Supplemental Data

Table S1, related to Figure 1, provided as an Excel file

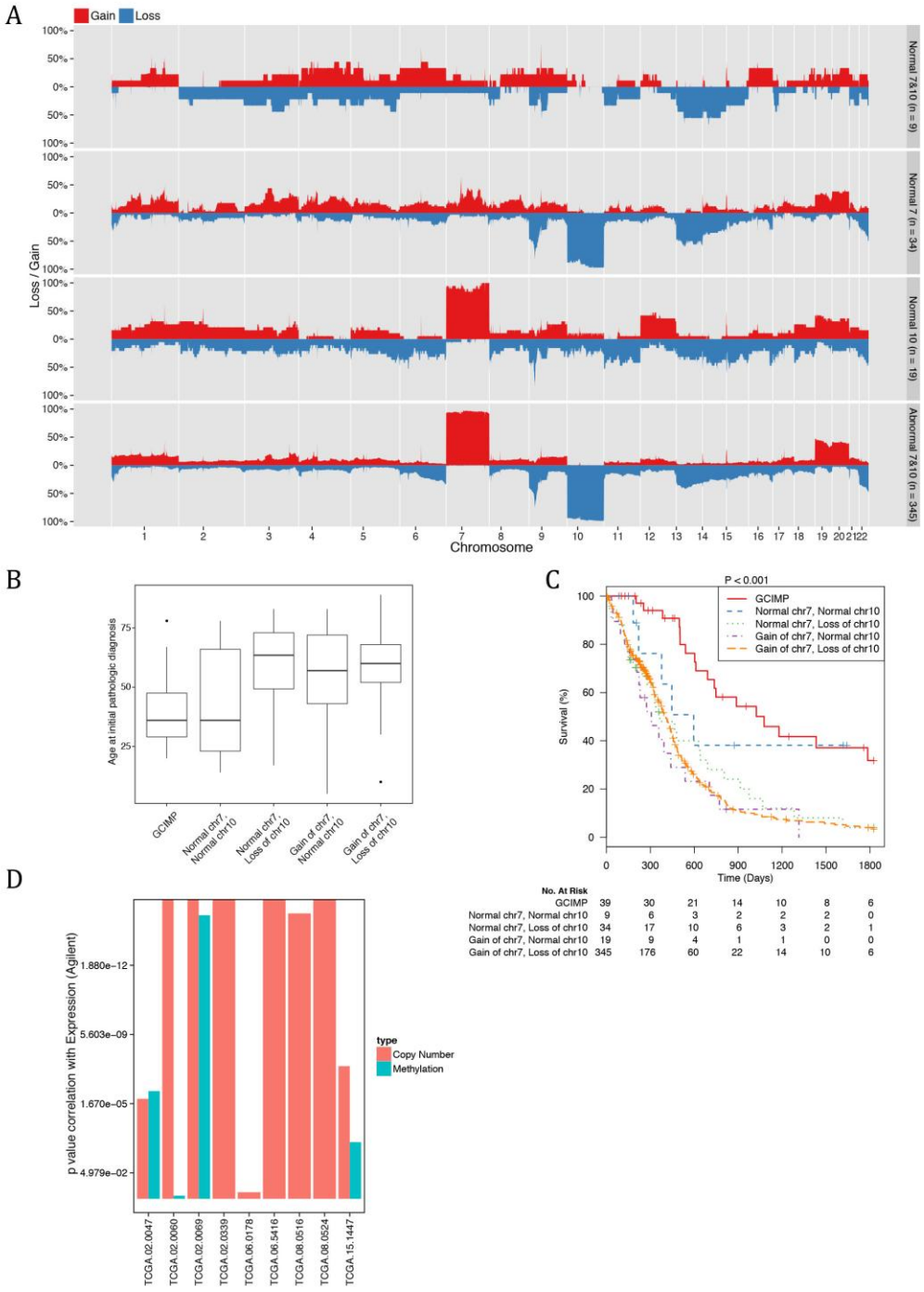
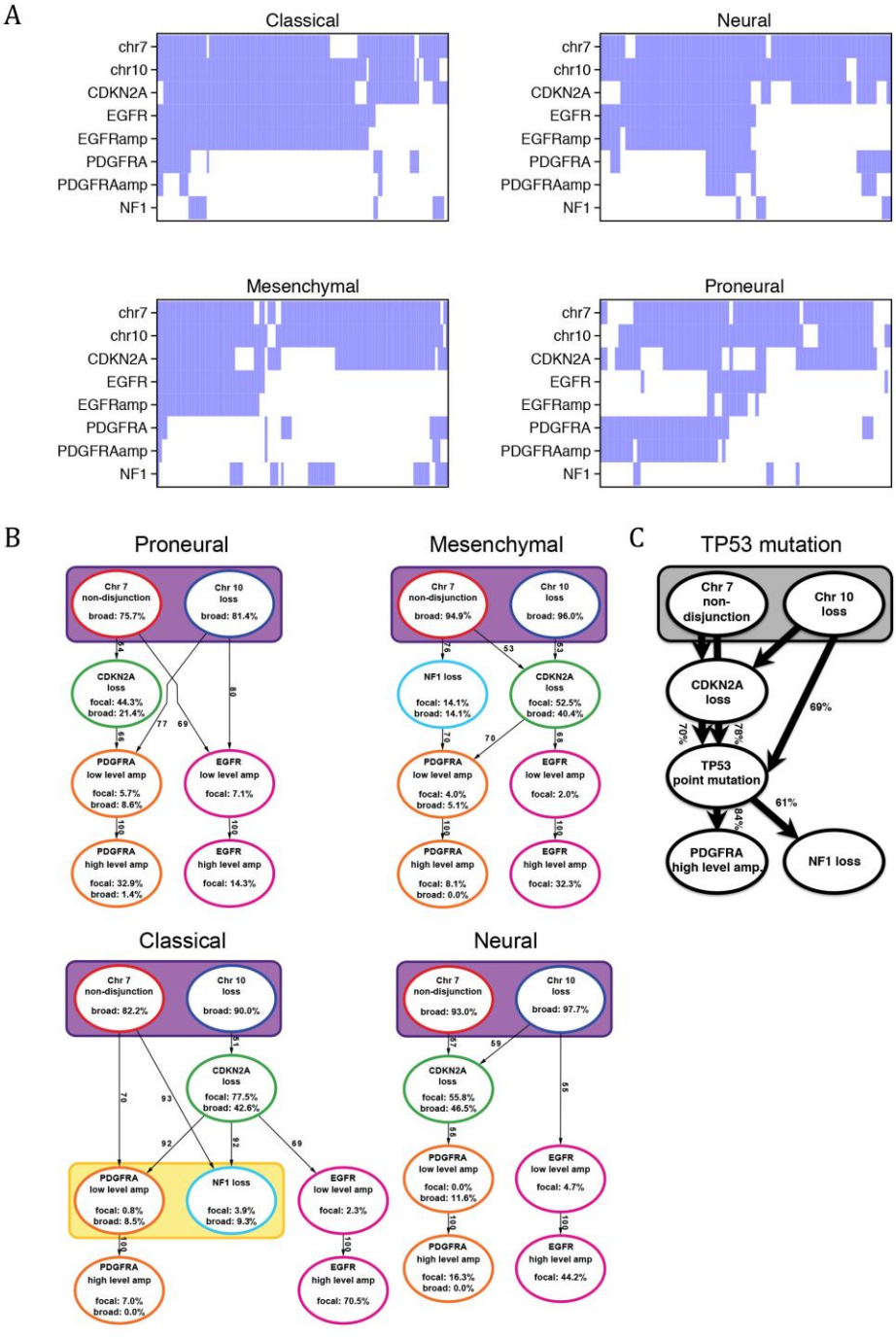


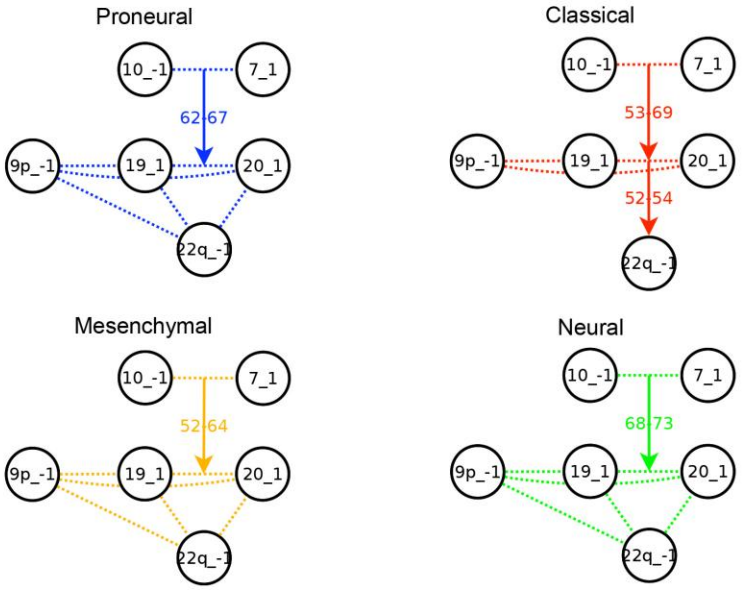
Figure S1, related to Figure 1. Stratification of non-GCIMP GBMs by the status of chr7 and chr10 (A) Somatic copy number alterations and their frequencies in non-GCIMP GBMs. The genome plot visualizes the frequencies of copy number gains (red) and losses (blue) along the genome in non-GCIMP GBMs with or without abnormal chr7 and/or chr10. See also **Table S2 and 3**. (B) Age distribution in GCIMP and non-GCIMP GBMs. Age distribution of each group was analyzed at initial pathologic diagnosis. Whiskers of all box plots extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. (C) Overall survival in GBM subgroups. Kaplan-Meier plots are shown for GCIMP and non-GCIMP GBMs stratified by the status of chr7 and chr10. Survival times were censored at 5 years. (D) Correlation of promoter methylation and copy number data with expression data. Two samples showed very low correlation coefficients ($p > 0.05$), indicative of a misalignment of data types (sample swaps).

Table S2, related to Figure S1, provided as an excel file

Table S3, related to Figure S1, provided as an excel file


A
Classical
Neural
Mesenchymal
Proneural
chr7
chr10
CDKN2A
EGFR
EGFRamp
PDGFRA
PDGFRAamp
NF1
B
Proneural
Chr 7 non-disjunction
broad: 75.7%
Chr 10 loss
broad: 81.4%
CDKN2A loss
focal: 44.3%
broad: 21.4%
PDGFRA low level amp
focal: 5.7%
broad: 8.6%
EGFR low level amp
focal: 7.1%
PDGFRA high level amp
focal: 32.9%
broad: 1.4%
EGFR high level amp
focal: 14.3%
Mesenchymal
Chr 7 non-disjunction
broad: 94.9%
Chr 10 loss
broad: 96.0%
NF1 loss
focal: 14.1%
broad: 14.1%
CDKN2A loss
focal: 52.5%
broad: 40.4%
PDGFRA low level amp
focal: 4.0%
broad: 5.1%
EGFR low level amp
focal: 2.0%
PDGFRA high level amp
focal: 8.1%
broad: 0.0%
EGFR high level amp
focal: 32.3%
Classical
Chr 7 non-disjunction
broad: 82.2%
Chr 10 loss
broad: 90.0%
CDKN2A loss
focal: 77.5%
broad: 42.6%
PDGFRA low level amp
focal: 0.8%
broad: 8.5%
NF1 loss
focal: 3.9%
broad: 9.3%
EGFR low level amp
focal: 2.3%
PDGFRA high level amp
focal: 7.0%
broad: 0.0%
EGFR high level amp
focal: 70.5%
Neural
Chr 7 non-disjunction
broad: 93.0%
Chr 10 loss
broad: 97.7%
CDKN2A loss
focal: 55.8%
broad: 46.5%
PDGFRA low level amp
focal: 0.0%
broad: 11.6%
EGFR low level amp
focal: 4.7%
PDGFRA high level amp
focal: 16.3%
broad: 0.0%
EGFR high level amp
focal: 44.2%
C
TP53 mutation
Chr 7 non-disjunction
Chr 10 loss
CDKN2A loss
TP53 point mutation
PDGFRA high level amp.
NF1 loss
69%
61%

D


Proneural
10_-1
7_1
62
67
9p_-1
19_1
20_1
22q_-1
Classical
10_-1
7_1
53
69
9p_-1
19_1
20_1
52
54
22q_-1
Mesenchymal
10_-1
7_1
52
64
9p_-1
19_1
20_1
22q_-1
Neural
10_-1
7_1
68
73
9p_-1
19_1
20_1
22q_-1

E

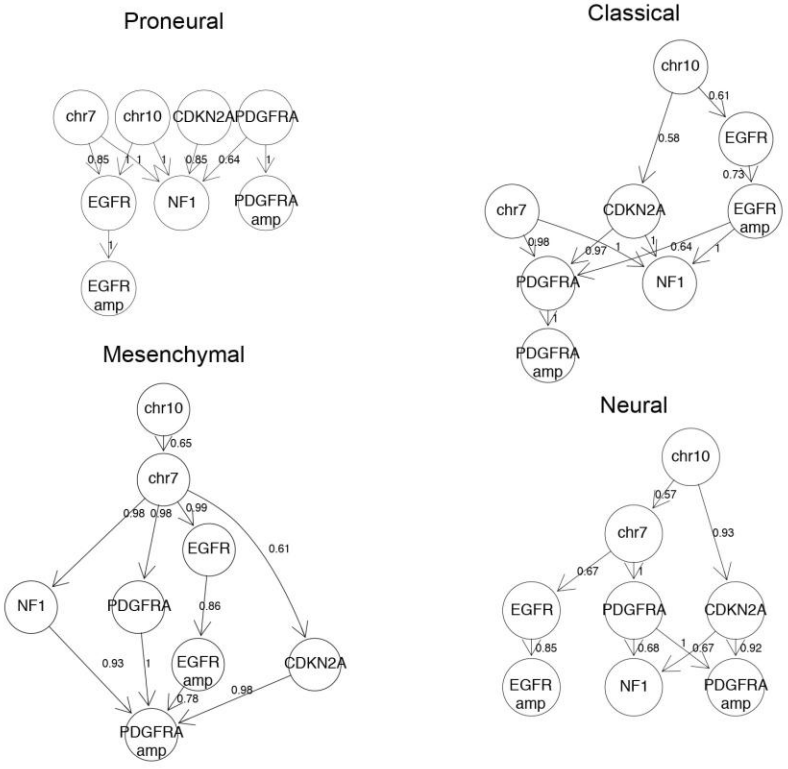

Proneural
chr7
chr10
CDKN2A
PDGFRA
0.85
1
1
1
0.85
0.64
1
EGFR
NF1
PDGFRA amp
1
EGFR amp
Classical
chr10
0.61
0.58
EGFR
0.73
chr7
CDKN2A
EGFR amp
0.98
0.97
1
1
0.64
1
PDGFRA
NF1
1
PDGFRA amp
Mesenchymal
chr10
0.65
chr7
0.98
0.98
0.99
EGFR
0.61
NF1
PDGFRA
0.86
0.93
1
EGFR amp
CDKN2A
0.78
0.98
PDGFRA amp
Neural
chr10
0.57
0.93
chr7
0.67
1
EGFR
PDGFRA
CDKN2A
0.85
0.68
1
0.67
0.92
EGFR amp
NF1
PDGFRA amp

Figure S2, related to Figure 2. Temporal sequence of somatic copy number alterations within GBM subtypes. (A) The heatmap of copy number alterations in all four subtypes. Two whole chromosome events as well as subtype-defining events (focal *EGFR* gain and amplification, *PDGFRA* gain and amplification, and *NF1* loss) were plotted in the figure and were then used for subsequent RESIC analysis. (B) RESIC order of events for subtype-specific copy number alterations per subtype. (C) RESIC order of events for all GBM samples for which *TP53* point mutation data was available (n=85). (D) RESIC order of chromosome level copy number alterations. The percent of samples with broad and focal amplifications is included per alteration. Percentages on arrows indicate percent of bootstrap iterations where the depicted order is dominant. Colored rectangles denote pairs of alterations where significant correlations occur, but the order of events could not be determined by RESIC. (E) Temporal sequence of somatic copy number alterations in all four subtypes by CT-CBN analysis.

