

Supplementary material

for

Network Analysis of Genomic Alteration Profiles Reveals Co-altered Functional Modules and Driver Genes in Glioblastoma

Yunyan Gu^{1*}, Hongwei Wang¹, Yao Qin¹, Yujing Zhang¹, Wenyuan Zhao¹, Lishuang Qi¹, Yuannv Zhang¹, Chenguang Wang¹, Zheng Guo^{1,2*}

¹ College of Bioinformatics Science and Technology, Harbin Medical University,
Harbin 150086, China

² Bioinformatics Center, Fujian Medical University, Fuzhou, 350004, China

* Corresponding author

Email: guoz@ems.hrbmu.edu.cn or guyunyan@ems.hrbmu.edu.cn

Phone: (86 451) 8661-5933

Fax: (86 451) 8666-9617

Table of Contents

A. Supplementary Methods (pg 3-4)

B. Supplementary Figures (pg 5-11)

C. Supplementary Tables (pg 12-19)

A. Supplementary Methods

The pseudo codes for searching co-altered module pair

Input: (1) $G(V, E)$, where G is the connected sub-network, V is the set of genes and E is the set of edges in the network;
(2) Alteration profiles;

Output: $R=(M_1, M_2, \dots, M_N)$, N module pairs, $M_i(G_1, G_2)$, $i=1\dots N$, will to be discovered and returned.

N is also the number of all seed pairs.

S is the co-alteration score of the module pair.

1: $R=\phi$. $S=0$. $N'=1$.

2: $M(G_1, G_2)$ is the module pair that we will discover.

3: **repeat**

4: Taking each pair of altered genes as a seed pair, also a module pair $M(G_1, G_2)$, $G_1=\{i\}$, $G_2=\{j\}$. $S=S_{\text{seed}}$.

5: **repeat**

6: $\quad\quad\quad$ #add one altered gene in module pair#

7: Find all edges $E(G_1, k_p)=1$ or $E(G_2, k_q)=1$ from E . $p, q=1, \dots, n$. we regard the nodes $k_1, \dots, k_p, \dots, k_q, \dots, k_n$ as the direct interaction neighbors with G_1 or G_2 .

8: $V'=\{v \mid v \in k_1, \dots, k_p, \dots, k_q, \dots, k_n, \text{ and } E(G_1, v)=1 \text{ or } E(G_2, v)=1, v \notin G_1 \cup G_2\}$

9: **if** V' is empty **then**

10: $\quad\quad\quad S$ unchanged

11: **Else**

12: **while** there is at least one interaction neighbor with G_1 or G_2 **do**

13: **for all** $v \in V'$ **do**

14: $\quad\quad\quad$ Calculate $S(v)=\max(S(G_1 \cup v, G_2) \mid E(G_1, v)=1, S(G_1, G_2 \cup v) \mid E(G_2, v)=1)$

15: $\quad\quad\quad$ **if** $S(v) > S$ **then**

16: $\quad\quad\quad$ Set $S = S(v)$, added v to $G_1 \mid E(G_1, v)=1$ or added v to $G_2 \mid E(G_2, v)=1$,

17: $\quad\quad\quad$ **end for**

18: **end while**

19: $\quad\quad\quad$ #delete one altered gene from module pair#

20: Find all nodes $x_1, \dots, x_p, \dots, x_n$ with only one interaction neighbor in the current module pair $M(G_1, G_2)$. And, they are not the current seeds and not the new added node.

21: $X'=\{x \mid x \in x_1, \dots, x_p, \dots, x_n, x \text{ is not the current seed and not the new added node}\}$

