

Ribosome Profiling Reveals a Cell-Type-Specific Translational Landscape in Brain Tumors

Christian Gonzalez,¹ Jennifer S. Sims,^{2*} Nicholas Hornstein,^{1,3*} Angeliki Mela,⁴ Franklin Garcia,^{4,5}  Liang Lei,⁴ David A. Gass,⁶ Benjamin Amendolara,² Jeffrey N. Bruce,² Peter Canoll,⁴ and Peter A. Sims^{1,7}

Departments of ¹Systems Biology, ²Neurological Surgery, ³Columbia University MD/PhD Program, ⁴Pathology and Cell Biology, ⁵Graduate PhD Program in Pathobiology and Molecular Medicine, ⁶Pediatrics, and ⁷Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, New York 10032

Glioma growth is driven by signaling that ultimately regulates protein synthesis. Gliomas are also complex at the cellular level and involve multiple cell types, including transformed and reactive cells in the brain tumor microenvironment. The distinct functions of the various cell types likely lead to different requirements and regulatory paradigms for protein synthesis. Proneural gliomas can arise from transformation of glial progenitors that are driven to proliferate via mitogenic signaling that affects translation. To investigate translational regulation in this system, we developed a RiboTag glioma mouse model that enables cell-type-specific, genome-wide ribosome profiling of tumor tissue. Infecting glial progenitors with Cre-recombinant retrovirus simultaneously activates expression of tagged ribosomes and delivers a tumor-initiating mutation. Remarkably, we find that although genes specific to transformed cells are highly translated, their translation efficiencies are low compared with normal brain. Ribosome positioning reveals sequence-dependent regulation of ribosomal activity in 5′-leaders upstream of annotated start codons, leading to differential translation in glioma compared with normal brain. Additionally, although transformed cells express a proneural signature, untransformed tumor-associated cells, including reactive astrocytes and microglia, express a mesenchymal signature. Finally, we observe the same phenomena in human disease by combining ribosome profiling of human proneural tumor and non-neoplastic brain tissue with computational deconvolution to assess cell-type-specific translational regulation.

Key words: cell-type-specific expression; glioblastoma; glioma; ribosome profiling; translational regulation

Introduction

Gliomas are responsible for the majority of malignant neoplasms in the brain. A large-scale effort by The Cancer Genome Atlas (TCGA) characterized gene expression in glioblastoma (GBM), an aggressive form of glioma, and identified four subtypes: proneural, neural, classical, and mesenchymal (Verhaak et al., 2010). Proneural GBM closely resembles glial progenitors, which are capable of transforming and proliferating in the brain (Assanah et al., 2006), resulting in tumors that recapitulate the proneural signature in animal models (Lei et al., 2011; Liu et al., 2011). Glioma growth is driven by pathways that converge on the pro-

tein synthesis machinery through mTOR, AKT, and other signaling molecules (Rajasekhar et al., 2003; Parsa and Holland, 2004; Takeuchi et al., 2005; Fan et al., 2007; Jiang and Liu, 2009; Helmy et al., 2012). However, proneural glioma cells maintain a relatively immature phenotype, which has been associated with tight translational control in other systems (Signer et al., 2014; Tahmasebi et al., 2014). Despite its potential importance to protein expression, relatively little is known about translational regulation in this disease.

Gliomas are diffusively infiltrative, presenting a challenge to translational profiling. Previously, investigating cell-type-specific translation would require physical separation of cells with techniques that could perturb this dynamic process. Recent efforts have overcome this problem by cell-type-specific expression of epitope-tagged ribosomes for immunoprecipitation of ribosome-bound RNA from homogenized tissue (Heiman et al., 2008; Sanz et al., 2009). For example, the RiboTag system uses Cre-mediated activation of hemagglutinin-fused Rpl22. Here, we retrovirally deliver Cre recombinase and platelet-derived growth factor (PDGF-B) to the adult brain (Lei et al., 2011) of a mouse harboring the RiboTag and floxed *Trp53* alleles. Expression of HA-tagged ribosomes is therefore restricted to transformed cells that arise from the originally infected cells, allowing isolation of tumor-specific RNA by immunoprecipitation from homogenized tissue.