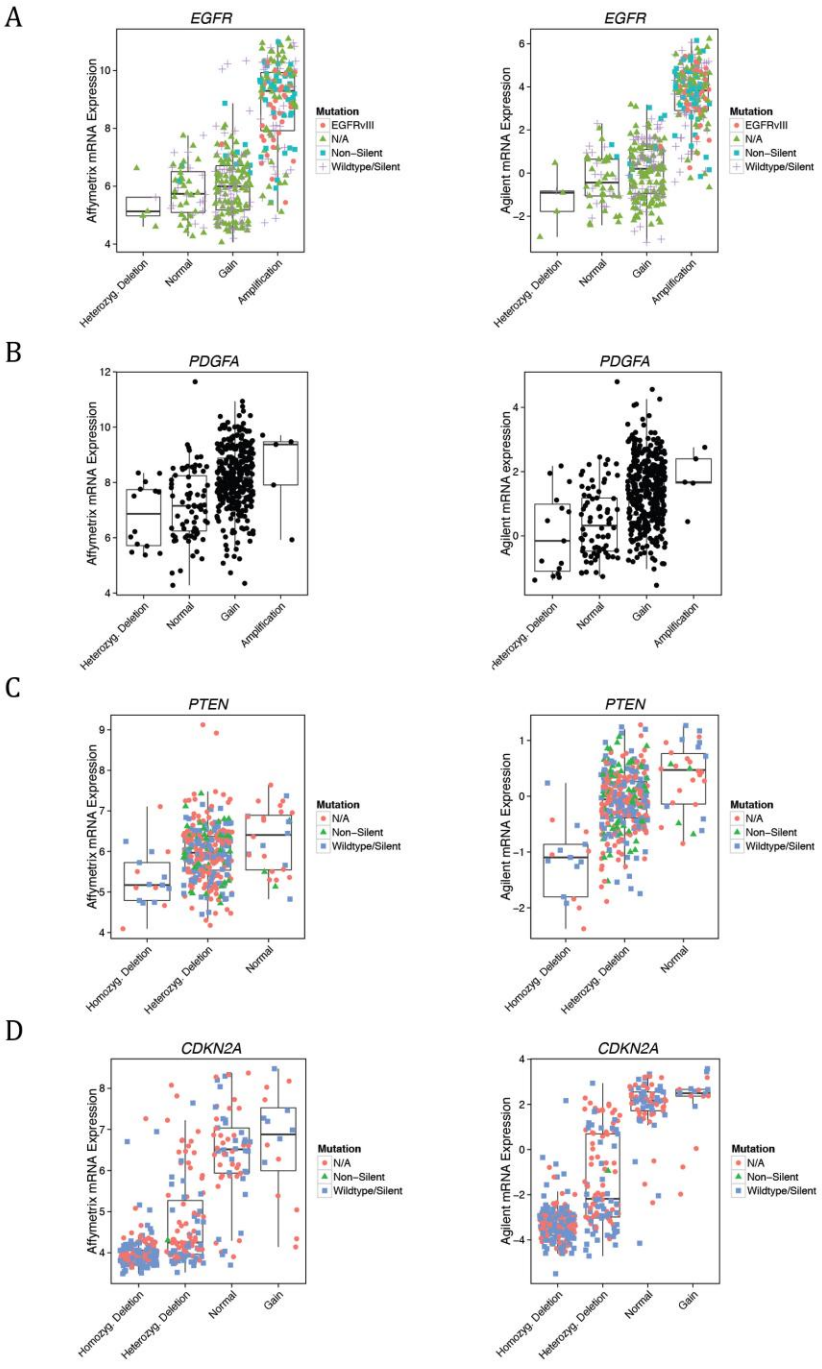

A
EGFR
Affymetrix mRNA Expression
Agilent mRNA Expression
Mutation
EGFRvIII
N/A
Non−Silent
Wildtype/Silent
Heterozyg. Deletion
Normal
Gain
Amplification
B
PDGFA
Agilent mRNA expression
C
PTEN
Homozyg. Deletion
D
CDKN2A

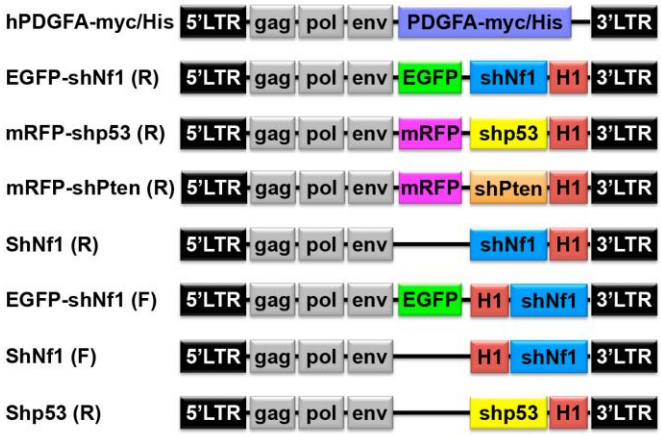
Figure S3, related to Figure 3. (A-D) Correlation between mRNA expression and copy number of the *EGFR* (A), *PDGFA* (B), *PTEN* (C) and *CDKN2A* (D) gene in non-GCIMP GBMs. For each copy number category (homozygous loss, heterozygous loss, normal, gain and amplification) on the x-axis, the mRNA expression measured on the Affymetrix (left panel) and Agilent (right panel) platform is plotted for each patient (symbols) on the y-axis. For the *EGFR* gene (**Fig. S3A**), the colors and shapes of the symbols visualize somatic mutation information (legend). Not available (N/A) refers to samples without somatic mutation information and negative or unavailable (due to missing aCGH data) *EGFRvIII* status. Two samples had a silent mutation of *EGFR* and were combined with the wild-type samples for easier readability. The expression distributions are further plotted in the background as box plots for each copy number category. See also **Table S1**. For the *PTEN* gene (**Fig. S3C**), the status of PTEN mRNA was analyzed with a two-sided student t-test between samples with normal, hetero- and homozygous loss in all non-GCIMP GBM samples. The level of PTEN mRNA expression decreased depending on the gene dosage. The analysis was performed on the TCGA data from two mRNA expression platforms, Affymetrix; Homozyg vs Heterozyg: $p = 0.002258$, Heterozyg vs Normal: $p = 0.0429$, Agilent; Homozyg vs Heterozyg: $p = 4.18e\text{-}06$, Heterozyg vs Normal: $p = 0.0007209$. Similarly, for the *CDKN2A* gene (**Fig. S3D**) we find highly significant expression differences across copy number categories. Affymetrix; Homozyg vs Heterozyg: $p < 0.0001$, Heterozyg vs Normal: $p < 0.0001$, Agilent; Homozyg vs Heterozyg: $p < 0.0001$, Heterozyg vs Normal: $p < 0.0001$. Whiskers of all box plots extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

Table S4, related to Figure 3, provided as an Excel file

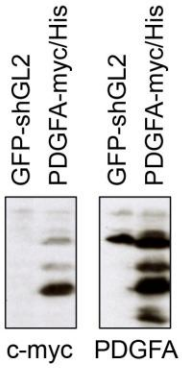
Table S5, related to Figure 3, provided as an Excel file

Table S6, related to Figure 3, provided as an Excel file

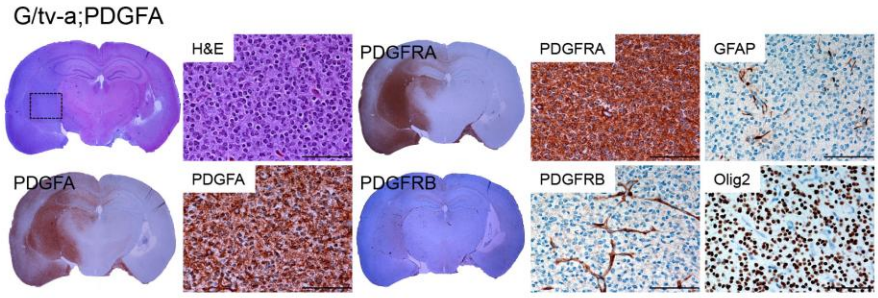
A

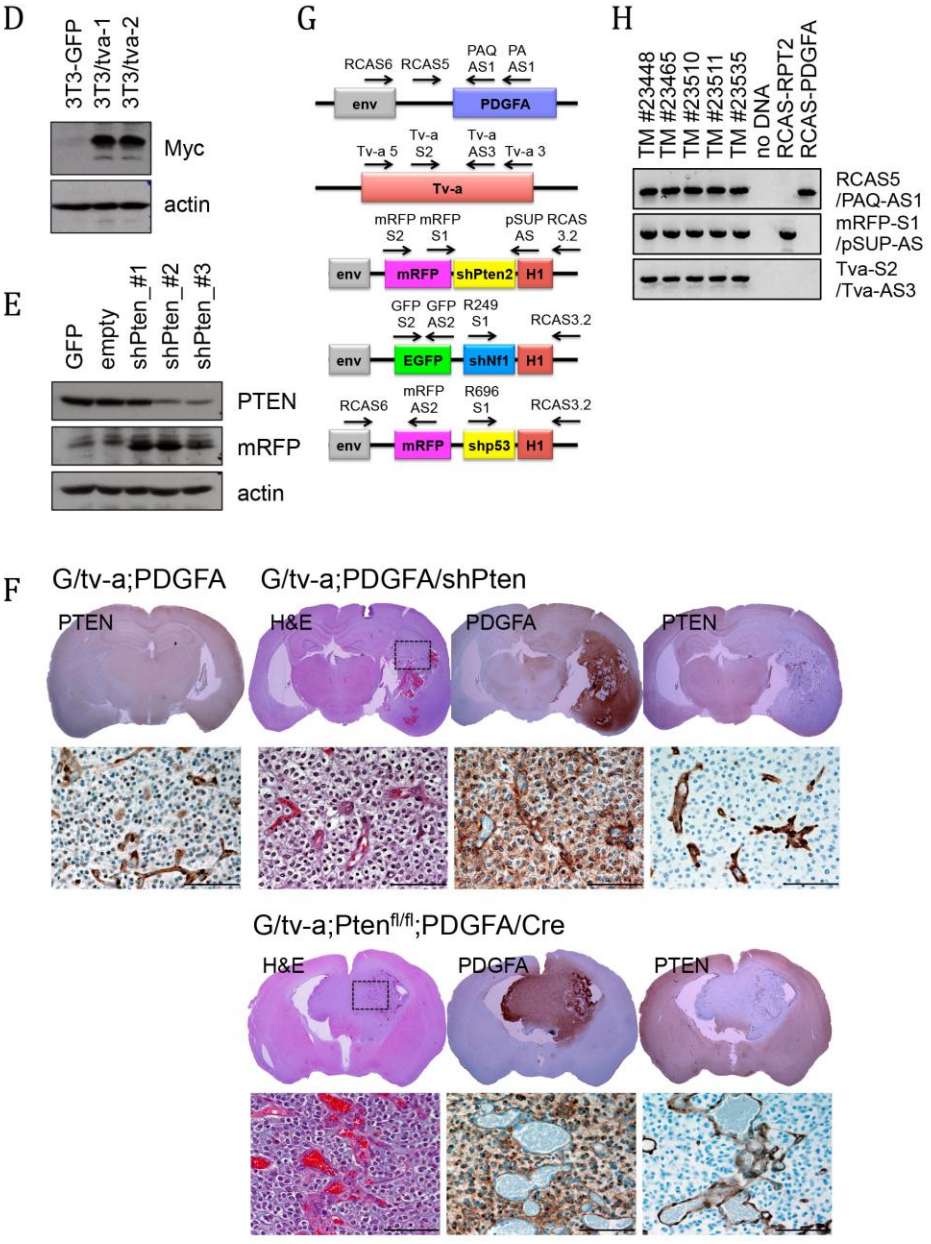

hPDGFA-myc/His
5'LTR
gag
pol
env
PDGFA-myc/His
3'LTR
EGFP-shNf1 (R)
EGFP
shNf1
H1
mRFP-shp53 (R)
mRFP
shp53
mRFP-shPten (R)
shPten
ShNf1 (R)
EGFP-shNf1 (F)
ShNf1 (F)
Shp53 (R)

