organization with 30 DEG hits. **Supplementary Figure 1B** shows the results for GO cellular component enrichment, which revealed dysregulation of genes that encode proteins for the extracellular matrix and synaptic membranes. The important enriched molecular function GO terms are channel activity and transmembrane transporter activity (**Supplementary Figure 1C**). The disruption in extracellular matrix organization is one of the important signatures in glioblastoma treatment response dealing with invasiveness and malignancy (De Vega et al., 2009). Deeper biological insights are required in this aspect. It is interesting to see enrichment of genes known to be involved in glioma (**Figure 1A**). Gene signature enrichment based on hallmark gene sets of MSigDB clearly signifies the enrichment of epithelial-to-mesenchymal transition depicted in **Figure 1B**. The process of epithelial-to-mesenchymal transition plays a very important role in GBM survival by driving tumor invasiveness and drug resistance (Iwade, 2016). Important pathways like Aurora signaling, G2/M phase transition, and TGF- β pathway are found

to be enriched according to TRANSPATH® (**Table 2**). The full list of enrichment results can be found in **Supplementary Table 1-B**.

Identifying the Master Regulators of Dysregulated Gene Networks

Reconstruction of the disease-specific regulatory networks can help to identify potential master regulators that may serve as mechanism-based biomarkers or as therapeutic targets to block a specific pathological regulatory cascade. Using the promoter analysis as a first step, we analyzed enrichment of TF binding sites in promoters of upregulated genes of STS using DNA-binding motifs from the TRANSFAC® library. Two hundred seventy-four TFs (**Supplementary Table 1-C**) enriched for CCKR signaling, interleukin signaling, PDGF signaling, and WNT signaling were found to have their binding sites enriched; full enrichment results can be found in **Supplementary Table 1-D**.

Next, we applied the Composite Module Analyst (CMA) and identified two modules involving 12 TF binding site combinations that regulate the expression of the genes of interest. CMA revealed the following modules comprising clustering binding sites for the following TFs: module 1: HNF3B, NANOG, NFKAPPAB, TAF1, TCF4, and FRA-1; module 2: PPARG, TAL1, REST, POU6F1, FOSJUN, and PBX. The modules and their significance are depicted in **Supplementary Figure 2**. Differential expression statistics for the 12 TFs are given in **Supplementary Table 2**. Among them, FRA-1 TF (also known as FOSL1) was found to be *p*-value significant and upregulated in STS of GBM (log₂FC = 0.023, *p*-value = 0.008, adjusted *p*-value (0.093) (**Supplementary Table 2**).

Figure 2 validates the predicted cluster of TF binding sites from the composite modules identified in the promoter of IGFBP2 gene. We can see that binding sites for the TFs – c-Fos/c-Jun, Nanog, Tal-1, and HNF3/FoxA1 in this cluster can be confirmed by publicly available ChIP-seq data of the GTRD database (Kolmykov et al., 2021). In addition, binding site of FRA-1 can be confirmed by a cluster of mapped reads of independent publicly available ChIP-seq data (FRA1 track in **Figure 2**) (full map is shown in the **Supplementary Figure 4**).

Finally, we reconstructed signaling network that activates the TFs revealed by CMA analysis and thereby identifying the top regulators in these networks using the TRANSPATH® database. With this approach, we identified five important master regulators that are plausible drivers of short survival in GBM: IGFBP2, VEGFA/VEGF165, platelet-derived growth factor A (PDGFA), AEBP1, and OSMR. All the master regulators were found to be significantly upregulated in STS. The genes that encode the master regulator proteins are controlled by the TFs revealed by CMA in their promoters, which maintains the multiple positive feedback loops in the system. It should be underlined here that, in such networks with positive feedback loops, the identified key TFs, such as FRA-1, are both upstream of their target genes, among them the IGFBP2, as well as downstream from the master regulator proteins, one of them the IGFBP2 protein. The regulatory network reconstructed with six master regulators is shown in **Figure 3**, and the master regulators and their log₂FC in STS are listed in **Table 3**. Since

TABLE 1 | The list of the top five significantly upregulated and downregulated genes in STS identified in the GSE dataset.

Gene symbol	Description	Log ₂ FC	<i>p</i> -Value	Adjusted <i>p</i> -value
Upregulated genes				
CHI3L1	Chitinase-3-like 1	1.371	9.73E–05	0.013
PDPN	Podoplanin	1.241	7.88E–07	0.002
MEOX2	Mesenchymal homeobox 2	1.159	6.45E–04	0.028
IGFBP2	Insulin-like growth factor binding protein 2	1.149	4.87E–05	0.010
COL6A2	Collagen type VI alpha 2 chain	1.0479	5.79E–05	0.011
Downregulated genes				
KLRC2	Killer cell lectin-like receptor C2	–1.2187	3.63E–04	0.022
KLRC1	Killer cell lectin-like receptor C1	–1.2187	3.63E–04	0.022
FUT9	Fucosyltransferase 9	–1.0709	1.15E–04	0.014
DPP10	Dipeptidyl peptidase-like 10	–1.02781	2.97E–05	0.008
GABRB3	Gamma-aminobutyric acid type A receptor subunit beta3	–0.96352	6.73E–05	0.011

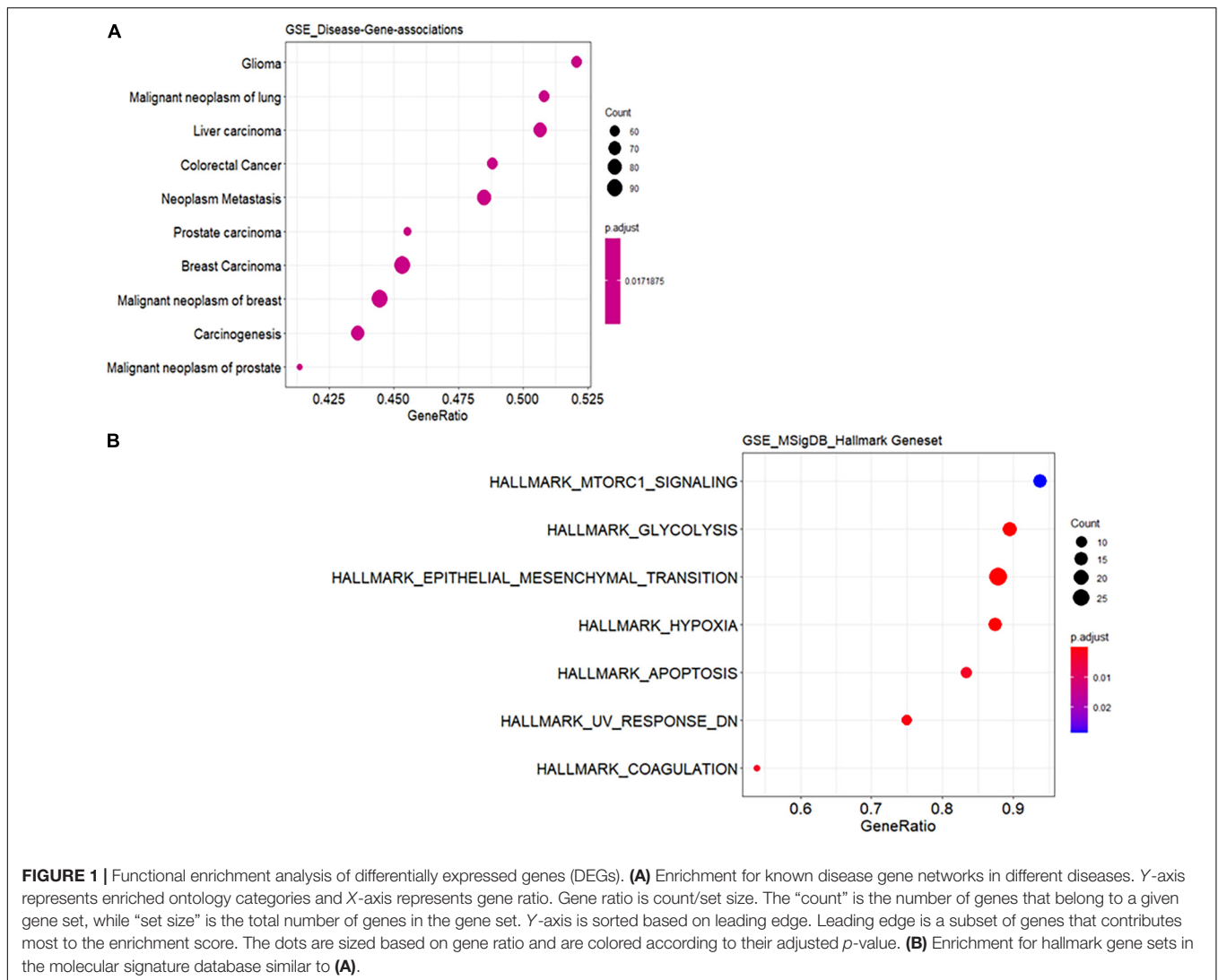


FIGURE 1 | Functional enrichment analysis of differentially expressed genes (DEGs). **(A)** Enrichment for known disease gene networks in different diseases. Y-axis represents enriched ontology categories and X-axis represents gene ratio. Gene ratio is count/set size. The “count” is the number of genes that belong to a given gene set, while “set size” is the total number of genes in the gene set. Y-axis is sorted based on leading edge. Leading edge is a subset of genes that contributes most to the enrichment score. The dots are sized based on gene ratio and are colored according to their adjusted *p*-value. **(B)** Enrichment for hallmark gene sets in the molecular signature database similar to **(A)**.