Received Jan. 8, 2014; revised June 17, 2014; accepted July 7, 2014.

Author contributions: C.G., J.N.B., P.C., and P.A.S. designed research; C.G., J.S.S., N.H., A.M., F.G., L.L., D.A.G., B.A., P.C., and P.A.S. performed research; C.G., N.H., A.M., P.C., and P.A.S. analyzed data; C.G., J.S.S., N.H., P.C., and P.A.S. wrote the paper.

N.H. is supported by NIH/NINDS fellowship 1F31NS089106. P.A.S. is supported by NIH/NIBIB Grant 1K01EB016071. P.C. is supported by NIH/NINDS Grant 1R01NS066955. P.A.S. and P.C. are part of the Brain Tumor Ecology Collaborative supported by the James S. McDonnell Foundation. We thank Alexander Tzagoloff and Ruben Gonzalez for sharing equipment, Thomas Ludwig for providing the *Trp53*-floxed mice, and Yufeng Shen and Dr. Xiaojun Feng for helpful technical discussions.

The authors declare no competing financial interests.

*J.S.S. and N.H. contributed equally to this work.

Correspondence should be addressed to either Dr Peter A. Sims or Dr Peter Canoll, Columbia University Medical Center, Irving Cancer Research Center, Room 1002B, 1130 St Nicholas Avenue, New York, NY 10032, E-mail: pas2182@columbia.edu or pc561@columbia.edu.

DOI:10.1523/JNEUROSCI.0084-14.2014

Copyright © 2014 the authors 0270-6474/14/3410924-13\$15.00/0

Previous efforts at cell-specific translational profiling involved quantification of intact, ribosome-bound RNA (Doyle et al., 2008; Heiman et al., 2008; Sanz et al., 2009). However, these measurements do not allow direct quantification of ribosome density or provide ribosome positioning information, complicating accurate estimates of translation rates and efficiencies and precluding determination of whether ribosome density originates from annotated coding or upstream sequences. In contrast, ribosome profiling, based on deep sequencing of ribosome-protected mRNA footprints, enables genome-wide analysis of protein synthesis and ribosome positioning (Ingolia et al., 2009). The approach has been applied broadly from studies of non-canonical translation in yeast (Brar et al., 2012) to translational control in cancer (Hsieh et al., 2012).

Here we describe a strategy for cell-type-specific measurements of protein synthesis by combining the tissue specificity of the RiboTag system with ribosome profiling. We measured genome-wide ribosomal positioning and translation rates, identified genes that are selectively translated by transformed cells, and discovered non-cell autonomous effects on translation in the tumor microenvironment. Using computational deconvolution, we assessed how these genes are distributed among cell types in murine and human tumors. Finally, we found that translation efficiency is cell-type-specific in proneural glioma, with transformed glial progenitors showing a significant decrease in translation efficiency compared with other cells in the tumor microenvironment.

