Journal of Pathology
J Pathol 2013; **229:** 449–459
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.4109

ORIGINAL PAPER

Transcriptome and small RNA deep sequencing reveals deregulation of miRNA biogenesis in human glioma

Lynette M Moore,[1] Virpi Kivinen,[2] Yuexin Liu,[1] Matti Annala,[2] David Cogdell,[1] Xiuping Liu,[3] Chang-Gong Liu,[3] Raymond Sawaya,[4] Olli Yli-Harja,[2] Ilya Shmulevich,[5] Gregory N Fuller,[1] Wei Zhang[1]* and Matti Nykter[2]*

[1] *Department of Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, 77030, USA*
[2] *Department of Signal Processing, Tampere University of Technology, FI-33101, Tampere, Finland*
[3] *Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, 77030, USA*
[4] *Department of Neurosurgery, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, 77030, USA*
[5] *Institute for Systems Biology, 401 Terry Avenue North, Seattle, WA, 98109, USA*

**Correspondence to: Wei Zhang, PhD, Department of Pathology, Unit 85, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA. e-mail: wzhang@mdanderson.org*
Matti Nykter, Institute of Signal Processing, Tampere University of Technology, PO Box 553, FI-33101 Tampere, Finland. e-mail: matti.nykter@tut.fi

Abstract

Altered expression of oncogenic and tumour-suppressing microRNAs (miRNAs) is widely associated with tumourigenesis. However, the regulatory mechanisms underlying these alterations are poorly understood. We sought to shed light on the deregulation of miRNA biogenesis promoting the aberrant miRNA expression profiles identified in these tumours. Using sequencing technology to perform both whole-transcriptome and small RNA sequencing of glioma patient samples, we examined precursor and mature miRNAs to directly evaluate the miRNA maturation process, and examined expression profiles for genes involved in the major steps of miRNA biogenesis. We found that ratios of mature to precursor forms of a large number of miRNAs increased with the progression from normal brain to low-grade and then to high-grade gliomas. The expression levels of genes involved in each of the three major steps of miRNA biogenesis (nuclear processing, nucleo-cytoplasmic transport, and cytoplasmic processing) were systematically altered in glioma tissues. Survival analysis of an independent data set demonstrated that the alteration of genes involved in miRNA maturation correlates with survival in glioma patients. Direct quantification of miRNA maturation with deep sequencing demonstrated that deregulation of the miRNA biogenesis pathway is a hallmark for glioma genesis and progression.

Keywords: microRNA; biogenesis; glioma

Received 15 May 2012; Revised 12 September 2012; Accepted 13 September 2012

No conflicts of interest were declared.

Introduction

MicroRNAs (miRNAs) are a class of conserved, small, non-coding RNAs that control gene expression by binding to complementary sequences at the 3' untranslated regions (UTRs) of target messenger RNAs (mRNAs), resulting in translational repression or mRNA degradation [1]. MiRNAs have been shown to play important roles in mammalian systems by influencing genes involved in processes such as cell proliferation, apoptosis, and tumourigenesis [2]. Many miRNAs have been designated as oncogenes ('oncomiRs') or tumour suppressors based on the effects of the miRNAs on cells and the functions of the mRNA target genes [3–6].

Similar to mRNAs, miRNAs can be regulated at the transcriptional level by DNA-binding transcription factors or epigenetic mechanisms [4,7] and post-transcriptionally by a multistep processing pathway [8]. The lack of correlation between primary miRNA (pri-miRNA) transcripts and mature miRNAs in tumours and the association of miRNA processing factors with tumourigenesis in cell culture and mouse model studies indicate that deregulation of the biogenesis pathway is likely to be a key player in the aberrant miRNA expression profiles observed in cancer [9].

MiRNA biogenesis is controlled by the multistep miRNA processing pathway [10]. The pri-miRNA transcript is transcribed by RNA polymerase II (in some cases, polymerase III) in the nucleus. For intergenic miRNA, the pri-miRNA is cleaved into a short, 60 to 70-nucleotide (nt), hairpin precursor miRNA (pre-miRNA) by a microprocessor unit containing the nuclease, *RNASEN* (also known as *DROSHA*), and other factors [11]. MiRNA located in intronic regions or within exons of protein-coding genes can be cleaved by splicing machinery to generate the pre-miRNA [12]. This nuclear processing is followed

by the transport of pre-miRNA from the nucleus into the cytoplasm via exportin-5 (*XPO5*) [13]. In the cytoplasm, the pre-miRNA is cleaved into an approximately 19–25-nt-long mature form of the miRNA by the ribonuclease Dicer (*DICER1*) [14] or by a Dicer-independent maturation process that is beginning to be revealed [15]. The mature miRNA is loaded into the RNA-induced silencing complex (RISC), where it initiates translational repression or degradation of target mRNAs [16,17]. Thus, miRNA biogenesis is tightly controlled by a set of protein-coding genes that eventually lead to the production of functionally mature miRNAs in the cytoplasm.

The deregulation of the miRNA biogenesis pathway has been associated with various cancers. This is supported by the large number of differentially expressed mature miRNAs that have been identified in multiple cancers [18]. Cell culture studies and animal models have demonstrated that deregulation of miRNA expression by knockdown of key miRNA biogenesis factors enhances tumourigenesis [19]. The expression levels of these factors have also been associated with prognosis. For example, elevated expression of RNASEN accelerates the proliferation of oesophageal squamous cell carcinoma cells and is a negative prognostic marker in oesophageal cancer patients [20], and reduced expression of *DICER1* is associated with poor prognosis in lung and in ovarian cancer patients [21,22], whereas overexpression has been linked to poor prognosis in colon cancer [23]. Mutations have also been identified in biogenesis components, including *XPO5* and the Dicer cofactor *TARBP2* in colon cancers [24,25]. It should be noted that the role of these processing genes in tumourigenesis remains complex. For example, while some studies indicate that for *DICER1*, loss of one copy is advantageous to tumour growth while complete loss is disadvantageous, others have demonstrated that in some instances, *DICER1* null cells can maintain tumourigenic capacity [26,27].

A number of studies have reported changes in the steady-state levels of mature miRNAs in glioma [28–30]. One of the best-characterized events is the elevation of oncomiRs miR-21 and miR-221 [30–35]. However, the mechanisms underlying deregulation of miRNA biogenesis in glioma are unknown. As mutations in biogenesis components such as XPO5 and TARBP2 have not been identified in glioma, the clinical significance of deregulated biogenesis genes remains uncertain.