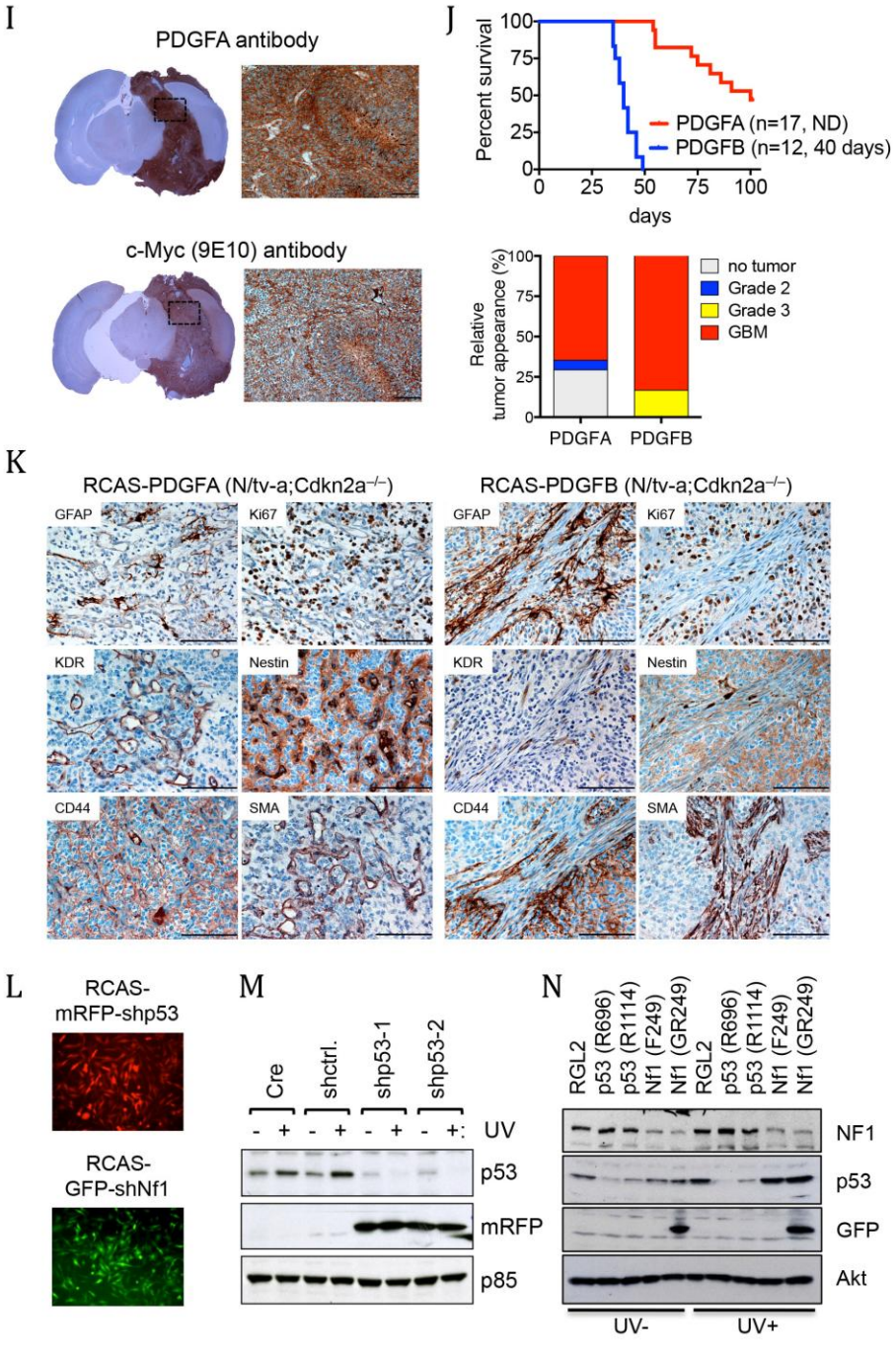
B


GFP-shGL2
PDGFA-myc/His
GFP-shGL2
PDGFA-myc/His
c-myc
PDGFA

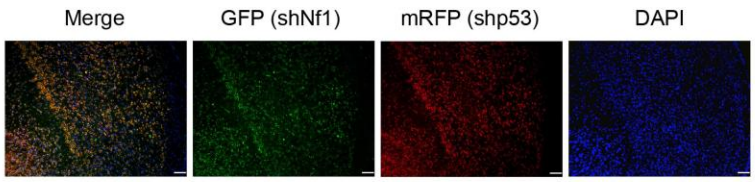
C


G/tv-a;PDGFA
H&E
PDGFRA
PDGFRA
GFAP
PDGFA
PDGFA
PDGFRB
PDGFRB
Olig2

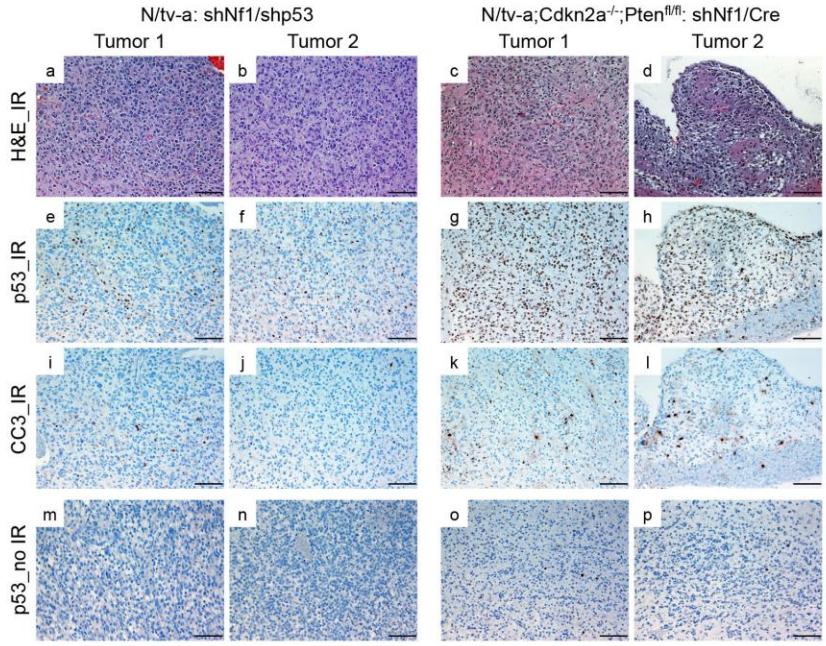

D
3T3-GFP
3T3/tva-1
3T3/tva-2
Myc
actin
E
GFP
empty
shPten_#1
shPten_#2
shPten_#3
PTEN
mRFP
actin
G
RCAS6
RCAS5
PAQ AS1
PA AS1
env
PDGFA
Tv-a 5
Tv-a S2
Tv-a AS3
Tv-a 3
Tv-a
mRFP S2
mRFP S1
pSUP AS
RCAS 3.2
mRFP
shPten2
H1
GFP S2
GFP AS2
R249 S1
RCAS3.2
EGFP
shNf1
mRFP AS2
R696 S1
shp53
H
TM #23448
TM #23465
TM #23510
TM #23511
TM #23535
no DNA
RCAS-RPT2
RCAS-PDGFA
RCAS5 /PAQ-AS1
mRFP-S1 /pSUP-AS
Tva-S2 /Tva-AS3
F
G/tv-a;PDGFA
G/tv-a;PDGFA/shPten
PTEN
H&E
PDGFA
PTEN
G/tv-a;Ptenfl/fl;PDGFA/Cre
H&E
PDGFA
PTEN


I
PDGFA antibody
c-Myc (9E10) antibody
J
Percent survival
PDGFA (n=17, ND)
PDGFB (n=12, 40 days)
days
Relative tumor appearance (%)
no tumor
Grade 2
Grade 3
GBM
PDGFA
PDGFB
K
RCAS-PDGFA (N/tv-a;Cdkn2a−/−)
RCAS-PDGFB (N/tv-a;Cdkn2a−/−)
GFAP
Ki67
KDR
Nestin
CD44
SMA
L
RCAS-mRFP-shp53
RCAS-GFP-shNf1
M
Cre
shctrl.
shp53-1
shp53-2
UV
p53
mRFP
p85
N
RGL2
p53 (R696)
p53 (R1114)
Nf1 (F249)
Nf1 (GR249)
NF1
p53
GFP
Akt
UV-
UV+

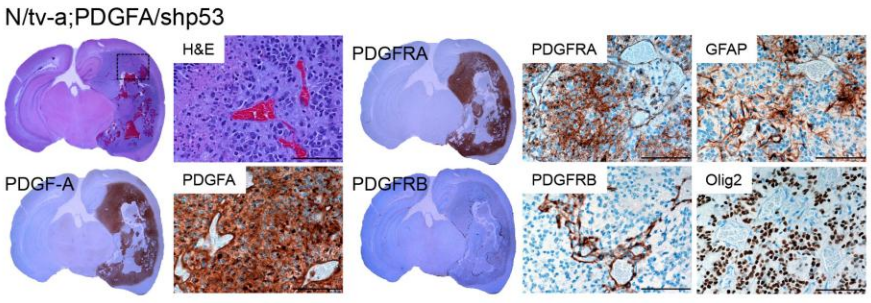
O


Merge
GFP (shNf1)
mRFP (shp53)
DAPI

P


N/tv-a: shNf1/shp53
N/tv-a;Cdkn2a-/-;Ptenfl/fl: shNf1/Cre
Tumor 1
Tumor 2
Tumor 1
Tumor 2
H&E_IR
p53_IR
CC3_IR
p53_no IR
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p

Q


N/tv-a;PDGFA/shp53
H&E
PDGFRA
PDGFRA
GFAP
PDGF-A
PDGFA
PDGFRB
PDGFRB
Olig2

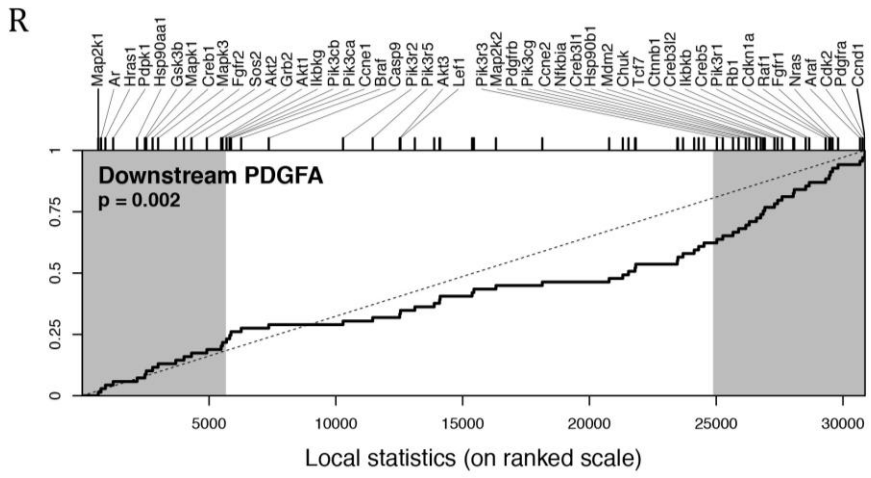
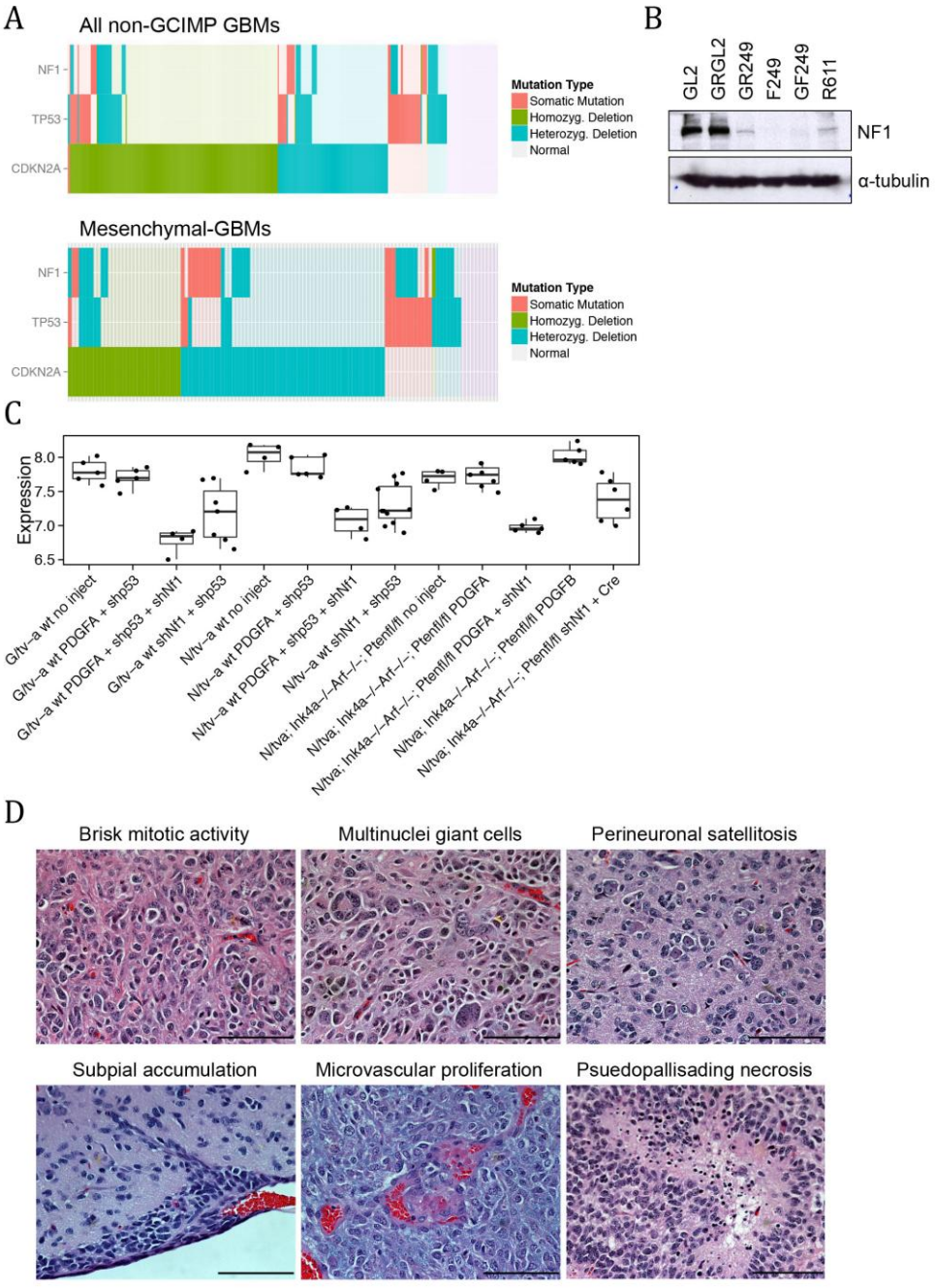


Figure S4, related to Figure 4. **Functional validation of RCAS constructs** (A) Schematic of the RCAS constructs. R and F denote reverse and forward direction of H1 promoter toward 5'LTR of the RCAS vector, respectively. (B) RCAS-hPDGFA-myc/6xHis vector expression in DF-1 cells. DF-1 cells were transfected with RCAS-GFP-shGL2 or RCAS-hPDGFA-myc/6xHis vector. Western blots for c-myc tag and PDGFA are shown. (C) Immunohistochemical analysis of the RCAS-PDGFA-induced glioma in *G/tv-a* mice. Tumors were generated with the injection of RCAS-PDGFA virus into neonatal pups brain. Representative H&E staining and immunostaining for the indicated antibody are shown in the figure. Boxes denote the enlarged-region. Scale bars, 100 μm. (D) Tv-a-myc/6xHis protein expression in NIH-3T3 cells. NIH-3T3 cells were lentivirally transduced with pLJM1-GFP or pLJM1-tv-a-myc/6xHis vector. Western blots for c-myc tag and actin are shown. (E) RCAS-mRFP-shPten vector-mediated *Pten* knockdown in NIH-3T3/tv-a-myc/6xHis cells. The cells were retrovirally infected with RCAS-shPten viruses targeting three different sequences (#1-3), respectively. GFP and empty denote NIH-3T3-GFP and NIH-3T3/tv-a-myc/6xHis cells, respectively. Western blots for PTEN, mRFP and actin are shown. (F) Immunohistochemical analysis of the RCAS-PDGFA/shPten#2 (RPT2)-induced glioma in *G/tv-a* mice (upper panel) and the RCAS-PDGFA/Cre-induced glioma in *G/tv-a;Pten*$^{fl/fl}$ mice (lower panel). Tumors were generated with the co-injection of the RCAS-PDGFA and shPten#2 or Cre virus into neonatal pups brain. Representative H&E staining and immunostaining for the indicated antibody are shown in the figure. Boxes denote the enlarged-region. For a comparison of PTEN protein expression, PTEN staining of same *G/tv-a;PDGFA* tumor as **Fig. S4C** is also shown in the figure. The PTEN expression is restricted in the endothelium in the *Pten* knockdown and *Pten* gene-deleted tumors. By contrast, the nuclear and cytoplasmic staining in the tumor cells as well as the endothelium was detected in the *Pten* intact tumor as previously demonstrated (Huse et al., 2009). Scale bars, 100 μm. (G) Schematic of primer design for the RCAS integration analysis. (H) DNA integration of RCAS-PDGFA and RCAS-GW-mRFP-shPten#2 (RPT2) virus in the RCAS-PDGFA/RPT2-driven tumors (*G/tv-a* mice). Tumor DNAs were extracted from FFPE tissue sections and then subjected to nested-PCR analysis. Primer pairs: [RCAS6/PA-AS1 and RCAS5/PAQ-AS1], [Tv-a5/Tv-a3 and Tv-a-S2/Tv-a-AS3], [mRFP-S2/RCAS3.2 and mRFP-S1/pSUP-AS] were used for 1st and 2nd PCR of the PDGFA, Tv-a and mRFP-shPten#2 detection, respectively. RCAS-PDGFA and RCAS-mRFP-shPten#2 plasmids serve as a control. (I) RCAS-PDGFA expression in murine tumor. The PDGFA