Materials and Methods

RiboTag mouse glioma model. For experimental induction of murine glioma, transgenic C57BL/6 mice carrying loxP recognition sites at exon 7 of *Trp53* were crossed with RiboTag mice (JAX ID 011029), which carry the HA-affinity tag adjacent to the ribosomal protein *Rpl22*, separated from the natively expressed terminal exon by loxP recognition sites. Hence, the tagged version of *Rpl22* is only expressed following Cre-mediated recombination. These mice were bred to *Trp53*^{lox/lox} and RiboTag homozygosity. Proneural gliomas were induced *de novo* by stereotactic injection into subcortical white matter of the right frontal lobe of $\sim 5 \times 10^4$ replication incompetent, retroviral particles expressing human PDGF-B and Cre recombinase, as described previously (Lei et al., 2011). Two of three mice in which tumors were induced were 43-d-old and the third mouse was 64-d-old. Age-matched control mice were injected with an equal volume of serum-free media. All six of the mice were female. Mice were monitored for tumor morbidity by behavior and weight, and killed at 30 d according to Columbia University IACUC protocol no. AC-AAAF1710. At this time point, all three mice exhibited symptoms of tumor morbidity and tumors were clearly visible upon removal of the brain. The survival curve in Figure 1 indicates a median survival time of 47 ± 7 d postinjection, and so we killed the animals at 30 d postinjection to avoid death due to tumor morbidity at an uncontrolled time so that we could harvest fresh polysomal RNA from the tumor tissue. The right frontal lobe tissue (containing the injection site) and distal tissue from the contralateral hemisphere were snap-frozen in liquid nitrogen immediately following death. Tissue immediately adjacent to the experimental sample, containing tumor, was fixed in 4% paraformaldehyde (PFA) for 48 h before immunofluorescence. The survival curves depicted in Figure 1C were generated by injecting nine *Trp53*^{lox/lox} mice and 11 wild-type C57BL/6 mice of either sex (all 6- to 8-weeks-old) with 5×10^4 viral particles expressing PDGF-B and Cre recombinase (Lei et al., 2011). We note that we have previously reported the use of this specific initiating alteration in conjunction with PDGF-B overexpression (Sonabend et al., 2014).

Immunofluorescence of mouse brain tissue sections. Gliomas were induced in three 9-week-old male RiboTag mice harboring *Trp53*^{lox/lox} by retroviral delivery of Cre recombinase and human PDGF-B as described above. At 28 d postinjection, mouse brains were fixed in 4% PFA for 48 h.

Brains were then cryoprotected in 30% sucrose for 4 d, and then stored in OCT at -80°C . Cryosections (10 μm) were fixed in 4% PFA for 10 min at room temperature, washed in PBS, blocked with 5% horse serum (Sigma-Aldrich) for 30 min, and then labeled with primary antibodies overnight at 4°C . Sections were then washed three times with PBS and incubated with AlexaFluor-conjugated secondary antibodies (1:1000, Invitrogen) for 1 h at room temperature and counter-stained with DAPI.

Antibodies. The following primary antibodies were used: mouse anti-HA (1:1000, Covance), rabbit anti-PDGFR α (1:500, Cell Signaling Technology), rabbit anti-OLIG2 (1:100, Millipore), rabbit anti-GFAP (1:500, Dako), rabbit anti-RBFOX3 (1:1000, Millipore), rat anti-CD44 (1:150, Calbiochem), and rabbit anti-AIF1 (1:1000, Wako).

Human brain tumor and non-neoplastic brain tissue specimens. The five adult patients included in this study presented for surgical resection of either malignant glioma or of non-neoplastic brain tissue to relieve epilepsy symptoms. The three epilepsy patients have no oncological history. Resected tissue specimens were snap-frozen with liquid nitrogen in the operating room to maximize RNA preservation.

Microscopy. Stained tissue sections were imaged using a Nikon TE2000 epifluorescence microscope equipped with MetaMorph software (Molecular Devices). Micrographs were merged using MetaMorph and ImageJ. Immunofluorescence images of the RiboTag cell line were obtained using a Nikon Ti-U epifluorescence microscope equipped with an EM-CCD camera (Andor iXon) and a 532 nm diode-pumped solid-state laser.

Extraction and preparation of total RNA. Tissue samples were homogenized in 1 ml of polysome lysis buffer (20 mM Tris-HCl pH 7.3, 250 mM NaCl, 15 mM MgCl₂, 1 mM DTT, 0.5% Triton X-100, 0.024 U/ml TurboDNase, and 0.1 mg/ml cycloheximide) at 4°C using a glass Dounce homogenizer. Homogenates were centrifuged for 10 min at $14,000 \times g$, and the supernatant was snap frozen in liquid nitrogen. Total RNA from homogenates was purified using the RNeasy Mini Kit (Qiagen), and RNA integrity was assessed using a Bioanalyzer (Agilent).