We sought to thoroughly examine the miRNA processing pathway by taking advantage of deep sequencing technology to directly measure the maturation process of all the miRNAs and to identify key changes in biogenesis gene expression. This analysis provides a comprehensive view of the deregulation of miRNA biogenesis in glioma and the subsequent miRNA expression-profile alterations. This analysis shows that miRNA maturation is tightly regulated and that the multilevel regulatory process is altered at each step during glioma genesis and progression. Correlation of miRNA biogenesis gene expression with survival information from a major glioma database further demonstrates the significance of these findings.

Materials and methods

Patient samples

Human glioma samples were obtained from The University of Texas MD Anderson Cancer Center Brain Tumor Center tissue bank and were collected under an institutional review board-approved protocol. Commercially available normal brain RNA pooled from multiple donors was used as a reference (Ambion, Carlsbad, CA, USA).

Library preparation and RNA sequencing

Library preparation for both whole-transcriptome sequencing and small RNA sequencing was performed using Applied Biosystems Incorporated's (ABI, Carlsbad, CA, USA) whole-transcriptome and small RNA sequencing protocols. Sequencing runs were performed using ABI's SOLiD System version 3.5 for both whole-transcriptome sequencing and small RNA sequencing. From the whole-transcriptome sequencing and small RNA sequencing, over 610 million and 230 million 50-nt-long sequencing reads were obtained, respectively. A detailed description is provided in the Supplementary methods.

Sequencing data analysis

Sequencing reads were aligned against transcript sequences from the National Center for Biotechnology Information (NCBI) reference sequence build version 38 using Bowtie version 0.12.5 [36]. Gene expression levels were quantified as reads per kilobase per million (RPKM)-normalized expression values and quantile-normalized across sample pools.

Small RNA sequencing reads were aligned against 721 pre-miRNA and 904 mature miRNA sequences from miRBase build 14 [37]. We excluded from further analysis miRNAs with a median number of reads less than 20 in mature miRNAs or pre-miRNAs across all sample pools, which left us with 505 miRNAs for subsequent analysis. Expression levels were normalized by the total number of mappable reads per sequencing experiment.

Prior to sequencing, we used an Agilent miRNA microarray (Agilent Technologies, Santa Clara, CA, USA) to characterize each individual sample. For each sequencing pool, the mean Pearson correlation between the samples was above 0.91, indicating good consistency of the pooled samples (data are described in the Supplementary methods).

Deep sequencing data have been deposited in short read archives under accession number SRA057775.

Microarray data analysis

Raw gene expression array data were downloaded from the Repository of Molecular Brain Neoplasia Data (https://caintegrator.nci.nih.gov/rembrandt/) and summarized into gene expression values using the robust multichip average (RMA) algorithm [38]. Differential expression between high- and low-grade was tested with a t-test under means are equal null hypothesis. Bonferroni corrected p values less than 0.05 were considered significant. Enrichment was tested with a hypergeometric distribution, using all the genes on the array as a background.

Clustering analysis

Gene expression profiles of biogenesis genes were scaled to an interval of 0–1. A K-means clustering algorithm with squared Euclidean distance was applied to identify gene clusters. Based on a silhouette plot, three clusters provided a good separation for the data.

Survival analysis

Kaplan–Meier analysis was used to determine the survival effect based on gene expression data. Two groups of samples were defined by applying K-means clustering to the expression profiles of a group of genes. The statistical significance of Kaplan–Meier curves was evaluated with the log-rank test.

Cell culture

SNB19 human glioblastoma cells were cultured in Dulbecco's modified Eagle's medium nutrient mixture F-12 (DMEM:F12) (Life Technologies, Carlsbad, CA, USA) with 10% fetal bovine serum (FBS). Small interfering RNA (siRNA) SMART pools (Dharmacon, Lafayette, CO, USA) targeting *DICER1*, *EIF2C1*, *EIF2C2*, and *SMAD5* were transfected using RNAiMax (Life Technologies) according to the manufacturer's instructions. Cells were harvested after 48 h and total RNA was isolated using the mirVana Kit (Ambion).

Quantitative RT-PCR

Total RNA from individual patient samples represented in the pooled samples for the sequencing analysis was used for these experiments. The precursor forms of miRNAs were detected using SYBR green and normalized to GAPDH. The mature forms of miRNAs were assayed using the TaqMan MicroRNA Assay Kit (ABI) according to the manufacturer's protocol and normalized to RUN6B.

The primer sequences were as follows: human pre-miR-21: 5' TGTCGGGTAGCTTATCAGAC and 5' TGTCAGACAGCCCATCGACT; human pre-miR-1912: 5' CTCTAGGATGTGCTCATTGC and 5' AAT CTCTATTATGTTCACAC; human GAPDH: 5' ACCA CAGTCCATGCCATCAC and 5' TCCACCACCCT-GTTGCTGTA.

MiRNA northern blot analysis

MiRNA northern blot analysis was performed as previously described [39]. Briefly, 5–20 µg of total RNA was resolved on a 12% acrylamide/urea gel and transferred to a nylon membrane. Blots were hybridized overnight with 10 pmol of ^{32}P-labelled locked nucleic acid (LNA) miRNA probes (Exiqon, Woburn, MA, USA).

Results

Alteration of miRNA maturation and miRNA biogenesis genes in gliomas

We performed deep sequencing analysis of both pre-miRNAs and mature miRNAs from 40 glioma patient samples and normal brain reference RNA. The glioma samples included patients diagnosed with both low-grade gliomas [oligodendroglioma (Oligo)] and high-grade gliomas [anaplastic oligodendroglioma (AO), anaplastic astrocytoma (AA), and glioblastoma (GBM)]. In the interest of cost, we pooled 20 GBM samples into four groups and pooled each of the other glioma subtypes into individual pools (one AA, one AO, and two Oligo pools; Table 1).