expression was evaluated with immunohistochemisty for anti-PDGFA or c-Myc antibody in the RCAS-hPDGFA-myc/6xHis-driven GBM (*N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$ mice). Given that the PDGFA antibody detected the RACS-PDGFA expression more clear than the c-myc antibody, the PDGFA antibody was mainly used for the probe in this study. Boxes denote the enlarged-region. Scale bars, 100 μm. (J) Kaplan-Meier curves demonstrating symptom-free survival (Log rank test; $p < 0.0001$) (upper panel) and relative tumor grade (lower panel) for *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$ adult mice injected with the RCAS-PDGFA or PDGFB. The percentages of tumors exhibiting WHO grade II, grade III, and grade IV histological features are shown for each genotype. (K) Immunohistochemical analysis of the PDGFA (left panel) and PDGFB (right panel)-induced gliomas in *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$ mice. Representative H&E staining and immunostaining for the indicated antibodies are shown in the figure. Vascular histology of the RCAS-PDGFA-induced GBMs is more similar to that seen in human counterpart than the RCAS-PDGFB. The images were same enlarged-regions as the boxes in **Fig. 4D**. Scale bars, 100 μm. (L) mRFP or GFP expression in DF-1 cells infecting the RCAS-mRFP-shp53 (mR696) or RCAS-GFP-shNf1 (GR249) virus. Cells were observed with fluorescence microscope. (M) RCAS-mRFP-shp53 (mR696) vector-mediated *Tp53* knockdown in NIH-3T3/tv-a cells. The cells were retrovirally infected with RCAS-Cre, shControl or mRFP-shp53 (mR696, duplicate samples) virus. For p53 induction, cells were harvested after 6 hours of ultra violet (UV) exposure for 3 min and were then subjected to western blot analysis for p53 and mRFP. p85PI3K regulatory subunit serves as a loading control. (N) RCAS-shRNA vector-mediated *Tp53* or *Nf1* knockdown in NIH-3T3/tv-a cells. NIH-3T3/tv-a cells retrovirally infecting RCAS-shGL2 (RGL2), shp53 (R696, R1114) or shNf1 (F249, GR249) virus were harvested after UV exposure for p53 induction as well and were then subjected to western blot analysis for NF1, p53, GFP and Akt. Akt serves as a loading control. See also List of vector constructs (O) GFP and mRFP expression in *G/tv-a;GFP-shNf1/mRFP-shp53* tumor. GFP and mRFP expression was observed with immunofluorescent analysis. Representative images were shown in the figure. Scale bars, 100 μm. (P) Functional validation of RCAS-shp53 vector *in vivo*. Symptomatic mice from *N/tv-a;RCAS-shNf1/shp53* (a, b, e, f, i, j) and *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$*;RCAS-shNf1/Cre* (c, d, g, h, k, l) injection were sacrificed after 3 hours of irradiation. H&E (a-d) and immunohistochemistry for p53 (e-h) and Cleaved-Caspase 3 (CC3)(i-l) were shown in the figure. p53 and CC3 expression are obviously lower in tumors incorporating the RCAS-shp53 virus. Untreated-*N/tv-a;shNf1/shp53* (m and n) and *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$*;shNf1/Cre* (o and p) tumors serves as a control for p53 expression. Scale bars, 100 μm. (Q) Immunohistochemical analysis of the RCAS-PDGFA/shp53 (mR696)-induced gliomas in *N/tv-a* mice. Tumors were generated with the co-injection of RCAS-PDGFA and shp53 virus into neonatal pups brain. Representative H&E staining and immunostaining for the indicated antibody are shown in the figure. Boxes denote the enlarged-region. Scale bars, 100 μm. (R) GSEA of murine PDGFA-induced tumors. We examined if the downstream genes of PDGFA, which was identified as most possible driver gene on chr7 gain in computational analysis (**Fig 3C**), were enriched in the murine PDGFA-induced tumors. To this end, the gene expression profiles of these PDGFA downstream genes were analyzed with a GSEA in PDGFA/shp53 and shNf1/shp53-induced tumors in the combined *G* and *N/tv-a* mice data. X-axis denotes ranked-genes. The genes on the right in the plot present up-regulated genes in the murine PDGFA-induced glioma data, genes on the left-hand side the down-regulated genes. Grey areas represent the significantly differentially expressed genes. Y-axis denotes enrichment score. The dashed diagonal visualized the expected value for random gene rankings. Panel shows a highly non-random ($p = 0.002$) ranking of PDGFA downstream genes when comparing the PDGFA/shp53 and shNf1/shp53-induced tumors.

Table S7, related to Figure 4, provided as an excel file


A
All non-GCIMP GBMs
NF1
TP53
CDKN2A
Mutation Type
Somatic Mutation
Homozyg. Deletion
Heterozyg. Deletion
Normal
Mesenchymal-GBMs
NF1
TP53
CDKN2A
Mutation Type
Somatic Mutation
Homozyg. Deletion
Heterozyg. Deletion
Normal
B
GL2
GRGL2
GR249
F249
GF249
R611
NF1
α-tubulin
C
Expression
8.0
7.5
7.0
6.5
G/tv-a wt no inject
G/tv-a wt PDGFA + shp53
G/tv-a wt PDGFA + shp53 + shNf1
G/tv-a wt shNf1 + shp53
N/tv-a wt no inject
N/tv-a wt PDGFA + shp53
N/tv-a wt PDGFA + shp53 + shNf1
N/tv-a wt shNf1 + shp53
N/tva; Ink4a−/−Arf−/−; Ptenfl/fl no inject
N/tva; Ink4a−/−Arf−/−; Ptenfl/fl PDGFA
N/tva; Ink4a−/−Arf−/−; Ptenfl/fl PDGFA + shNf1
N/tva; Ink4a−/−Arf−/−; Ptenfl/fl PDGFB
N/tva; Ink4a−/−Arf−/−; Ptenfl/fl shNf1 + Cre
D
Brisk mitotic activity
Multinuclei giant cells
Perineuronal satellitosis
Subpial accumulation
Microvascular proliferation
Psuedopallisading necrosis

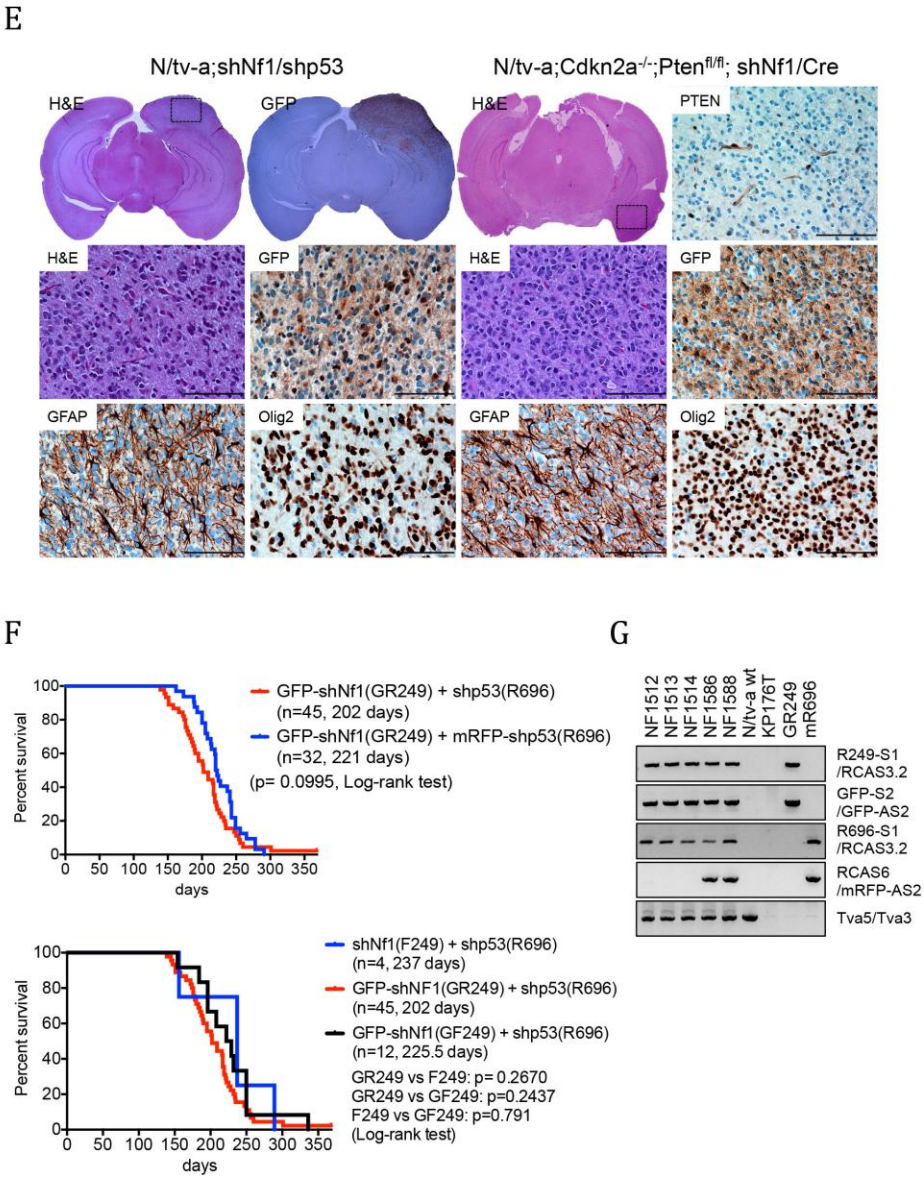
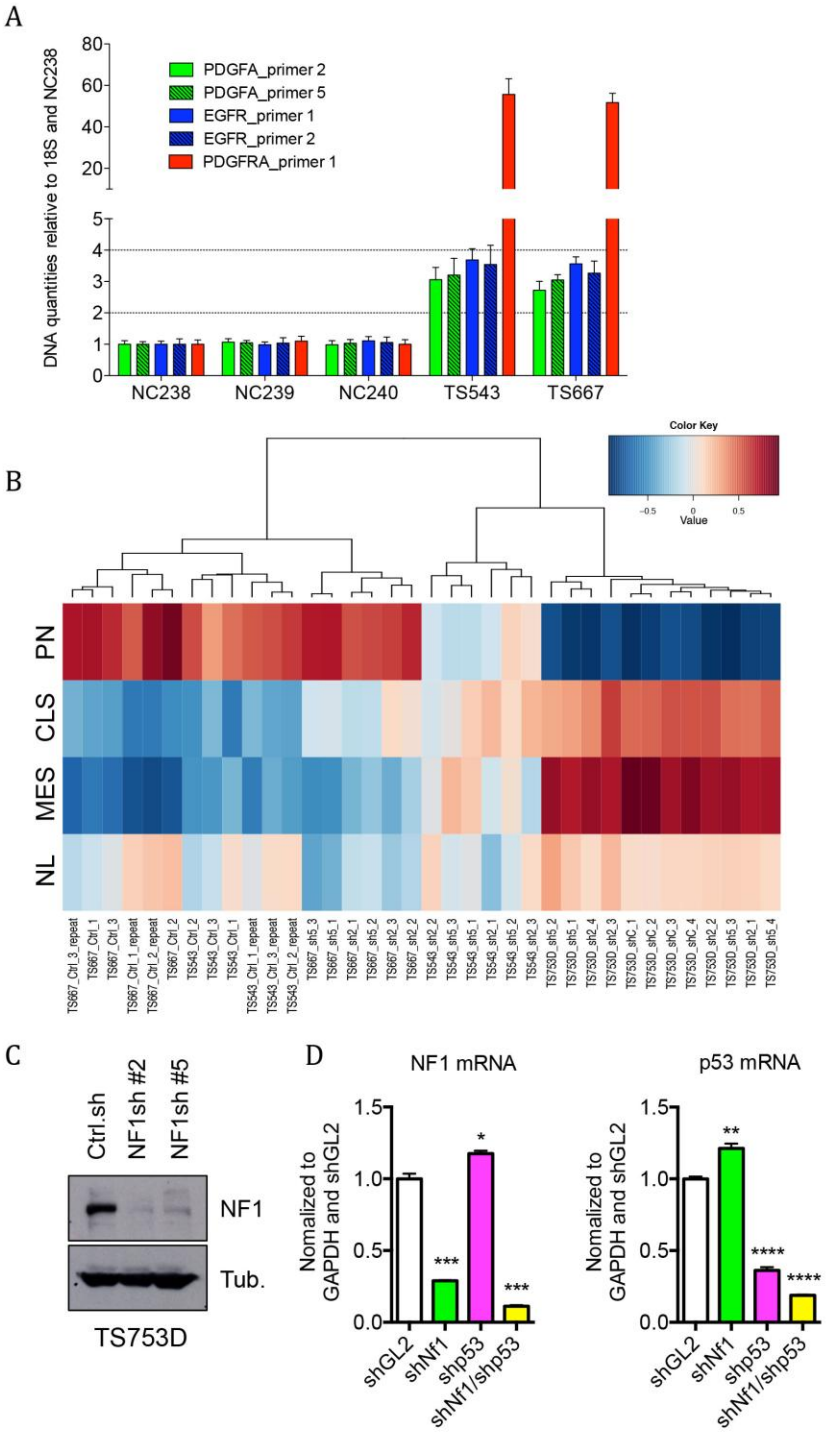
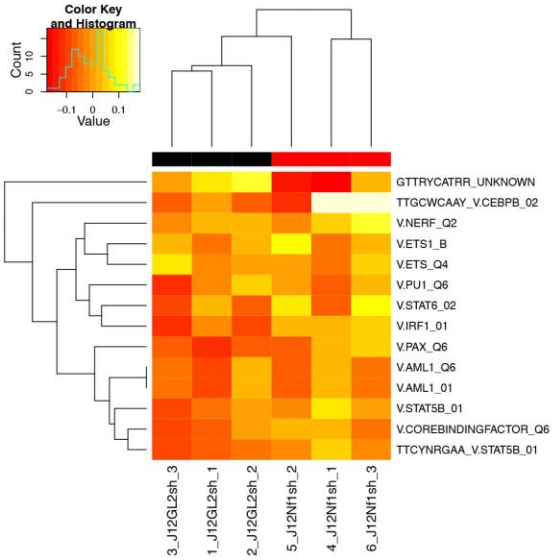


