Supporting Information

Kim et al. 10.1073/pnas.0909896107

SI Methods

Identification of microRNAs Whose Expression Is Driven by Copy-Number Alteration. Array comparative genomic hybridization (aCGH) measures DNA copy-number alterations. The aCGH, microRNA, and mRNA data were obtained from the recently published Cancer Genome Atlas (TCGA) data for glioblastoma (http://cancergenome.nih.gov/dataportal/). For quality control, we excluded samples in which the maximal percentage of necrosis was >40% or the maximal normal tissue percentage was >90%. When a sample had more than one portion, only one of the portions was selected for each patient. To identify miRNAs whose expression is driven by copy-number alteration, we performed correlation analysis between miRNA expression and copy-number log ratios across patients for each miRNA. A total of 176 samples having both aCGH and miRNA data available were used for this analysis. To avoid undue influence of outliers, we also computed the Spearman rank correlation coefficient. A list of top-ranked miRNAs are presented in Table S1.

Some high correlation coefficients were derived by copy-number deletion at the genome location of the miRNAs. Because we were interested in miRNAs that are overexpressed because of high copy-number amplification, we focused on the three miRNAs that are overexpressed due to copy-number amplification: miR-26a, miR-339, and miR-148a. The aCGH segment $\log_2$ values at the genome location of miR-339 and miR-148a were generally in the range of (0, 1), whereas the $\log_2$ values were much higher for miR-26a (i.e., aCGH segment $\log_2$ value at the miR-26 location >1.0 for many samples). Therefore, we chose miR-26a for further study. Whereas the correlation coefficient for miR-26a is high and significant, lack of stronger correlation can be attributed to several factors, including experimental noise and the presence of two host genes for hsa-miR-26a (CTDSPL on chromosome 3 and CTDSP2 on chromosome 12).

Correlation of Expression Profiles for miR-26a, CENTG1, CDK4, and CTDSP2. We investigated the relationship between the expression of miR-26a and the mRNA abundance of CENTG1, CDK4, and CTDSP2. Data analyses were performed with MATLAB (2009, The Mathworks) and the R software program (http://www.R-project.org). Pearson's correlation between miR-26a and these genes was significant: corr(miR-26a, CENTG1) = 0.39 with P value $<1 \times 10^{-6}$, corr(miR-26a, CDK4) = 0.51 with P value $<1 \times 10^{-6}$, and corr(miR-26a, CTDSP2) = 0.68 with P value $<1 \times 10^{-6}$, where CTDSP2 is a host gene of miR-26a. These correlations can be explained by coamplification of these elements for a number of samples. The correlations between genes were also significant; i.e., corr(CENTG1, CDK4) = 0.65 with P value $<1 \times 10^{-6}$, corr (CENTG1, CTDSP2) = 0.51 with P value $<1 \times 10^{-6}$, and corr (CDK4, CTDSP2) = 0.57 with P value $<1 \times 10^{-6}$.

Accurate Identification of microRNA Targets by Integrative Analysis. There are several state-of-the-art microRNA target prediction algorithms, including miRanda (1), miRBase (2), and TargetScan (3). These algorithms are primarily based upon the strength of predicted binding between microRNA and the 3′-UTR sequence of the target messenger RNAs. However, tissue-specific miRNA targets cannot be predicted accurately by binding energy calculation alone.

In our work, we took advantage of the multiplatform TCGA data set to investigate the downstream effects of miR-26a overexpression in glioblastoma. Most importantly, correlation between miR-26a expression and mRNA expression of its targets was used to capture the downstream effects of miR-26a. However, mRNA expression of the targets depends not only on miR-26a but also on DNA copy number alterations and DNA methylation that affect many genes in cancer. Thus, when we computed Pearson's correlation coefficient between the miR-26a expression profile and an mRNA expression profile, we excluded samples that fit the following criteria: (*i*) DNA copy number at the mRNA genomic location is amplified or deleted (i.e., |segment value| > 0.3), (*ii*) DNA methylation occurs at CpG islands near the mRNA genomic location (DNA methylation beta value > 0.2), (*iii*) maximal necrosis percentage >40%, and (*iv*) maximal normal percentage >90%.