Analysis of both pre-miRNA and mature miRNA expression levels revealed widespread alterations in the abundance of many miRNAs. The 20 miRNAs with the highest number of sequencing reads are shown in Figure 1A. While the abundance of many mature miRNAs, including miR-21, changed with glioma grade (Figure 1A, top), the abundance of the corresponding precursor forms remained unchanged, suggesting that the alterations in mature miRNA levels may be due to deregulated miRNA processing (Figure 1A, bottom). Similar discrepancies between mature and pre-miRNA levels have also been previously observed [40]. For the most abundant pre-miRNA (Figure 1A, bottom), we did not observe increased processing and thus, they do not appear among the most abundant mature miRNAs. We calculated the ratio of mature miRNA to pre-miRNA (M/P) for each miRNA as a measure of miRNA maturation (processing from precursor to mature form) (Supplementary Figure 1) and identified 35 miRNAs whose change in M/P ratio correlated with tumour grade (Spearman rank correlation; $p < 0.01$) (Figure 1B). The expression levels of mature and pre-miRNA forms of these 35 miRNAs are indicated in Supplementary Figure 2. The M/P ratios of three of the miRNAs were decreased in gliomas and were inversely associated with tumour grade. The M/P ratios of the other 32 miRNAs were increased in gliomas and were positively associated with increasing glioma grade. Among the identified differentially matured miRNAs, several have been previously associated with tumour development and prognosis in glioma [24], including miR-21 and miR-221 (Figures 1C–1E). We subsequently validated these findings by measuring

Table 1. Glioma patient sample information. Samples were grouped into sequencing pools based on available clinical information. The total number of sequencing reads for each pool is also shown

Sample No	Grade	Survival (months)	Sequencing pool	Average survival	Whole-transcriptome reads	Small RNA reads
1	GBM	6	A	<6 months	75 321 717	20 724 499
2		3				
3		4				
4		1				
5		6				
6	GBM	11	B	10–15 months	75 295 418	21 275 202
7		10				
8		11				
9		13				
10		15				
11	GBM	15	C	Intermediate > 1 year	52 438 870	18 597 970
12		18				
13		>12				
14		20				
15		>12				
16	GBM	>12	D	>1–2 years	67 066 205	22 944 568
17		>24				
18		>48				
19		>84				
20		29				
21	Oligo	48	E	NA	58 344 119	19 075 166
22		>48				
23		>48				
24		>72				
25	Oligo	>72	F	NA	66 169 704	18 095 113
26		>36				
27		36				
28		>36				
29	AO	>36	G	NA	87 724 772	34 399 996
30		2				
31		24				
32		>60				
33		24				
34		60				
35		NA				
36	AA	>36	H	NA	70 196 537	36 178 145
37		35				
38		NA				
39		>36				
40		>36				
41	Adult normal	NA	I	NA	57 701 862	42 319 647

GBM = glioblastoma; Oligo = oligodendroglioma; AO = anaplastic oligodendroglioma; AA = anaplastic astrocytoma; NA = no survival data available or not applicable.

the precursor and mature forms of miR-21 and miR-1912 by quantitative RT-PCR in a subset of patient samples which were represented in the sequencing pools (Supplementary Figure 3). These results indicate that the changes in mature miRNA expression are likely a function of the deregulated miRNA processing.

We reasoned that the extensive changes in miRNA maturation revealed by our deep sequencing analysis were a result of altered expression of the genes that regulate the miRNA biogenesis pathway. We therefore compiled a list of genes with known roles in the miRNA biogenesis pathway and examined their gene expression patterns by whole-transcriptome sequencing in the same glioma patient samples (Table 2). Gene expression analysis revealed that 14 genes examined were deregulated in the glioma progression, demonstrating deregulation at each of the three major steps of miRNA biogenesis (Figure 2 and Table 2). Statistical analysis with an independent cohort of samples from the Rembrandt glioma database validated eight of these 14 genes to be deregulated (Table 2). Enrichment analysis indicated that more biogenesis pathway genes are deregulated in glioma than expected by random ($p = 2.0711\text{e-}04$, hypergeometric test). Clustering analysis of miRNA biogenesis pathway genes identified three clusters: one that correlated positively with increasing tumour grade (cluster 1); one that correlated negatively with tumour grade (cluster 3); and one that was without a distinctive pattern among the glioma samples (cluster 2).