Figure S5, related to Figure 5: Simultaneous loss of *Nf1* and *Tp53* induces MES-gliomas in the RCAS/tv-a model. (A) Mutual exclusivity of *CDKN2A*, *NF1* and *TP53* gene alteration in human all non-GCIMP-GBMs and Mesenchymal-GBMs. The x-axis in the mutation heatmap visualizes patient samples and the y-axis indicates whether the corresponding gene is mutated, with colors visualizing

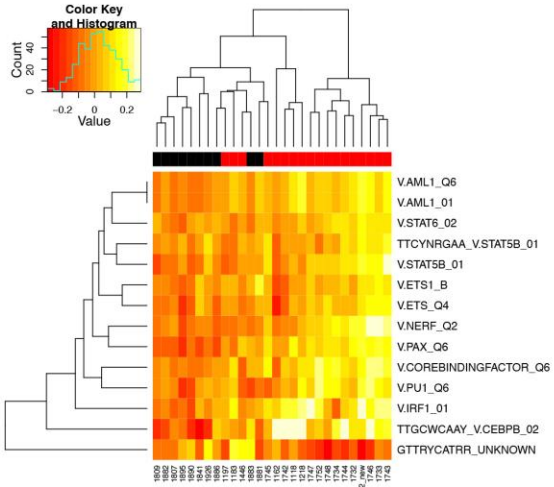
the type of mutation as shown in the legend. *NF1* loss tends to co-occur with *TP53* loss (odds ratio [OR] 9.23; 95% CI 3.50 to 26.21) and tends to be mutual exclusive to *CDKN2A* loss (OR 0.38; 95% CI 0.15 to 0.96) in the MES-GBMs as well as all non-GCIMP-GBM (TP53 loss: OR 6.29; 95% CI 3.60 to 11.17; CDKN2A loss: OR 0.53; 95% CI 0.30 to 0.94). (B) RCAS-shNf1 vector-mediated *Nf1* knockdown in *N/tv-a* neurosphere. Cells were retrovirally infected with RCAS-shControl (GL2 and GRGL2) or RCAS-shNf1 (GR249, F249, GF249 and R611) viruses. Western blots for NF1 and α-tubulin are shown. See also List of vector constructs (C) Functional validation of the RCAS-shNf1 vectors *in vivo*. The levels of NF1 mRNA expression were analyzed with microarray probe from Illumina MouseWG-6 v2.0 array in various types of the RCAS vector-driven murine gliomas. The NF1 expression was significantly ($p < 0.01$) lower in tumors incorporating the RCAS-shNf1 virus than gliomas without the RCAS-shNf1 as well as the normal counterpart. Y-axis denotes log2 ratio. Lane1: *G/tv-a*; normal brain, Lane2: *G/tv-a;PDGFA/shp53*, Lane3: *G/tv-a;PDGFA/shp53/shNf1* ($p = 0.0002583$, compared to Lane2), Lane4: *G/tv-a;shNf1/shp53* ($p = 0.006949$, compared to Lane1) , Lane5: *N/tv-a*; normal brain, Lane6: *N/tv-a;PDGFA/shp53*, Lane7: *N/tv-a;PDGFA/shp53/shNf1* ($p = 0.001512$, compared to Lane6), Lane8: *N/tv-a;shNf1/shp53* ($p < 0.001$, compared to Lane5), Lane9: *N/tv-a;Ink4a-Arf(Cdkn2a)$^{-/-}$;Pten$^{fl/fl}$*; normal brain, Lane10: *N/tv-a;Ink4a-Arf$^{-/-}$;Pten$^{fl/fl}$;PDGFA*, Lane11: *N/tv-a;Ink4a-Arf$^{-/-}$;Pten$^{fl/fl}$;PDGFA/shNf1* ($p < 0.001$, compared to Lane10), Lane12: *N/tv-a;Ink4a-Arf$^{-/-}$;Pten$^{fl/fl}$;PDGFB*, Lane13: *N/tv-a;Ink4a-Arf$^{-/-}$;Pten$^{fl/fl}$;shNf1/Cre* ($p = 0.006391$, compared to Lane 9). Whiskers of all box plots extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. (D) RCAS-shNf1/shp53 induced-tumors present diagnostic histologic features of gliomas. Representative H&E staining of the *N/tv-a;shNf1(GR249 or F249)/shp53(R696)* tumors is shown in the figure. Scale bars, 100 μm. (E) Immunohistochemistry of the *N/tv-a;shNf1(GF249)/shp53(R696)* (Left panel) and *N/tv-a;Cdkn2a$^{-/-}$;Pten$^{fl/fl}$;shNf1(GR249)/Cre* (Right panel) tumors. Representative H&E and immunohistochemistry for the indicated antibody are shown in the figure. Scale bars, 100 μm. Additional loss of *Pten* in *N/tv-a;Cdkn2a$^{-/-}$;Pten$^{fl/fl}$* mice enhanced shNf1-induced tumor formation but still led to lower penetrance than knockdown of *Tp53* (see also **Table S7**). (F) Functional comparison of various RCAS-shNf1 and shp53 constructs *in vivo*. Murine tumors were generated by co-injection of DF-1 cells producing the relevant RCAS-shRNA virus into *N/tv-a* neonatal pups brain. Additional mRFP (upper panel) and GFP (lower panel) sequence as well as direction of the shRNA (lower panel) in the RCAS vector did not affect the tumors formation as shown by Kaplan-Meier survival curves demonstrating symptom-free survival. While the target sequence of each shRNA was more important. The effect of knockdown on the target mRNA did not necessarily correlate between *in vitro* (**Fig. S4E, M, N and S5B**) and *in vivo* analysis (**Table S7**) as demonstrated by the RCAS-shNf1 (249), (611), RCAS-shp53 (696), (1114) as well as the RCAS-shPten#1(8138), #2 (1524), #3(1339). Therefore we addressed the different sequence effect by comparing our shRNA models to the relevant genetic loss mouse models (**Fig. 4B, S4F and Table S7**) (Reilly et al., 2000; Zhu et al., 2005). Survival curve of *N/tv-a;GFP-shNf1(GR249)/mRFP-shp53(mR696)* tumors from **Fig. 5A** was also shown in the figure for the comparison. (G) DNA integration of RCAS-GFP-shNf1 (GR249), RCAS-shp53 (R696) or RCAS-mRFP-shp53 (mR696) virus in murine tumor sphere cells. DNAs were extracted from tumor sphere lines derived from the RCAS-shNf1/shp53-driven gliomas in *N/tv-a* mice and then subjected to genomic PCR analysis with the relevant primer pair as shown in **Fig. S4G**. Tumor sphere lines: NF1512, NF1513 and NF1514 were generated from *N/tv-a;GFP-shNf1(GR249)/shp53(R696)* tumors. NF1586 and NF1588 lines were from *N/tv-a;GFP-shNf1(GR249)/mRFP-shp53(mR696)* tumors. DNA samples: *N/tv-a* wt and *KP176* were prepared from *N/tv-a wild-type* and non *tv-a* mouse tail, respectively. GR249: RCAS-GFP-shNf1 and mR696: RCAS-mRFP-shp53 constructs serves as a control.


A
DNA quantities relative to 18S and NC238
PDGFA_primer 2
PDGFA_primer 5
EGFR_primer 1
EGFR_primer 2
PDGFRA_primer 1
NC238
NC239
NC240
TS543
TS667
B
Color Key
Value
PN
CLS
MES
NL
C
Ctrl.sh
NF1sh #2
NF1sh #5
NF1
Tub.
TS753D
D
NF1 mRNA
p53 mRNA
Nomalized to GAPDH and shGL2
shGL2
shNf1
shp53
shNf1/shp53

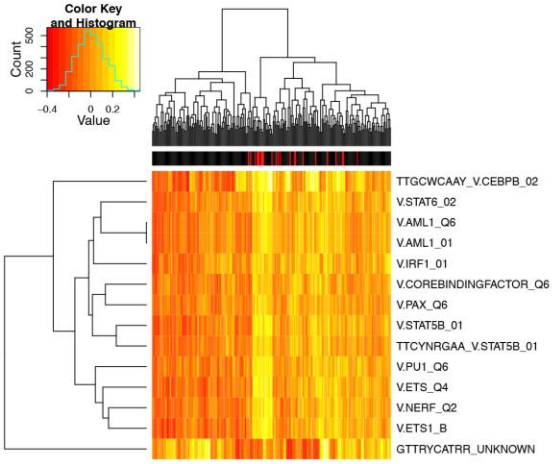
E


Color Key and Histogram
Count
Value
GTTRYCATRR_UNKNOWN
TTGCWCAAY_V.CEBPB_02
V.NERF_Q2
V.ETS1_B
V.ETS_Q4
V.PU1_Q6
V.STAT6_02
V.IRF1_01
V.PAX_Q6
V.AML1_Q6
V.AML1_01
V.STAT5B_01
V.COREBINDINGFACTOR_Q6
TTCYNRGAA_V.STAT5B_01
3_J12GL2sh_3
1_J12GL2sh_1
2_J12GL2sh_2
5_J12Nf1sh_2
4_J12Nf1sh_1
6_J12Nf1sh_3

F


Color Key and Histogram
Count
Value
V.AML1_Q6
V.AML1_01
V.STAT6_02
TTCYNRGAA_V.STAT5B_01
V.STAT5B_01
V.ETS1_B
V.ETS_Q4
V.NERF_Q2
V.PAX_Q6
V.COREBINDINGFACTOR_Q6
V.PU1_Q6
V.IRF1_01
TTGCWCAAY_V.CEBPB_02
GTTRYCATRR_UNKNOWN

G


Color Key and Histogram
Count
Value
TTGCWCAAY_V.CEBPB_02
V.STAT6_02
V.AML1_Q6
V.AML1_01
V.IRF1_01
V.COREBINDINGFACTOR_Q6
V.PAX_Q6
V.STAT5B_01
TTCYNRGAA_V.STAT5B_01
V.PU1_Q6
V.ETS_Q4
V.NERF_Q2
V.ETS1_B
GTTRYCATRR_UNKNOWN

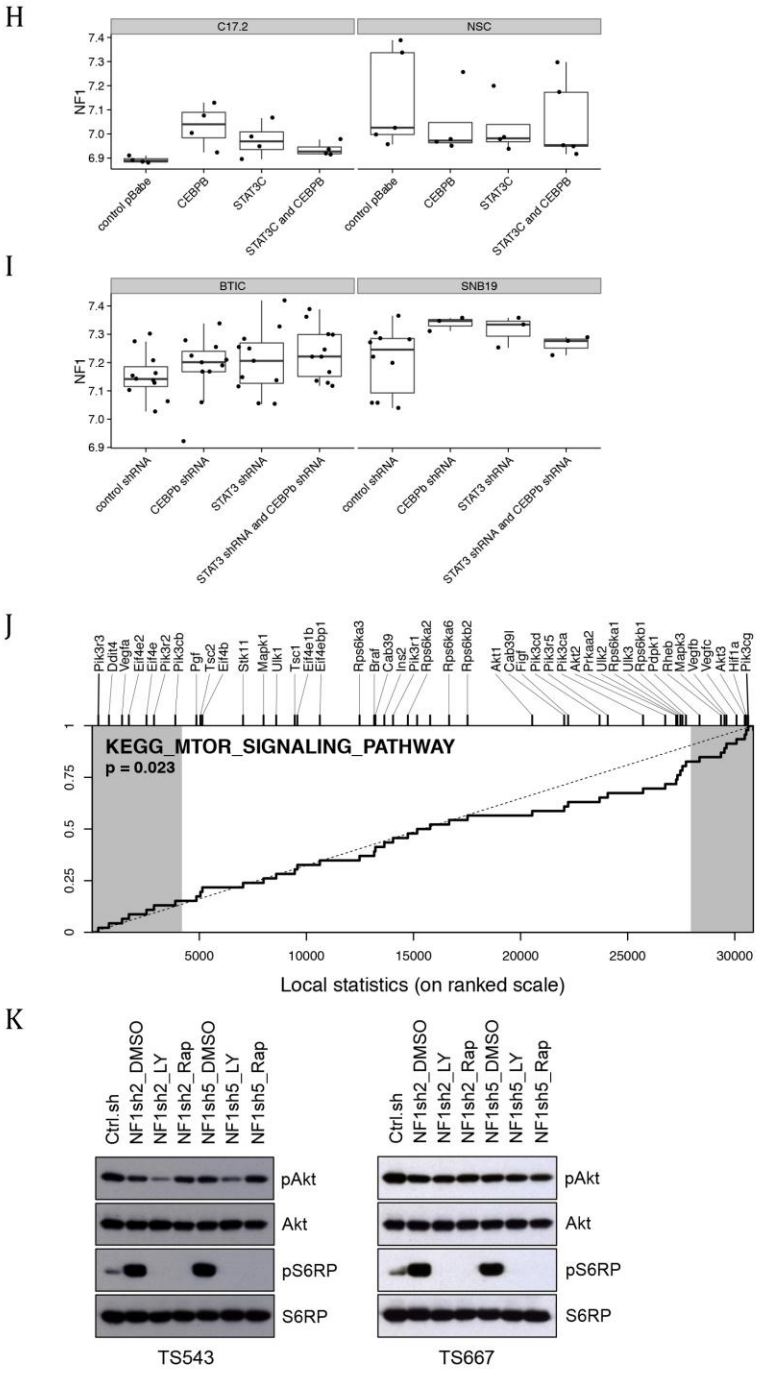


Figure S6, related to Figure 6. *Nf1* loss induces proneural to mesenchymal conversion *in vitro*.
(A) *PDGFA*, *EGFR* and *PDGFRA* gene quantitation in two human GBM cell lines (TS543 and TS667).

