

RESEARCH

Open Access



Developing the novel bioinformatics algorithms to systematically investigate the connections among survival time, key genes and proteins for Glioblastoma multiforme

Yujie You¹, Xufang Ru², Wanjing Lei¹, Tingting Li³, Ming Xiao¹, Huiru Zheng⁴, Yujie Chen^{2*} and Le Zhang^{1*}

From The 18th Asia Pacific Bioinformatics Conference
Seoul, Korea. 18-20 August 2020

* Correspondence: yujiechen6886@foxmail.com; zhangle06@scu.edu.cn

²Department of Neurosurgery, Southwest Hospital, Third Military Medical University, Chongqing, P.R. China

¹College of Computer Science, Sichuan University, Chengdu 610065, China

Full list of author information is available at the end of the article

Abstract

Background: Glioblastoma multiforme (GBM) is one of the most common malignant brain tumors and its average survival time is less than 1 year after diagnosis.

Results: Firstly, this study aims to develop the novel survival analysis algorithms to explore the key genes and proteins related to GBM. Then, we explore the significant correlation between *AEBP1* upregulation and increased *EGFR* expression in primary glioma, and employ a glioma cell line LN229 to identify relevant proteins and molecular pathways through protein network analysis. Finally, we identify that *AEBP1* exerts its tumor-promoting effects by mainly activating mTOR pathway in Glioma.

Conclusions: We summarize the whole process of the experiment and discuss how to expand our experiment in the future.

Keywords: Glioblastoma multiforme (GBM), Survival analysis, Adipocyte enhancer-binding protein 1 (AEBP1), Data mining, mTOR pathway

Background

Glioblastoma multiforme (GBM) is one of the most common malignant brain tumors and its average survival time is less than 1 year after diagnosis [1]. The occurrence of GBM is considered as a complicated biological phenomenon with multiple external simulating factors, genes and stages. The major challenge in the field is to translate the almost unique progress in deciphering the highly complex molecular genetic nature of GBM into advances that allow for better prognosis rate and survival of affected patients [2].



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Currently, cancer researchers usually employ the survival analysis to explore genes that are closely related to GBM [2–5]. For example, Xia et al. [2] developed a CoxSisLasso survival analysis algorithm to identify gene signature for GBM by processing $P \gg N$ (the dimension of the factors P are less than the number of samples N) type of data [2]. However, since CoxSisLasso survival analysis algorithm does not consider the interactions among genes, its predictive accuracy is very limited. Therefore, we propose our first research question: Can we develop a more efficient survival analysis algorithm to explore the gene signature of GBM for $P \gg N$ type of data?

Secondly, current commonly used genomics [6] or proteomics [7] biotechnologies have limitations in quantitatively measuring gene expression and protein contents, which prevent us from developing the precise mathematical model. Thus, we propose our second research question: In order to explore the key proteins by using quantitative gene/protein data, can we employ cutting-edge biotechnologies, such as the CRISPR (clustered regularly interspaced short palindromic repeat sequences) [7–9] to knock out the key gene and then use RPPA (reverse phase protein arrays) [10–15] to high-throughput screen the corresponding protein data?

Thirdly, although a few GBM studies ([2, 3, 16–22]) considered the multi-scale data from intracellular, cellular and tissue scales, they neither develop the high efficient computational biology algorithm nor use experimental data to validate their findings in genomic and proteomic level. Thus, we propose our third research question: can we build a precise multi-scale mathematical model that can be used to understand the origin of the GBM from systematic view?

This study aimed to address the above research questions and the following novel approaches have been taken: (1) Firstly, we improved previous additive survival analysis algorithm (CoxSisLasso) [2, 9, 23] to investigate the key genes for GBM survival by considering genes' interaction as well as used immune-chemistry experiments and TCGA data [24] to validate the findings. (2) Secondly, in order to study the effect of key gene on the expression of which proteins, we employed the CRISPR [7–9] and RPPA [10–15] biotechnologies to knock out the explored key gene and obtain the quantitative gene/protein data. (3) Thirdly, we used the aforementioned quantitative gene/protein data to develop a precise multi-scale mathematical model to find key proteins, which can be used to investigate the origin of the GBM from the genetic, protein and tissue level starting from explored key gene as well as use the related experiments to evaluate our findings.

In summary, this study has developed an efficient survival analysis algorithm to identify GBM related *Adipocyte enhancer-binding protein 1 (AEBP1)* gene, and then used the related quantitative gene/protein data to explore the key proteins and molecular mechanism for GBM. Results show significant correlation between *AEBP1* upregulation and increased *EGFR* expression in primary glioma. A glioma cell line LN229 was used to identify major protein players and molecular pathways through RPPA analysis after *AEBP1* overexpression and *AEBP1* knockdown. We reveal that *AEBP1* exerts its tumor-promoting effects by mainly activating mTOR pathway in Glioma.

Methods

The pipeline of the study is illustrated in Fig. 1.