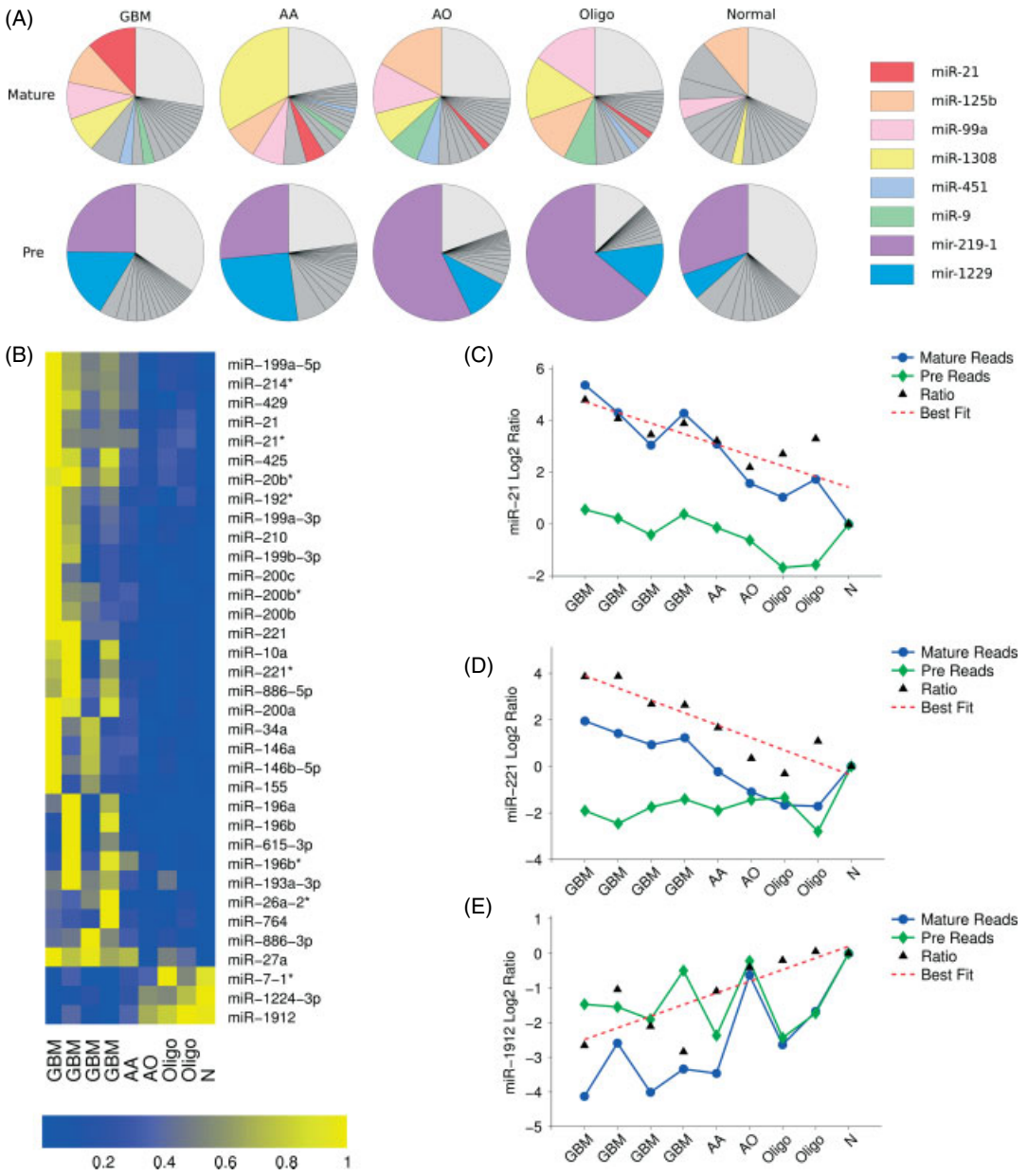

Figure 1. Discovery of miRNAs with increased maturation in glioma. (A) Abundance of mature miRNA and pre-miRNA in different tumour grades revealed by small RNA sequencing. Twenty miRNAs with the highest number of sequencing reads are shown as individual sectors with the sector size representing the number of reads for a given miRNA. The remaining miRNAs are contained in a single sector (light grey). Some of the miRNAs with large changes in their abundance are highlighted in colour. (B) M/P miRNA ratio across all samples. MiRNAs with M/P ratios that are significantly correlated with glioma grade (correlation, $p < 0.01$) and display differential maturation (M/P ratio > 2-fold) between gliomas and normal brain are shown. For visualization, the M/P ratio of each miRNA is scaled to a 0 (lowest expression ratio) to 1 (highest expression ratio) interval, as indicated by the colour bar. (C–E) Log_2 fold change of pre-miRNA (green) and mature miRNA (blue) read counts relative to read counts of normal brain, M/P expression ratio relative to the M/P expression ratio of normal brain, and linear fit to the M/P expression ratios are shown for typical examples: (C) miR-21, (D) miR-221, and (E) miR-1912.

Glioma miRNA maturation network signatures are predictive of survival

To take advantage of both mRNA and miRNA expression from the same samples, we constructed a network representation of miRNA maturation by correlating the M/P ratio of each of the 35 miRNAs with the gene expression of miRNA biogenesis genes. In Figure 3A, statistically significant correlations (Pearson correlation; $p < 0.01$) between expression levels of the biogenesis genes and the M/P ratios of miRNAs are shown. Biogenesis genes were grouped based on the biogenesis pathway steps with which they are primarily associated. This network indicates that several genes are correlated with the maturation of each miRNA. Furthermore, we found that the M/P ratio of each miRNA is correlated with genes from different steps of the biogenesis process, indicating the regulation of maturation at multiple steps of the miRNA processing pathway. For example, the transforming growth factor β (*TGF*β)/bone morphogenetic protein (*BMP*)/*SMAD* signalling pathway

Table 2. miRNA biogenesis pathway genes and statistical analysis. The list is compiled by literature review. Roles of the genes in the pathway are annotated in the table. Bonferroni corrected *p* values from Student's *t*-test for miRNA biogenesis genes using Rembrandt glioma data are shown

Gene name	Entrez ID	Role in pathway	Whole-transcriptome sequencing data set *p* value (*U*-test H0: medians are equal)	Rembrandt glioma data set *p* value (*t*-test H0: means are equal)
ADAR	103	Nuclear processing	0.785714	1
BMP1	649	Nuclear processing	0.035714	0.000137
BMP2	650	Nuclear processing	0.035714	0.007533
DCP1A	55802	Cytoplasmic processing	0.142857	1
DCP2	167227	Cytoplasmic processing	0.571429	1
DDX17	10521	Nuclear processing	0.035714	1
DDX20	11218	Cytoplasmic processing	0.785714	0.00959
DDX5	1655	Nuclear processing	0.035714	1
DDX6	1656	Cytoplasmic processing	0.035714	0.000055
DGCR8	54487	Nuclear processing	0.571429	1
DHX9	1660	Cytoplasmic processing	0.071429	1
DICER1	23405	Cytoplasmic processing	0.142857	1
EIF2C1	26523	Cytoplasmic processing	0.571429	1
EIF2C2	27161	Cytoplasmic processing	1	1
EIF2C3	192669	Cytoplasmic processing	0.392857	1
EIF2C4	192670	Cytoplasmic processing	0.142857	1
GEMIN4	50628	Cytoplasmic processing	0.25	1
HNRPA1	3178	Nuclear processing	0.035714	0.002531
MOV10	4343	Cytoplasmic processing	0.035714	0.000003
PRKRA	8575	Cytoplasmic processing	0.785714	NaN
PRMT5	10419	Cytoplasmic processing	0.142857	0.128026
RANGAP1	5905	Transport	0.25	1
RNASEN	29102	Nuclear processing	0.035714	0.000953
SMAD1	4086	Nuclear processing	1	1
SMAD2	4087	Nuclear processing	0.035714	1
SMAD3	4088	Nuclear processing	0.035714	1
SMAD4	4089	Nuclear processing	0.035714	0.242406
SMAD5	4090	Nuclear processing	0.035714	1
SMAD6	4091	Nuclear processing	1	1
SMAD7	4092	Nuclear processing	0.035714	0.046877
SMAD9	4093	Nuclear processing	0.035714	0.006042
TARBP2	6895	Cytoplasmic processing	1	1
TGFB1	7040	Nuclear processing	0.142857	3.04E-008
TNRC6A	27327	Cytoplasmic processing	0.321429	1
TNRC6B	23112	Cytoplasmic processing	0.25	0.000297
XPO5	57510	Transport	0.571429	1

NaN: probe set for gene was not included in microarray.

genes, which are part of the nuclear processing pathway, are connected to 20 out of the 35 differentially matured miRNAs. Overall, the network analysis uncovered complex relationships between miRNA maturation and genes involved in the biogenesis process. This suggests that the processing of pre-miRNA to mature miRNA can be affected by changes in the expression of genes in all steps of the biogenesis pathway.