TABLE 2 | Pathway enrichment using the TRANSPATH® pathway (2019.3) for differentially expressed genes.

ID (TRANSPATH)	Title	Group size	Expected hits	Nominal <i>p</i> -value	ES	Rank at max	NES	FDR	Number of hits
CH000001004	Aurora-A cell cycle regulation	68	67.262	0	0.422	8,347	4.138	0	68
CH000000919	Cyclosome regulatory network	77	76.164	0	0.349	7,336	3.728	0	77
CH000000694	G2/M phase (cyclin B: Cdk1)	66	65.284	0	0.375	6,641	3.587	0	66
CH000000879	Caspase network	83	82.099	0	0.333	8,414	3.523	0	83
CH000000711	TGFbeta pathway	153	151.340	0	0.232	8,431	3.346	0	151

VEGF165 is a splice variant of VEGFA, only the latter will be considered further on.

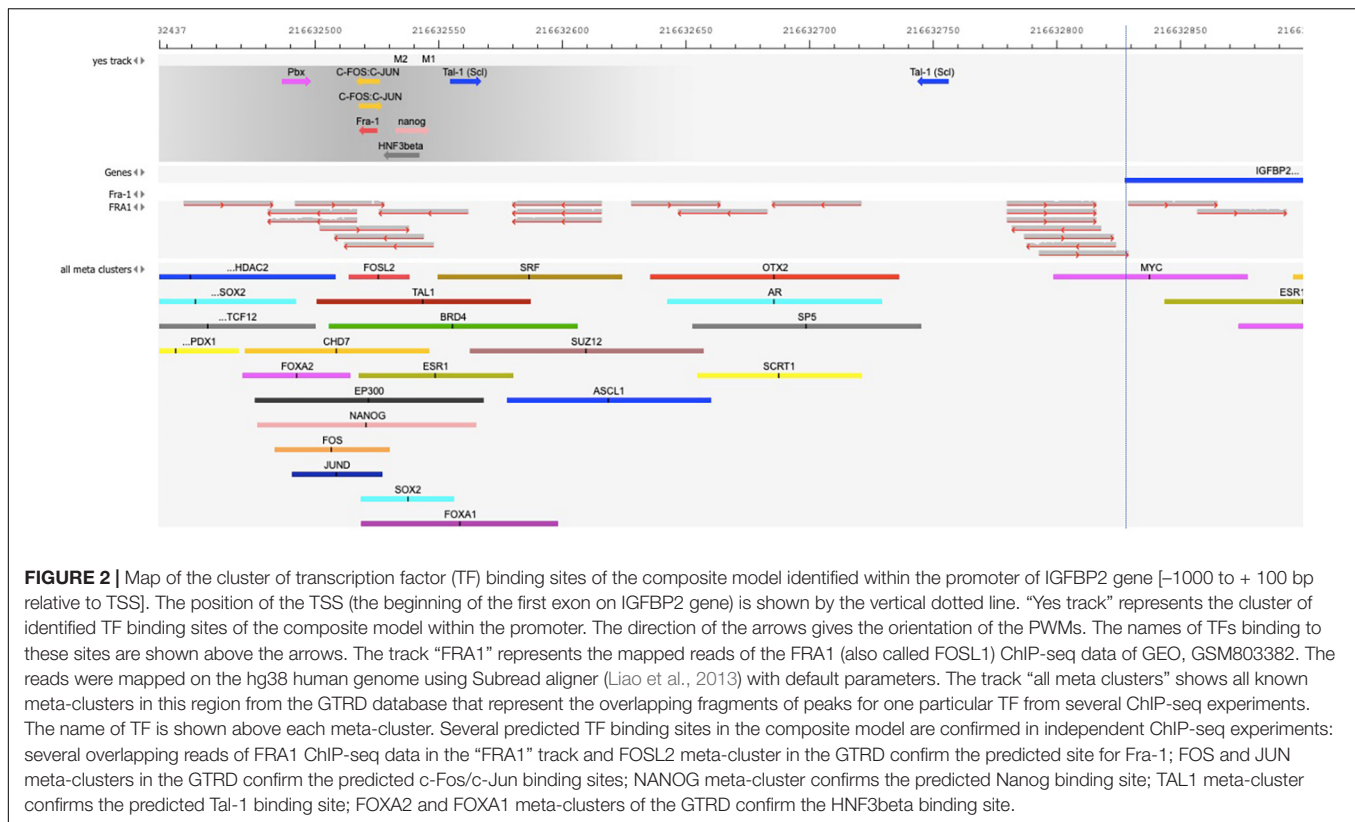
Validating the Expression of Master Regulators in Other Cohorts

The expression patterns of the master regulators identified above have been validated in two different cohorts: (A) TCGA-GBM microarray data (Grossman et al., 2016) and (B) GSE16011 (Gravendeel et al., 2009). The expression patterns were similar, and there is a significant upregulation of all master regulators

except for VEGFA (GSE16011: adjusted *p*-value = 0.069 and TCGA-GBM: adjusted *p*-value = 0.075) (Supplementary Tables 1-E, 1-F). The differential expression values are given in Table 4.

Validating the Master Regulators in the TCGA-GBM Cohort

The TCGA-GBM microarray data containing 271 STS and 49 LTS is used to validate the above-identified drivers of short survival. The data is preprocessed and adjusted for batch effects



(Supplementary Figure 3), and a differential gene expression analysis is performed. Same cutoffs for log2FC and adjusted p -value are used. We identified 171 genes upregulated in STS of GBM (log2FC > 0.5 and adjusted p -value < 0.05) (full list in Supplementary Table 1-E). Forty-nine of them were in common between the GSE dataset and TCGA-GBM; the full differential gene expression analysis results are given in Supplementary Table 1-G. Composite models selected by the CMA algorithm across the two datasets were expected to vary. We identified a model that includes a set of 16 TFs (Supplementary Table 3) and 12 master regulators upstream of them (Supplementary Table 4) regulating the signal transduction and gene regulatory network in STS.

As a result, the TCGA-GBM dataset validates IGFBP2, AEBP1 (ACLP), and PDGFA as master regulators driving the dysregulated gene network in STS. We also found that binding sites for FRA-1 TF are statistically significantly enriched at the regulatory regions of the dysregulated genes including IGFBP2 in the TCGA-GBM cohort (Supplementary Table 5).

Impact of Master Regulators on Survival in GBM

Univariate survival analysis was used to study the impact of these master regulators and the TFs they regulate on the overall survival in GBM based on TCGA-RNA-seq data. Patients are split into non-overlapping 50% upper and lower quantiles. Additionally, cox regression for the univariate survival analysis is performed, and hazard ratio (HR) and corresponding p -values are shown in

Figure 4. Univariate survival Cox regression analysis on other microarray datasets is given in Supplementary Table 1-I. All master regulators were found to have a significant impact upon survival except VEGFA. FRA-1 (FOSL1) was found to have a significant HR.