Ribosomal footprint isolation. Mouse tissue lysates were digested for 45 min at room temperature with RNase I. Monosomes were isolated using sucrose density gradient fractionation. The monosome fraction was split into two samples. RNA from the first sample was extracted with phenol chloroform and used for further purification of ribosomal footprints. HA-tagged monosomes were immunoprecipitated from the second sample to obtain cell-type-specific ribosomal footprints (Fig. 1A).

For ribosomal immunoprecipitation, we coupled 30 μl of mouse monoclonal anti-HA antibody (HA.11, ascites fluid, Covance) to 300 μl of protein G-coated Dynabeads (30 mg/ml, Life Technologies) for 1 h in citrate-phosphate buffer (24 mM citric acid, 52 mM dibasic sodium phosphate, pH 5.0). Beads were washed once in citrate-phosphate buffer and three times in polysome lysis buffer. Beads were added to the lysates and incubated with rotation at 4°C overnight. Beads were then washed four times with 500 μl of polysome lysis buffer. Ribosomes and footprints were released from the beads using EDTA. Beads were incubated with 140 μl of ribosome release buffer (20 mM Tris-HCl pH 7.3, 250 mM NaCl, 0.5% Triton X-100, 50 mM EDTA) for 5 min and the supernatant was set aside on ice. This was repeated three more times. The pooled supernatants were then extracted with phenol-chloroform to yield footprints and digested rRNA fragments.

Ribosome footprint isolation of human glioma and non-neoplastic brain tissue was accomplished as described above for the murine specimens without immunoprecipitation.

Ribosome profiling and RNA sequencing libraries. Ribosomal footprints were size selected as described previously (Ingolia et al., 2012) and sequencing libraries constructed. We note that, although the exact adapter scheme reported previously was used to construct ribosome profiling libraries for Tumor Mouse A and Normal Mouse A (Ingolia et al., 2012), a different reverse primer was used for amplification of all other ribosome profiling libraries. In particular, we used the following reverse primer that places the barcode sequence for demultiplexing at the 3' end of the linker sequence appended to each RNA footprint during single-stranded ligation:

5'-CAAGCAGAAGACGGCATAACGATNNNNNNATTGATG-TGCCTACAG-3'.