We next selected for further analysis those genes who mRNA expression was inversely correlated (corr < −0.15) with miR-26a expression after computing Pearson's correlation coefficients for all available genes in the Agilent 244K Custom Gene Expression platform. The predicted target genes of the transcription factors regulating expression of the miR-26a host gene (CTDSP2) were also filtered out. In addition, complementary binding sites between the miR-26a sequence and the 3′-UTR sequences of candidate mRNAs were identified by RNAup (4, 5). Using this tissue-specific approach, a total of 961 candidate miR-26a target transcripts were identified (Table S2).

We next searched for genes related to key GBM-related pathways among the remaining genes. PTEN and RB1 were selected for biological validation because they are tumor suppressors that play an important role in GBM. MAP3K2 (mitogen-activated protein kinase kinase kinase 2, alias MEKK2, MEKK2B) was also selected because it can activate JNK (6), and JNK activation can promote apoptosis in human glioblastoma cells (7). The total binding energies between the miR-26a sequence and the 3′-UTR sequences of PTEN, MAP3K2, and RB1 were −11.79, −10.7, and −8.95, respectively. We do not claim that these are direct targets of miR-26a on the basis of these computed binding energies. Instead, we investigated the net effect of elevated miR-26a expression regardless of direct binding. The correlation coefficients for PTEN, MAP3K2, and RB1 were −0.38, −0.17, and −0.19, respectively, and their P values were all <0.05. We also checked to see whether these three genes could be captured if we did not use the selection criteria (i.e., |segment value| >0.3 and DNA methylation beta value >0.2) for computing Pearson's correlation coefficients between miR-26 expression and each mRNA expression. When we did not use the sample selection filters, the correlation coefficients for PTEN, MAP3K2, and RB1 were −0.14, −0.17, and −0.05, respectively; more specifically, the coefficients for PTEN and RB1 were > −0.15 and their P values were not statistically significant (>0.05). Thus, PTEN and RB1 could be found only after we used the sample selection filters described.

Survival Analysis. One hundred and seventy eight samples that had both copy number and clinical data were used for survival analysis. To test the correlation between copy-number amplification of the genomic location of miR-26a, CENTG1, or CDK4 and patient survival, we performed log-rank tests between two groups (the first group consisted of samples with high copy-number amplification (aCGH $\log_2$ value >2.0) and the second group consisted of samples without high copy number (aCGH $\log_2$ value ≤2.0). For example, miR-26a had 17 samples with high copy-number amplification and 161 samples without. The P value for the null hypothesis that these two groups belong to the same group was 0.0233. The median survival for the two groups was 209 days and 383 days, respectively. That is, when the genomic location of miR-26a is highly amplified, the median survival is substantially

shorter. Table S3 shows a total of 23 samples whose genome locations of miR-26, CENTG1, or CDK4 were highly amplified. Fig. 1*A* illustrates 32 samples that were highly amplified within this genome location. For survival analysis, we excluded samples of which survival was <30 days (such a short survival could be for other reasons such as postoperative complications from the surgery). The data in Table S3 show that CDK4 was highly amplified whenever miR-26a was highly amplified. Whenever miR-26a was highly amplified, CENTG1 was also highly amplified except for one sample (06-0177).

EGFP and Luciferase Reporter Assays. ***EGFP fluorescent reporter assay.*** miR-26a binding sites in the 3′-UTRs of *PTEN*, *RB1*, and *MEKK2* were identified using a publicly available microRNA target prediction program (www.microrna.org). PCR primers flanking the miR-26a binding sites were designed and RT-PCR was performed using total mammalian mRNA to generate cDNA fragments for the 3′-UTRs of *PTEN*, *RB1*, and *MEKK2*. The correct sequences were verified by direct DNA sequencing of the cloned fragments. These fragments were then subcloned into the pGEM-T Easy vector and subsequently inserted into pEGFP-C2 to generate green fluorescent protein (EGFP) fused to the 3′-UTRs of *PTEN*, *RB1*, or *MEKK2*. Human 293T cells were transfected with these constructs or with control constructs containing only EGFP. After 24 h, the transfected cells were plated into separate wells, thus ensuring equivalent transfection efficiency across experimental groups. The cells were subsequently exposed to a miR-26a mimic or a control miR-26a oligonucleotide (100 nM, Dharmacon). After an additional 24 h, the cells were fixed, stained with DAPI to label nuclei, and viewed under epifluorescence. Automated image analysis was performed using a publicly available image analysis software (rsbweb.nih.gov/ij/). The percentage of cells expressing GFP above a preset threshold level of 100 was determined for cultures exposed to miR-26a mimic or controls. Experiments were repeated in triplicate. Statistical significance was determined by unpaired *t* test. Data shown are mean ± SEM. The 3′-UTRs of *PTEN* ($P < 0.031$), *RB1* ($P < 0.038$), and *MAP3K2*/*MEKK2* ($P < 0.0001$) conferred suppression of EGFP fusion protein expression by the miR-26a mimic.