22: **if** X' is empty **then**

23: $\quad\quad\quad S$ unchanged

24: **else**

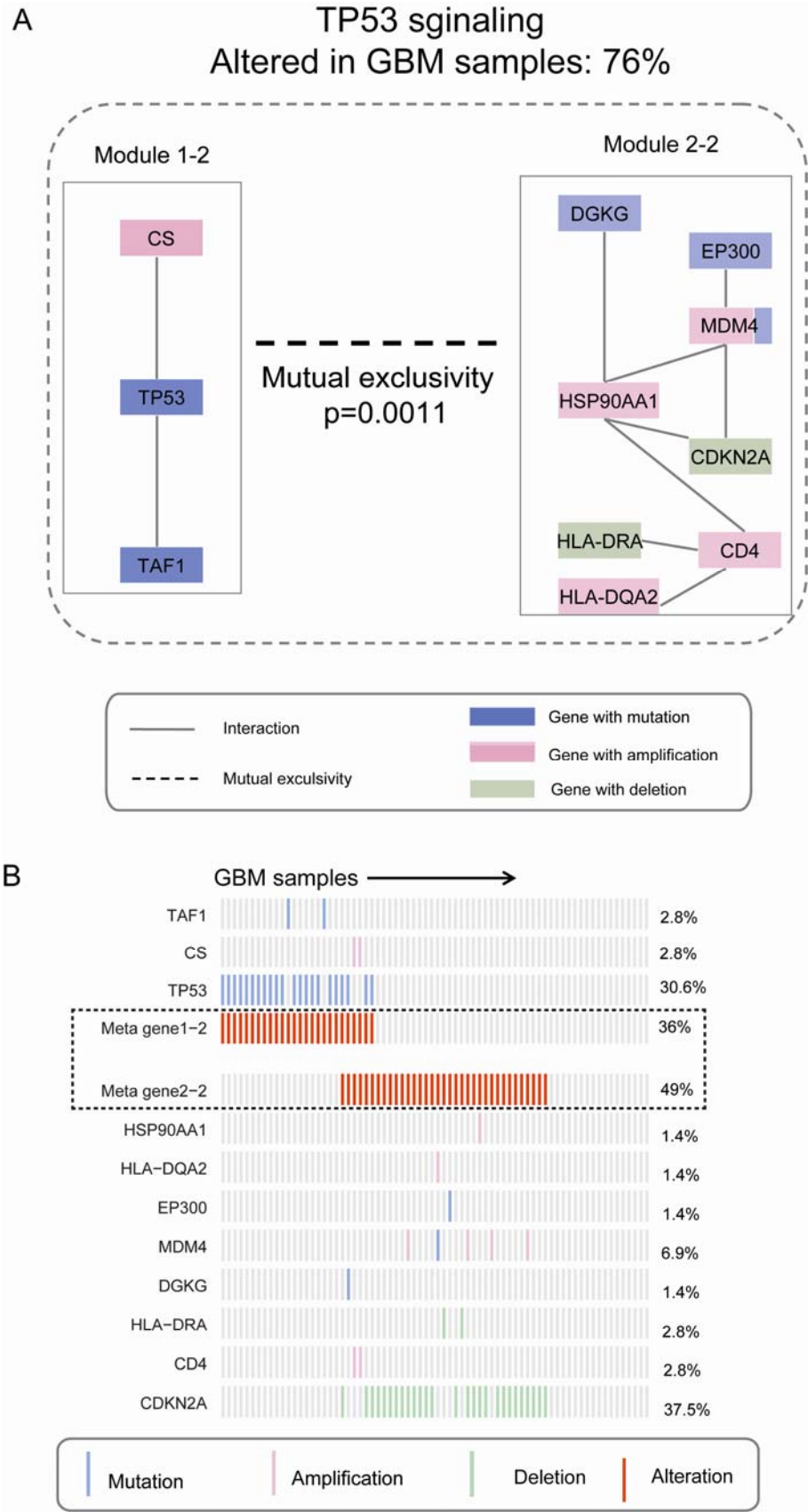
25: **while** there is at least one node that not the current seed and not the

```

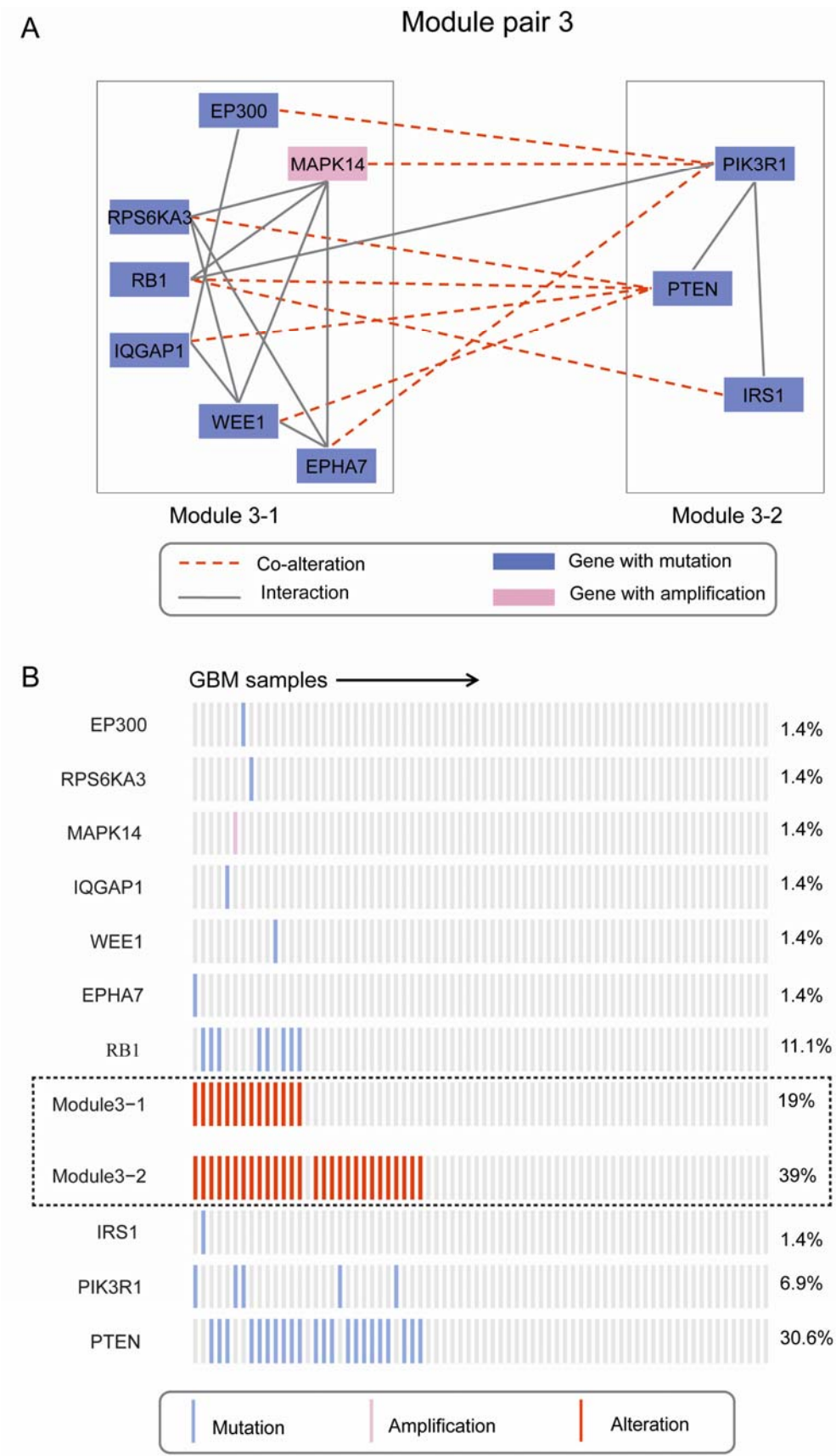
        new added node in  $M(G_1, G_2)$  do
26:    for all  $x \in X'$  do
27:        Calculate  $S(x) = \max(S(G_1 - x, G_2) \mid x \in G_1, S(G_1, G_2 - x) \mid x \in G_2)$ 
28:        if  $S(x) > S$  then Set  $S = S(x)$ ,  $G_1 = G_1 - x \mid x \in G_1$  or  $G_2 = G_2 - x \mid x \in G_2$ 
29:    end for
30:    end while
31: until  $S$  no longer increased
32:    add  $M(G_1, G_2)$  into  $R$ 
33:     $N' = N' + 1$ 
34: until  $N' > N$ 
    
```