In order to evaluate the clinical relevance of the deregulated miRNA biogenesis pathway, we used the genes in the network representation to define a gene expression signature for each of the three major biogenesis steps (nuclear processing, nucleocytoplasmic transport, and cytoplasmic processing). Using the Rembrandt glioma data as an independent data set, we examined the relationship between these signatures and patient survival. All three signatures were associated with survival (Figure 3B). The strongest correlation was observed for cytoplasmic processing, where overall low expression indicates poor survival ($p = 5*10e-9$). Nuclear processing showed the opposite trend, but the result was only marginally significant ($p = 0.02$). This survival correlation is in good agreement with the clusters identified in Figure 2. Cluster 1 includes primarily genes involved in the nuclear processing of pri-miRNA and cluster 3 includes genes involved in the cytoplasmic processing.

Knockdown of miRNA biogenesis genes inhibits miRNA maturation in glioma cell lines

Our analysis suggested that changes in biogenesis gene expression resulted in changes in the maturation of a set of miRNAs. To confirm this regulatory relationship in glioma, we examined the effect of altered biogenesis gene expression on the miRNA M/P ratio in the SNB19 glioma cell line. Focusing on biogenesis factors involved in cytoplasmic and nuclear miRNA processing, we performed knockdown studies of *DICER1*, *EIF2C1*, *EIF2C2*, and *SMAD5* genes to demonstrate a direct relationship between gene

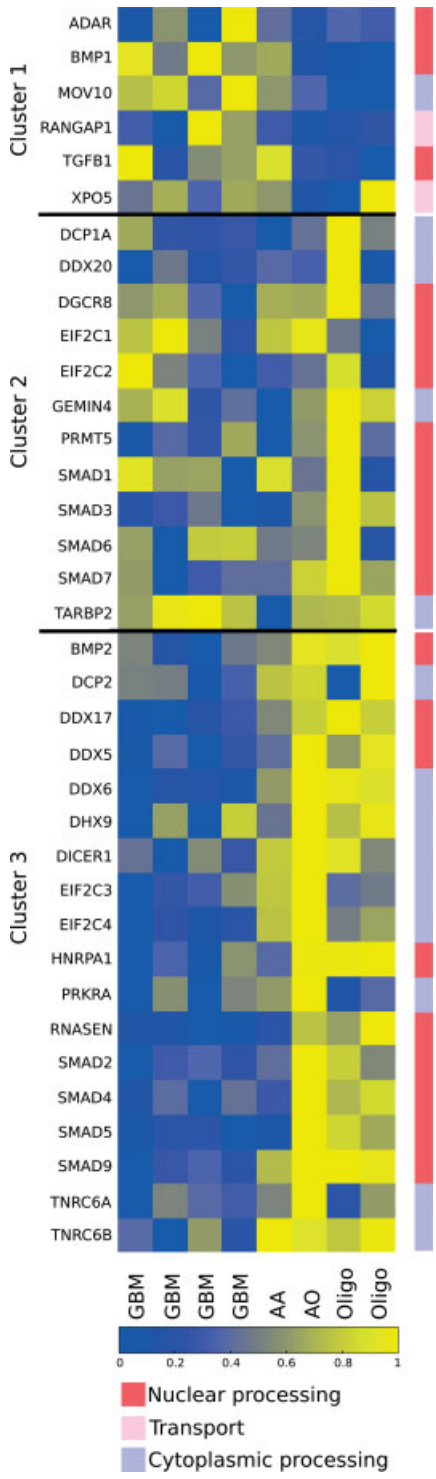

Figure 2. Altered gene expression in the miRNA biogenesis pathway. Expression levels of key miRNA biogenesis genes revealed by whole-transcriptome sequencing are shown in the heat map. Data are shown as $\log_2$ expression ratios, relative to normal brain. Data of each gene are scaled such that the minimum value is shown in blue and the maximum in yellow. Differential expression among different grades of glioma is indicated by the colour intensity, represented by the accompanying colour bar. The best-known association of each gene with the specific miRNA biogenesis step is also indicated. Clustering analysis identified three groups. Genes within cluster 1 have a positive correlation with tumour grade, and genes within cluster 3 have a negative correlation with tumour grade. Cluster 2 does not demonstrate a distinct pattern relative to glioma grade.

expression and miRNA maturation. Knockdown of *DICER1*, *EIF2C1*, *EIF2C2*, and *SMAD5* expression by siRNA in the SNB19 GBM cells was confirmed using real-time PCR (Figure 4A). We then examined the expression of both pre-miRNAs and mature miRNAs by northern blot analysis using probes specific for miR-21 and miR-221 (Figure 4B). Inhibition of each of these miRNA processing genes promoted a change in the M/P ratio (Figure 4C). Specifically, decreasing the levels of miRNA processing genes that were found to be deregulated in glioma (*DICER1*, *EIF2C1*, *EIF2C2*, and *SMAD5*) directly resulted in decreased M/P ratios, validating our observations of M/P ratio changes in the sequencing data. Knockdown of the cytoplasmic processing genes (*DICER1*, *EIF2C1*, and *EIF2C2*) had a greater impact on the M/P ratio than knockdown of a nuclear processing gene (*SMAD5*), which is indicative of the more direct effects of the cytoplasmic processing genes in this validation experiment. However, this study confirms that deregulation of either the cytoplasmic or the more upstream nuclear processing steps can ultimately affect the M/P ratio. These studies confirmed that the miRNA M/P ratios are modulated by changes in expression of the processing genes.

Discussion

MiRNAs represent key nodes in a regulatory network that modulates diverse biological processes. Thus, it is not surprising that deregulation of miRNAs markedly contributes to tumourigenesis [42]. The published literature has focused on changes in the steady-state levels of oncogenic and tumour-suppressing miRNAs and their protein-coding targets. Recent studies have begun to reveal upstream alterations of genes that regulate miRNA production [43].

The studies presented here demonstrate that the altered miRNA-expression profiles of glioma are at least in part due to widespread gene expression changes that affect miRNA processing. Using deep sequencing technologies, we dissected part of the miRNA biogenesis process by directly measuring the miRNAs at the pre- and mature steps in a set of 40 glioma samples representing different grades. We calculated the ratios of mature to pre-miRNAs to obtain the M/P ratio, which represents the biogenesis, or maturation process, of each miRNA. The M/P ratios revealed that the miRNA maturation process is markedly activated for a number of miRNAs in glioma genesis, including miR-21 and miR-221, in such a way that mature forms are actively produced. We also identified a small number of miRNAs whose maturation process is suppressed in glioma compared with normal brain, suggesting a role in glioma suppression.