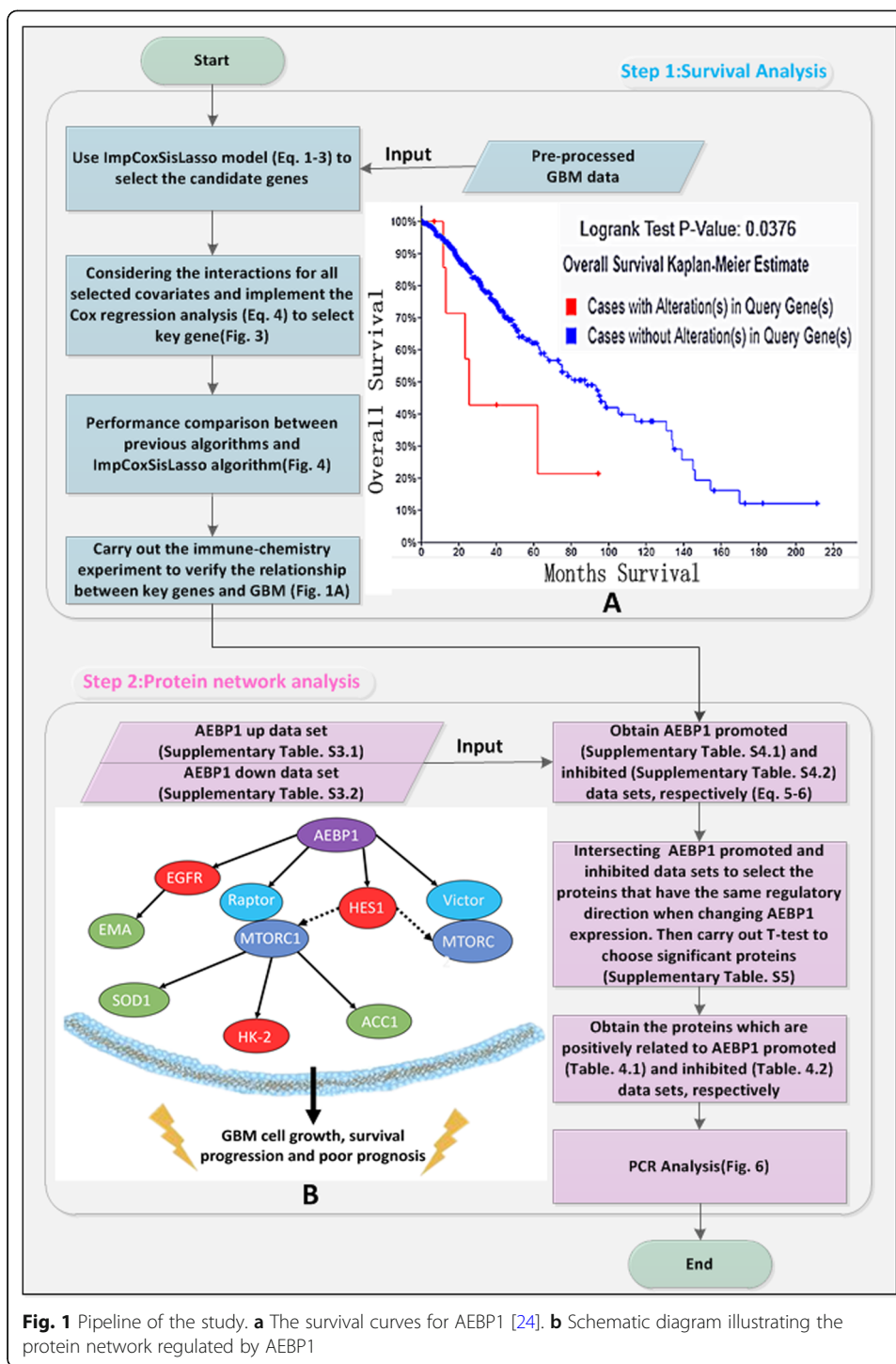


Fig. 1 Pipeline of the study. **a** The survival curves for AEBP1 [24]. **b** Schematic diagram illustrating the protein network regulated by AEBP1

Survival data

We use a multi-study microarray database of GBM expression profiles ($n = 227$) from the Georgetown Database of Cancer G-DOC [2], based on the Affymetrix U133 plus 2.0 GeneChip microarray platform. The original microarray datasets are normalized and preprocessed by R software package. After preprocessing step, there are 227 samples and 54,675 genes left in the data matrix. Next, the interquartile range (IQR)

threshold is employed to screen out the genes with small variance value. After that, there are only 227 samples and 10,992 genes left in the GBM gene expression and survival time data matrix.

Improved survival analysis algorithm (ImpCoxSisLasso)

We extend our previously well-developed CoxSisLasso strategy [2] by considering the gene interactions to develop ImpCoxSisLasso algorithm(Eq. 1–4).

$$\hat{\beta}_{LASSO} = \arg \min \left\{ - \sum_{k \in D} [x_k^T \beta - \log(\sum_{j \in R_k} \exp(x_j^T \beta))] + \lambda \sum_{j=1}^p |\beta_j| \right\} \tag{1}$$

where β is an unknown p -dimensional regression coefficient vector and x_i is a vector of potential predictors for the i^{th} individual. Based on the samples, $\hat{\beta}_{LASSO}$ is the LASSO [2] estimator of the unknown parameter coefficients β . D is the set of indices of the events and R_k denotes the set of indices of the individuals at risk in time t_k . k and i represent the index for the set D and set R_k , respectively. The tuning parameter λ is used to control the sparsity of the estimator.

$$\hat{\beta}_m = \arg \max_{\beta_m} \left\{ \sum_{k \in D} [x_{k,C_0}^T \beta_{C_0} + x_{k,m} \beta_m - \log(\sum_{j \in R_k} \exp(x_{j,C_0}^T \beta_{C_0} + x_{j,m} \beta_m))] \right\} \tag{2}$$

where the index of selected covariates for the nonzero components in $\hat{\beta}_{LASSO}$ is denoted by C_0 and each remaining covariate except C_0 is denoted by x_m , where $m \in \{1, 2, \dots, p\}$. We choose the significant covariates with P -Value smaller than a threshold value, $1/p$ for example.

$$\min_{\beta_{\Theta}} \left\{ - \sum_{k \in D} [x_{k,\Theta}^T \beta_{\Theta} - \log(\sum_{j \in R_k} \exp(x_{j,\Theta}^T \beta_{\Theta}))] + \lambda \sum_{j \in \beta_{\Theta}} |\beta_j| \right\} \tag{3}$$

where Θ is the collection of the augmented selected predictors $C_0 \cup C_1$ with C_0 denoting the index of selected covariates with Lasso, and C_1 denoting the chosen covariates by Eq. 2.