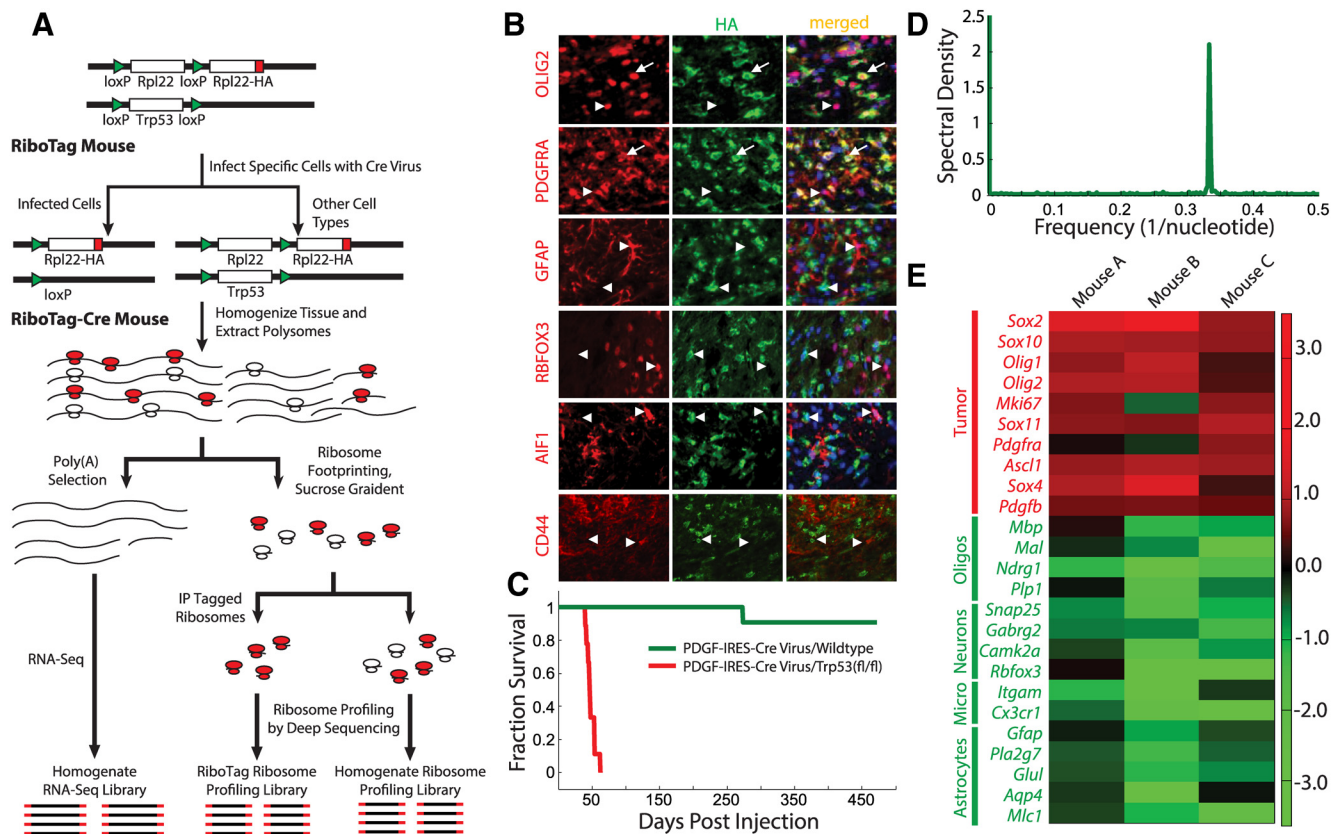


Figure 1. RiboTag mouse glioma model and cell-type-specific ribosome profiling. **A**, Schematic of the RiboTag glioma mouse model and experimental workflow. Cells infected by a retrovirus that expresses Cre recombinase and PDGF-B express the RiboTag (Rpl22-HA) and harbor a transforming genetic lesion—loss of *Trp53*. Polysomes are extracted from homogenate tumor tissue. Poly(A) RNA is selected from a portion of this for RNA-Seq. The remaining polysomes are digested to monosomes and purified on a sucrose gradient. The purified monosome sample is split in half. One-half is converted into a ribosome-profiling library. HA-tagged (RiboTag) monosomes originating from the transformed cells are immunoprecipitated from the other half and converted into a ribosome-profiling library. Translation rates from the homogenate and RiboTag ribosome profiles are compared in order to identify genes that are enriched or depleted in the transformed population. **B**, Immunofluorescence staining of tissue sections from an end-stage RiboTag glioma mouse showing the diversity of cell types present in the tumor. Cells expressing HA (the RiboTag epitope) overlap significantly with OLIG2- and PDGFRA-expressing cells. However, there is essentially no overlap between cells expressing HA and cells expressing GFAP (astrocytes), RBFOX3 (neurons), AIF1 (microglia), or CD44 (reactive astrocytes). **C**, Survival curves for *Trp53^{fllox/fllox}* and wild-type mice after injection with PDGF-B-IRES-Cre virus indicating a median survival time of 47 ± 7 d postinjection for our mouse glioma model. **D**, Power spectrum of the 5'-end read positions along CDSs for the first 500 bases of the CDS for all genes with a CDS length of at least 500 bases. This power spectrum was computed from the RiboTag profile of Mouse A, demonstrating that RiboTag immunoprecipitation preserves the expected three-base periodicity arising from codons as indicated by the clear peak at a frequency of $\sim 0.33 \text{ nt}^{-1}$. **E**, Heat map displaying the translation rate enrichment scores (plotted as score -1 where a score >1 indicates enrichment in the RiboTag profile and a score <1 indicates depletion) for several canonical markers of different cell types across three mice. The enrichment score is calculated by dividing the translation rate in the RiboTag profile by that in the homogenate profile.