Luciferase reporter assay. As an alternative to the GFP-fusion protein reporter assays described above, luciferase reporter assays were also performed. RT-PCR was used to clone the 3′-UTR cDNA fragments for *RB1* and *MEKK2* into the pMIR-reporter luciferase vector (Applied Biosystems) to generate luciferase fused to the 3′-UTR of *RB1* and *MEKK2*. 293T cells were then transfected with these constructs or a control luciferase construct. A β-galactosidase vector was cotransfected to facilitate normalization for differences in transfection efficiency. After 24 h, the cells were exposed to either miR-26a mimic or miR-26a control oligonucleotides (100 nM). After an additional 24 h, the cells were lysed and luminescence was detected using a luminometer according to the manufacturer's protocol. Between four and eight wells were run for each condition. Experiments were repeated in triplicate. Data shown are mean ± SEM. The 3′-UTRs of *RB1* and *MAP3K2*/*MEKK2* conferred suppression of luciferase activity by the miR-26a mimic.

In Vitro Growth and Proliferation Assays. ***CCK-8 and MTT cell growth assays.*** A total of 1×10^3 GBM cells were plated into 96-well culture plates in triplicate, and cell growth was determined daily for 6 days using a tetrazolium salt-based colorimetric assay (Dojindo Molecular Technologies) according to the manufacturer's protocol.

The MTT cell growth assay was also used. LN229 or U87 GBM cells were plated in 96-well plates (1×10^4 cells/well) and maintained in serum-supplemented medium. MTT was added to the cultures and incubated at 37°C for 4 h. MTT reaction product was measured at 590 nm according to the manufacturer's instructions. Eight wells were run for each condition, and the mean and SD were calculated. To account for differences in plating density, the results were normalized using MTT values obtained around the time of first plating. Data were analyzed by *t* test.

BrdU proliferation assay. Quantitative measurement of DNA synthesis was performed using a bromo-deoxyuridine (BrdU) ELISA (Roche Applied Science). Briefly, GBM cells were plated at a density of 1×10^4 cells/well in 96-well plates. BrdU was added to the medium at a final concentration of 10 μM, and the cells were incubated for an additional 1 h. Newly synthesized BrdU-DNA was then assayed using colorimetric detection. Eight wells were run for each condition, and the mean and SD were calculated. In other experiments, BrdU incorporation into DNA was visualized using immunohistochemistry, and cell counts of DAPI-labeled nuclei were performed. Data were analyzed by *t* test.

Annexin V and live cell/dead cell apoptosis assays. Cultured U87 or LN229 GBM cells were plated at a density of 1×10^4 cells/well in 96-well plates and pretreated as indicated. To induce apoptosis, the cells were exposed overnight to camptothecin (50 μM) or to vehicle control. In some experiments, cells were exposed to the specific JNK inhibitor SP600125 (50 μM) or to a control inhibitor before camptothecin exposure. Annexin V-alexa 568 was then added to the cultures for 15 min, washed, and visualized according to the manufacturer's instructions.

For live cell/dead cell assays, cell viability was assayed using a commercially available kit (Invitrogen-Molecular Probes) that is based upon the accumulation of calcein acetomethoxy (AM) in the cytoplasm of healthy cells and ethidium homodimer-1 in the nuclei of dead cells. Direct cell counts were made under epifluorescence and the data were analyzed by *t* test.

Intracranial GBM Growth Assay. Human U87 GBM cells were first transduced with a lentivirus to generate a stable cell line expressing firefly luciferase. These cells were then transduced with a miR-26a lentivirus or a control virus and were subsequently transplanted (2×10^5cells/animal) into the frontal cortex of nude mice ($n = 10$ animals/group). The animals were then followed for ≈4 weeks. Intracranial tumor growth was imaged noninvasively at periodic time points following intraperitoneal administration of luciferin substrate using a Xenogen camera. The mean of eight to ten luminescence measurements was calculated for each tumor at each time point. The mean and SEM of the data from the animals in each group were calculated for each time point. Statistical significance was determined using the unpaired *t* test.

1. Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA.org resource: Targets and expression. *Nucleic Acids Res* 36 (Database issue):D149–D153.
2. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: Tools for microRNA genomics. *Nucleic Acids Res* 36 (Database issue):D154–D158.
3. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* 115:787–798.
4. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36:W70–W74.
5. Mückstein U, et al. (2006) Thermodynamics of RNA-RNA binding. *Bioinformatics* 22: 1177–1182.
6. Su B, Cheng J, Yang J, Guo Z (2001) MEKK2 is required for T-cell receptor signals in JNK activation and interleukin-2 gene expression. *J Biol Chem* 276:14784–14790.
7. Li L, et al. (2008) Glioma pathogenesis-related protein 1 exerts tumor suppressor activities through proapoptotic reactive oxygen species-c-Jun-NH2 kinase signaling. *Cancer Res* 68:434–443.

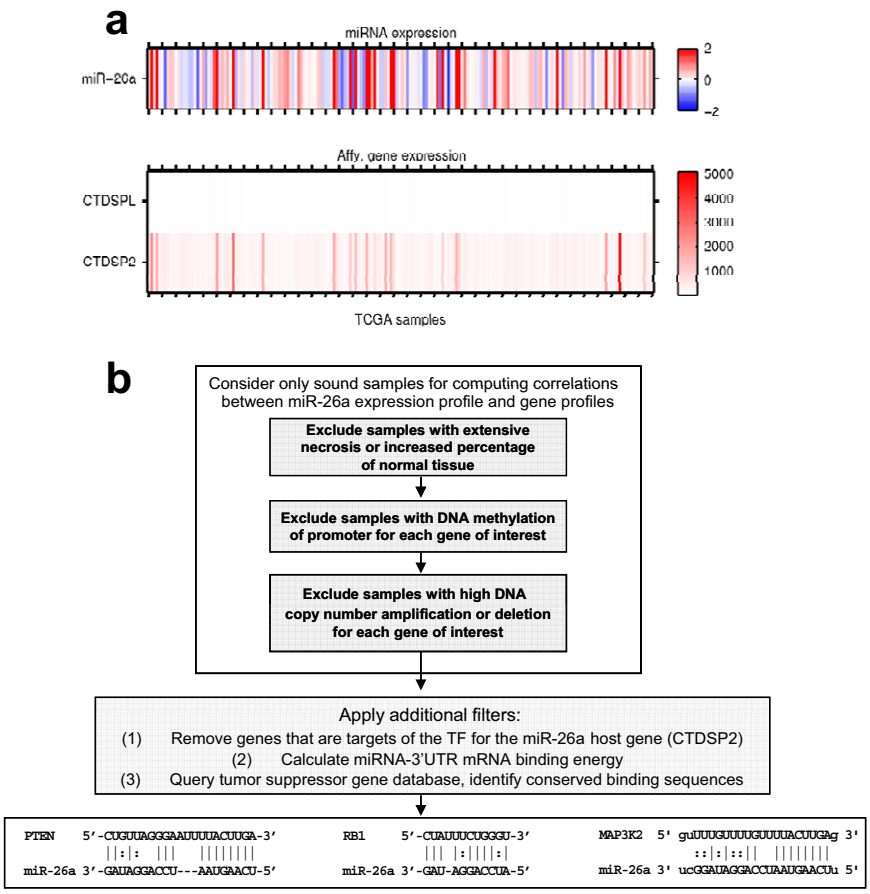

Fig. S1. (*a*) Correlation of CTDSP2 and CTDSPL mRNA with hsa-miR-26a expression. Diagram illustrates the correlation between hsa-miR-26a expression and mRNA expression for the *CTDSP2* gene (which is frequently amplified in GBM and located on chromosome12q) and the *CTDSPL* gene (which is not frequently amplified in GBM and is located on chromosome 3). The data reveal a close correlation between the mRNA expression for CTDSP2 and hsa-miR-26a. (*b*) Strategy for tissue-specific identification of microRNA targets. A computational method that integrates genome-scale DNA, mRNA, microRNA, and epigenetic and clinical data from specific tissues was developed to identify probable mRNA targets of miR-26a in human glioblastomas. The method relies upon calculating the correlations between the miR-26a expression profile and the mRNA expression profiles of all genes after corrections for specimen quality and copy-number- and methylation-driven changes in gene expression and selected transcription factor-mediated effects on gene expression. Once this filtering process was completed, RNA sequence analysis and mRNA/miR-26a binding energy calculations for each candidate transcript were performed. For additional details, see *SI Methods*.