$$\max_{\beta_C, \beta_{i_1, i_2}} \sum_{k \in D} \left[- \log \left(\sum_{j \in R_k} \exp \left(x_{j,C}^T \beta_C + \sum_{i_1 \neq i_2, i_1, i_2 \in C} x_{ki_1} x_{ki_2} \beta_{i_1, i_2} \right) \right) \right] \tag{4}$$

where C denotes the index of finally selected covariates by Eq. 3. and the product $x_{i_1} x_{i_2}$ with $i_1 \neq i_2$, i_1 and $i_2 \in C$ denote the interactions between the selected genes. Here, i_1 and i_2 represent the index for the interactions' terms. Then, Fig. 2 describes ImpCoxSisLasso algorithm as the following.

```

1:  Input: gene data  $G$  with  $X$  vectors and  $p$  dimensions
2:   $C_0 \leftarrow \text{Lasso}(G)$  Obtain the index of selected gene data by Eq.1
3:  for each  $X_m$  in  $G$  and  $m \notin C_0$  do by Eq.2
       $(\text{Coef}_m, P\text{-Value}_m) \leftarrow \text{CoxRegression}(X_m)$ 
    end for
    for each index  $i$  in  $m$ 
      if  $P\text{-Value}_i \leq 1/p$ 
         $C_1 \leftarrow i$ 
      end if
    end for
4:   $C \leftarrow \text{Lasso}(X_{C_0 \cup C_1})$  Obtain the index of selected gene data by Eq.3
5:   $(\text{Coef}_n, P\text{value}_n) \leftarrow \text{CoxRegression}(X_C \setminus (X_{i_1} * X_{i_2}))_{i_1 \neq i_2, i_1, i_2 \in C}$  by Eq.4
    for each index  $i$  in  $n$ 
      if  $P\text{value}_i \leq 0.05$  Obtain the name of selected gene
         $\text{ImpCoxSisLassoGene} \leftarrow \text{GeneName}(i)$ 
      end if
    end for
6:  Output:  $\text{ImpCoxSisLassoGene}$ 

```

Fig. 2 The ImpCoxSisLasso algorithm

Immune chemistry experiment

Immunofluorescence staining was performed on GBM tissue sections as previously described. The specimens were rapidly isolated and post fixed in 4% PFA for 24 h and then soaked in 40% sucrose for 1 day. GBM tissue sections (10 μm) were obtained using a cryostat (Leica CM3050S-3-1-1, Bannockburn, IL, USA) and permeabilized with 0.3% Triton X-100 in PBS for 30 min. Sections were blocked with 5% donkey serum for 1 h and slices were incubated with a solution of 0.6 mg/ml diaminobenzidine and 0.05% H_2O_2 for 2 min. After that, incubation was terminated with three 10-min phosphate-buffered saline washes. Finally, slices were mounted onto gelatin-coated slides and dried overnight before placing a coverslip on them. Representative sections for each GBM tissue were then photographed.

Cell lines, plasmids

The LN229 cell line was purchased from ATCC [25]. The *pcDNA-AEBP1* Plasmid was provided by Professor Ro from Dalhousie University, Canada [26]. Plasmid for *sgRNA-AEBP1* was constructed by inserting two guide RNA from exon 12 of *AEBP1* into pPB-Cas9-puro at the cloning site. We named the construct with insert as *pPB-Cas9-puro-AEBP1-f1/r4*.

AEBP1 overexpression experiment

Cell line LN229 was seeded at 4×10^5 cells/well in a 60 mm plate and incubated in standard culture medium at 37°C overnight. Culture plates designated for pcDNA vector or pcDNA-AEBP1 transfection were done in triplicates. Transfection was done the next day using Fugene 6 transfection reagent from Roche. We used Reagent to DNA ratio of 3:1 during the transfection. Cells were harvested 48 h after transfection. Growing Cells in culture plates were trypsinized, spun down and washed with PBS for

two times. Cell pellet was stored at -80°C for RPPA analysis. Cells from one plate of vector transfection and another plate of pcDNA-AEBP1 transfection were lysed for Western blotting analysis to confirm the expression of *AEBP1* at the same time of RPPA cell pellet preparation.

AEBP1 CRISPR-cas9 knockdown experiment

Cell line LN229 was seeded at 4×10^5 cells/well in a 60 mm plate as above. Culture plates designated for vector or sgRNA-AEBP1 transfection were done in triplicates. Transfection was performed the next day after seeding using Fugene 6 transfection reagent from Roche and followed the same procedure as above. 48 h after transfection, puromycin was added to each culture dish at a final concentration of $1.5\mu\text{g/ml}$. Transfected LN229 cells survived puromycin selection and grew well. LN229 cell pellet was prepared in the same way as above and stored at -80°C for RPPA analysis. RPPA processing of frozen cell pellets: Frozen cell pellets were submitted to the RPPA core facility at M.D. Anderson Cancer Center. At the RPPA core, protein was extracted by applying RPPA lysis buffer to the cell pellets. Protein lysates were serially-diluted in lysis buffer and printed on nitrocellulose-coated RPPA slides. Slides were incubated with around 300 validated primary antibodies followed by binding with corresponding Biotinylated secondary antibodies and Avidin-Biotinylated Peroxidase (Vectastain Elite ABC kit, Vector Lab). Signals were detected by DAB colorimetric reaction. Signals on the slides were scanned and quantified as per protocol at the core facility.

Polymerase chain reaction (PCR)

Total RNA was extracted from cultured cells using Trizol reagent (Invitrogen, Carlsbad, CA, USA). Isolated RNA was reverse-transcribed into cDNA using a cDNA synthesis kit (Vazyme, Jiangsu, China), in accordance with the manufacturer's protocols. qPCR was performed using synthetic primers and SYBR Green (Thermo, Rockford, IL, USA) with an IQ5 Detection System. After incubating at 50°C for 2 min and 95°C for 10 min, the samples were subjected to 40 cycles of 95°C for 15 s and 60°C for 1 min. The sequences of the primers specific for target genes are listed in Table 1.