B. Supplementary Figures

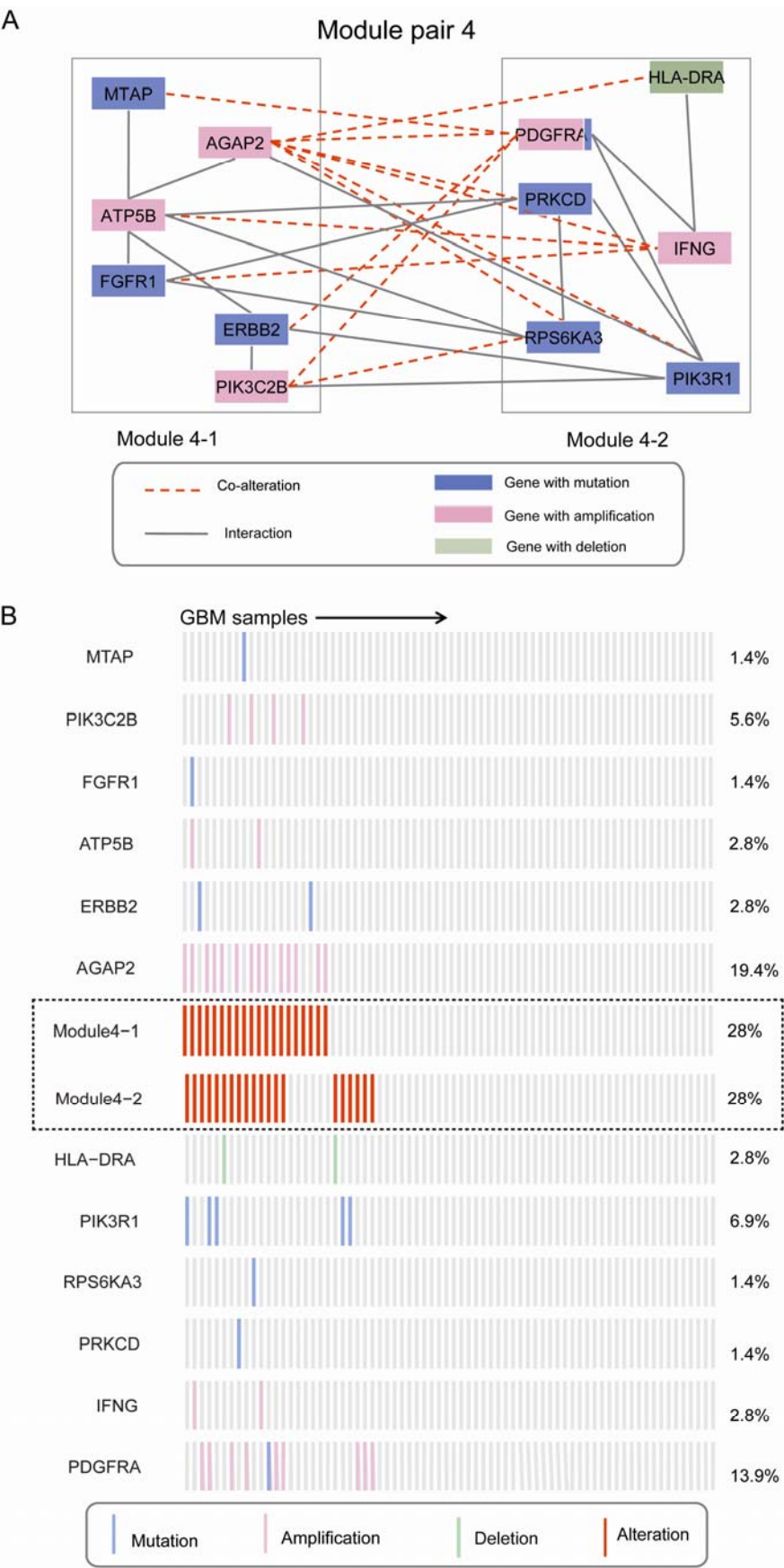
Supplementary Figure 1- Top mutual exclusive module pair