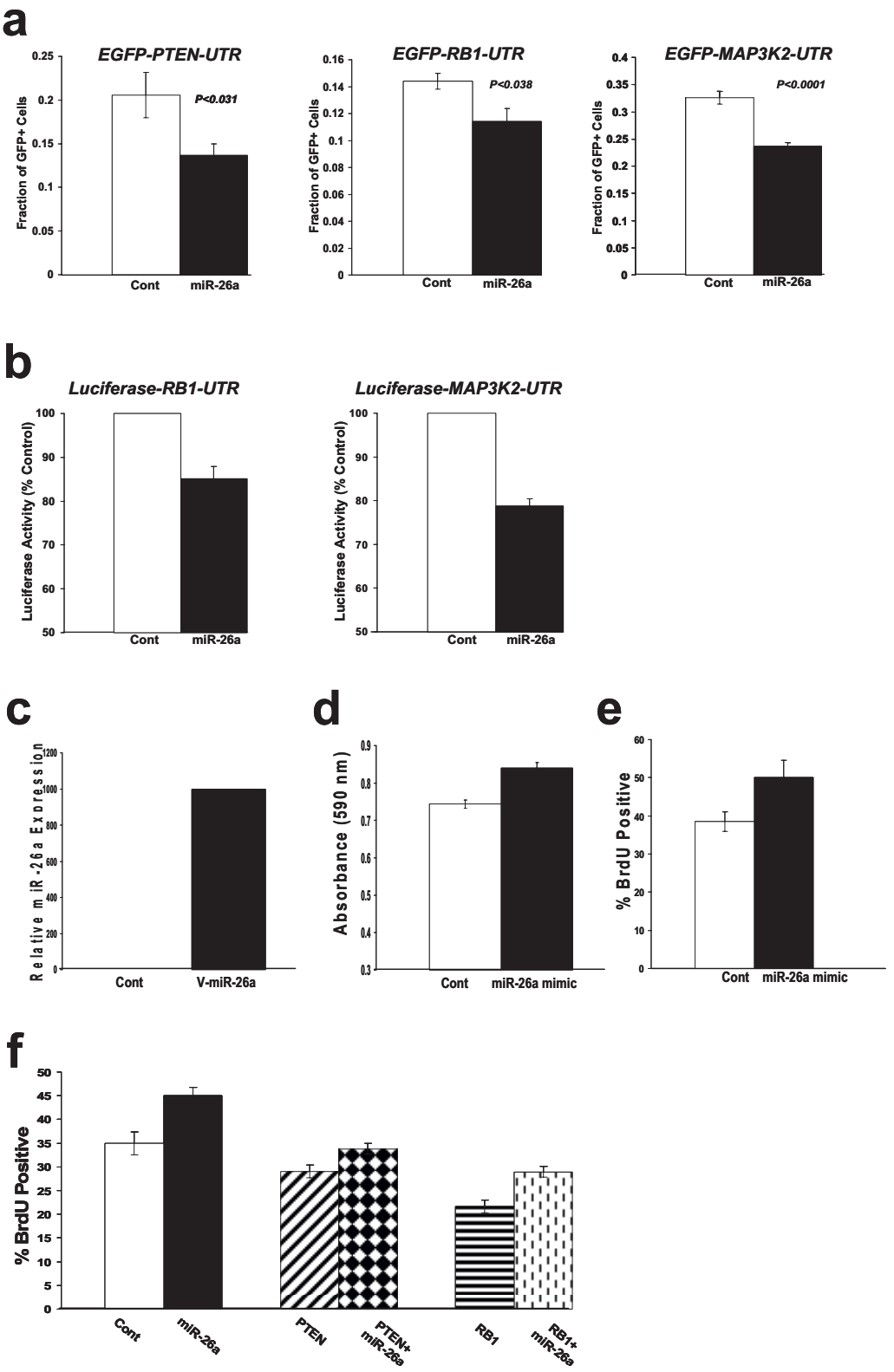

Fig. S2. miR-26a targets the 3′-UTRs of *PTEN*, *RB1*, and *MAP3K2* to promote GBM growth. (*a*) EGFP fluorescent reporter assay. Human 293T cells were transfected with a vector containing EGFP fused to the 3′-UTR of *PTEN*, *RB1*, or *MAP3K2*. Control cells were transfected with a control vector containing only EGFP. After 24 h, they were exposed to a miR-26a mimic or a control oligonucleotide (100 nM, Dharmacon). After an additional 24 h, the cells were fixed, stained with DAPI to label nuclei, and viewed under epifluorescence. Automated image analysis was performed using a publicly available image analysis software (rsbweb.nih.gov/ij/). Statistical significance was determined by an unpaired *t* test. Data shown are mean ± SEM. Experiments were repeated three times and showed that the 3′-UTRs of *PTEN* ($P < 0.031$), *RB1* ($P < 0.038$), and *MAP3K2*/*MEKK2* ($P < 0.0001$) conferred suppression of EGFP fusion protein expression by miR-26a. (*b*) Luciferase reporter assay. As an alternative to the GFP-fusion protein reporter assays described above, luciferase reporter assays

Legend continued on following page

were also performed. 293T cells were transfected with vectors containing luciferase fused to the 3′-UTR of *RB1* or *MEKK2* or with a control luciferase construct. A β-galactosidase vector was cotransfected to facilitate normalization for differences in transfection efficiency. After 24 h, the cells were exposed to either miR-26a mimic or control oligonucleotides (100 nM). After an additional 24 h, the cells were lysed and luminescence was detected using a luminometer according to the manufacturer's protocol. Between four and eight wells were run for each condition. Experiments were repeated in triplicate. Data shown are mean ± SEM. (*c*) Taqman real-time PCR analysis of miR-26a expression in U87 glioblastoma cells after transduction with a miR-26a lentivirus (V-miR-26a) or a control virus. Data shown are relative to control expression. (*d*) miR-26a promotes GBM cell growth. MTT growth assay is shown for LN229 human glioblastoma cells exposed to miR-26a mimic or an oligonucleotide inhibitor of miR-26a (200 nM) for 48 h. Data shown are mean ± SEM for eight replicates ($P \leq 0.0002$, unpaired *t* test). (*e*) BrdU proliferation assay for LN428 human glioblastoma cells after exposure to a miR-26a oligonucleotide mimic (100 nM) or a control oligonucleotide. Cells were incubated in BrdU (10 μM) for 1 h before fixation and staining for BrdU immunohistochemistry. Nuclei were counterstained with DAPI. BrdU-positive nuclei were counted and are represented as a percentage of the total number of nuclei present. Data shown are mean ± SEM for seven high-powered fields ($P \leq 0.044$, unpaired *t* test). (*f*) Overexpression of *PTEN* or *RB1* antagonizes the proliferative effect of miR-26a on GBM cell growth. LN229 GBM cells were transiently transfected with either a control vector or an expression vector for *PTEN* or *RB1*. The cells were then exposed overnight to a miR-26a oligonucleotide mimic (100 nM) or a control oligonucleotide. To assay proliferation, the cells were incubated in BrdU (10 μM) for 1 h before fixation and staining for BrdU immunoreactivity. Nuclei were counterstained with DAPI. BrdU-positive nuclei were counted and are represented as a percentage of the total number of nuclei present. Data shown are mean ± SEM for six high-powered fields. The miR-26a mimic increased proliferation ($P \leq 0.007$, unpaired *t* test), and this effect was abrogated by reexpression of *PTEN* or *RB1*.