By integrated analysis of both whole-transcriptome and small RNA sequencing results from the same set of samples, we gained insight into specific defects in the miRNA regulatory network in glioma and revealed

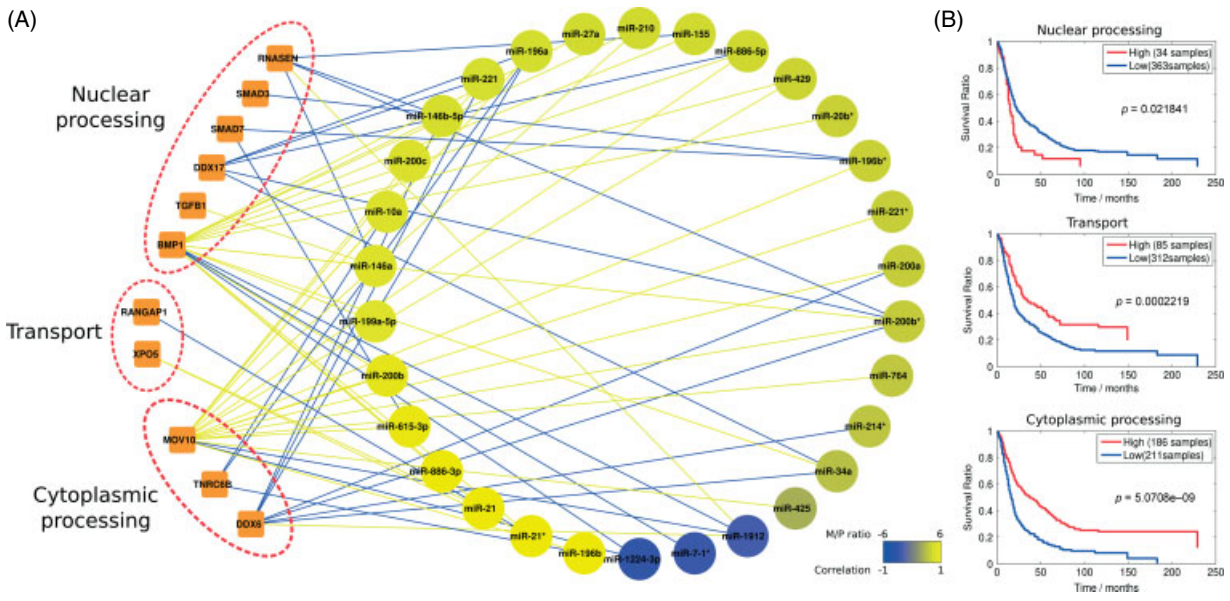

Figure 3. Network analysis identifies miRNA biogenesis gene signatures predictive of survival. (A) Statistically significant ($p < 0.01$) correlations between biogenesis genes and miRNA maturation (M/P ratio) are shown. The network is highly connected and includes 30 out of 35 miRNAs. Genes have been grouped and labelled in three groups based on their role in the miRNA biogenesis pathway. Link colours indicate positive (yellow) or negative (blue) correlation between the gene and miRNA. The same colour code is used for colouring the miRNA nodes in the networks and to show M/P ratio differences between glioma and normal. The network is visualized using Cytoscape version 2.8.1 [41]. (B) Survival analysis of the network-based gene signatures on the independent data set from Rembrandt. Shown from top to bottom are nuclear processing, transport from the nucleus, and cytoplasmic processing. High expression of genes in cytoplasmic processing and nucleo-cytoplasmic transport signatures is related to poor survival. Nuclear processing genes indicate an opposite pattern.

the signatures of genes that correlate with survival in a large and independent data set. This study implicates the miRNA biogenesis pathway as a key pathway targeted for deregulation in glioma cells.

One limitation of our analysis is that the M/P ratio is an indicator of pre-miRNA-to-mature miRNA processing only. Pri-miRNAs are problematic for the analysis as they have not been systematically annotated. In addition, it has been shown that pri-miRNAs are processed co-transcriptionally, which makes the expression quantification from RNA sequencing data problematic [44]. We were unable to examine the miRNA transport process where pre-miRNAs are exported into the cytoplasm. Dissection of this step requires the isolation of nuclear and cytoplasmic small RNA for library construction and deep sequencing. Whereas this can be achieved using cultured cells, it was not feasible with the small amounts of frozen tumour tissue available for our study. However, these limitations can be partially ameliorated by quantifying the expression of the genes that are known to regulate the nuclear processing and transport steps of miRNA biogenesis. In this study, we actually performed integrated analysis of miRNAs and genes that regulate the three major steps in miRNA biogenesis – nuclear processing; transport from the nucleus; and cytoplasmic processing. Interestingly, the network-based gene signatures for each of these steps are associated with grade and survival in glioma.

Some of the most striking gene expression changes identified in our study can be seen in the TGFβ/BMP/SMAD signalling pathway. The BMP-specific receptor SMADs (R-SMADs: SMAD1 and SMAD5) were first implicated in miRNA biogenesis by the observation that they interacted with the RNA helicase p68 (DDX5) to aid in the processing of pri-miR-21 to pre-miR-21 [45]. Later, the TGFβ/BMP pathway was found to regulate a larger subset of miRNAs via interaction with pri-miRNAs containing an R-SMAD consensus sequence (R-SBE) and enhancing Drosha-mediated processing [46]. Our network analysis identified this mechanism as part of the gene signature whose expression is correlated with most of the 35 miRNAs. Within the nuclear processing network, the *BMP1* and *TGFB1* genes were highly connected and correlated positively with the maturation of a large number of miRNAs. Other members of the SMAD family were included in the network, albeit with a smaller and more distinct set of correlated miRNAs. This indicates that the above-described Drosha-mediated processing, where individual cofactors such as the members of the TGFβ/BMP/SMAD signalling pathway control the processing in a miRNA-specific manner, is also likely to occur in glioma. Further investigation will be necessary to fully understand the specificity that occurs with SMAD-mediated processing.