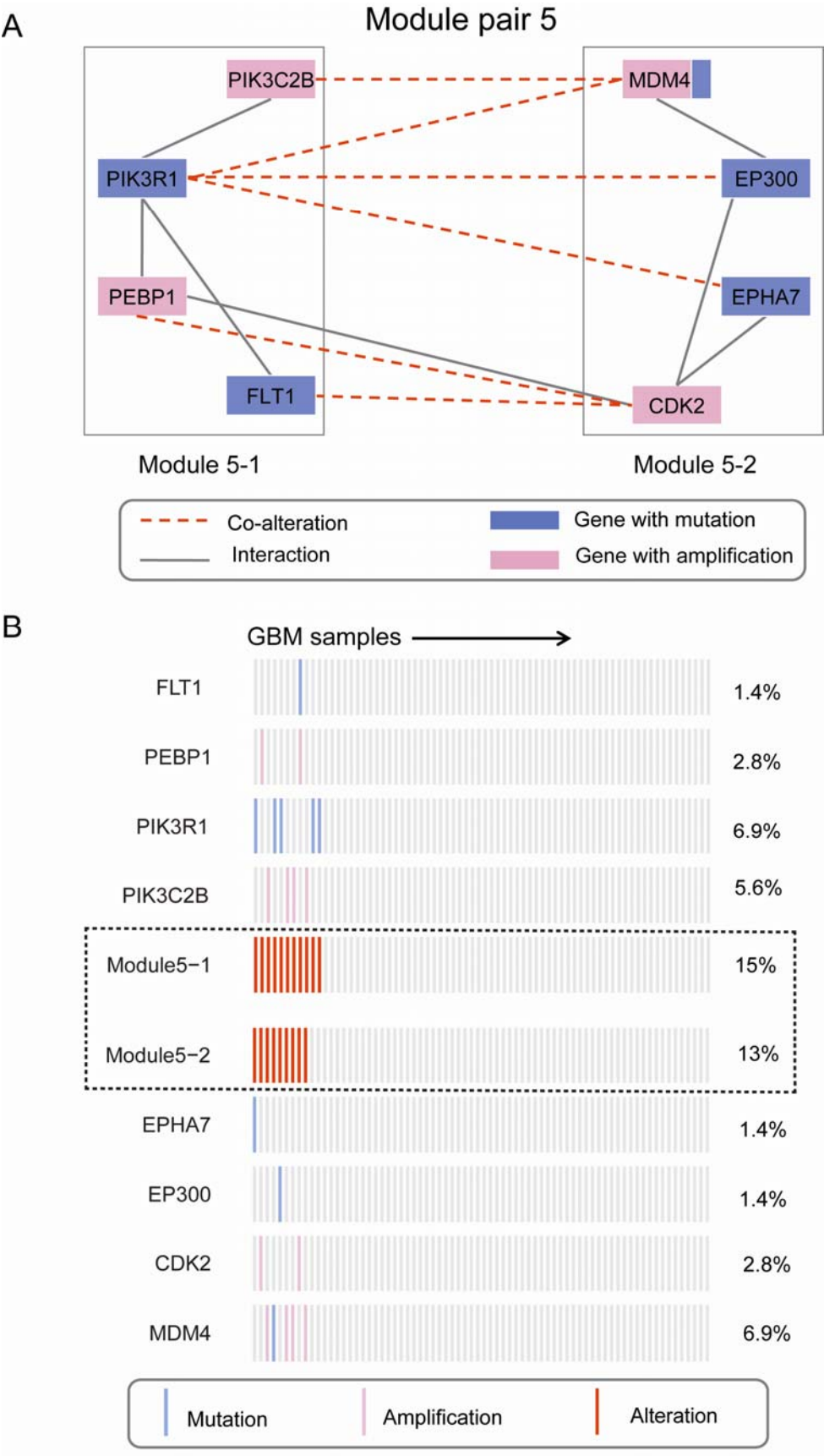
Supplementary Figure 2- Co-altered module pair 3



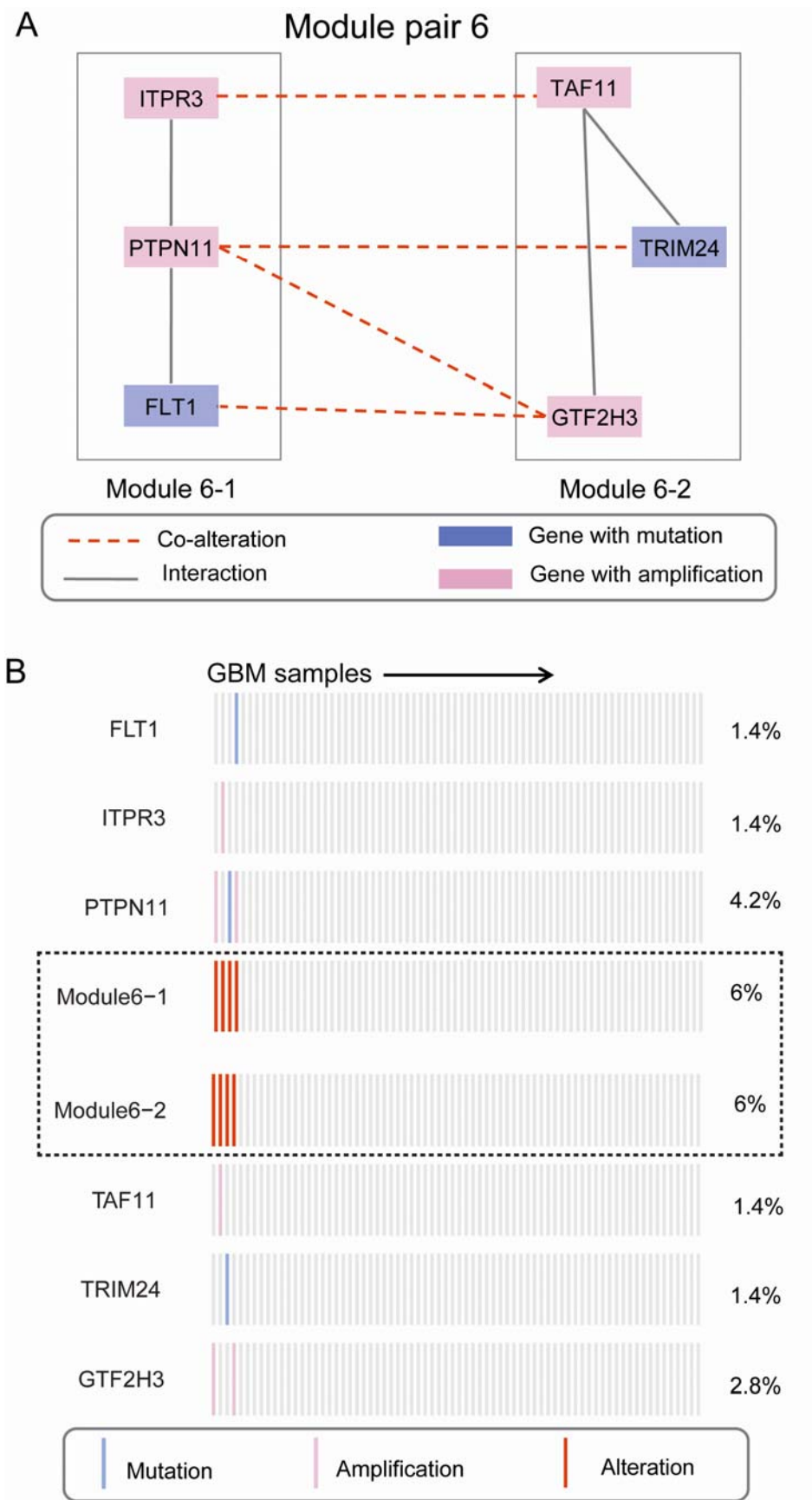
Supplementary Figure 3- Co-altered module pair 4