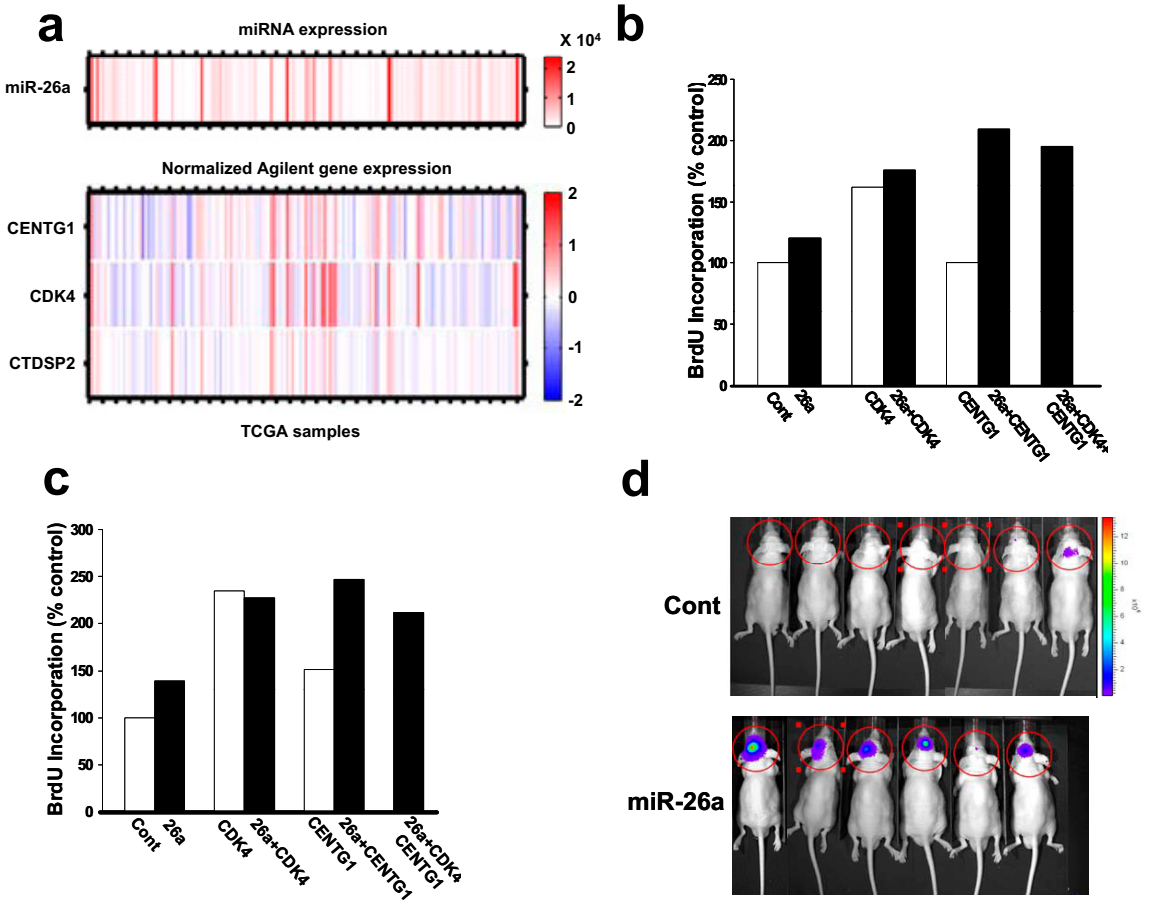

Fig. S3. Coordinated expression of 12q amplicon genes collaboratively promotes tumor growth. (*a*) Diagram illustrating correlation between hsa-miR-26a expression and mRNA expression for genes involved in the 12q amplicon, including *CENTG1*, *CDK4*, and *CTDSP2*. The data reveal a close correlation between the RNA expression for all four genes. (*b*) Expression vectors for hsa-miR-26a, *CDK4*, and *CENTG1* were transiently transfected alone or in combination into human HEK 293T cells. Equal amounts of control vector DNA were transfected as controls. After 48 h, BrdU proliferation assays were performed as described in *Methods*, and the results were expressed relative to those obtained using the appropriate control expression vectors. Note that miR-26a further increased DNA synthesis in 293T cells overexpressing either *CDK4* or *CENTG1*. (*c*) Expression vectors for hsa-miR-26a, *CDK4*, and *CENTG1* were transiently transfected alone or in combination into human U87 GBM cells lacking functional PTEN. Equal amounts of control vector DNA were transfected as controls. After 48 h, BrdU proliferation assays were performed as described in *Methods*, and the results were expressed relative to those obtained using the appropriate control expression vectors. Note that in the absence of PTEN, miR-26a further increased GBM cell proliferation when coexpressed with *CENTG1,* but not with *CDK4*. (*d*) Composite image of intracranial growth of human U87 GBM cells transduced with a miR-26a lentivirus or a control virus and subsequently transplanted into the mouse brain. The cells had also been transduced before this experiment to stably express firefly luciferase. After administration of luciferin substrate, intracranial tumor growth was imaged noninvasively using a Xenogen camera over a period of 4 weeks. Images shown here were obtained 4 weeks after transplantation. The luminescence data obtained from this experiment are displayed quantitatively in Fig. 5*B*. miR-26a significantly increased the intracranial growth of PTEN-deficient U87 GBM cells in vivo ($P < 0.0089$, unpaired *t* test).

Other Supporting Information Files

Table S1 (DOC)
Table S2 (DOC)
Table S3 (DOC)