Master Regulator Expression Patterns Across GBM Subtypes

Based on the regulatory landscape of GBM, there are four subtypes—classical, mesenchymal, proneural, and neural (Verhaak et al., 2010). There is a significant level of intertumoral as well as intra-tumoral heterogeneity within each of them (Verhaak et al., 2010; Bradshaw et al., 2016). Molecular subtypes of GBM in the GSE dataset is given in Supplementary Table 6. DEGs between STS and LTS within each subtype are given in Supplementary Table 7. The expression patterns of master regulators across subtypes and across survival groups are depicted as boxplot in Supplementary Figure 5. None of the master regulators were found to be significantly differentially expressed between survivor groups in any subtypes.

DISCUSSION

Gene regulatory networks represent the causal regulatory relationships between TFs and their gene targets, which enables us to discover dysregulated genes in certain biological states (Marbach et al., 2012). Comparative studies of STS and LTS of

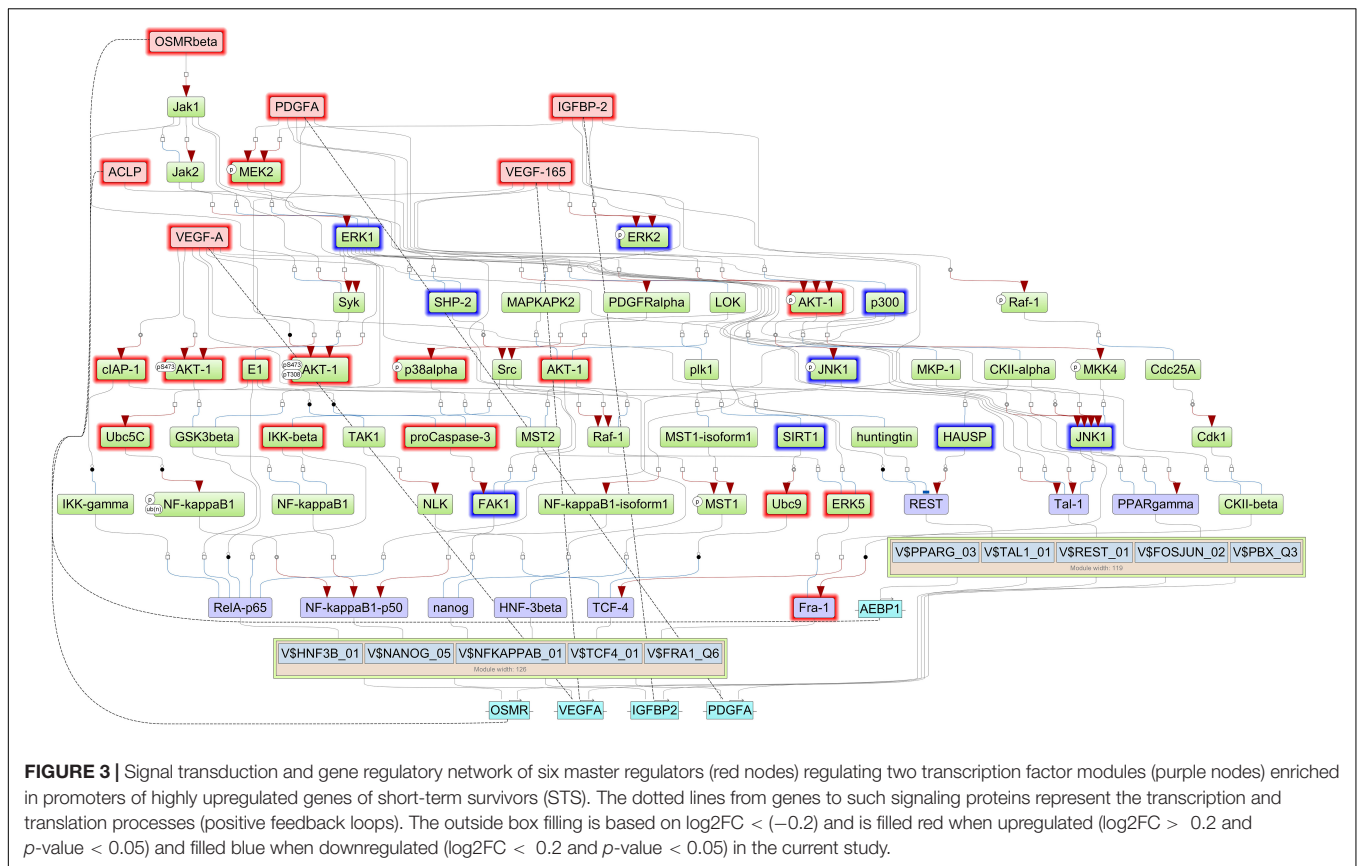


FIGURE 3 | Signal transduction and gene regulatory network of six master regulators (red nodes) regulating two transcription factor modules (purple nodes) enriched in promoters of highly upregulated genes of short-term survivors (STS). The dotted lines from genes to such signaling proteins represent the transcription and translation processes (positive feedback loops). The outside box filling is based on $\log_2FC < -0.2$ and is filled red when upregulated ($\log_2FC > 0.2$ and $p\text{-value} < 0.05$) and filled blue when downregulated ($\log_2FC < 0.2$ and $p\text{-value} < 0.05$) in the current study.

GBM showed that gene expression programs executed across survival groups vary significantly. In the light of these findings, we sought to apply an upstream analysis approach to gain an insight about gene regulatory networks driving the short survival.

In the promoter analysis, we identified a set of 12 TFs in composite clusters that are enriched in the promoter regions of dysregulated genes in STS (upregulated in STS). For several of these TFs, a connection to GBM has previously been established. The TFs NANOG and REST are critical for self-renewal and maintenance of oncogenic signatures in glioblastoma stem-like cells (Kamal et al., 2012; Bradshaw et al., 2016); PPARG has emerged as a promising therapeutic target as its agonists increased median survival in GBM patients (Ellis and Kurian, 2014); NF- κ B is implicated in several processes like invasion, epithelial–mesenchymal transition (Yamini, 2018), resistance to radiotherapy (Avci et al., 2020), and maintenance of cancer stem-like cells (da Hora et al., 2019); and FRA-1/FOSL1 has been reported to be important in maintenance/progression of malignant glioma (Debinski and Gibo, 2005). FRA-1 along with JUN-B modulates a malignant feature of GBM by regulating the expression of the metalloproteinases like MMP-2 and MMP-9 (Kesari and Bota, 2011). Among these 12 TFs, we found that FRA-1 has a significant impact upon survival and has a higher expression in STS. Debinski and Gibo (2005) hypothesized that any AP1-stimulating signals like epidermal growth factor (EGF), leukemia inhibitory factor, OSMR, or FGF-2 can positively

regulate FRA-1. VEGF-D is regulated by FRA-1 (supporting the feedback loop found in our work) and is a known prognostic factor in other aggressive cancers (Debinski et al., 2001; Azar et al., 2014).

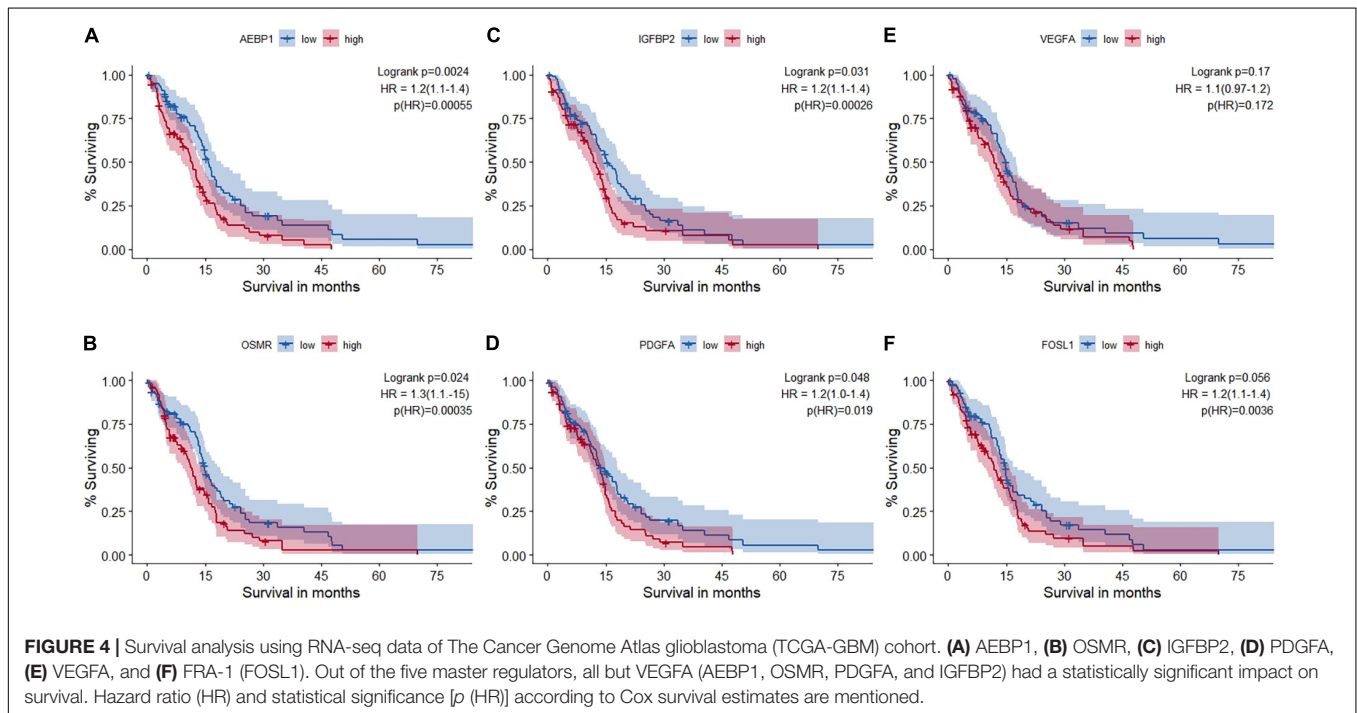
A graph analysis of the signal transduction network upstream of these TFs identified five potential master regulators that