Gene status was determined using q-PCR analysis as described previously (Martinho et al., 2009; Ozawa et al., 2010). The Y-axis indicates average gene quantities relative to the reference sample (NC238) and 18S gene. Average DNA quantities from 2 to 4 and greater than 4 were defined as gain and amplification, respectively. The results suggested that both cell lines harbored chr 7p gain and *PDGFRA* amplification, presumably being proneural cell lines. Error bars represent standard deviation. (B) Single sample GSEA of human GBM cell lines. This heatmap visualizes continuous subtype scores for all cell line samples, estimated by single sample GSEA. Red indicates an enrichment of a subtype signature, blue indicates anti-correlated expression, i.e. genes known to be up-regulated in the subtype are down-regulated and vice versa. As expected by above copy number analysis, both TS543 and TS667 cell lines presented proneural gene enrichment. *NF1* loss induced clearly a sift of the expression pattern from proneural to mesenchymal/classical in TS543 cells. While TS667 cell line still preserved a proneural expression pattern after *NF1* knockdown but decreased the proneural expression and tended to sift toward mesenchymal/classical expression pattern. TS753D cell lines serve as a mesenchymal control. Therefore *NF1* knockdown didn't affect the expression pattern. (C) *NF1* knockdown in TS753D cells. TS753D human GBM cell line was lentivirally transduced with pLKO.1-NF1shRNA (#2 and #5) or empty control vector. Western blots for NF1 and α-tubulin are shown. (D) NF1 and p53 mRNA expression in murine neurosphere lines. Neurospheres derived neonatal *N/tv-a wt (J12)* pups brain were retrovirally infected with relevant RCAS virus, and *Nf1* and *Tp53* knockdown was then examined using q-PCR analysis. The ΔΔCt method was used to calculate the gene expression in the samples normalized to the reference gene (GAPDH) and calibrated to control samples (shGL2). Error bars represent standard error of the mean. (*) $p < 0.05$, (**) $p < 0.005$, (***) $p < 0.0005$, (****) $p < 0.0001$. Each sample was analyzed in quadruplicates. (E-G) Transcription Factor enrichment analysis. Utilizing a fixed-effects meta-analysis and a single sample GSEA, we identified 14 TF gene sets (MSigDB 4.0), corresponding to 9 different TFs, which were significantly (FDR < 0.05) associated with *NF1* loss in three datasets. Panels E-G show heatmaps of these 14 TFs in these datasets. Below the column dendrogram, *NF1* wild-type samples are marked in black, samples with *NF1* loss in red. Heatmap colors indicate the expression level of the putative TF target genes as estimated by single sample GSEA (high expression=yellow, low expression=red). Rows represent the TF gene sets and their MSigDB id names, the columns represent the samples. (E) murine neurosphere data; (F) murine glioma data; (G) TCGA data. (H and I) Effect of perturbations of the STAT3 and/or C/EBPβ expression on the *NF1* gene expression. *NF1* gene expression after overexpression (H) or knockdown (I) of the STAT3 and/or C/EBPβ was examined in two different mouse cells (GSE19113) and two different GBM-derived cells (GSE19114) using a publicity available database, respectively. Neither perturbation of the STAT3 and/or C/EBPβ expression affected the NF1 expression in both models. Y-axis denotes log2 ratio. Whiskers of all box plots extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. (J) Gene set enrichment analysis of murine shNf1/shp53 tumors. To examine if *NF1* loss associates with an up-regulation of mTOR signaling pathway, the gene expression profiles were compared between murine PDGFA/shp53 and shNf1/shp53 tumors in *G* and *N/tv-a* mice. Direction for many genes was consistent with mTOR signaling pathway in the murine shNf1/shp53 tumors, implying an enrichment of the pathway in the shNf1-induced tumors more than PDGFA-induced tumors. X-axis denotes ranked-genes. The genes on the right in the plot present up-regulated genes in the murine shNf1/shp53 tumors (on the left down-regulated genes). Grey areas represent the significantly differentially expressed genes. Y-axis denotes enrichment score. (K) The effects of mTOR and PI3K inhibitor on the elevated phospho-S6RP of the cells expressing NF1-shRNAs. Cells were treated with 0.1% DMSO, 1nM Rapamycin, or 10 μM LY294002 for 5 hours and the whole-cell lysates were then subjected to immunoblot analysis with the indicated antibodies. Untreated cells expressing the control-shRNA serve as control of the treatment.

Supplemental Experimental Procedures

Datasets

We obtained segmented copy-number data (Level 3, Affymetrix Genome-Wide Human SNP Array 6.0, 496 samples; Level 2, Agilent Human Genome aCGH 244A, 433 samples), raw and normalized mRNA expression data (Level 1, Affymetrix GeneChip Human Genome U133A, 519 samples; Level 3, Agilent G4502A, 535 samples), methylation data (Level 3, Human Methylation 27 and 450, 268 and 113 samples, respectively) and somatic mutation data (Level 3, 259 samples) from The Cancer Genome Atlas (TCGA) (The Cancer Genome Atlas Research Network, 2008). All analyses were done in R/Bioconductor (Gentleman et al., 2004). CEL files were normalized with the GC-Robust Multiarray Averaging (GCRMA) algorithm (Wu Z, 2004). We used the custom CDFs HTGU133A_HS_Entrez and HGU133Plu2_Hs_ENTREZG in Version 15.0 (Dai et al., 2005) in order to summarize probe set intensities for each gene. Probe sets with an interquartile range (IQR) of 0 were removed. If not mentioned otherwise, mRNA expression refers to expression estimates obtained from the Agilent data, as more samples and genes were profiled on this platform compared to the Affymetrix data. Copy number always refers to SNP Array 6.0 data; probe level aCGH data was only used to determine *EGFRvIII* mutation status. Matched copy number and Agilent expression data was available for 486 patients.

Subtypes

Subtype classifications (classical, mesenchymal, neural, proneural) for 203 samples were obtained from the supplementary material from Verhaak et al (Verhaak et al., 2010). Remaining samples in the TCGA dataset were classified using the support vector machine (SVM) implementation in the MLInterfaces R package with default radial kernel (ksvm function). Only the genes included in the Verhaak 840-gene signature (Verhaak et al., 2010) were considered for classification. We classified twice using both Affymetrix and Agilent data. Samples with a classification posterior probability smaller than 0.5 or with conflicting classification between platforms were excluded from the analysis (**Table S1**). The same approach was used to identify non-GCIMP tumors, with training classification for 541 samples obtained from the cBio Cancer Genomics Portal (http://cbio.mskcc.org/gdac-portal/index.do). GCIMP classification was repeated with Affymetrix, Agilent and Human Methylation 27 data; all previously unclassified samples were predicted to be non-GCIMP tumors.

Recurrent copy number changes

The GISTIC algorithm in version 2.0.12 (Mermel et al., 2011) was used with default parameters to identify recurrent focal and broad copy number aberrations. To test the predictive utility of chromosome 7 and 10 copy numbers, we constructed univariate Cox models with the broad copy numbers as obtained by the GISTIC software (Mermel et al., 2011). In all our analyses, we used default GISTIC cutoffs, i.e., cutoffs for homozygous loss, heterozygous loss, gain and amplification were set to log base 2 ratios of -1.3, 0.1, 0.1 and 0.9, respectively. Furthermore, throughout the manuscript, broad and focal copy numbers always refer to GISTIC estimates of these numbers.

Pathway analysis

We developed a computational method for gene prioritization in large copy number alterations. The algorithm identifies genes from (i) disease-related pathways for which (ii) copy number gain has

profound impact on mRNA levels. While (ii) is relatively straightforward, the former (i) problem is rather subjective and, for optimal results, thus has to be tailored to the data at hand. We utilized survival association of downstream genes as surrogate for diseases-relatedness due to the observed survival association of chr7 gain. Other approaches, for example a simple expression correlation of downstream genes with copy number, or testing for a significantly increased number of alterations, might be more suitable to identify disease-associated pathways in the absence of a survival association.

In detail, we tested the predictive utility of genes that are potentially affected by a chr7 gain and chr10 loss. For all genes on chr7 and 10, we tested the association of downstream genes in their pathways for association with overall survival. We first constructed gene sets with pathways that harbor genes located on these two chromosome arms. Pathway topologies for the KEGG (Kanehisa M, 2000), Reactome (Matthews et al., 2009), BioCarta and NCI pathway (Schaefer et al., 2009) databases were obtained from the graphite R package (Sales et al., 2012) in version 1.0. Since these pathway topologies did not provide the direction of gene-gene interactions (activating vs. deactivating), we did not integrate expression direction into our algorithm. This means we only tested for survival association and did not test whether the direction (up-regulation associated with good vs. bad outcome) was consistent with the literature. The Floyd-Warshall algorithm was applied to the pathway networks to identify genes that are downstream of chr7 and chr10 genes. A gene is defined as downstream of another gene if there exists a directed path in the network connecting both genes. Genes in feedback loops are thus included. Only genes were considered that were arrayed on both Affymetrix and Agilent platforms and that had 10 or more downstream genes in at least one pathway database. We further excluded genes with a fold-change smaller than 1.25 when we compared patients with normal and altered copy numbers. This cutoff was chosen instead of the more common 1.5 cutoff, because it resulted in a still manageable list of approximately 200 genes; stringent cutoffs potentially exclude driver genes. P values and false discovery rates of association of downstream genes with survival were then estimated by gene set analysis (Subramanian et al., 2005) with the SAFE R package in version 3.0 (Barry et al., 2005) as described in the SAFE documentation. To obtain a final ranking of genes, we ranked the corresponding gene sets by enrichment p value. Thus, our approach has a few limitations and further work is required to demonstrate a robust utility of this algorithm for routine gene prioritization. The main limitations are: requirement of survival association of downstream genes; dependence of pathway databases and their biases (e.g. well studied pathways and genes are highly overrepresented); no incorporation of expression direction.

RESIC analysis

The RESIC algorithm considers, for each patient, a population of cells at risk of accumulating the genetic alterations driving tumorigenesis. Cells proliferate according to a stochastic process (Moran PAP, 1962): during each time step, a cell is chosen for division proportional to its fitness. During each cell division, a genetic alteration arises with a certain probability. Following the division event, another cell is chosen at random to die. A genetic alteration may reach fixation (i.e., 100% frequency) in the population or go extinct due to stochastic fluctuations. Depending on the order in which individual alterations arise, the system might follow different evolutionary paths to tumorigenesis.

We developed a mathematical model describing the evolutionary dynamics of this system (Attolini et al., 2010). When applied to cross-sectional genomics data such as the TCGA GBM sample set, this algorithm allows for the inference of the temporal sequence of emergence of genomic events. Our method is based on genomic data only since epigenetic changes lead to potentially reversible and fluctuating expression changes, which are not currently captured by our algorithm.

Using the copy number alterations and subtypes as defined earlier, we applied the RESIC algorithm (Attolini et al., 2010; Cheng et al., 2012) to the GBM dataset. We selected two whole chromosome events, chr7 non-disjunction and chr10 loss, as well as subtype-defining events (focal *EGFR* gain and amplification, *PDGFRA* gain and amplification, and *NF1* loss) as the alterations of interest for the analysis. Whole chromosome events were defined as gain (chr7) or loss (chr10) of at least one chromosome arm, identified by GISTIC (GISTIC output file broad_values_by_arm.txt) (**Fig. S2A**). For gene and chromosomal deletion, loss of at least one chromosome arm was included; therefore both hetero- and homozygous loss was grouped in this analysis and denoted as loss. Using RESIC, we then determined the relative order of copy number alterations in all four subtypes. All four subtypes were similar, with chr7 non-disjunction and chr10 loss identified as early events (**Fig. 2A and S2B**). *CDKN2A* loss was placed as a middle event. However, the subtypes differed more significantly in the subtype-defining events involving *EGFR*, *NF1* and *PDGFRA*, which were categorized primarily as late events. *NF1* loss was placed earlier in the mesenchymal subtype than in the classical subgroup and was missing in the proneural and neural subtypes. *PDGFRA* amplification occurred most often in the proneural subtype, and was placed more closely in time to the *CDKN2A* event in this subtype. In order to place *TP53* mutational events, we analyzed a data set consisting of 85 samples from all subtypes for which TCGA had analyzed somatic mutations. Analysis across all samples instead of independently for each subtype was necessary given the lesser availability of somatic mutational data. By doing this, we identified *TP53* mutation as occurring after the two whole chromosome events and *CDKN2A* loss but before the subtype-defining events (**Fig. 2B and S2C**). We further compared our RESIC results with those obtained by an alternative software, CT-CBN (**Fig. S2E**) (Beerenwinkel and Sullivant, 2009). CT-CBN was used with default parameters and 1000 bootstrap replicates. Both tools generated similar orders of events, with chr7 and chr10 being again initial events. Due to slightly higher prevalence of chr10 loss compared to chr7 gain, CT-CBN placed chr10 before chr7 in the mesenchymal and neural subtypes, albeit with low bootstrap confidence (65% and 57%, respectively). In the proneural subtype, CT-CBN placed *CDKN2A* and *PDGFRA* alteration independently of the whole chromosome events, due to a small number of patients without these whole chromosome events.

Gene expression analysis for murine brain tumors, neurosphere and human GBM cell lines

Total RNAs were extracted with miRNeasy Mini kit (QIAGEN) according to manufacturer's protocol and then labeled by the Illumina protocol followed by hybridization to the MouseWG-6 v2.0 or Human HT-12 Expression BeadChip (Illumina, Inc. San Diego, CA). Raw data were processed at Genomics Core Lab of Memorial Sloan-Kettering Cancer Center and then normalized with the lumi R package (Du et al., 2008). For probe sets mapping to the same HGNC symbols, we selected the probe set with highest mean across all samples used in this study.

Gene Set Enrichment Analysis (GSEA)

The Verhaak et al (Verhaak et al., 2010) subtype signatures were translated to mouse gene symbols using Biomart (version May 2012) (Durinck et al., 2005). This approach mapped on average 91% of the HGNC symbols in the original TCGA subtype signatures to mouse gene symbols. Differential expression between test and control groups were estimated with the limma package (Smyth, 2005) and genes were then ranked by moderated t-statistic. For cell line data, we adjusted for the global differences in expression of cell lines with a linear model. In these models, the cell line differences were subtracted from the observed differences between test and control. We identified a small batch effect in the mouse expression data and thus incorporated the batch in the linear model as described in the limma manual. As the direction of expression of the subtype genes were known, we flipped the moderated t-statistic of the down-regulated genes and performed a one-sided gene set analysis, i.e., we tested whether the genes in the subtype gene sets were ranked higher than expected by chance using the limma package (wilcoxGST function). The cell line data showed some heterogeneity across cell lines and we thus used a single sample GSEA to calculate and visualize subtype scores for each sample (Hanzelmann et al., 2013). Subtypes scores were calculated for down- and up-regulated genes separately; the scores of the down-regulated genes were then subtracted from the scores of the up-regulated genes to obtain final subtype scores.

Mutual exclusivity of altered genes in human mesenchymal GBMs

CDKN2A, *NF1* and *TP53* somatic mutation data were retrieved from The Cancer Genome Atlas (Level 2 data downloaded March 17, 2013). Then the mutual exclusivity for each gene was analyzed for cases with homozygous deletion, heterozygous deletion and somatic point mutation in Mesenchymal GBMs subgroup and all non-GCIMP tumors using Fisher's exact test. Reported odds ratios represent the odds for a patient having non-wild-type *TP53* or *CDKN2A* when *NF1* is mutated (**Fig. S5A**).

Phylogenetic analysis

We applied the Neighbor-Joining (NJ) algorithm (Saitou and Nei, 1987) independently to the TCGA GBM Agilent expression, SNP copy number, promoter methylation and somatic mutation data. As described in detail earlier (Riester et al., 2010), we performed a bootstrapping approach in which patients and genes were simultaneously sampled with replacement. For the patient bootstrapping, subtypes were sampled independently, i.e., for a subtype with n patients, we sampled n patients of this subtype with replacement. For each data type, we generated 1000 bootstrap replicates. Subtype centroids were then calculated for each subtype by averaging over all the sampled genes and patients. NJ was then used to construct phylogenetic trees of the subtype centroids. We used Pearson correlation distance for all the continuous data types. For the somatic mutations, we used the Euclidean distance. Bootstrap replicates were then summarized and visualized with majority consensus trees (Margush and McMorris, 1981) as implemented in Dendroscope 3.2.2 (Huson et al., 2007). For copy number data, we included an artificial normal group with normal diploid genome in every copy number segment (**Fig. 1B**).

***EGFRvIII* mutation status**

Status of *EGFRvIII* mutation was determined as previously described (Szerlip et al., 2012). We compared the mean aCGH copy number ratios of probes in the deleted region of *EGFR* (Agilent probe ids A_14_P102368, A_16_P17952602, A_16_P38034515, A_16_P01718265), with the mean ratios of the probes in the 3' prime end of *EGFR*, downstream of exon 7 (A_14_P133869, A_16_P17952748, A_16_P01718353, A_16_P17952840, A_16_P38034765, A_14_P106592). Difference of > 1 standard deviation was considered significant (**Table S1**).