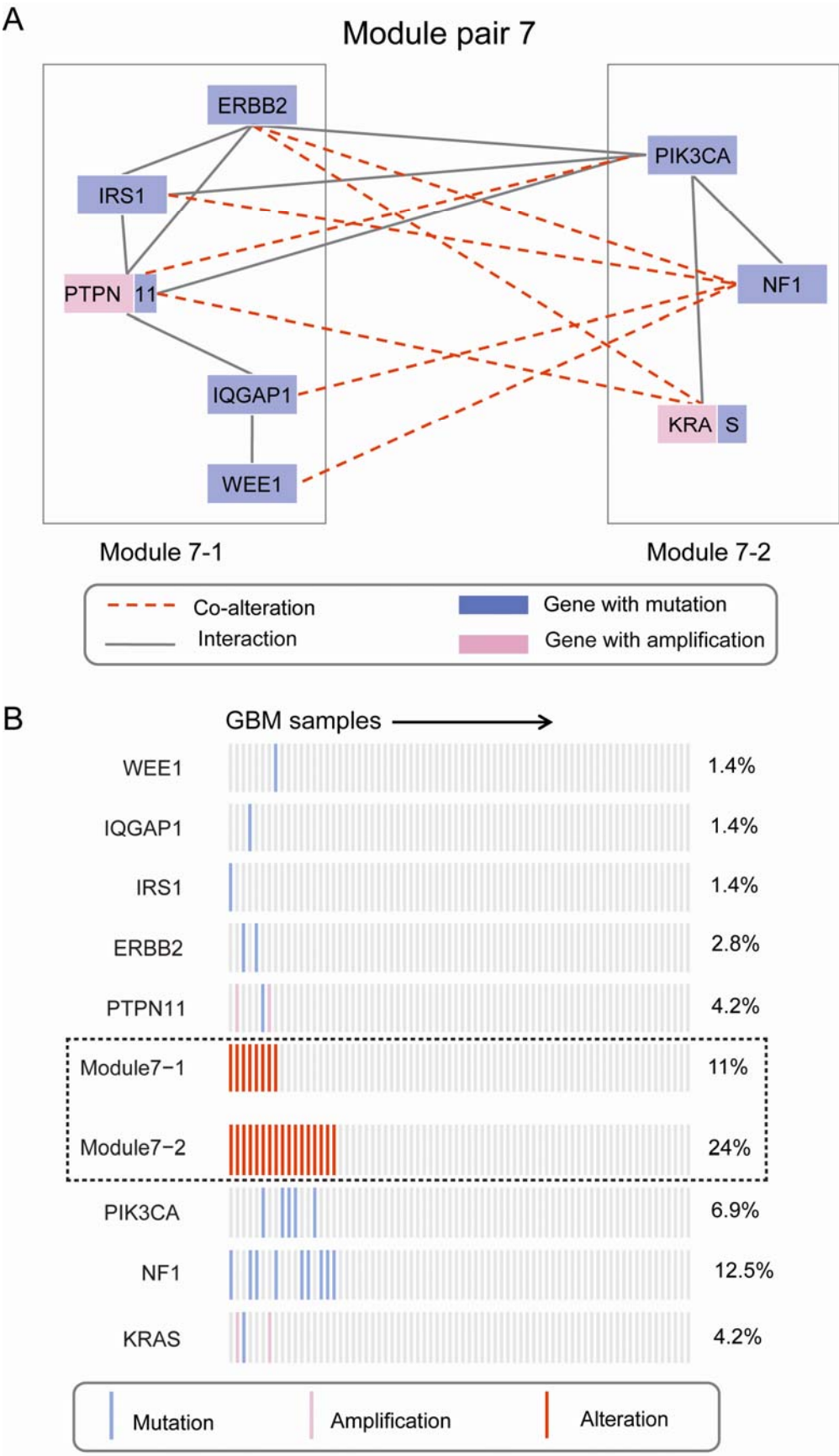
Supplementary Figure 4- Co-altered module pair 5



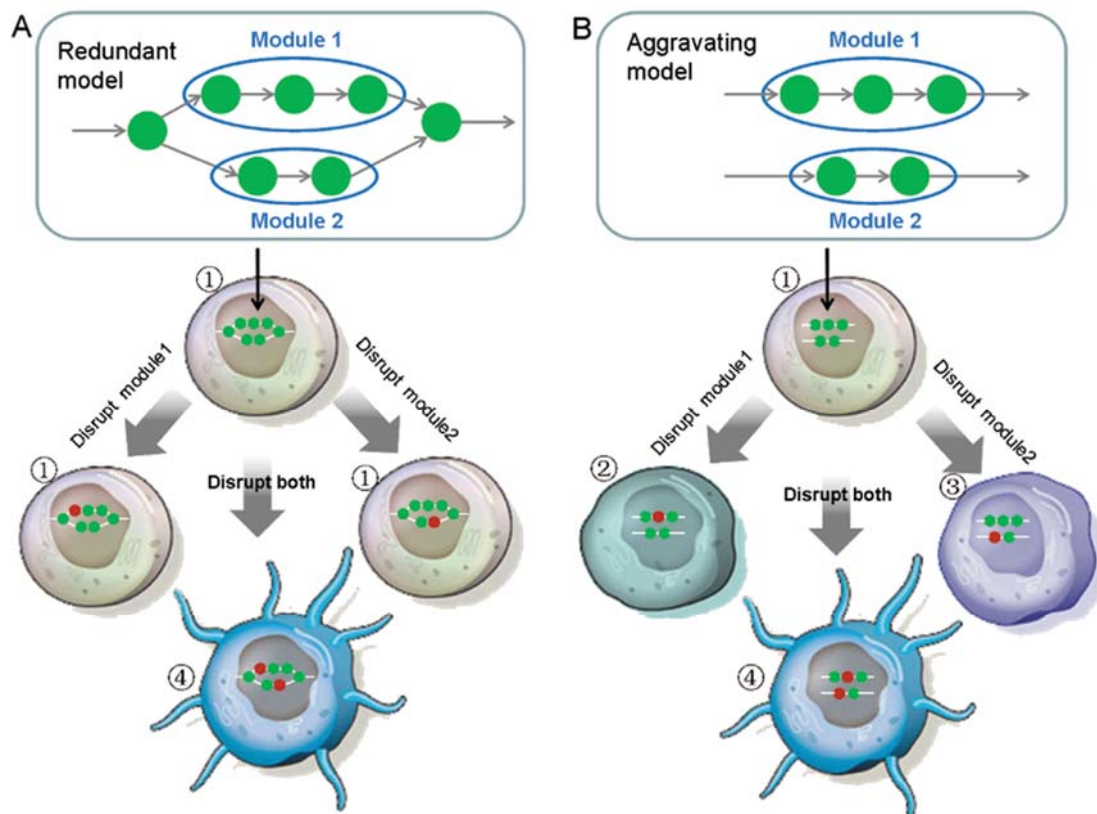
Supplementary Figure 5- Co-altered module pair 6