TABLE 3 | Table of the master regulators identified, their description, \log_2FC in STS, and number of transcription factors regulated.

Molecule name	Gene description	HGNC gene symbol	Log2FC in STS	Number of TFs regulated
IGFBP2	Insulin-like growth factor binding protein 2	IGFBP2	1.149	9
ACL	AE-binding protein 1	AEBP1	0.782	9
VEGFA	Vascular endothelial growth factor A	VEGFA	0.778	9
VEGF165	Vascular endothelial growth factor A	VEGFA	0.778	9
OSMRbeta	Oncostatin M receptor	OSMR	0.634	8
PDGFA	Platelet-derived growth factor subunit A	PDGFA	0.529	9

TABLE 4 | Expression of the master regulators identified across survival groups (STS and LTS, respectively) and across three datasets (GSE, GSE16011 and TCGA-GBM microarray).

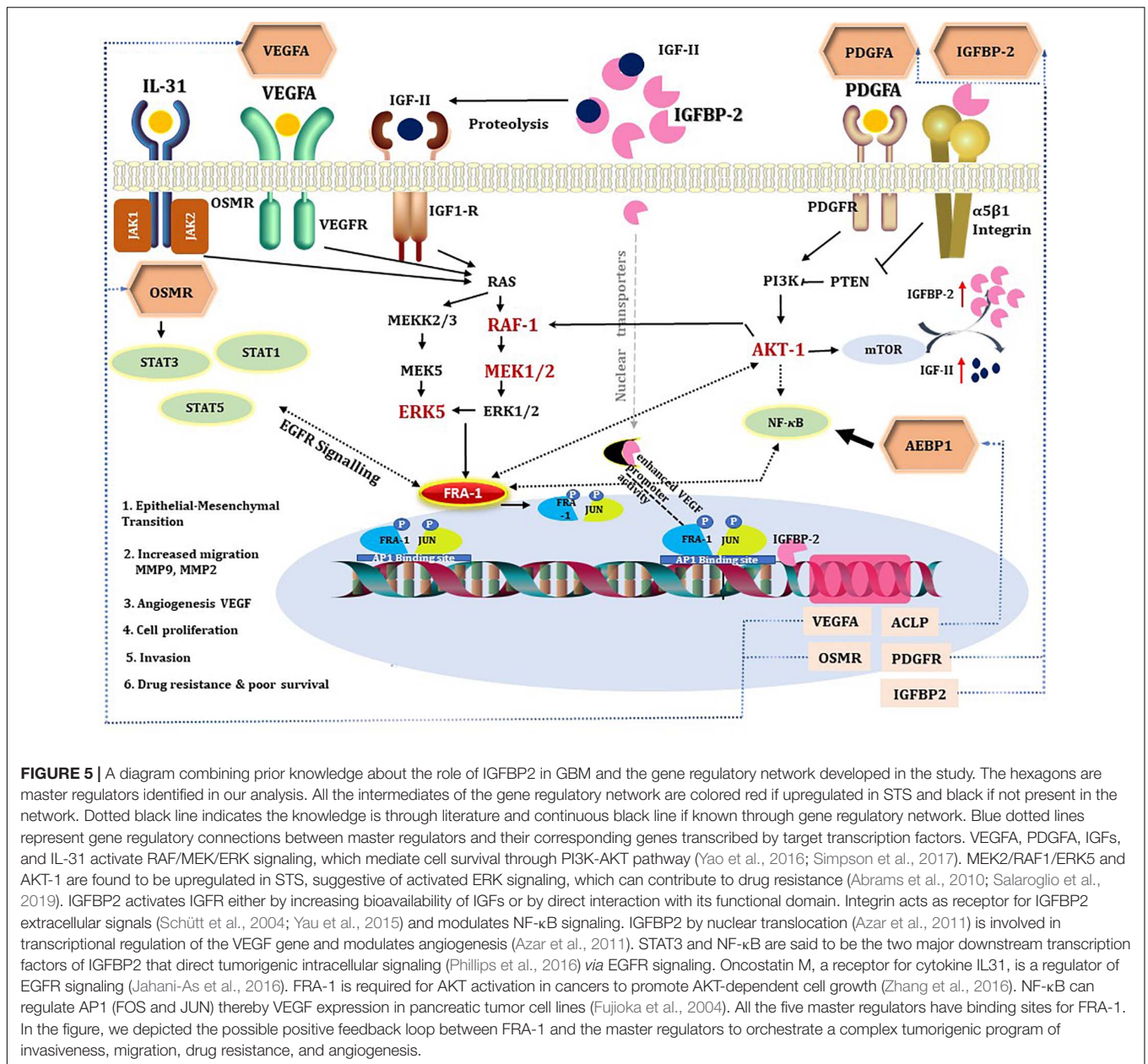
Master regulator	GSE		GSE16011		TCGA	
	Log2FC (STS vs LTS)	Adjusted p-value	Log2FC (STS vs LTS)	Adjusted p-value	Log2FC (STS vs LTS)	Adjusted p-value
IGFBP2	1.149	4.87E-05	2.030	4.598E-04	1.098	5.00E-06
AEBP1	0.782	7.75E-05	1.723	0.001	0.971	3.96E-06
PDGFA	0.529	4.55E-04	1.680	4.709E-09	0.825	2.07E-05
VEGFA	0.778	5.20E-04	0.884	0.069	0.500	0.0752
OSMR	0.634	8.65E-04	1.957	4.24E-05	0.486	0.0318



might explain gene dysregulation in STS, namely, insulin-like growth factor binding protein 2 (IGFBP2), VEGFA, its isoform VEGF165, PDGFA, OSMR, and AEBP1. All the identified master regulators were upregulated in STS, and their expression patterns were validated computationally in two other independent cohorts. We found that the expression of all master regulators, with the exception of VEGFA, was correlated with overall survival in the GBM patients. IGFBP2, AEBP1, and PDGFA master regulators driving short survival were validated as master regulators of short survival in the TCGA-GBM microarray cohort. Out of them, IGFBP2 had higher expression in STS. The IGFBP2 is said to be one of most potential glioma oncogenes and functions as a hub of oncogenic signaling pathways by regulating pro-tumorigenic signals of tumor initiation and progression. Earlier studies have suggested IGFBP2 to drive EMT and as a potential therapeutic target in mesenchymal GBM (Yamini, 2018; Liu et al., 2019). It is established that exogenous IGFBP2 promotes proliferation, invasion, and chemoresistance to temozolomide in glioma cells *via* integrin $\beta 1$ by promoting ERK phosphorylation and nuclear translocation (Schütt et al.,

2004; Yau et al., 2015). IGFBP2 is considered as one of the strongest biomarkers of aggressive behavior in GBM (Holmes, 2012; Phillips et al., 2016) and also a prognostic marker for survival (McDonald et al., 2007; Phillips et al., 2016).