Transcription factor analysis

We tested whether *NF1* loss was associated with up- or down-regulation of certain transcription factors (TFs). We utilized MSigDB 4.0, which provides 615 putative TF/target gene sets (c3 category) (Subramanian et al., 2005). To make data from different platforms comparable, we used a single sample GSEA as implemented in the GSVA package (Hanzelmann et al., 2013). This approach transforms a gene expression matrix (columns samples, rows genes) into a TF expression matrix (columns samples, rows 615 TFs). In brief, when a TF's target genes have a higher expression than expected by chance in a particular sample, then this method will set a high value in this matrix for this sample (and vice versa). This method makes the data more easily comparable across platforms, because all datasets will have the same set of features (here TFs) instead of measuring different genes with different probe sets. We used the GSVA function and set the minimum gene set size to 10, the maximum to 500 and all other parameters to their default values.

We performed the single sample GSEA for the (i) mouse neurosphere lines expressing shGL2 or shNf1 (**Fig. 6C and S6D**) (ii) mouse gliomas (N and G/tv-a;PDGF-A+shp53 tumors and N and G/tv-a;shp53+shNf1 tumors) and (iii) TCGA data. For TCGA, only samples with available *NF1* mutation status were analyzed and *NF1* loss was defined as non-silent mutation. To test the association of *NF1* loss with these 3 TF expression matrices, we used a logistic regression independently for each TF and dataset and then pooled the regression coefficients with a fixed-effects meta-analysis (weighting by inverse of squared standard error), implemented in the metafor R package (Viechtbauer, 2010). A random-effects meta-analysis was not feasible to due the small numbers of matrices. TFs with pooled regression coefficient significantly different from 0 (FDR < 0.05) were visualized in a heatmap. Overlaps of statistically significant gene sets were visualized with a heat map. Overlap of two gene sets was defined as number of overlapping genes divided by the size of the smaller gene set.

We next tested whether a recently reported mesenchymal gene signature (Carro et al., 2010) was enriched in the combined 14 TF gene sets. A p value was calculated with a hypergeometric distribution, using all genes in MSigDB c3 as background. We observed a statistically significant overlap ($p = 0.005$). The corresponding overlapping genes were visualized with Cytoscape 2.8. (Smoot et al., 2011). Utilizing a two-sided t-test, we finally tested whether overexpression or knockdown of C/EBPβ and/or STAT3 affects the *NF1* gene expression using publicly available murine and human cell line data (GSE19113 and GSE19114, respectively) (Carro et al., 2010). After controlling for multiple testing, no comparison was statistically significant, i.e. NF1 displayed no statistically significant differences in gene expression after MES-TF perturbation.

Chemicals, cell culture, transfections, retroviral and lentiviral infection

NIH-3T3 (ATCC: CRL-1658), Platinum-E (CELL BIOLABS, INC.), human GBM sphere lines (TS543, TS667 and TS753D) and DF-1 cells (ATCC: CRL-12203) were maintained with minor modification as previously described (Charles et al., 2010; Hambardzumyan et al., 2009; Ozawa et al., 2010; Szerlip et al., 2012). Murine neurosphere (*N/tv-a*) and RCAS-shNf1/shp53 driven-tumor sphere lines were prepared as previously described (Bleau et al., 2009). HEK293T cells were grown in DMEM with 10% fetal bovine serum (PAA Laboratories Inc.). Rapamycin (R0395) and LY294002 (#9901) were purchased from Sigma-Aldrich (St. Louis, MO) and Cell Signaling Technology (Danvers, MA), respectively.

Retrovirus for stable NIH-3T3 cell lines expressing Tv-a was produced with Fugene-6 (Roche Diagnostics) in the Platinum-E cells and was harvested at 48 hours after transfection of pBabe-puro-tv-a vector. NIH-3T3 cells were then subjected to 3 8 hours cycles infection using virus-containing supernatant supplemented with 8 μg/ml polybrene (Sigma) and later selected with 2.5 μg/ml puromycin (GIBCO). RCAS virus was produced in DF-1 packaging cells with minor modification as described previously (Holland et al., 2000; Holland et al., 1998; Holland and Varmus, 1998; Hu et al., 2005). The NIH-3T3/tv-a or tv-a-myc/6xHis cells (see below) were subjected to retroviral infection using RCAS virus-containing supernatant of DF-1 cells as well. For retroviral infection of murine neurosphere, RCAS virus was produced in DF-1 packaging cells maintained with serum free neurosphere medium and was then diluted 1:1 with collection media (fresh murine neurosphere media).

Lentivirus for stable NIH-3T3 cell lines expressing Tv-a-myc/6xHis was produced with X-treamGENE9 (Roche Diagnostics) by cotransfection of pLJM1-Tv-a-myc/His constructs (pLJM1-EGFP vector was used as the control.) with 2 packaging plasmids (pMD2.G and psPAX2) into HEK293T cells. The viral particles were collected at every 24 hours after the transfection. NIH-3T3 cells were then subjected to 3 8 hours cycles infection using virus-containing supernatant supplemented with 8 μg/ml polybrene and later selected with 2.5 μg/ml puromycin. Lentivirus for transduction of NF1 shRNA into human GBM sphere lines was produced with Fugene 6 by cotransfection of pLKO.1-NF1 shRNA constructs (pLKO.1 empty vector was used as the control.) with 2 packaging plasmids (pMD2.G and psPAX2) into HEK293T cells maintained with serum free neurosphere medium. The viral particles were collected at every 24 hours after the transfection and were diluted 1:1 with collection media (fresh human neurosphere media). GBM sphere lines were then subjected to 3 8 hours cycles infection using virus-containing supernatant supplemented with 8 μg/ml polybrene and later selected with 2.5 μg/ml puromycin as well.

Vector constructs

The human PDGFA cDNA ORF Clone (pCMV6Entry) was purchased from OriGene Technologies, Inc.. The PDGFA cDNA was initially PCR-amplified from the pCMV6Entry-PDGFA vector and was then subcloned into pcDNA3.1/myc-HisB vector (Invitrogen) as between the NotI and XhoI restriction sites. Subsequently, the PDGFA containing myc-6xHis tag in the C-terminus was further subcloned into RCAS-Y vector (Dunn et al., 2000) as between the NotI and PmeI restriction sites.

For the generation of Tv-a expression vector, the Tv-a cDNA was PCR-amplified from Nestin Tv-a vector (Holland et al., 1998) and was then subcloned into pBabe-puro vector as between BamHI and EcoRI restriction sites. As well, the Tv-a cDNA was PCR-amplified from pBabe-Tv-a vector and was then subcloned into pcDNA3.1/myc-HisB vector (Invitrogen) as between the NotI and XhoI restriction sites. Subsequently, the Tv-a containing myc-6xHis tag in the C-terminus was further PCR-amplified and was then cloned into pLJM1-EGFP vector as between the NheI and EcoRI restriction sites. NIH-3T3/tv-a cell lines stably transduced with either Tv-a expression vector were used for a knockdown analysis of RCAS-shRNA vectors. The RCAS-PDGFB-HA and RCAS-Cre for *Pten* gene deletion was previously described (Hu et al., 2005; Shih et al., 2004).

For the generation of RCAS-shRNA vector, shRNA sequences containing H1 promoter were PCR-amplified from pSUPER.retro.puro vector (OligoEngine, Inc.) and then subcloned into RCAS-Y vector as between the Not I and Pac I restriction sites. To examine if the H1 promoter interferes with 5'LTR promoter of the RCAS vector for the shRNA expression, the H1 promoter containing the shRNA was inserted in both directions toward 5'LTR of the RCAS vector (**Fig. S4A**). For the generation of RCAS-EGFP-GL2 and -Nf1 (249) shRNA vector, the EGFP was PCR-amplified from pMXs-IRES-EGFP retroviral vector (CELL BIOLABS, INC. San Diego, CA) and was then inserted in the Not I restriction site which located on the upstream of shRNA or H1 promoter sequence. For the generation of RCAS-mRFP-p53 (696) shRNA vector, the mRFP was PCR-amplified from pCX-mRFP1 vector and was then inserted in the Not I restriction site which located on the upstream of shRNA sequence as well.

For the generation of RCAS-mRFP-shPten vector, RCAS-Gateway (GW)-mRFP-shRNA (rev.) vector system was created. Initially, mRFP and H1 promoter sequence was PCR-amplified as described in above and then subcloned into pENTR1A-Dual Selection vector (Invitrogen) as between the Not I and Xho I for mRFP sequence, and Xho I and EcoRV restriction sites for H1 promoter sequence. Subsequently oligonucleotides for the shPten were assembled between the Bgl II and Hind III restriction enzyme sites in the pENTR-mRFP-H1 vector and the mRFP-shPten sequence was finally transferred into RCAS-Destination vector (DV) (Loftus et al., 2001) by LR recombination according to the manufactures protocol (Invitrogen). Structural diagram and the target sequences of individual RCAS-shRNA construct were described in **Fig. S4A and below**, respectively. The DNA sequence of all vector constructs was confirmed with ABI 3730 capillary sequencer (Applied Biosystems). Lentiviral shRNA constructs targeting human *NF1* and empty control were purchased from High-Throughput Drug Screening Facility of Memorial Sloan-Kettering Cancer Center [pLKO.1-NF1shRNA (#2), SKI-RSI-192342; pLKO.1-NF1shRNA (#5), SKI-RSI-192345; pLKO.1-empty vector, SKI-RSI-249999]. pLJM1-EGFP (#19319)(Sancak et al., 2008), pMD2.G (#12259) and psPAX2 (#12260) constructs were purchased from Addgene, Inc.. All vectors used in this study were listed in **below**.

Target sequence for the RCAS-shRNA vectors

Gene	Sequence (5' -> 3')	Reference

Nf1 (249)	CAAGGAGTGTCTGATCAAC	Mus musculus neurofibromatosis 1 (Nf1), mRNA, NM_010897.2, (Patrakitkomjorn et al., 2008)
Nf1 (611)	GGTTACAGGAGTTGACTGT	Mus musculus neurofibromatosis 1 (Nf1), mRNA, NM_010897.2, (Patrakitkomjorn et al., 2008)
Trp53 (696)	GTACATGTGTAATAGCTCC	Mus musculus transformation related protein 53 (Trp53), transcript variant 1, mRNA, NM_011640.3
Trp53 (1114)	CAGTCTACTTCCCGCCATA	Mus musculus transformation related protein 53 (Trp53), transcript variant 1, mRNA, NM_011640.3
GL2	CGTACGCGGAATACTTCGA	Firefly (Photinus pyralis) luciferase gene
Pten #1 (8138)	GGACATAAGACTAGAAATA	Mus musculus phosphatase and tensin homolog (Pten), mRNA, NM_008960.2
Pten #2 (1524)	AGCTAAAGGTGAAGATATA	Mus musculus phosphatase and tensin homolog (Pten), mRNA, NM_008960.2
Pten #3 (1339)	AGTAAGGACCAGAGACAAA	Mus musculus phosphatase and tensin homolog (Pten), mRNA, NM_008960.2

List of vector constructs

Vector Constructs	**References**	**Source**	**Abbrevi-ations**
pSUPER-retro-puro-shNf1 (249)		Dr. Saya, Keio university	
pSUPER-retro-puro-shNf1 (611)		Dr. Saya, Keio university	
pSUPER-retro-puro-shp53 (696)		Dr. Saya, Keio university	
pSUPER-retro-puro-shp53 (1114)		Dr. Saya, Keio university	
pSUPER-retro-puro-shGL2		Dr. Saya, Keio university	
pBabe-Tv-a-puro			
Nestin-Tv-a	(Holland et al., 1998)		
pLJM1-EGFP	(Sancak et al., 2008)	Addgene 19319	
pcDNA3.1-Tv-a-myc/6xHis			

pLJM1-Tv-a-myc/6xHis			
RCAS-shGL2 (reverse)			RGL2
RCAS-EGFP-shGL2 (reverse)			GRGL2
RCAS-shNf1 (249) (forward)			F249
RCAS-EGFP-shNf1 (249) (forward)			GF249
RCAS-EGFP-shNf1 (249) (reverse)			GR249
RCAS-shNf1 (611) (reverse)			R611
RCAS-shp53 (696) (reverse)			R696
RCAS-mRFP-shp53 (696) (reverse)			mR696
RCAS-shp53 (1114) (reverse)			R1114
RCAS-hPDGFA-myc/6xHis			PA
RCAS-hPDGFB-HA	(Shih et al., 2004)		PB
RCAS-Cre	(Hu et al., 2005)		Cre
RCAS-Y	(Dunn et al., 2000)		
RCAS-Y-DV (destination vector)	(Loftus et al., 2001)		
pENTR 1A Dual Selection vector		Invitrogen #A10462	
pENTR-mRFP-H1 (reverse)			
pENTR-mRFP-shPten #1			
pENTR-mRFP-shPten #2			
pENTR-mRFP-shPten #3			
RCAS-GW-mRFP-shPten #1			RPT1
RCAS-GW-mRFP-shPten #2			RPT2
RCAS-GW-mRFP-shPten #3			RPT3
pCMV6Entry-hPDGFA (RC223058)		OriGene Technologies, Inc.	

pcDNA3.1/myc-His B		Invitrogen #V800-20	
pcDNA3.1-hPDGFA-myc/6xHis			
pLKO.1		HTDS facility, MSKCC	
pLKO.1-NF1shRNA (#2)		HTDS facility, MSKCC	
pLKO.1-NF1shRNA (#5)		HTDS facility, MSKCC	
pMD2.G		Addgene 12259	
psPAX2		Addgene 12260	
pMXs-IRES-EGFP		CELL BIOLABS, INC.	
pCX-mRFP1	(Helmy et al., 2012) (Campbell et al., 2002)		

Antibodies for Western blot analysis

The following primary antibodies were used for Western blots: c-Myc (M4439, Sigma-Aldrich); PDGFA (sc-128, Santa Cruz Biotechnology); Actin (sc-1616, Santa Cruz Biotechnology); Neurofibromin (sc-67, Santa Cruz Biotechnology); p53 (NCL-p53-CM5P, Novacastra); GFP (A11122, Invitrogen); RFP (ab62341, abcam); PI3 Kinase p85 (06-195, Millipore); α-tubulin (T-9026, Sigma Aldrich); Akt (#9272, Cell Signaling); phospho-Akt (Ser473) (#4060, Cell Signaling); phospho-Erk (#4370, Cell Signaling); Erk (#9102, Cell Signaling); phospho-S6RP (#2211, Cell Signaling); S6RP (#2217, Cell Signaling) ; PTEN (Cell signaling, #9188). For an analysis of *Tp53* knockdown, cells were collected after 6 hours of UV exposure on ultra violet trans-illuminator for 3 min.