Supplementary Figure 6- Co-altered module pair 7



Supplementary Figure 7- Schematic diagram of the redundant and aggravating models for co-altered modules



Supplementary Figure 7– (A) The redundant model for co-altered modules. Alterations in modules 1 or 2 alone could not transform normal cells into cancer cells; cells become cancerous only when both of the modules are disrupted. (B) The aggravating model for co-altered modules. Alterations in modules 1 and 2, which individually have limited effects on promoting cancer, cooperate to transform normal cells into cancer cells. ① denotes normal cells. ② and ③ denote the cells with characteristics of cancer cells. ④ denotes the cancer cells. The red and green nodes within the modules represent genes with and without alterations, respectively.

C. Supplementary Tables

Supplementary Table 1- The co-altered module pairs

ID	Module ID	Entrez Gene ID	Gene Symbol ¹	Pathway	#N ²	#M ³	#S ⁴	P-value
1	Module 1-1	5925	RB1	RB signaling pathway	8	13	12	6.41E-06
		1027	CDKN1B		2			
		8826	IQGAP1		1			
		7465	WEE1		1			
		2045	EPHA7		1			
	Module 1-2	7157	TP53	TP53 signaling pathway	22	26		
		1431	CS		2			
		6872	TAF1		2			
2	Module 2-1	1956	EGFR	RTK signaling pathway	34	41	31	8.57E-08
		1432	MAPK14		1			
		506	ATP5B		2			
		5287	PIK3C2B		4			
		5290	PIK3CA		5			
	Module 2-2	1029	CDKN2A	TP53 signaling pathway	27	35		
		920	CD4		2			
		3122	HLA-DRA		2			
		1608	DGKG		1			
		4194	MDM4		5			
		2033	EP300		1			
		3118	HLA-DQA2		1			
		3320	HSP90AA1		1			
3	Module 3-1	5925	RB1	RB signaling pathway	8	14	14	1.34E-07
		2045	EPHA7		1			
		7465	WEE1		1			
		8826	IQGAP1		1			
		1432	MAPK14		1			
		6197	RPS6KA3		1			
		2033	EP300		1			
	Module 3-2	5728	PTEN	RTK signaling pathway	22	28		
		5295	PIK3R1		5			
		3667	IRS1		1			
4	Module 4-1	116986	AGAP2	RTK signaling pathway	14	20	14	2.66E-06
		2064	ERBB2		2			
		506	ATP5B		2			
		2260	FGFR1		1			
		5287	PIK3C2B		4			
		4507	MTAP		1			
	Module 4-2	5156	PDGFRA	RTK signaling	10	20		
		3458	IFNG		2			

		5580	PRKCD	pathway	1			
		6197	RPS6KA3		1			
		5295	PIK3R1		5			
		3122	HLA-DRA		2			
5	Module 5-1	5287	PIK3C2B	RTK signaling pathway	4	11	9	6.46E-10
		5295	PIK3R1		5			
		5037	PEBP1		2			
		2321	FLT1		1			
	Module 5-2	4194	MDM4	TP53 signaling pathway	5	9		
		1017	CDK2		2			
		2033	EP300		1			
		2045	EPHA7		1			
6	Module 6-1	5781	PTPN11	/	3	4	4	9.72E-07
		3710	ITPR3		1			
		2321	FLT1		1			
	Module 6-2	2967	GTF2H3	/	2	4		
		8805	TRIM24		1			
		6882	TAF11		1			
7	Module 7-1	5781	PTPN11	RTK signaling pathway	3	8	8	2.03E-06
		2064	ERBB2		2			
		3667	IRS1		1			
		8826	IQGAP1		1			
		7465	WEE1		1			
	Module 7-2	3845	KRAS	RTK signaling pathway	3	17		
		4763	NF1		9			
		5290	PIK3CA		5			

Note:

¹Gene symbols marked in blue indicate that the gene is annotated in the GBM signaling pathway (column “Pathway”) identified by the original TCGA analysis.

²#N denotes the number of glioblastoma samples that contain alterations of the individual genes in the module.

³#M denotes the number of glioblastoma samples that contain the alterations of the Module.

⁴#S denotes the number of glioblastoma samples that contain the alterations from both modules.