Here, we propose that IGFBP2 can be a potential regulator of FRA-1 TF. IGFBP2-induced RAF/MAPK signaling can activate FRA-1 (**Figure 3**). It has been shown earlier that IGFBP2 and FRA-1 regulate transcription of VEGF (Debinski et al., 2001; Azar et al., 2011, 2014), which is the second most dysregulated master regulator in our network. Enhanced ERK signaling, triggered by these master regulators, may lead to mitogen-induced FRA-1 transcription (Adiseshaiah et al., 2005) as well as its protection from proteasomal degradation (Vial and Marshall, 2003). The gene regulatory network deduced here suggests that FRA-1 mediates a positive feedback loop where it activates transcription of master regulator genes in cooperation with other TFs, which in turn cause an increase in FRA-1 activity. Promoters of the genes of all five master regulators reported in the study contain potential binding sites for FRA-1. Experimental evidences that IGFBP2 can drive GBM invasion by enhancing MMP2 expression



(Wang et al., 2003) support our computational prediction of IGFBP2 as a therapeutic target. Hence, the gene regulatory networks proposed by our computational analysis suggest a novel molecular mechanism associated with GBM survival in which FRA-1 acts as a transcription regulator of IGFBP2. The study of Kesari and Bota (2011) confirmed our hypothesis that IGFBP2 can enhance GBM invasion *via* TF AP1 (FOS-JUN). Metalloproteinases like MMP-2/MMP-9 have been reported earlier to be regulated by FRA-1 in several cancers including GBM (Debinski and Gibo, 2005; Adiseshiaiah et al., 2008; Kimura et al., 2011; Prywes and Henckels, 2013). Taking these findings together, our work proposes that the regulation of IGFBP2 gene expression *via* AP1 (FOS-JUN) can be an important mechanism of GBM invasion. An overview of the gene regulatory network developed

in this work and supporting literature evidence is illustrated in Figure 5.

In summary, our work proposes a gene regulatory network associated with STS in GBM, which is regulated by five master regulators, namely, IGFBP2, VEGFA, PDGFA, OSMR, and AEBP1. Furthermore, these five master regulators may present biomarkers of GBM prognosis and/or as therapeutic targets for enhancing survival in GBM. This work also proposes a novel mechanism of gene dysregulation by IGFBP2 by modulating a key molecule of tumor invasiveness and progression—FRA-1 TF. All the genes encoding these five master regulators have binding sites for FRA-1 in their promoters. FRA-1 and the master regulators cooperate in a positive feedback loop to orchestrate a complex tumorigenic program leading to poor survival in GBM.

TABLE 5 | Statistics of datasets under study.

	Platform	Short-term survivors	Long-term survivors
GSE53733 (Reifenberger et al., 2014)	HU133 plus 2.0 arrays	16	23
GSE108474 (Gusev et al., 2018)	HU133 plus 2.0 arrays	97	35

The datasets with labels GSE¹ were collected from the GEO database.

MATERIALS AND METHODS

Data Collection

The genome-wide expression profiles based on Human Genome U133 plus 2.0 array and clinical information of patients with GBM were collected from the public repository of GEO database—GSE108474 (Gusev et al., 2018)² and GSE53733 (Reifenberger et al., 2014)³. The two datasets were pooled together leading to 113 and 58 samples corresponding to STS (survival <12 months) and LTS (survival >36 months) with GBM, respectively (Table 5). Duplicates were not removed. Sample information and cleaned datasets are given in GitHub.

Affymetrix Microarray Data Pre-processing

The raw data files (.CEL format) for GSE108474 and GSE53733 were collected from the GEO database—from here on called as GSE dataset. RMA algorithm is used in R (affy package) for background correction, quality check, and normalization to obtain log₂-transformed expression values (Gautier et al., 2004). Batch correction of the pooled expression data was performed using empirical Bayes framework (Leek et al., 2012). This batch-corrected file is used for further analysis. Multiple Affymetrix IDs were summarized to gene IDs by choosing the maximum out of the probe intensities of multiple probes belonging to a single gene. The final expression matrix comprised 21,526 probes and 171 samples.

Differential Gene Expression Analysis

The LIMMA method was applied to identify DEGs (Ritchie et al., 2015). It is an efficient tool that is stable even for experiments with small samples. A differential gene expression analysis of 171 samples of the GSE dataset was performed with Benjamini–Hochberg adjusted *p*-value. Nine hundred fifty-seven genes were significantly (adjusted *p*-value < 0.05) differentially expressed (DEGs). One hundred fifteen of them were significantly upregulated (adjusted *p*-value < 0.05 and log₂FC > 0.5) and 83 were significantly downregulated [adjusted *p*-value < 0.05 and log₂FC < (−0.5)].

²<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108474>

³<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53733>

Databases Used in the Study

Transcription factor binding sites in promoters and enhancers of DEGs were analyzed using known DNA-binding motifs described in the TRANSFAC[®] library, release 2019.3 (geneXplain GmbH, Wolfenbüttel, Germany)⁴ (Wingender et al., 1996). The master regulator search uses the TRANSPATH[®] database, release 2019.3 (geneXplain GmbH, Wolfenbüttel, Germany)⁵ (Krull et al., 2003). A comprehensive signal transduction network of human cells is built by the Genome Enhancer software based on reactions annotated in TRANSPATH[®]. The information about drugs corresponding to identified drug targets and clinical trials references were extracted from the HumanPSDTM database (Wingender et al., 2007), release 2020.2⁶. The Ensembl database build 99.38⁷ (Aken et al., 2016) was used for gene ID representation and GO⁸ (Ashburner et al., 2000) was used for functional classification of the studied gene set.

Functional Annotation

To explore the biological importance of gene signatures, a gene set enrichment analysis is performed. All the adjusted *p*-value significant genes were used. GSEA is an efficient method to determine whether the genes of interest show statistically significant enrichment between different biological states. GO enrichments for cellular component, biological process, and molecular functions were performed. To investigate the top enriched ontology terms, 1,000 random permutations were done and an adjusted *p*-value cutoff of 0.05 is used. The dysregulated gene network enrichment also gives a useful insight about known disease signatures (Subramanian et al., 2005). The hallmark gene set of MSigDB (Liberzon et al., 2011) defines specific biological states or processes. Enrichment analysis is performed in R using DOSE package (Yu et al., 2015). PANTHER pathway enrichment of the identified TFs was performed using the EnrichR tool (Chen et al., 2013). TRANSPATH[®] (Krull et al., 2003) pathway enrichment was performed using the geneXplain platform.

Genome Enhancer Pipeline

The approaches mentioned above help us in understanding the impact of the DEGs in GBM biology. To understand the reason behind this dysregulation, the genome enhancer pipeline of geneXplain is used. The genome enhancer is a multi-omics analysis service (see text footnote 1) that is built using an open-source programming environment BioUML (Kolpakov et al., 2019)⁹ and incorporates an automated pipeline for the previously published “upstream analysis” (Koschmann et al., 2015; Boyarskikh et al., 2018) and the advanced approach “walking pathways” (Kel et al., 2019). Significantly upregulated genes in STS were used in this workflow.

The workflow works in 2 steps.

⁴<https://genexplain.com/transfac>

⁵<https://genexplain.com/transpath>

⁶<https://genexplain.com/humanpsd>

⁷<http://www.ensembl.org>

⁸<http://geneontology.org>

⁹www.biouml.org

A. Analysis of enriched transcription factor binding sites and composite modules

Binding of TFs to the specific sites in promoters and enhancers is the key to the transcriptional regulation of genes. Identifying clusters of binding sites for TFs (composite modules) in the upstream regulatory regions [−1,000 bp upstream of transcription start site (TSS)] of the genes of interest is a determining step to understand the gene regulatory mechanism (composite regulatory modules) (Kel-Margoulis et al., 2002).

We use the CMA (Waleev et al., 2006) to detect such potential enhancers, as targets of multiple TFs bound to the regulatory regions of the genes of interest. The TFs are ranked based on (a) the yes/no ratio: given a set of promoter sequences of dysregulated genes, denoted as a yes set, and promoter sequences of unchanged genes under the same experimental condition, denoted as a no set, motifs are considered important if they have a high yes/no ratio, the ratio of motif occurrences per promoter in yes and no sets, and a statistically significant enrichment of occurrences in yes sequences assessed by the binomial *p*-value. (b) A regulatory score, which is a measure of involvement of a TF in controlling the expression of genes that encode master regulators. CMA identifies the TFs that, through their cooperation, provide a synergistic effect and thus have a great influence on the gene regulation process.