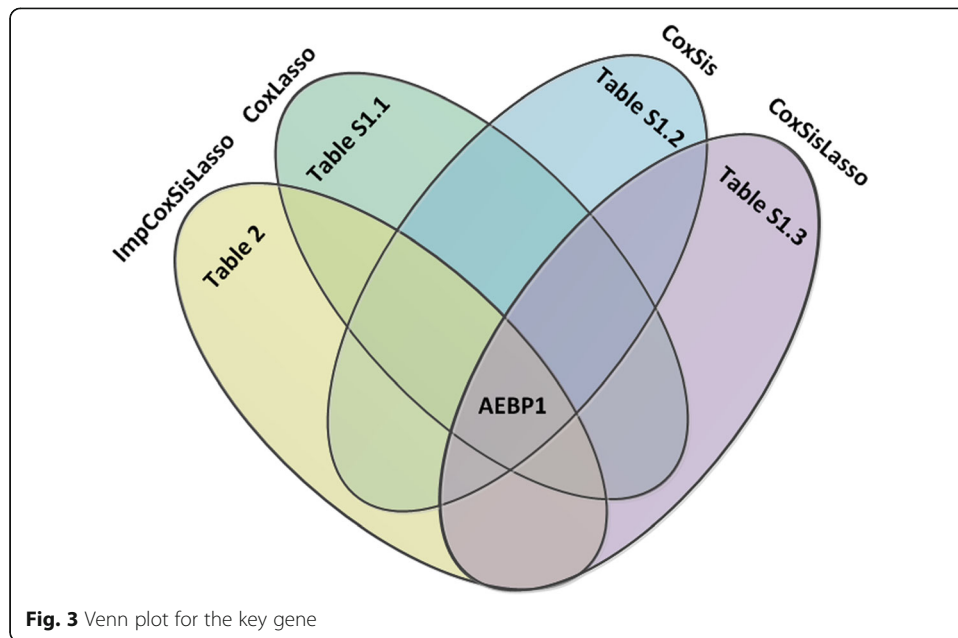
Results

Key genes selection

By integrating the genes selected by ImpCosSisLasso and previous studies [2], Fig. 3 demonstrates that *AEBP1* is the only mutual explored gene for ImpCoxSisLasso, Cox Lasso, Coxsis and CoxSisLasso, which implies that *AEBP1* is very potential for the survival time of GBM. The key genes selected by ImpCoxSisLasso are listed in Table 2 and

Table 1 The sequences of the primers specific for target genes

Target gene	The sequences of the primers specific
HES1-F	5'GGCGGCTAAGGTGTTTGAGG3'
HES1-R	5'GGGCCGCTGGAAGGTGACAC3'
HK2-F	5'AGGGGACTTTGATATCGACATTG3'
HK2-R	5'GCCCCCACTCCATATTGATAC3'



Supplementary Tables S1.1, 1.2, 1.3. The genes selected by previous algorithms [2] are shown in Supplementary Table S1.4.

Performance comparison for different survival analysis algorithms

Here, we use Receiver operating Characteristic (ROC) and Area Under Curve (AUC) [21] curves to compare the performance of the aforementioned algorithms [2]. Both ROC (Fig. 4a) and AUC (Fig. 4b) curves show that ImpCoxSisLasso outperforms the CoxLasso, CoxSis and CoxSisLasso algorithms [2].

Immune-chemistry experiment

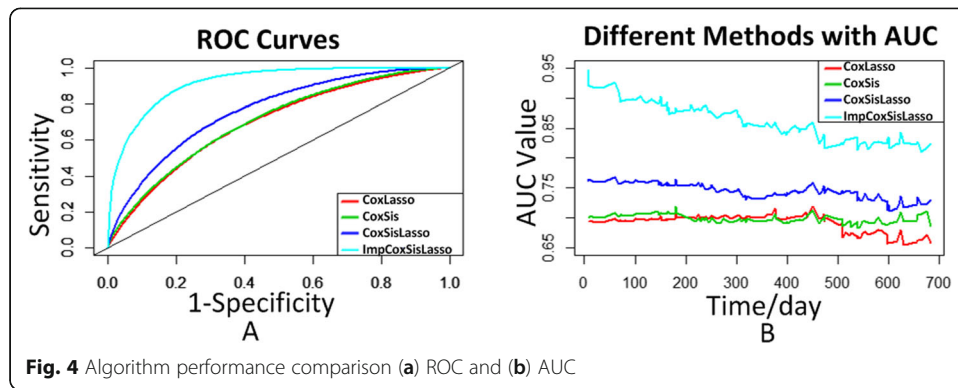
Figure 5a, b and c demonstrate that *AEBP1* is positively detectable in GBM. Figure 5d, e and f also show that *EGFR* is positively expressed in GBM. Next, Table 3 shows that there is a strong correlation between *AEBP1* and *EGFR*, since the *p* value of Chi square test [16] is less than 0.05. Lastly, Supplementary Table S2 lists the protein expressions for *AEBP1* and *EGFR* in immune-chemistry experiment, as well as we use Fig. 1a to demonstrate the impact of *AEBP1* on the survival time for GBM patients.

Protein network analysis

Firstly, we increase and keep the *AEBP1* expression for LN229 cell lines in the experimental group and control group, respectively. Then, we carry out RPPA experiment twice for these two groups. Finally, the *AEBP1* Up data set (Supplementary Table S3.1) shows 287 related proteins' expression.

Table 2 Selected genes by ImpCoxSisLasso strategy

Key gene
AEBP1, CDCA7L, SNTB1, TELO2, SLC35D1, FOXG1, ARIH2_INTS1, ZNF786_AEBP1, ZNF786_INTS1, ZNF786_SGCD, ZNF786_EIF3A, ZNF786_CDCA7L, ZNF786_TELO2, AEBP1_IL17RC, INTS1_SGCD, INTS1_IL17RC, INTS1_SLC35D1, GDNF_IL17RC, SGCD_TELO2, IL17RC_TELO2, CBLN1_TELO2, SLC35D1_TELO2



Secondly, we knock out the *AEBP1* by CRISPR-cas9 and keep the *AEBP1* expression for LN229 cell lines in the experimental group and control group, respectively. Then, we carry out RPPA experiment twice for these two groups. Finally, the *AEBP1* Down data set (Supplementary Table S3.2) shows 302 related proteins' expression.

Thirdly, we use Eq. 5. to compute the fold change between experimental and control group for both *AEBP1* Up and Down data set, respectively.

$$FC_i = \frac{\sum_{j=1}^m Exp_{ij}}{\sum_{j=1}^m CoP_{ij}} \tag{5}$$

Here, FC_i is the fold change between experimental (Exp_{ij}) and control group (CoP_{ij}). i represents the index of the proteins. j represents the index for the experiment, and m represents the replicates number. Exp_{ij} and CoP_{ij} are the proteins' expression measured by RPPA experiment.