Supplementary Table 2- Candidate driver genes of GBM

Entrez Gene ID	Gene symbol	Sample [#]	Cancer [*]	GBM ^{**}	F-census [§]	Core pathway ^{&}
506	ATP5B	2	22	1	No	No
920	CD4	2	18330	65	No	No
1017	CDK2	2	2515	6	Yes	No
1027	CDKN1B	2	3325	11	Yes	No
1029	CDKN2A	27	7141	60	Yes	Yes
1431	CS	2	21557	36	No	No
1432	MAPK14	1	76	2	Yes	No
1608	DGKG	1	1	0	No	No
1956	EGFR	34	13379	273	Yes	Yes
2033	EP300	1	200	2	Yes	No
2045	EPHA7	1	29	1	Yes	No
2064	ERBB2	2	14558	16	Yes	Yes
2260	FGFR1	1	590	5	Yes	No
2321	FLT1	1	204	1	Yes	No
2967	GTF2H3	2	1	0	No	No
3118	HLA-DQA2	1	0	0	No	No
3122	HLA-DRA	2	121	0	No	No
3320	HSP90AA1	1	6	1	Yes	No
3458	IFNG	2	83	1	Yes	No
3667	IRS1	1	434	1	No	No
3710	ITPR3	1	2	0	No	No
3845	KRAS	3	2514	5	Yes	Yes
4194	MDM4	5	168	5	Yes	Yes
4507	MTAP	1	136	1	Yes	No
4763	NF1	9	2155	8	Yes	Yes
5037	PEBP1	2	94	0	Yes	No
5156	PDGFRA	10	1565	19	Yes	Yes
5287	PIK3C2B	4	14	2	No	No
5290	PIK3CA	5	781	3	Yes	Yes
5295	PIK3R1	5	41	2	Yes	Yes
5580	PRKCD	1	231	0	Yes	No
5728	PTEN	22	4542	129	Yes	Yes
5781	PTPN11	3	515	2	Yes	No
5925	RB1	8	1163	5	Yes	Yes
6197	RPS6KA3	1	0	0	Yes	No
6872	TAF1	2	21	0	Yes	No
6882	TAF11	1	1	1	No	No
7157	TP53	22	5123	58	Yes	Yes
7465	WEE1	1	195	0	Yes	No

8805	TRIM24	1	13	0	Yes	No
8826	IQGAP1	1	85	1	No	No
116986	AGAP2	14	18	1	Yes	No

Note:

[#]Sample indicates the number of samples harboring alterations in the gene.

^{*}Cancer indicates the number of times the altered gene co-appeared with the term “cancer” in the articles from PubMed.

^{**}GBM indicates the number of times the altered gene co-appeared with the term “GBM” in the articles from PubMed.

[§]If the altered gene is recorded in the cancer gene database “F-census”, it is marked with “Yes”; otherwise, it is marked with “No”.

[&] Core pathway indicates the three critical GBM pathways analyzed in the original TCGA GBM study. If the altered gene appears in any of the three pathways, it is marked with “Yes”; otherwise, it is marked with “No”.

Supplementary Table 3- Literature evidence for candidate cancer driver genes

Gene Symbol	Description from papers	PMID*
ATP5B	ATP5B co-regulated with ANT1 in glioblastoma.	22978616
CD4	It has been reported that CD4 may function as an important mediator of immune-mediated diseases of the central nervous system, and CD4 ⁺ T cells were required for GBM regression and immunological memory	19143470/18490770/2872572/22485134/16631933/8592229
CDK2	Assessment of gene expression in TCGA-derived GBMs revealed overexpression of MRC cancer genes AURKB, BIRC5, CCNB1, CCNB2, CDC2, CDK2, and FOXM1, which form a transcriptional network important for G2/M progression and/or checkpoint activation.	19139420/21807073/17289901/15531918
CDKN1B	CDKN1B decreased with anaplasia and almost disappeared in glioblastomas.	12133571/16805985
CDKN2A	The most common homozygously deleted region contained CDKN2A/CDKN2B (p15 and p16) occurring in 29% of GBM cases. The CDKN2A/CDKN2B locus was deleted in 46.4% of the combined cases.	20212223/22736234/22711607/22086906/22046342/21987724/21713760/21472719/20822523/18981259/17465990/17121137
CS	RA, together with Cs, Cl and CP, might be the new leaders in the evaluation of brain tumors.	23085868
MAPK14	The upregulation of PTEN gene was correlated with the downregulation of numerous genes including Akt, JUN, MAPK14 in glioma cells.	20436671/22185703/21663587
DGKG	/	/
EGFR	Our experiments found that miR-34a was often deleted and epidermal growth factor receptor (EGFR) was frequently amplified in genomic DNA of 55 GBMs using single-nucleotide polymorphism DNA microarray. Notably, we found that the mean survival time was significantly shortened for patients whose GBMs had both EGFR amplification and miR-34a deletion.	22580610/23043252/23029035/23012408/222930388/22737970/22736234/22588883/22472960/22348136/22323597/22232519/22162832/22089350/21975932
EP300	The gene EP300 is a driver in GBM, which could predict survival for patients with GBM.	21489305/21807073/21304179
EPHA7	Overexpression of EPHA7 protein was predictive of the adverse outcome in GBM patients, independent of MVD expression. Moreover, high density of MVD as well as higher EphA7 expression predicted the disease outcome more accurately than EphA7 variable alone	18366728
ERBB2	Compensatory activation of related ERBB family	22745588/21827413/