B. Finding master regulators in networks

The second step involves the signal transduction database TRANSPATH® and special graph search algorithms to identify common regulators of the revealed TFs. These master regulators appear to be the key candidates for therapeutic targets as they have a master effect on the regulation of intracellular pathways that activate the pathological process of our study. Master regulators regulating the TFs revealed in step A are ranked based on (a) logFC, (b) CMA score, which signifies how strong is the potential for this gene to be regulated by TFs of interest, and (c) master regulator score, which signifies how strong is the potential of this gene product to regulate the activity of those TFs. Selected master regulators can also be visualized and with the possibility to map the logFC and *p*-value on the created regulatory network.

Validation of Observed Gene Signatures

The raw microarray data of 560 TCGA-GBM samples were downloaded from TCGA legacy. The GSE16011 raw.CEL data

was downloaded from the GEO repository. Both raw datasets were processed and analyzed independently following same steps as mentioned earlier. These two datasets are used to observe and validate the expression pattern of master regulators across the two survival groups (see **Table 6**). GSE16011 comprises of data generated at a single center and is used in several studies (Prasad et al., 2020), unlike TCGA. TCGA-GBM microarray data PCA plots are given **Supplementary Figure 3**, and no significant batch effects in the context of survival groups were found.

Validation of Master Regulators

The TCGA-GBM microarray data downloaded from TCGA legacy archive is processed in the same fashion as GSE. Similar cutoffs (log2FC and *p*-value) and parameters are used to identify enriched TFs and network analysis in order to understand drivers of gene regulatory networks in short survival.

Impact on Survival

Master regulators and their target TFs affect the whole regulatory network and therefore can have an independent impact on survival in GBM patients. Level 3 RNA-seq data and clinical data for 152 TCGA-GBM cohort is downloaded using the TCGAbiolinks package in R. Survival and survminer libraries in R were used to perform a univariate survival analysis. A univariate survival analysis was used to understand the impact of individual master regulator on survival in GBM with non-overlapping 50% upper and lower quantiles. Additionally, a univariate Cox regression for survival analysis was performed using the coxph function of the survival package to calculate the HR with *p*-value cutoff of 0.05 for significance.

CONCLUSION

In the work presented, we have identified candidate master regulators responsible for gene dysregulation in STS. These candidates have sufficient experimental evidence toward their role in GBM. Out of reported five master regulators, IGFBP2 is established as the most promising master regulator. Through the gene regulatory network analysis, we propose that IGFBP2 and FRA-1 are in a positive feedback loop that may lead to a pathological self-enhancing process responsible for poor survival in GBM.

DATA AVAILABILITY STATEMENT

The results of the analysis performed using the Genome Enhancer in geneXplain platform are available here: https://github.com/genexplain/Manasa_KP_et_al_IGFBP2_regulatory_networks_in_Glioblastoma.

AUTHOR CONTRIBUTIONS

AK involved in conceptualization, providing resources, supervision, and manuscript reviewing. MK had conceptualized

TABLE 6 | Statistics of the two validation datasets.

Datasets	Platform	Short-term survivors	Long-term survivors
GSE16011 (Gravendeel et al., 2009)	HU133 plus 2.0 arrays	93	16
TCGA-GBM microarray (Grossman et al., 2016)	HU133	271	49

the work, performed the data collection and data analysis, interpreted the results, and wrote the manuscript. DW supervised the data analysis pipeline and manuscript writing. EW and TB involved in supervision of work and in reviewing the draft. All authors contributed to the article and approved the submitted version.

FUNDING

This project had received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 766069 and by German Ministry of Education and Research (BMBF) e:Med project MyPathSem (031L0024).

REFERENCES

- Abrams, S. L., Steelman, L. S., Shelton, J. G., Wong, E. W. T., Chappell, W. H., Bäsecke, J., et al. (2010). The Raf/MEK/ERK pathway can govern drug resistance, apoptosis and sensitivity to targeted therapy. *Cell Cycle* 9, 1781–1791. doi: 10.4161/cc.9.9.11483
- Adisheshaiah, P., Peddakama, S., Zhang, Q., Kalvakolanu, D. V., and Reddy, S. P. (2005). Mitogen regulated induction of FRA-1 proto-oncogene is controlled by the transcription factors binding to both serum and TPA response elements. *Oncogene* 24, 4193–4205. doi: 10.1038/sj.onc.1208583
- Adisheshaiah, P., Vaz, M., Machireddy, N., Kalvakolanu, D. V., and Reddy, S. P. (2008). A Fra-1-dependent, matrix metalloproteinase driven EGFR activation promotes human lung epithelial cell motility and invasion. *J. Cell. Physiol.* 216, 405–412. doi: 10.1002/jcp.21410
- Aken, B. L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., et al. (2016). The ensembl gene annotation system. *Database* 2016:baw093. doi: 10.1093/database/baw093
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Athar, A., Füllgrabe, A., George, N., Iqbal, H., Huerta, L., Ali, A., et al. (2019). ArrayExpress update - From bulk to single-cell expression data. *Nucleic Acids Res.* 47, D711–D715. doi: 10.1093/nar/gky964
- Avci, N. G., Ebrahimzadeh-Pustchi, S., Akay, Y. M., Esquenazi, Y., Tandon, N., Zhu, J. J., et al. (2020). NF-κB inhibitor with Temozolomide results in significant apoptosis in glioblastoma via the NF-κB(p65) and actin cytoskeleton regulatory pathways. *Sci. Rep.* 10:13352. doi: 10.1038/s41598-020-70392-5
- Azar, W. J., Azar, S. H. X., Higgins, S., Hu, J. F., Hoffman, A. R., Newgreen, D. F., et al. (2011). IGFBP-2 enhances VEGF gene promoter activity and consequent promotion of angiogenesis by neuroblastoma cells. *Endocrinology* 152, 3332–3342. doi: 10.1210/en.2011-1121
- Azar, W. J., Zivkovic, S., Werther, G. A., and Russo, V. C. (2014). IGFBP-2 nuclear translocation is mediated by a functional NLS sequence and is essential for its pro-tumorigenic actions in cancer cells. *Oncogene* 33, 578–588. doi: 10.1038/onc.2012.630
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets - Update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193
- Bi, W. L., and Beroukhi, R. (2014). Beating the odds: extreme long-term survival with glioblastoma. *Neuro. Oncol.* 16, 1159–1160. doi: 10.1093/neuonc/nou166
- Boyerskikh, U., Pintus, S., Mandrik, N., Stelmashenko, D., Kiselev, I., Evshin, I., et al. (2018). Computational master-regulator search reveals mTOR and PI3K pathways responsible for low sensitivity of NCI-H292 and A427 lung cancer cell lines to cytotoxic action of p53 activator Nutlin-3. *BMC Med. Genomics* 11(Suppl. 1):12. doi: 10.1186/s12920-018-0330-5
- Bradshaw, A., Wickremsekera, A., Tan, S. T., Peng, L., Davis, P. F., and Itinteang, T. (2016). Cancer stem cell hierarchy in glioblastoma multiforme. *Front. Surg.* 3:21. doi: 10.3389/fsurg.2016.00021
- Cai, J., Chen, Q., Cui, Y., Dong, J., Chen, M., Wu, P., et al. (2018). Immune heterogeneity and clinicopathologic characterization of IGFBP2 in 2447 glioma samples. *Oncoimmunology* 7:e1426516. doi: 10.1080/2162402X.2018.1426516
- Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14:128. doi: 10.1186/1471-2105-14-128
- Chen, J. R., Yao, Y., Xu, H. Z., and Qin, Z. Y. (2016). Isocitrate dehydrogenase (IDH)1/2 mutations as prognostic markers in patients with glioblastomas. *Med. (United States)* 95:e2583. doi: 10.1097/MD.0000000000002583
- da Hora, C. C., Pinkham, K., Carvalho, L., Zinter, M., Tabet, E., Nakano, I., et al. (2019). Sustained NF-κB-STAT3 signaling promotes resistance to Smac mimetics in glioma stem-like cells but creates a vulnerability to EZH2 inhibition. *Cell Death Discov.* 5:72. doi: 10.1038/s41420-019-0155-9
- Das, P., Puri, T., Jha, P., Pathak, P., Joshi, N., Suri, V., et al. (2011). A clinicopathological and molecular analysis of glioblastoma multiforme with long-term survival. *J. Clin. Neurosci.* 18, 66–70. doi: 10.1016/j.jocn.2010.04.050
- De Vega, S., Iwamoto, T., and Yamada, Y. (2009). Fibulins: multiple roles in matrix structures and tissue functions. *Cell. Mol. Life Sci.* 66, 1890–1902. doi: 10.1007/s00018-009-8632-6
- Debinski, W., and Gibo, D. M. (2005). Fos-related antigen 1 modulates malignant features of glioma cells. *Mol. Cancer Res.* 3, 237–249. doi: 10.1158/1541-7786.MCR-05-0004
- Debinski, W., Slagle-Webb, B., Achen, M. G., Stacker, S. A., Tulchinsky, E., Gillespie, G. Y., et al. (2001). VEGF-D is an X-linked/AP-1 regulated putative onco-angiogen in human glioblastoma multiforme. *Mol. Med.* 7, 598–608. doi: 10.1007/bf03401866
- Ellis, H. P., and Kurian, K. M. (2014). Biological rationale for the use of PPAR γ agonists in glioblastoma. *Front. Oncol.* 4:52. doi: 10.3389/fonc.2014.00052
- Franceschi, S., Mazzanti, C. M., Lessi, F., Aretini, P., Carbone, F. G., La Ferla, M., et al. (2015). Investigating molecular alterations to profile short- and long-term recurrence-free survival in patients with primary glioblastoma. *Oncol. Lett.* 10:3599–3606. doi: 10.3892/ol.2015.3738
- Fujioka, S., Niu, J., Schmidt, C., Scwab, G. M., Peng, B., Uwagawa, T., et al. (2004). NF-κB and AP-1 connection: mechanism of NF-κB-dependent regulation of AP-1 activity. *Mol. Cell. Biol.* 24, 7806–7819. doi: 10.1128/mcb.24.17.7806-7819.2004
- Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315. doi: 10.1093/bioinformatics/btg405
- Gravendeel, L. A. M., Kouwenhoven, M. C. M., Gevaert, O., De Rooij, J. J., Stubbs, A. P., Duijm, J. E., et al. (2009). Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. *Cancer Res.* 69, 9065–9072. doi: 10.1158/0008-5472.CAN-09-2307
- Grossman, R. L., Heath, A. P., Ferretti, V., Varmus, H. E., Lowy, D. R., Kibbe, W. A., et al. (2016). Toward a shared vision for cancer genomic data. *N. Engl. J. Med.* 375, 1109–1112. doi: 10.1056/nejmp1607591