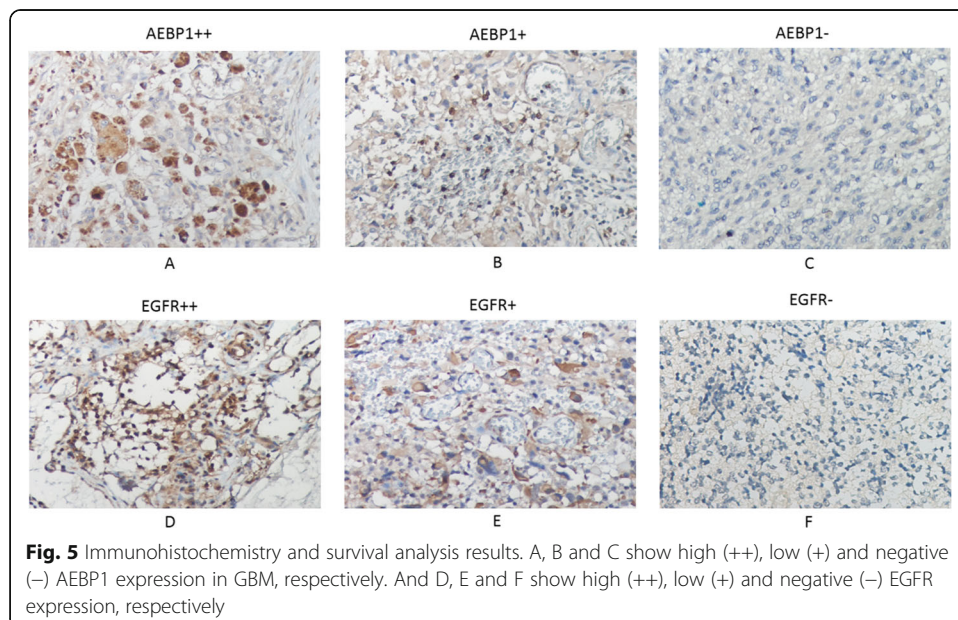


Table 3 Statistical analysis of IHC data for AEBP1 and EGFR protein in GBM

	AEBP1 positive (++&+)	AEBP1 negative (-)	Row Total
EGFR+	18	11	29
EGFR-	2	7	9
Column total	20	18	38
P-Value	0.036485093		

$$\text{Judge}_i(\text{FC}_i) = \begin{cases} \text{UP} & \text{FC}_i > 1 \\ \text{DOWN} & \text{FC}_i \leq 1 \end{cases} \quad (6)$$

Next, we use Eq. 6. to determine which proteins are promoted or inhibited by AEBP1. For *AEBP1* Up data set, the promoted proteins are listed in Supplementary Table S4.1 as AEBP1 promoted data set. For AEBP1 Down data set, the inhibited proteins are listed in Supplementary Table S4.2 as AEBP1 inhibited data set.

Lastly, we compare the experimental group (Exp_{ij}) and control group (CoP_{ij}) for both AEBP1 Promoted and Inhibited data sets by T test [3]. The null hypothesis is that the average expression level of a protein under experimental condition is equal to the level under control conditions. After we compute the P-value for both *AEBP1* promoted data set and *AEBP1* inhibited data set, 7 proteins’ expression are statistically significant in AEBP1 promoted data set (Table 4) and AEBP1 inhibited data set (Table 5), respectively. The detail data is listed in Supplementary Table S5.

Since manually reviewed experimental evidences already turn out that *ACCI-R-C* (*ACCI*) [27], *EMA-M-C* (*EMA*) [28], *Raptor-R-V* (*Raptor*) [29], *Rictor_pT1135-R-V* (*Rictor*) [30] and *SOD1-M-V* (*SOD1*) [31] are related to *AEBP1*, we then carry out PCR experiment for *HES1-R-V* (*HES1*) and *Hexokinase-II-R-V* (*HK2*) in Cell Line LN-18, LN-229, and U251 to confirm the correlation between the expression of these two proteins and *AEBP1*, respectively.

PCR verification

Figure 6 and Supplementary Table S6 indicate that the expressions of both *HES1* and *HK2* are positively related to the expression of *AEBP1*. Therefore, we conclude that all the proteins listed in Tables 4 and 5 are related to the expression of *AEBP1*.

Discussion

This study aims to investigate the origin of the GBM from systematic view. Three research questions were proposed and addressed.

Since previous additive survival analysis algorithms (CoxLasso, CoxSis and CoxSisLasso [2]) do not consider the gene’s interaction, they neither agree with the nature of biology [32–35] nor have the high predictive accuracy. Here, after we integrate interaction item into Eq. 4 of ImpCoxSisLasso algorithm, we not only can explore the gene

Table 4 AEBP1 promoted dataset

Protein Name	FC _i	P-value
ACC1-R-C	1.06701864	0.01841297
EMA-M-C	1.36047171	0.02989766
HES1-R-V	1.15596	0.00029086