The TGFβ pathway has also been shown to enhance epidermal growth factor (EGF)-mediated effects in gliomas and heighten the expression of SMAD2 and SMAD4 in glioma cell lines [47]. Furthermore, TGFβ induces leukaemia inhibitory factor (LIF) through the SMAD-dependent pathway, which increases the self-renewing capacity of glioma-initiating cells [48]. In gliomas with an unmethylated *PDGFB* gene, the

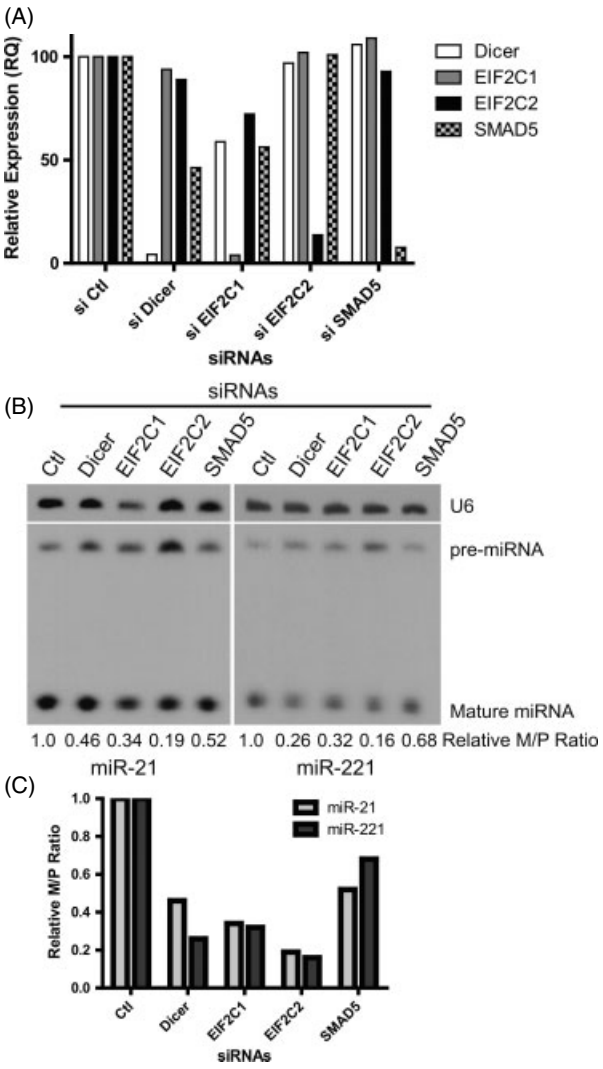

Figure 4. Loss of several biogenesis genes decreases the miRNA maturation. (A) Quantitative, real-time PCR demonstrates siRNA-mediated knockdown of miRNA processing genes. The graph represents the average of two independent experiments. (B) Northern blot analysis of pre-miRNA and mature miRNAs. Knockdown of these genes inhibits the maturation of miR-21 and miR-221. (C) Quantitation of northern blots in B.

TGFβ/SMAD pathway increases the expression of *PDGFB*, which can induce the proliferation of glioma cells [49]. High TGFβ/SMAD activity is associated with poor prognosis in patients with glioma [49], which was confirmed in our current study. The emerging picture suggests that this pathway may be a key regulator of glioma development and progression through miRNA-mediated effects.

In summary, by utilizing the comprehensive data provided by deep sequencing, we have uncovered a complicated network of gene expression changes within the miRNA biogenesis pathway that impact miRNA maturation and correlate with glioma progression. This study not only provides a newly characterized mechanism of miRNA alterations in glioma, but also leads to a number of potentially new oncogenic and tumour-suppressing miRNAs that are important for glioma. Further studies on these miRNAs will likely yield important insights as well as opportunities for the prognosis of and therapeutics for glioma.

Acknowledgments

This study was supported in part by the Paul and Joann Oreffice Fund for Brain Tumor Research (RS, GNF, and WZ), U24 CA143835 from the National Institutes of Health (IS and WZ), the Academy of Finland project No 132877 (MN), the Finnish Funding Agency for Technology and Innovation Finland Distinguished Professor Programme (MN and OY-H), the Sigrid Juselius Foundation (MN), and The University of Texas MD Anderson Cancer Center core grant CA016672 from the National Institutes of Health. We would like to thank Sue Moreau from the Department of Scientific Publications at The University of Texas MD Anderson Cancer Center for editing this manuscript.

Author contribution statement

LMM, VK, YL, MA, DG, MN, and WZ analysed data. LMM, DC, XL, and CGL performed experiments. MA implemented the sequence analysis. LMM, MN, WZ, and IS conceived the study and OYH, RS, and GNF participated in its design and coordination. LMM, VK, MN, and WZ drafted the manuscript. All authors read and approved the final manuscript.

References

1. Pillai RS. MicroRNA function: multiple mechanisms for a tiny RNA? *RNA* 2005; **11:** 1753–1761.
2. Ambros V. The functions of animal microRNAs. *Nature* 2004; **431:** 350–355.
3. Bommer GT, Gerin I, Feng Y, *et al*. p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Curr Biol* 2007; **17:** 1298–1307.
4. He L, He X, Lim LP, *et al*. A microRNA component of the p53 tumour suppressor network. *Nature* 2007; **447:** 1130–1134.
5. He L, Thomson JM, Hemann MT, *et al*. A microRNA polycistron as a potential human oncogene. *Nature* 2005; **435:** 828–833.
6. Ling H, Zhang W, Calin GA. Principles of microRNA involvement in human cancers. *Chin J Cancer* 2011; **11:** 739–748.
7. Chuang JC, Jones PA. Epigenetics and microRNAs. *Pediatr Res* 2007; **61:** 24R29R.
8. Thomson JM, Newman M, Parker JS, *et al*. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev* 2006; **20:** 2202–2207.
9. Farazi TA, Spitzer JI, Morozov P, *et al*. miRNAs in human cancer. *J Pathol* 2011; **223:** 102–115.
10. Winter J, Jung S, Keller S, *et al*. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature Cell Biol* 2009; **11:** 228–234.
11. Lee Y, Ahn C, Han J, *et al*. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 2003; **425:** 415–419.
12. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. *Nature* 2007; **448:** 83–86.