ACKNOWLEDGMENTS

It is the authors' pleasure to acknowledge Ravi Kumar Nadella who has read and given comments on every version of this paper and their colleague Philip Stegmaier for his expertise, discussions, and assistance in improving the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.670240/full#supplementary-material>

- Gusev, Y., Bhuvaneshwar, K., Song, L., Zenklusen, J. C., Fine, H., and Madhavan, S. (2018). Data descriptor: the REMBRANDT study, a large collection of genomic data from brain cancer patients. *Sci. Data* 5:180158. doi: 10.1038/sdata.2018.158
- Han, S., Meng, L., Han, S., Wang, Y., and Wu, A. (2014). Plasma IGFBP-2 levels after postoperative combined radiotherapy and chemotherapy predict prognosis in elderly glioblastoma patients. *PLoS One* 9:e93791. doi: 10.1371/journal.pone.0093791
- Holmes, K. M. (2012). *Elucidating the IGFBP2 Signaling Pathway in Glioma Development Elucidating the IGFBP2 Signaling Pathway in Glioma Development and Progression and Progression*. Available online at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations (accessed September 5, 2020).
- Iwadate, Y. (2016). Epithelial-mesenchymal transition in glioblastoma progression. *Oncol. Lett.* 11, 1615–1620. doi: 10.3892/ol.2016.4113
- Jahani-As, A., Yin, H., Soleimani, V. D., Haque, T., Luchman, H. A., Chang, N. C., et al. (2016). Control of glioblastoma tumorigenesis by feed-forward cytokine signaling. *Nat. Neurosci.* 19, 798–806. doi: 10.1038/nn.4295
- Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., et al. (2020). The reactome pathway knowledgebase. *Nucleic Acids Res.* 48, D498–D503. doi: 10.1093/nar/gkz1031
- Kamal, M. M., Sathyan, P., Singh, S. K., Zinn, P. O., Marisetty, A. L., Liang, S., et al. (2012). REST regulates oncogenic properties of glioblastoma stem cells. *Stem Cells* 30, 405–414. doi: 10.1002/stem.1020
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2020). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 49, D545–D551. doi: 10.1093/nar/gkaa970
- Kel-Margoulis, O. V., Kel, A. E., Reuter, I., Deineko, I. V., and Wingender, E. (2002). TRANSCompel \S : a database on composite regulatory elements in eukaryotic genes. *Nucleic Acids Res.* 30, 332–334. doi: 10.1093/nar/30.1.332
- Kel, A., Boyarskikh, U., Stegmaier, P., Leskov, L. S., Sokolov, A. V., Yevshin, I., et al. (2019). Walking pathways with positive feedback loops reveal DNA methylation biomarkers of colorectal cancer. *BMC Bioinformatics* 20:119. doi: 10.1186/s12859-019-2687-7
- Kel, A. E., Gößling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., and Wingender, E. (2003). MATCHTM: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* 31, 3576–3579. doi: 10.1093/nar/gkg585
- Kel, A., Voss, N., Jauregui, R., Kel-Margoulis, O., and Wingender, E. (2006). Beyond microarrays: finding key transcription factors controlling signal transduction pathways. *BMC Bioinformatics* 7:S13. doi: 10.1186/1471-2105-7-S2-S13
- Kesari, S., and Bota, D. A. (2011). Fos-related antigen-1 (Fra-1) is a regulator of glioma cell malignant phenotype. *Cancer Biol. Ther.* 11, 307–310. doi: 10.4161/cbt.11.3.14718
- Kimura, R., Ishikawa, C., Rokkaku, T., Janknecht, R., and Mori, N. (2011). Phosphorylated c-Jun and Fra-1 induce matrix metalloproteinase-1 and thereby regulate invasion activity of 143B osteosarcoma cells. *Biochim. Biophys. Acta Mol. Cell Res.* 1813, 1543–1553. doi: 10.1016/j.bbamcr.2011.04.008
- Kolmykov, S., Yevshin, I., Kulyashov, M., Sharipov, R., Kondrakhin, Y., Makeev, V. J., et al. (2021). Gtrd: an integrated view of transcription regulation. *Nucleic Acids Res.* 49, D104–D111. doi: 10.1093/nar/gkaa1057
- Kolpakov, F., Akberdin, I., Kashapov, T., Kiselev, L., Kolmykov, S., Kondrakhin, Y., et al. (2019). BioUML: an integrated environment for systems biology and collaborative analysis of biomedical data. *Nucleic Acids Res.* 47, W225–W233. doi: 10.1093/nar/gkz440
- Koschmann, J., Bhar, A., Stegmaier, P., Kel, A., and Wingender, E. (2015). “Upstream Analysis”: an integrated promoter-pathway analysis approach to causal interpretation of microarray data. *Microarrays* 4, 270–286. doi: 10.3390/microarrays4020270
- Krex, D., Klink, B., Hartmann, C., von Deimling, A., Pietsch, T., Simon, M., et al. (2007). Long-term survival with glioblastoma multiforme. *Brain* 130, 2596–2606. doi: 10.1093/brain/awm204
- Krull, M., Voss, N., Choi, C., Pistor, S., Potapov, A., and Wingender, E. (2003). TRANSPATH \S : an integrated database on signal transduction and a tool for array analysis. *Nucleic Acids Res.* 31, 97–100. doi: 10.1093/nar/gkg089
- Lee, Y., Scheck, A. C., Cloughesy, T. F., Lai, A., Dong, J., Farooqi, H. K., et al. (2008). Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age. *BMC Med. Genomics* 1:52. doi: 10.1186/1755-8794-1-52
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi: 10.1093/bioinformatics/bts034
- Liao, Y., Smyth, G. K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41:e108. doi: 10.1093/nar/gkt214
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260
- Lindström, M. S. (2019). Expanding the scope of candidate prognostic marker IGFBP2 in glioblastoma. *Biosci. Rep.* 39:BSR20190770. doi: 10.1042/BSR20190770
- Liu, Y., Song, C., Shen, F., Zhang, J., and Song, S. W. (2019). IGFBP2 promotes immunosuppression associated with its mesenchymal induction and Fc γ RIIB phosphorylation in glioblastoma. *PLoS One* 14:e0222999. doi: 10.1371/journal.pone.0222999
- Majdalawieh, A. F., Massri, M., and Ro, H. S. (2020). AEBP1 is a novel oncogene: mechanisms of action and signaling pathways. *J. Oncol.* 2020:8097872. doi: 10.1155/2020/8097872
- Marbach, D., Costello, J. C., Küffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., et al. (2012). Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804. doi: 10.1038/nmeth.2016
- Martinho, O., and Reis, R. M. (2011). “Malignant gliomas: role of platelet-derived growth factor receptor A (PDGFRA),” in *Tumors of the Central Nervous System*, ed. M. Hayat (Dordrecht: Springer), 109–118. doi: 10.1007/978-94-007-0344-5_12
- Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., et al. (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* 34, D108–D110. doi: 10.1093/nar/gkj143
- McDonald, K. L., O’Sullivan, M. G., Parkinson, J. F., Shaw, J. M., Payne, C. A., Brewer, J. M., et al. (2007). IQGAP1 and IGFBP2. *J. Neuropathol. Exp. Neurol.* 66, 405–417. doi: 10.1097/nen.0b013e31804567d7
- Phillips, L. M., Zhou, X., Cogdell, D. E., Chua, C. Y., Huisinga, A., Hess, K. R., et al. (2016). Glioma progression is mediated by an addiction to aberrant IGFBP2 expression and can be blocked using anti-IGFBP2 strategies. *J. Pathol.* 239, 355–364. doi: 10.1002/path.4734
- Prasad, B., Tian, Y., and Li, X. (2020). Large-scale analysis reveals gene signature for survival prediction in primary glioblastoma. *Mol. Neurobiol.* 57, 5235–5246. doi: 10.1007/s12035-020-02088-w
- Prywes, R., and Henckels, E. (2013). Fra-1 regulation of Matrix Metalloproteinase-1 (MMP-1) in metastatic variants of MDA-MB-231 breast cancer cells. *F1000Research* 2:229. doi: 10.12688/f1000research.2-229.v1
- Reifenberger, G., Weber, R. G., Riehmer, V., Kaulich, K., Willscher, E., Wirth, H., et al. (2014). Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. *Int. J. Cancer* 135, 1822–1831. doi: 10.1002/ijc.28836
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Salaroglio, I. C., Mungo, E., Gazzano, E., Kopecka, J., and Riganti, C. (2019). ERK is a pivotal player of chemo-immune-resistance in cancer. *Int. J. Mol. Sci.* 20:2505. doi: 10.3390/ijms20102505
- Schütt, B. S., Langkamp, M., Rauschnabel, U., Ranke, M. B., and Elminger, M. W. (2004). Integrin-mediated action of insulin-like factor binding protein-2 in tumor cells. *J. Mol. Endocrinol.* 32, 859–868. doi: 10.1677/jme.0.0320859
- Scott, J. N., Rewcastle, N. B., Brasher, P. M. A., Fulton, D., MacKinnon, J. A., Hamilton, M., et al. (1999). Which glioblastoma multiforme patient will become a long-term survivor? A population-based study. *Ann. Neurol.* 46, 183–188. doi: 10.1002/1531-8249
- Shinawi, T., Hill, V. K., Krex, D., Schackert, G., Gentle, D., Morris, M. R., et al. (2013). DNA methylation profiles of long- and short-term glioblastoma survivors. *Epigenetics* 8, 149–156. doi: 10.4161/epi.23398
- Simpson, A., Petnga, W., Macaulay, V. M., Weyer-Czernilofsky, U., and Bogenrieder, T. (2017). Insulin-like growth factor (IGF) pathway targeting in cancer: role of the IGF axis and opportunities for future combination studies. *Target. Oncol.* 12, 571–597. doi: 10.1007/s11523-017-0514-5