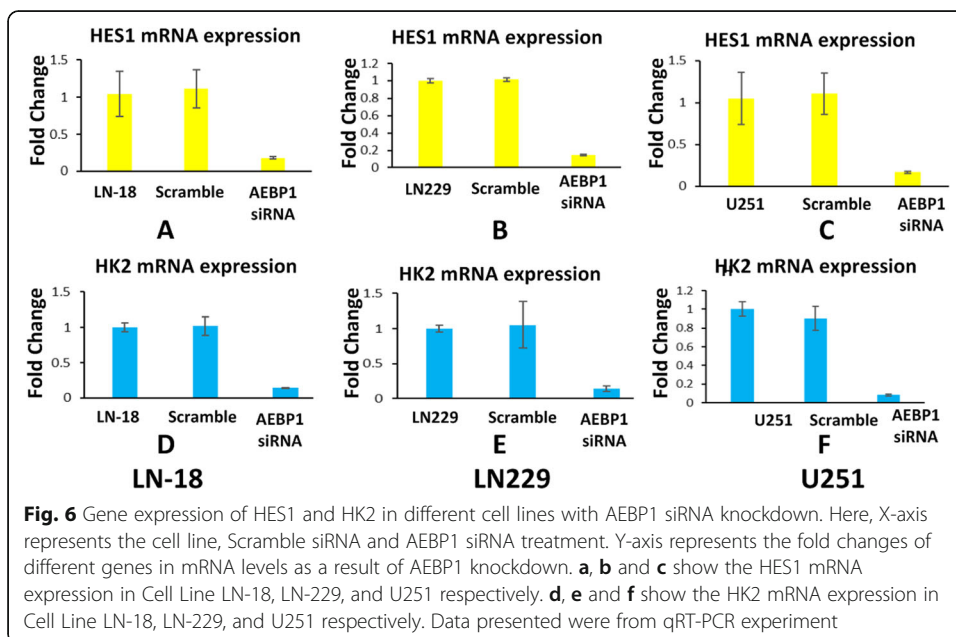
Table 5 AEBP1 inhibited dataset

Protein Name	FC _i	P-value
Hexokinase-II-R-V	0.98022947	0.00639911
Raptor-R-V	0.9132965	0.03564656
Rictor_pT1135-R-V	0.95940456	0.02726222
SOD1-M-V	0.97722104	0.03054562

signature of GBM for P > N type of data (Table 2), but also Fig. 4 turns out that ImpCoxSisLasso algorithm is much better than the previous [2]. Moreover, Fig. 3 shows that both ImpCoxSisLasso algorithm and our previous survival algorithms (Cox-Lasso, CoxSis and CoxSisLasso [2]) mutually find *AEBP1* as the key gene related to the survival time of GBM patients. Lastly, we use TCGA data [24] (Figs. 1a) and 5 to cross validate the impact of *AEBP1* gene on GBM.

In this report, Table 3 and Fig. 5 demonstrated that there is a significant correlation between the expression of *AEBP1* and *EGFR* in patient samples of GBM. Through over-expression or knockdown of *AEBP1* gene in glioma cell line LN229 and RPPA analysis, we can show that a panel of proteins (Tables 4 and 5) are significantly affected by the manipulation of *AEBP1* expression. By a stringent cutoff of $p < 0.05$, *HK2*(Fig. 6), *SOD1* [36], *Raptor* [37], *Rictor* [38] were markedly downregulated when *AEBP1* expression was knocked down. Using the same cutoff standard, *ACC1* [39], *HES1*(Fig. 6) and *EMA* [28] were significantly upregulated when *AEBP1* was overexpressed.

Through extensive literature research [36, 37, 40, 41], we found that affected proteins by either *AEBP1* overexpression or knockdown converge to the regulation of mainly one signaling pathway: the MTOR pathway. MTOR forms two different complexes, MTORC1 and *MTORC2*. *HK2* is a downstream effector of MTORC1 through the transcriptional regulation of transcription factor *HIF1* [41]. It integrates energy metabolism with cancer cell survival. MTORC1 directly regulates *Superoxide dismutase 1 (SOD1)* through phosphorylation [36]. *SOD1* can in turn act as a nuclear transcription factor to



fend off oxidative stress [14] in cancer cells [31]. Raptor is an interacting partner of MTORC1 [40] and phosphorylation of Raptor by MTORC1 is essential for the kinase activity of MTORC1 on its substrates [37]. Dissociation of Raptor from MTORC1 resulted in the inhibition of MTOR signaling activity [29].

On the other hand, *Rictor* is an essential partner of MTORC2 [42, 43]. Blocked association of Rictor from MTORC2 led to the inhibition of MTORC2 activity and cancer cell death in glioma cell line in vitro [30]. In addition, overexpression of Rictor was associated with increased MTORC2 activity and tumor growth in glioma patients [44].

ACC1 is a *SREBP1* target gene [39]. MTORC1 signaling controls the expression of *SREBP1* and consequently regulates *SREBP1* target gene *ACC1* expression [27]. *HES1* (Hairy and Enhancer of Split 1) is a downstream target of Notch signaling pathway [45, 46]. *HES1* antagonizes PTEN-induced inhibition of MTOR signaling activity [47]. *EMA* (Epithelial Membrane Antigen) is a glycosylated protein encoded by *MUC1* gene [28]. Although the cytoplasmic tail of *EMA* has roles in signal transduction, the relationship between *EMA* and MTOR signaling pathway is not clear at this moment. But *EGFR* signaling activation increased *MUC1* gene expression in some cancer cell lines [48].

Taken together, our results indicate that *AEBP1* regulates the expression of *Raptor* and *Rictor*, key binding partners of MTORC1 and MTORC2. *AEBP1* also has an important role in regulating the expression of *HES1* which promotes MTOR activity by inhibiting PTEN function [47, 49]. Enhanced MTOR signaling by *AEBP1* stimulates downstream expression of MTOR targets *HK2* and *ACC1* or increases phosphorylation of MTOR target *SOD1*. The net effect of *AEBP1* overexpression in glioma is the activation and promotion of MTOR signaling pathway in cancer cells. We illustrate the effects of *AEBP1* with different molecular players on MTOR pathway in a schematic model (Fig. 1b).

AEBP1 overexpression in glioma is associated with higher tumor grade and worse prognosis [50]. *AEBP1* as therapeutic target for cancer treatment has not drawn much attention. However, clinical trials are under way to test the effectiveness of MTOR inhibitors in the treatment of different types of cancers [51]. Indeed, it has been reported that MTOR inhibitors are effective in a subset of patients with *Rictor* gene amplification in lung cancer [52] and gastric cancer [38]. MTOR inhibitors are actively investigated in preclinical cancer cell line and mouse models of GBM with various success [51, 53]. Experimental results from these preclinical studies place MTOR as the pivotal target in treating GBM [53]. A search of clinical trials in the USA revealed 34 trials using MTOR inhibitors in glioma patients (www.clinicaltrials.gov last accessed on July 20th, 2019). Most of these clinical trials are still under way. Completed trials showed limited successes. In terms of our findings, we may need to take into consideration the expression levels of *AEBP1* in patients' tumor samples when planning clinical trials of MTOR inhibitors on GBM or evaluating clinical responses of MTOR inhibitors in the treatment of GBM.