13. Yi R, Qin Y, Macara IG, *et al*. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 2003; **17:** 3011–3016.
14. Bernstein E, Caudy AA, Hammond SM, *et al*. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; **409:** 363–366.
15. Cifuentes D, Xue H, Taylor DW, *et al*. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 2010; **328:** 1694–1698.
16. Sevignani C, Calin GA, Siracusa LD, *et al*. Mammalian microRNAs: a small world for fine-tuning gene expression. *Mamm Genome* 2006; **17:** 189–202.
17. Zhang B, Wang Q, Pan X. MicroRNAs and their regulatory roles in animals and plants. *J Cell Physiol* 2007; **210:** 279–289.
18. Garzon R, Calin GA, Croce CM. MicroRNAs in cancer. *Annu Rev Med* 2009; **60:** 167–179.
19. Kumar MS, Lu J, Mercer KL, *et al*. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genet* 2007; **39:** 673–677.
20. Sugito N, Ishiguro H, Kuwabara Y, *et al*. RNASEN regulates cell proliferation and affects survival in esophageal cancer patients. *Clin Cancer Res* 2006; **12:** 7322–7328.
21. Karube Y, Tanaka H, Osada H, *et al*. Reduced expression of Dicer associated with poor prognosis in lung cancer patients. *Cancer Sci* 2005; **96:** 111–115.
22. Merritt WM, Lin YG, Han LY, *et al*. Dicer, Drosha, and outcomes in patients with ovarian cancer. *N Engl J Med* 2008; **359:** 2641–2650.
23. Faber C, Horst D, Hlubek F, *et al*. Overexpression of Dicer predicts poor survival in colorectal cancer. *Eur J Cancer* 2011; **9:** 1414–1419.
24. Melo SA, Moutinho C, Ropero S, *et al*. A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell* 2010; **18:** 303–315.
25. Melo SA, Ropero S, Moutinho C, *et al*. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nature Genet* 2009; **41:** 365–370.
26. Kumar MS, Pester RE, Chen CY, *et al*. *Dicer1* functions as a haploinsufficient tumor suppressor. *Genes Dev* 2009; **23:** 2700–2704.
27. Ravi A, Gurtan AM, Kumar MS, *et al*. Proliferation and tumorigenesis of a murine sarcoma cell line in the absence of DICER1. *Cancer Cell* 2012; **21:** 848–855.
28. Ciafrè SA, Galardi S, Mangiola A, *et al*. Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochem Biophys Res Commun* 2005; **334:** 1351–1358.
29. Kim TM, Huang W, Park R, *et al*. A developmental taxonomy of glioblastoma defined and maintained by microRNAs. *Cancer Res* 2011; **71:** 3387–3399.
30. Sana J, Hajduch M, Michalek J, *et al*. MicroRNAs and glioblastoma: roles in core signalling pathways and potential clinical implications. *J Cell Mol Med* 2011; **15:** 1636–1644.
31. Chan JA, Krichevsky AM, Kosik KS. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res* 2005; **65:** 6029–6033.
32. Conti A, Aguennouz M, La Torre D, *et al*. miR-21 and 221 upregulation and miR-181b downregulation in human grade II–IV astrocytic tumors. *J Neurooncol* 2009; **93:** 325–332.
33. Papagiannakopoulos T, Shapiro A, Kosik KS. MicroRNA-21 targets a network of key tumor-suppressive pathways in glioblastoma cells. *Cancer Res* 2008; **68:** 8164–8172.
34. Zhang J, Han L, Ge Y, *et al*. miR-221/222 promote malignant progression of glioma through activation of the Akt pathway. *Int J Oncol* 2010; **36:** 913–920.
35. Moore LM, Zhang W. Targeting miR-21 in glioma: a small RNA with big potential. *Expert Opin Ther Targets* 2010; **11:**1247–1257.
36. Langmead B, Trapnell C, Pop M, *et al*. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; **10:** R25.
37. Griffiths-Jones S, Grocock RJ, van Dongen S, *et al*. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006; **34:** D140D144.
38. Irizarry RA, Bolstad BM, Collin F, *et al*. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003; **31:** e15.
39. Varallyay E, Burgyan J, Havelda Z. MicroRNA detection by northern blotting using locked nucleic acid probes. *Nature Protoc* 2008; **3:** 190–196.
40. Chan JA, Krichevsky AM, Kosik KS. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res* 2005; **65:** 6029–6033.
41. Smoot ME, Ono K, Ruscheinski J, *et al*. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011; **27:** 431–432.
42. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nature Rev Cancer* 2006; **6:** 857–866.
43. Davis BN, Hata A. Regulation of MicroRNA biogenesis: a miRiad of mechanisms. *Cell Commun Signal* 2009; **7:** 18.
44. Morlando M, Ballarino M, Gromak N, *et al*. Primary microRNA transcripts are processed co-transcriptionally. *Nature Struct Mol Biol* 2008; **15:** 902–909.
45. Davis BN, Hilyard AC, Lagna G, *et al*. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* 2008; **454:** 56–61.
46. Davis BN, Hilyard AC, Nguyen PH, *et al*. Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. *Mol Cell* 2010; **39:** 373–384.
47. Held-Feindt J, Lütjohann B, Ungefroren H, *et al*. Interaction of transforming growth factor-beta (TGF-beta) and epidermal growth factor (EGF) in human glioma cells. *J Neurooncol* 2003; **63:** 117–127.
48. Peñuelas S, Anido J, Prieto-Sánchez RM, *et al*. TGF-beta increases glioma-initiating cell self-renewal through the induction of LIF in human glioblastoma. *Cancer Cell* 2009; **15:** 315–327.
49. Bruna A, Darken RS, Rojo F, *et al*. High TGFbeta-Smad activity confers poor prognosis in glioma patients and promotes cell proliferation depending on the methylation of the PDGF-B gene. *Cancer Cell* 2007; **11:** 147–160.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Figure 1. M/P miRNA ratios across all samples. 292 miRNAs were identified as differentially maturated (M/P ratio > 2-fold) between gliomas and normal brain. For visualization, the M/P ratio of each miRNA is scaled to interval of zero (lowest expression ratio)-to-one (highest expression ratio), as indicated by the colour bar.

Figure 2. Mature miRNA (A) and pre-miRNA (B) expression levels. Levels are shown across all samples for the 35 differentially matured miRNA which correlate with glioma grade. The expression of each miRNA is scaled to interval of zero-to-one.

Figure 3. Quantitative RT-PCR validation of pre- and mature miR-21 (A) and miR-1912 (B). Individual patient samples from different sequencing pools were selected at random for validation by qRT-PCR. Fold change of pre-miRNA (green) and mature miRNA (blue) relative quantity normalized to normal brain control, M/P expression ratio relative to the M/P expression ratio of normal brain.

Figure 4. Distribution of small RNA sequencing reads in the first GBM sequencing pool across four different pre-miRNA sequences. The bar height at each base represents the number of reads overlapping that position. Locations of the pre-miRNA loop and mature sequences are labelled. While mature sequences are substantially more abundant, there are a sufficient number of reads also for the pre-miRNA allowing the quantification of the expression level.

Appendix S1. Supplementary methods