	receptors (ERBB2 and ERBB3) was observed in GBM CSCs deprived of EGFR signal. Dual inhibition of EGFR and ERBB2 with lapatinib significantly reduced GBM proliferation in colony formation assays compared to cetuximab-mediated EGFR-specific inhibition.	19793689/17571214/17536308/17457042/15256472/14965445/12902879/12532415/7760096
FGFR1	Studied key genes through GO-analysis, pathway-analysis and in the Me-CCNU-related signal transduction networks, 25 core genes that influenced chemosensitivity of GBM to Me-CCNU were obtained, including TP53, FGFR1, MAP2K2, EP300, PRKCA, CCND1, AKT2, RBL1, RAF1, CDKN2C and so on.	21807073/22837387/19340397
FLT1	A prerequisite for tumor angiogenesis, besides the expression of the angiogenic VEGF by tumor and stromal cells, is the expression of the VEGF receptors FLT-1 and KDR in rat C6 glioblastoma cells.	9393770
GTF2H3	/	/
HLA-DQA2	/	/
HLA-DRA	/	/
HSP90AA1	We investigated protein expression between the four regions of glioblastoma on clinically relevant biopsies from 5 patients. We identified 584 non-redundant proteins and 31 proteins were found to be up-regulated in the tumor region compared to the peri-tumoral brain tissue, among which, 24 proteins belong to an interaction network linked to 4 biological processes. The core of this network is mainly constituted of interactions between beta-actin (ACTB) with heat shock proteins (HSP90AA1, HSPA8) and 14-3-3 proteins (YWHAZ, YWHAG, YWHAB).	22575386/22952576
IFNG	ATRA plus IFN- γ could significantly decrease cell viability and increase morphological features of apoptosis in glioblastoma cells lines.	17960384/22185703
IRS1	For glioblastoma, we identified alterations of critical genes in the TP53 pathway (TP53, MDM2, and MDM4), the RB1 pathway (RB1, CDK4, and CDKN2A), and the PI3K/PTEN pathway (PIK3CA, PIK3R1, PTEN, and IRS1).	18772396
ITPR3	/	/
KRAS	We show here that Ink4a-Arf deficiency allows for GBM formation from astrocytes and that it enhances tumor incidence in neural progenitor cells. Furthermore, KRAS alone can cooperate with deletion of the Ink4a-Arf locus in tumor formation from both neural	12359767/21949886

	progenitor cells and astrocytes.	
MDM4	Illumina Bead Arrays were used to assay 22 GBMs and Digital Karyotyping was used on 8 GBM cell lines and one primary sample. The common amplifications we observed for all 31 samples was GLI/CDK4 (22.6%), MDM2 (12.9%) and PIK3C2B/MDM4 (12.9%).	19609742/21807073/16319692
MTAP	This finding suggests that all three genes (CDKN2A/p16, CDKN2B/p15, and MTAP) may be inactivated in glioblastomas by a large deletion event.	21884817
NF1	Humans with mutations in NF1 develop neurofibromatosis type I (NF1) and have increased risk of optic gliomas, astrocytomas and glioblastomas	10973261/23045694/2539962/22086906/21889780/20442305/20129251/19915670/18975243
PEBP1	/	/
PDGFRA	Using primary GBM cells maintained under neurosphere conditions, we then demonstrated that miR-34a specifically affects growth of proneural glioma cells in vitro and in vivo. Further bioinformatic analysis identified PDGFRA as a direct target of miR-34a and this interaction was experimentally validated.	22479456/23029035/22747609/22661320/22661233/22323597/21880180/22086906/20129251/19915670/19609742/18816605/19967449/17504929/17002787
PIK3C2B	PIK3C2B has a crucial role in the PI3K signaling pathway involved in the regulation of cell proliferation in GBM patients	21861842 /19189657/19609742
PIK3CA	The highest frequencies of gains were detected on PIK3CA (64.3%), EGFR (57.1%), CSE1L (57.1%), NRAS (50%), MYCN (42.9%), FGR (35.7%), ESR (35.7%), PGY1 (35.7%), and D17S167 (35.7%). These genes are suggested to be involved in the GBM tumorigenesis.	11351043/22930388/2064833/22026810
PIK3R1	We mapped alterations in each of these pathways and found that they included the catalytic PIK3CA and regulatory PIK3R1 subunit genes of the class IA PI3K. Knockdown of either of these genes separately in GBM cell lines by lentiviral-mediated shRNA expression resulted in decreased proliferation, migration, and invasion in all lines tested.	22064833/21663587
PRKCD	/	/
PTEN	PTEN appears to be the major target of inactivation on chromosome 10q in glioblastoma multiforme.	9331071/22479427/20736378/20462843/19150964/19276385

PTPN11	Inhibition of SHP-2 expression by Shp-2 siRNA inhibited cell growth, transformation and altered morphology of these EGFRvIII transformed GBM cells.	19427850/21934682
RB1	We found that abnormalities in any of the four genes (CDKN2A, CDKN2B, RB1, and CDK4) coding for components of the Rb1 pathway were associated with shorter survival in glioblastomas patients.	14519639/18816605/2157621/21555372
RPS6KA3	/	/
TAF1	/	/
TAF11	TAF11 was overexpressed in glioblastomas relative to non-neoplastic brain tissue as indicated by microarray analysis	18398573
TP53	TP53 mutations were found in 12 out of 54 (22%) GBMs of short-term survivors and 8 out of 35 (23%) tumors of long-term survivors.	11519857/22661320/23034333/22886134/2287028/21528672/21483692/20593219/18948956/18383819/11083071/12622447
WEE1	We analyze the kinase gene expression profiles of various tumor types and reveal the WEE1 kinase to be overexpressed in glioblastomas. We demonstrate that WEE1 is a major regulator of the G2 checkpoint in glioblastoma cells. Our results suggest that inhibition of WEE1 kinase holds potential as a therapeutic approach in treatment of glioblastoma	20832752
TRIM24	/	/
IQGAP1	We confirmed that IQGAP1 is a reliable marker that may help to distinguish oligodendroglioma from glioblastoma. Although in both tumors IQGAP1 is expressed by endothelial cells, only in glioblastoma it specifies a population of amplifying tumor cells.	16982749/21196113
AGAP2	We also identify new candidate drivers in GBM, including AGAP2/CENTG1, a putative oncogene and an activator of the PI3K pathway;	20169195

Note:

* The column of PMID lists some of the PubMed IDs of papers supporting that the candidate driver genes may play roles in carcinogenesis of glioblastomas (GBM). The PubMed ID marked with blue color denotes that the detailed description in this paper is provided in the second column.