Conclusions

In conclusion, in step one, our improved survival analysis model not only finds that the key gene, but also has better performance and efficiency. In step two, through CRISPR and RPPA techniques, we get reliable protein data, and then search for the seven key proteins by a precise multi-scale mathematical model. And last, we analyze the

relationship between AEBP1 and key proteins to identify that AEBP1 exerts its tumor-promoting effects by mainly activating mTOR pathway in Glioma.

Although the results show that we can find the mechanism of AEBP1 in mTOR, we can still improve it. For instance, in protein network, we can consider adding machine learning algorithm to find the key protein. In the future, we can even improve the survival time of patients through the explored pathway mechanism.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12859-020-03674-4>.

Additional file 1. This file contains all the supplementary tables and figures.

Abbreviations

GBM: Glioblastoma multiforme; AEBP1: Adipocyte enhancer-binding protein 1; CRISPR: Clustered regularly interspaced short palindromic repeat sequences; RPPA: Reverse phase protein arrays; IQR: Interquartile range; PCR: Polymerase Chain Reaction; ROC: Receiver operating Characteristic; AUC: Area Under Curve

Acknowledgements

Not applicable.

About this supplement

This article has been published as part of *BMC Bioinformatics Volume 21 Supplement 13, 2020: Selected articles from the 18th Asia Pacific Bioinformatics Conference (APBC 2020): bioinformatics*. The full contents of the supplement are available online at <https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-21-supplement-13>.

Authors' contributions

LZ participated in the design of the study. YJY and WJL collected and analyzed the data. XFR did the experiment. YJY, WJL, TTL, MX, HRZ and LZ wrote and edited the manuscript. All authors read and approved the final manuscript.

Funding

Publication costs are funded by the National Natural Science Foundation of China [Nos.61372138 and Nos.11701469] and the National Science and Technology Major Project [2018ZX10201002]. Top-notch Talent Cultivation Plan of Southwest Hospital (SWH2018BJKJ-05).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and the additional file.

Ethics approval and consent to participate

the experimental protocol and the retrospective analysis of GBM specimens were approved by the Ethical Committee of Southwest Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Computer Science, Sichuan University, Chengdu 610065, China. ²Department of Neurosurgery, Southwest Hospital, Third Military Medical University, Chongqing, P.R. China. ³College of Mathematics and Statistics, Southwest University, Chongqing 400715, P.R. China. ⁴School of Computing, Ulster University, Coleraine, Londonderry, Northern Ireland, UK.

Published: 17 September 2020

References

1. Ahmadloo N, et al. Treatment outcome and prognostic factors of adult glioblastoma multiforme. *J Egypt Natl Canc Inst.* 2013;25(1):21–30.
2. Xia Y, et al. Exploring the key genes and signaling transduction pathways related to the survival time of glioblastoma multiforme patients by a novel survival analysis model. *BMC Genomics.* 2017;18(Suppl 1):950.
3. Zhang L, et al. EZH2-, CHD4-, and IDH-linked epigenetic perturbation and its association with survival in glioma patients. *J Mol Cell Biol.* 2017;9(6):477–88.
4. Ladha J, et al. Identification of genomic targets of transcription factor Aebp1 and its role in survival of Glioma cells. *Mol Cancer Res.* 2012;10(8):1039–51.
5. Wangaryattawanich P, et al. Survival analysis of pre-operative GBM patients by using quantitative image features. In: *International Conference on Control*; 2014.