- Sonoda, Y., Kumabe, T., Watanabe, M., Nakazato, Y., Inoue, T., Kanamori, M., et al. (2009). Long-term survivors of glioblastoma: clinical features and molecular analysis. *Acta Neurochir. (Wien)*. 151, 1349–1358. doi: 10.1007/s00701-009-0387-1
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Thomas, P. D., Thomas, P. D., Campbell, M. J., Kejariwal, A., Al, E., Thomas, P. D., et al. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* 13, 2129–2141. doi: 10.1101/gr.772403
- Verhaak, R. G. W., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., et al. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *PLoS One*. 5, e12217. doi: 10.1016/j.ccr.2009.12.020
- Vial, E., and Marshall, C. J. (2003). Elevated ERK-MAP kinase activity protects the FOS family member FRA-1 against proteasomal degradation in colon carcinoma cells. *J. Cell Sci.* 116, 4957–4963. doi: 10.1242/jcs.00812
- Waleev, T., Shtokalo, D., Konovalova, T., Voss, N., Cheremushkin, E., Stegmaier, P., et al. (2006). Composite module analyst: identification of transcription factor binding site combinations using genetic algorithm. *Nucleic Acids Res.* 34:W541. doi: 10.1093/nar/gkl342
- Wang, H., Wang, H., Shen, W., Huang, H., Hu, L., Ramdas, L., et al. (2003). *Insulin-like Growth Factor Binding Protein 2 Enhances Glioblastoma Invasion by Activating Invasion-enhancing Genes 1*. Available online at: www.mdanderson.org/genomics (accessed January 13, 2021).
- Wen, P. Y., and Kesari, S. (2008). Malignant gliomas in adults. *N. Engl. J. Med.* 359, 492–507. doi: 10.1056/NEJMra0708126
- Wingender, E., Dietze, P., Karas, H., and Knüppel, R. (1996). TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* 24, 238–241. doi: 10.1093/nar/24.1.238
- Wingender, E., Hogan, J., Schacherer, F., Potapov, A. P., and Kel-Margoulis, O. (2007). Integrating pathway data for systems pathology. *In Silico Biol.* 7(2 Suppl), S17–S25.
- Xu, C., Wu, X., and Zhu, J. (2013). VEGF promotes proliferation of human glioblastoma multiforme stem-like cells through VEGF receptor 2. *Sci. World J.* 2013:417413. doi: 10.1155/2013/417413
- Yamini, B. (2018). NF- κ B, Mesenchymal differentiation and glioblastoma. *Cells* 7:125. doi: 10.3390/cells7090125
- Yao, X., Sun, S., Zhou, X., Guo, W., and Zhang, L. (2016). IGF-binding protein 2 is a candidate target of therapeutic potential in cancer. *Tumor Biol.* 37, 1451–1459. doi: 10.1007/s13277-015-4561-1
- Yau, S. W., Azar, W. J., Sabin, M. A., Werther, G. A., and Russo, V. C. (2015). IGFBP-2 - taking the lead in growth, metabolism and cancer. *J. Cell Commun. Signal.* 9, 125–142. doi: 10.1007/s12079-015-0261-2
- Yu, G., Wang, L.-G., Yan, G.-R., and He, Q.-Y. (2015). DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 31, 608–609. doi: 10.1093/bioinformatics/btu684
- Zhang, X., Wu, J., Luo, S., Lechler, T., and Zhang, J. Y. (2016). FRA1 promotes squamous cell carcinoma growth and metastasis through distinct AKT and c-Jun dependent mechanisms. *Oncotarget* 7:34371–34383. doi: 10.18632/oncotarget.9110
- Zhang, X., Zhang, W., Cao, W. D., Cheng, G., and Zhang, Y. Q. (2012). Glioblastoma multiforme: molecular characterization and current treatment strategy (Review). *Exp. Ther. Med.* 3, 9–14.

Conflict of Interest: MK, DW, and TB are from the Department of Medical Bioinformatics, University Medical Center Göttingen. MK, AK, and EW are employees of geneXplain GmbH.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Kalya, Kel, Wlochowitz, Wingender and Beißbarth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.