Antibodies for immunohistochemistry

The following antibodies were used for immunohistochemical staining: HA (Roche, 11867423001) at 1:150; c-Myc (Sigma-Aldrich, M4439) at 1:1,000; Ki67 (Vector, VP-RM04) at 1:200; GFP; (Invitrogen, A11122) at 1:100; GFP; (Cell signaling, #2956) at 1:300; Nestin (BD Biosciences, 556309) at 1:100; PDGFRA (Cell signaling, #3174) at 1:300; PDGFRB (Cell signaling, #3169) at 1:300; GFAP (DAKO, Z0334) at 1:8000; Olig2 (MILLIPORE, AB9610) at 1:400; SMA (Abcam, ab5694) at 1:100; eNOS (BD Biosciences, 610296) at 1:100; CD44 (BD Biosciences, 550538) at 1:1,000; Cleaved-Caspase 3 (Cell signaling, #9661) at 1:300; p53 (NCL-p53-CM5P, Novacastra) at 1:500; PDGFA (Santa Cruz Biotechnology, sc-128) at 1:200; phospho-STAT3 (Tyr705) (Cell signaling, #9145) at 1:250; C/EBPβ (Santa Cruz Biotechnology, sc-746) at 1:250; PTEN (Cell signaling, #9188) at 1:100 dilution.

Generation of murine brain tumors

In vivo gene transfer in humans to verify that PDGFA is capable of initiating glioma formation is

obviously not possible, and genetic modification of cultured human cells followed by implantation into immunodeficient mice may not adequately recapitulate the events of glioma initiation *in vivo* (Hambardzumyan et al., 2011). Therefore we tested the ability of genes predicted to drive human GBM evolution to induce gliomas *in vivo* in immunocompetent mice using the RCAS/tv-a system. The RCAS/tv-a system used in this work has been described previously (Holland et al., 2000; Holland et al., 1998; Holland and Varmus, 1998; Hu et al., 2005; Uhrbom et al., 2004). Briefly, DF-1 cells were transfected with the relevant RCAS viral plasmid using Fugene 6 Transfection reagent (Roche) accordingly to manufacturer's protocol. The cells were regularly maintained for three passages at least for a propagation of the RCAS virus to entire cells. Then cells after passage 4 were used for an injection of the DF-1 cells into murine brain. The new-born pups were injected intracranially with 1 μL of approximately $1x10^5$ DF-1 cells infected with and producing relevant RCAS viruses. Simultaneous delivery of two or three RCAS viruses was done by the injection of 1 μL of approximately $2x10^5$ DF-1 cells mixed with equal ratio. Then mice were monitored daily until they developed symptoms of disease such as lethargy, poor grooming, weight loss, dehydration, macrocephaly, seizure, jumping and paralysis. Cases with severe hydrocephalus presumably due to an injection trauma and an inflammatory response against the DF-1 cells were excluded from survival analysis in this study. The RCAS-injection into cortex or SVZ area of adult mice was performed with minor modification as described previously (Marumoto et al., 2009). Relevant DF-1 cells were injected into 1-1.9 month old *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$ mice (cortex) for survival analysis of the RCAS-PDGFA and B injection, 2.2-3.2 month old *G/tv-a* and 2.9 month old *G/tv-a*$^{+/-}$*;Tg-Ef-luc* mice (SVZ) for sequential injection analysis on a stereotactic device. For the sequential injection, proneural tumors were initially generated by the co-injection of DF-1 cells producing the RCAS-PDGFA and RCAS-shp53 virus into putative SVZ area (2.3 mm depth) of *G/tv*-a or *G/tv-a*$^{+/-}$*;Tg-Ef-luc* adult mouse brain. Subsequently, DF-1 cells producing the RCAS-GFP-shNf1 virus were secondary injected in the brain at 2.7 mm depth using the same stereotactic coordinates after 2 to 4 weeks of first injection. *Cdkn2a(Ink4a-Arf)*$^{-/-}$ mice developed various spontaneous tumors such as lymphomas and sarcomas with faster latency than the RCAS-shNf1 alone or with the RCAS-Cre-induced brain tumors (Kamijo et al., 1999; Serrano et al., 1996). Therefore some of RCAS-shNf1 and RCAS-shNf1/Cre- induced tumors in this genetic background were incidentally identified when were sacrificed at the appearance of non-brain tumor-related symptoms. Summary of RCAS-mediated brain tumors in this study was listed in **Table S7**. Kaplan-Meier analysis demonstrating symptom-free survival in murine gliomas was performed using the Log-rank test in Prism 6 software (GraphPad). Log rank p values have been measured with the Mantle-Cox test. A value of $p < 0.05$ was considered significant in this study.

PCR

For an analysis of RCAS-DNA integration, DNAs were extracted from murine tumor sphere or paraffin sections of murine tumors using the DNeasy Blood and Tissue Kit (QIAGEN Inc. Valencia, CA). PCR amplification was performed with 1 μL of 100 ng DNA template, 1x PCR buffer minus Mg++, 0.2 mM each dNTP mixture, 1.5 mM MgCl2, 0.5 μM each primer, 1 μL of DMSO and 2.5 units of Taq DNA polymerase (Invitrogen) in a final volume of 20 μL on a MJ Research PTC-200 Thermo Cycler. PCR condition for murine tumor sphere was 5 min at 94°C, denaturation for 45 sec at 94°C, annealing for 30 sec at 55°C, and extension for 90 sec at 72°C for 40 cycles, followed by 10 min of final extension. For PCR amplification of paraffin sections, nested-PCR was used as shown in **Fig. S4G and H**. Primary reactions were performed with 25 cycles in same PCR condition and mixture as described above. Then 1 μL of the primary PCR product with x50 dilution was used for second amplification with 35

cycles for RACS-PDGFA detection or 40 cycles for tv-a detection as well. All primers are listed in **below**.

PCR; RCAS-DNA integration

Primer name	Direction	Sequence (5' -> 3')
R249-S1	F	AACTCTCTTGAAGTTGATCAG
R696-S1	F	TCCTCTCTTGAAGGAGCTATT
RCAS6	F	CAGCCTGAAAGCAGAATAGTA
RCAS5	F	CTCTGCTGGTGGCCTCGCGTACCACTG
RCAS3.2	R	CCCGTACATCGCATCGAT
GFP-S2	F	TACGGCAAGCTGACCCTGAAG
GFP-AS2	R	GTGTTCTGCTGGTAGTGGTC
mRFP-S1	F	CCCCGTAATGCAGAAGAAGA
mRFP-S2	F	CGAGGACGTCATCAAGGAGT
mRFP-AS2	R	GGAGCCGTACTGGAACTGAG
pSUP-AS	R	CAGTGTCACTAGGCGGGAACA
PA-AS1	R	CCTGACGTATTCCACCTTGG
PAQ-AS1	R	CAAAGAATCCTCACTCCCTACG
Tva 5	F	CTGCTGCCCGGTAACGTGACCGG
Tva 3	R	GCCCTGGGGAAGGTCCTGCCC
Tva-S2	F	TCCGGTAACGGTTCTTTGTC
Tva-AS3	R	CTCACCAGCTCACAGCAAAA

Q-PCR and QRT–PCR

DNA was extracted from human GBM cells using the DNeasy Blood and Tissue Kit (QIAGEN) and was then subjected to copy number analysis as described previously (Martinho et al., 2009; Ozawa et al., 2010). Total RNA was extracted from murine neurosphere infecting relevant RCAS virus using miRNeasy Mini kit (QIAGEN) and was used to synthesize cDNA by using the SuperScript system (Invitrogen) according to the manufacturer's protocol. SYBR Green real-time PCR was performed using relevant gene-specific primer sets, reagents, and protocols from Applied Biosystems in a 7900

HT Fast Real-Time PCR System (Applied Biosystems). Each sample was analyzed in quadruplicate. All primers are listed in **below.**

Q-PCR; DNA copy number analysis

Primer name	Direction	Sequence (5' -> 3')	Reference, Notes
hPDGFA-2	F	TGGTCGGTGGTCTCAGTGT	Intron 6, NG_029727.1
hPDGFA-2	R	GGCTCATCCTCACCTCACAT	Exon 7, NG_029727.1
hPDGFA-5	F	ATTATCGGGAAGAGGACACG	Exon 5, NG_029727.1
hPDGFA-5	R	TGTTCTCCAACACCGATGC	Intron 5, NG_029727.1
hEGFR-1	F	AACTGTGAGGTGGTCCTTGG	Exon 2, NG_007726.1
hEGFR-1	R	CCAACCTTTAAGAAGGAAAGATCA	Exon 2/Intron 2 junction, NG_007726.1
hEGFR-2	F	GTGGATGGCATTGGAATCA	Exon 24, NG_007726.1
hEGFR-2	R	TTAGCATCAGGATTATGACTCACC	Exon 24/Intron 24 junction, NG_007726.1
PDGFRA-F	F	TCAGCTACAGATGGCTTGATCC	(Martinho et al., 2009)
PDGFRA-R	R	GCCAAAGTCACAGATCTTCACAAT	(Martinho et al., 2009)
18S-F	F	GTAACCCGTTGAACCCCATT	(Martinho et al., 2009)
18S-R	R	CCATCCAATCGGTAGTAGCG	(Martinho et al., 2009)

QRT-PCR; murine neurosphere

Primer name	Direction	Sequence (5' -> 3')	Reference
mNf1-F1	F	ATGTGTTGGTAAACTCACTGCAC	
mNf1-R1	R	TCTTAGGCCACCAATCCAAC	
mTp53-F1	F	ATGCCCATGCTACAGAGGAG	
mTp53-R1	R	AGACTGGCCCTTCTTGGTCT	
mGAPDH-S1	F	ATGTTCCAGTATGACTCCACTCAC	(Bleau et al., 2009)

mGAPDH-AS1	R	GAAGACACCAGTAGACTCCACGACA	(Bleau et al., 2009)

Immunofluorescencent staining

5-µm formalin-fixed and paraffin-embedded (FFPE) tissue sections were stained as previously described (Charles et al., 2010). The following antibodies were used: GFP; (ab13970, abcam) at 1:250; RFP (ab62341, abcam) at 1:100; Alexa-fluor 488 goat anti-chicken IgG (A-11039) and 555 goat anti-rabbit IgG (A-21429) conjugated secondary antibodies (Invitrogen) at 1:250. For nuclear staining DAPI (Sigma Aldrich, USA) was used at 1:30,000. Analysis was performed by immunofluorescent microscopy (Leica DMI6000 microscope, FW4000 software).

Radiation

For the functional analysis of the RCAS-shp53 vector, tumor-bearing mice injected with the RCAS-shNf1 (GR249) with the RCAS-shp53 (R696) in *N/tv-a* mice or with the RCAS-Cre in *N/tv-a;Cdkn2a*$^{-/-}$*;Pten*$^{fl/fl}$ mice were sacrificed at 3 hours after 10Gy whole body irradiation, using a Cs-137 source at 115 cGy/min (Gammacell 40 Exactor, MDS Nordion). Immunohistochemical staining was performed on 5µm sections of formalin-fixed/paraffin-embedded tissues, on a Discovery XT automated staining processor (Ventana Medical Systems, Inc.), using anti-p53 (Leica) and anti-cleaved-caspase 3 (Cell Signaling) antibodies.

Supplemental References

Barry, W. T., Nobel, A. B., and Wright, F. A. (2005). Significance analysis of functional categories in gene expression studies: a structured permutation approach. Bioinformatics *21*, 1943-1949.

Bleau, A. M., Hambardzumyan, D., Ozawa, T., Fomchenko, E. I., Huse, J. T., Brennan, C. W., and Holland, E. C. (2009). PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. Cell Stem Cell *4*, 226-235.

Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. Proc Natl Acad Sci U S A *99*, 7877-7882.

Charles, N., Ozawa, T., Squatrito, M., Bleau, A. M., Brennan, C. W., Hambardzumyan, D., and Holland, E. C. (2010). Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. Cell Stem Cell *6*, 141-152.

Dai, M., Wang, P., Boyd, A. D., Kostov, G., Athey, B., Jones, E. G., Bunney, W. E., Myers, R. M., Speed, T. P., Akil, H.*, et al.* (2005). Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res *33*, e175.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439-3440.

Hambardzumyan, D., Parada, L. F., Holland, E. C., and Charest, A. (2011). Genetic modeling of gliomas in mice: new tools to tackle old problems. Glia *59*, 1155-1168.

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC bioinformatics *14*, 7.

Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: An interactive viewer for large phylogenetic trees. BMC bioinformatics *8*, 460.

Kamijo, T., Bodner, S., van de Kamp, E., Randle, D. H., and Sherr, C. J. (1999). Tumor spectrum in ARF-deficient mice. Cancer Res *59*, 2217-2222.

Kanehisa M, G. S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res *28*, 27-30.

Margush, T., and McMorris, F. R. (1981). Consensusn-trees. Bulletin of Mathematical Biology *43*, 239-244.

Martinho, O., Longatto-Filho, A., Lambros, M. B., Martins, A., Pinheiro, C., Silva, A., Pardal, F., Amorim, J., Mackay, A., Milanezi, F.*, et al.* (2009). Expression, mutation and copy number analysis of platelet-derived growth factor receptor A (PDGFRA) and its ligand PDGFA in gliomas. Br J Cancer *101*, 973-982.

Marumoto, T., Tashiro, A., Friedmann-Morvinski, D., Scadeng, M., Soda, Y., Gage, F. H., and Verma, I. M. (2009). Development of a novel mouse glioma model using lentiviral vectors. Nat Med *15*, 110-116.

Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., Garapati, P., Hemish, J., Hermjakob, H., Jassal, B.*, et al.* (2009). Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res *37*, D619-622.

Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol *12*, R41.

Moran PAP (1962). The Statistical Processes of Evolutionary Theory. Clarendon Press, Oxford, UK, p 200.

Patrakitkomjorn, S., Kobayashi, D., Morikawa, T., Wilson, M. M., Tsubota, N., Irie, A., Ozawa, T., Aoki, M., Arimura, N., Kaibuchi, K.*, et al.* (2008). Neurofibromatosis type 1 (NF1) tumor suppressor, neurofibromin, regulates the neuronal differentiation of PC12 cells via its associating protein, CRMP-2. J Biol Chem *283*, 9399-9413.

Sales, G., Calura, E., Cavalieri, D., and Romualdi, C. (2012). graphite - a Bioconductor package to convert pathway topology to gene network. BMC Bioinformatics *13*, 20.

Sancak, Y., Peterson, T. R., Shaul, Y. D., Lindquist, R. A., Thoreen, C. C., Bar-Peled, L., and Sabatini, D. M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. Science *320*, 1496-1501.

Schaefer, C. F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K. H. (2009). PID: the Pathway Interaction Database. Nucleic Acids Res *37*, D674-679.

Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., and DePinho, R. A. (1996). Role of the INK4a locus in tumor suppression and cell mortality. Cell *85*, 27-37.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics *27*, 431-432.

Smyth, G. K. (2005). Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.), Springer, New York, pages 397-420., 397-420.

Szerlip, N. J., Pedraza, A., Chakravarty, D., Azim, M., McGuire, J., Fang, Y., Ozawa, T., Holland, E. C., Huse, J. T., Jhanwar, S.*, et al.* (2012). Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. Proc Natl Acad Sci U S A *109*, 3041-3046.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor

Package. Journal of Statistical Software *36*.

Wu Z, I. R., Gentleman R, Martinez-Murillo F, Spencer F. (2004). A model-based background adjustment for oligonucleotide expression arrays. . Journal of the American Statistical Association *99*, 909–917.