6. Pudelko L, et al. Glioblastoma and glioblastoma stem cells are dependent on functional MTH1. *Oncotarget*. 2017;8(49):84671–84.
7. Wei N, et al. Efficient gene knockout in goats using CRISPR/Cas9 system. *PLoS One*. 2014;9(9):e106718.
8. Walters BJ, et al. Advanced In vivo Use of CRISPR/Cas9 and Anti-sense DNA Inhibition for Gene Manipulation in the Brain. *Front Genet*. 2016;6:362.
9. Xie Y, et al. An episomal vector-based CRISPR/Cas9 system for highly efficient gene knockout in human pluripotent stem cells. *Sci Rep*. 2017;7(1):2320.
10. Nishizuka SS, Mills GB. New era of integrated cancer biomarker discovery using reverse-phase proteinarrays ☆. *Drug Metab Pharmacokinet*. 2016;31(1):35–45.
11. Pushparaj PN, et al. siRNA, miRNA, and shRNA: in vivo applications. *J Dent Res*. 2008;87(11):992–1003.
12. Shull AY, et al. RPPA-based protein profiling reveals eIF4G overexpression and 4E-BP1 serine 65 phos-phorylation as molecular events that correspond with a pro-survival phenotype in chronic lymphocytic leukemia. *Oncotarget*. 2015; 6(16):14632–45.
13. Ludovini V, et al. Reverse phase protein array (RPPA) combined with computational analysis to unravel relevant prognostic factors in non- small cell lung cancer (NSCLC): a pilot study. *Oncotarget*. 2017;8(47): 83343–53.
14. Ummanni R, et al. Evaluation of reverse phase protein array (RPPA)-based pathway-activation profiling in 84 non-small cell lung cancer (NSCLC) cell lines as platform for cancer proteomics and biomarker discovery ☆. *Biochim Biophys Acta*. 2014;1844(5):950–9.
15. O., N., BerEP4 and AE1/3 are Reliable Markers of Epithelial Content for Biomarker Discovery Using Reverse Phase Protein Arrays (RPPA). 2016.
16. Zhang L, et al. Building up a robust risk mathematical platform to predict colorectal Cancer. *Complexity*. 2017;2017:14.
17. Zhang L, Zhang S. Using game theory to investigate the epigenetic control mechanisms of embryo development: comment on: "epigenetic game theory: how to compute the epigenetic control of maternal-to-zygotic transition" by Qian Wang et al. *Phys Life Rev*. 2017;20:140–2.
18. Zhang L, et al. Lineage-associated underrepresented permutations (LAUPs) of mammalian genomic sequences based on a jellyfish-based LAUPs analysis application (JBLA). *Bioinformatics*. 2018;34(21):3624–30.
19. Zhang L, et al. Investigation of mechanism of bone regeneration in a porous biodegradable calcium phosphate (CaP) scaffold by a combination of a multi-scale agent-based model and experimental optimization/validation. *Nanoscale*. 2016;8(31):14877–87.
20. Zhang L, et al. Revealing dynamic regulations and the related key proteins of myeloma-initiating cells by integrating experimental data into a systems biological model. *Bioinformatics*. 2019;26:btz542. <https://doi.org/10.1093/bioinformatics/btz542>.
21. Zhang L, et al. Computed tomography angiography-based analysis of high-risk intracerebral haemorrhage patients by employing a mathematical model. *BMC Bioinformatics*. 2019;20(Suppl 7):193.
22. Zhang L, et al. Comprehensively benchmarking applications for detecting copy number variation. *PLoS Comput Biol*. 2019;15(5):e1007069.
23. Tong X, et al. Development of an agent-based model (ABM) to simulate the immune system and integration of a regression method to estimate the key ABM parameters by fitting the experimental data. *PLoS One*. 2015;10(11): e0141295.
24. Katarzyna T, Patrycja C, Maciej W. The Cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol*. 2015;19(1A):68–77.
25. Maglott DR, Nierman WC. Clone and genomic repositories at the American type culture collection. *Genomics*. 1990;8(3): 601–5.
26. Majdalawieh A, Zhang L, Ro H-S. Adipocyte enhancer-binding protein-1 promotes macrophage inflammatory responsiveness by up-regulating NF-kappaB via I kappaBalpha negative regulation. *Mol Biol Cell*. 2007;18(3):930–42.
27. Jeon T, Osborne TF. SREBPs: metabolic integrators in physiology and metabolism. *Trends Endocrinol Metab*. 2012;23(2): 65–72.
28. C, Y., M. JL, and I. NK. MUC1 and cancer immunotherapy, vol. 1; 2017.
29. Oshiro N, et al. Dissociation of raptor from mTOR is a mechanism of rapamycin-induced inhibition of mTOR function. *Genes Cells*. 2004;9(4):359–66.
30. Angelica BS, et al. Specific blockade of Rictor-mTOR association inhibits mTORC2 activity and is cytotoxic in glioblastoma. *PLoS One*. 2017;12(4):e0176599.
31. Tsang CK, et al. Superoxide dismutase 1 acts as a nuclear transcription factor to regulate oxidative stress resistance. *Nat Commun*. 2014;5(1):3446.
32. Van Dijk D, Sharma R, Nainys J, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell*. 2018;174(3):716–729.e27.
33. Van Dijk D, et al. MAGIC: A diffusion-based imputation method reveals gene-gene interactions in single-cell RNA-sequencing data. *bioRxiv*. 2017:111591.
34. Howard TD, et al. Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. *Am J Hum Genet*. 2002;70(1):230–6.
35. Upstillgoddard R, et al. Machine learning approaches for the discovery of gene-gene interactions in disease data. *Brief Bioinform*. 2013;14(2):251–60.
36. Tsang CK, et al. SOD1 Phosphorylation by mTORC1 Couples Nutrient Sensing and Redox Regulation. *Mol Cell*. 2018; 70(3):502–15.
37. Foster KG, et al. Regulation of mTOR complex 1 (mTORC1) by raptor Ser863 and multisite phosphorylation. *J Biol Chem*. 2010;285(1):80–94.
38. Kim ST, et al. Rapamycin-insensitive companion of mTOR (RICTOR) amplification defines a subset of advanced gastric Cancer and is sensitive to AZD2014-mediated mTORC1/2 inhibition. *Ann Oncol*. 2016;28(3):547–54.
39. Yin G, et al. mTOR complex 1 signaling regulates the balance between lipid synthesis and oxidation in hypoxic lymphocytes. *Biosci Rep*. 2017;37(1):BSR20160479.

40. Kim D, et al. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell*. 2002;110(2):163–75.
41. Roberts DJ, Miyamoto S. Hexokinase II integrates energy metabolism and cellular protection: Akting on mitochondria and TORCing to autophagy. *Cell Death Differ*. 2015;22(2):248–57.
42. Kocalis H, et al. Rictor/mTORC2 facilitates central regulation of energy and glucose homeostasis. *Mol Metab*. 2014;3(4):394–407.
43. Sarbassov DD, et al. Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. *Curr Biol*. 2004;14(14):1296–302.
44. Masri J, et al. mTORC2 activity is elevated in Gliomas and promotes growth and cell motility via overexpression of Rictor. *Cancer Res*. 2007;67(24):11712–20.
45. Schnell SA, et al. Therapeutic targeting of HES1 transcriptional programs in T-ALL. *Blood*. 2015;125(18):2806–14.
46. Schreck KC, et al. The notch target Hes1 directly modulates Gli1 expression and hedgehog signaling: a potential mechanism of therapeutic resistance. *Clin Cancer Res*. 2010;16(24):6060–70.
47. Wong GW, et al. HES1 opposes a PTEN-dependent check on survival, differentiation, and proliferation of TCR -selected mouse thymocytes. *Blood*. 2012;120(7):1439–48.
48. Dharmaraj N, Engel BJ, Carson DD. Activated EGFR stimulates MUC1 expression in human uterine and pancreatic cancer cell lines. *J Cell Biochem*. 2013;114(10):2314–22.
49. Bhattacharya K, Maiti S, Mandal C. PTEN negatively regulates mTORC2 formation and signaling in grade IV glioma via Rictor hyperphosphorylation at Thr1135 and direct the mode of action of an mTORC1/2 inhibitor. *Oncogenesis*. 2016;5(5):e227.
50. Kohsaka S, et al. Epiregulin enhances tumorigenicity by activating the ERK/MAPK pathway in glioblastoma. *Neuro-oncology*. 2014;16(7):960.
51. Fan Q, Nicolaides T, Weiss WA. Inhibiting 4EBP1 in Glioblastoma. *Clin Cancer Res*. 2018;24(1):14–21.
52. Cheng H, et al. RICTOR amplification defines a novel subset of patients with lung cancer who may benefit from treatment with mTORC1/2 inhibitors. *Cancer Discov*. 2015;5(12):1262–70.
53. Fan Q, et al. A kinase inhibitor targeted to mTORC1 drives regression in Glioblastoma. *Cancer Cell*. 2017;31(3):424–35.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