where NNNNNN represents the barcode sequences used previously (Ingolia et al., 2012). Ribosome profiling libraries were sequenced using the Illumina HiSeq 2000 system at the Columbia Sulzberger Genome Center. Total RNA for each sample was also provided to the Columbia Sulzberger Genome Center for poly(A)-selection and RNA-Seq using the Illumina TruSeq kit. We note that a total of 10 RNA expression profiles were obtained for deconvolution analysis from 10 different mice with PDGF-driven, end-stage tumors. All 10 of these mice were female. Five of these mice were homozygous for both the RiboTag and *Trp53^{fllox/fllox}*. The other five mice were homozygous for *Trp53^{fllox/fllox}* and did not harbor the RiboTag allele. Sequencing of ribosome profiling and RNA-Seq libraries for the five human tissue specimens was accomplished using the same procedure.

Bioinformatic analysis of ribosome profiling data. Following Illumina sequencing of our ribosome profiling libraries, we first demultiplexed our raw data using a barcode that was attached to the sequencing template during PCR amplification. The 3'-adapter sequence (CTGTAG-GCACCAT) was clipped from each read using *fastx_clipper*. We discarded all reads that did not contain the adapter sequence or that were shorter than 25 bases after adapter clipping. We then used *Bowtie 2* to map the clipped reads to a mouse rRNA reference and discarded any reads that mapped to rRNA (Langmead and Salzberg, 2012). The re-

maining reads were then mapped to either the mouse (Illumina iGenomes mm10 reference) or human (Illumina iGenomes hg19 reference) transcriptomes using *Tophat 2* (Kim et al., 2013). We did not attempt to detect novel junctions and obtained ~ 0.5 – 11 M uniquely mapped reads per mouse sample and ~ 4 – 24 M uniquely mapped reads per human sample. We used *HTSeq* to compute read counts for reads that mapped uniquely to the mouse or human transcriptome in the coding sequence (CDS), 5' leader (5'-UTR), 3'-UTR, and complete exonic sequence of each gene.

Bioinformatic analysis of RNA-Seq data. For RNA-Seq, we received *fastq* files containing demultiplexed, single-end 100 base reads for each sample from the Columbia Sulzberger Genome Center. We used *Tophat 2* to map the reads to the mouse (Illumina iGenomes mm10 reference) or human (Illumina iGenomes hg19 reference) transcriptomes and did not attempt to detect novel junctions, obtaining ~ 30 million reads per sample. We used *HTSeq* to compute read counts for reads that uniquely mapped to the transcriptome for each gene.

Analysis of translational activity, enrichment, and differential translation rate. For each gene, we divided the CDS read counts determined above by CDS length and total number of uniquely mapped reads in the sample to determine the CDS ribosome footprint density which we refer to as the translation rate. Similarly, we computed the 5'-leader ribosome foot-

print density by dividing the 5′-leader read counts determined above by 5′-leader length and total number of uniquely mapped reads in each sample.

To calculate the enrichment scores for each gene, we divided the translation rate measured from each RiboTag ribosome profile by the translation rate measured from the corresponding homogenate ribosome profile. An enrichment score greater than one indicates that the gene is biased toward translation in the transformed cells. We generated a consensus list of enriched genes based on the criteria that the enrichment score is greater than one in all three biological replicates and that the number of uniquely mapped CDS reads in each mouse is >10 in the RiboTag ribosome profile. Similarly, we placed genes with enrichment scores <1 and with >10 CDS reads in the homogenate ribosome profile in all three biological replicates in the consensus list of depleted genes.

We calculated translation efficiency for a given gene by dividing its translation rate by the RNA read density (number of reads uniquely mapping to the complete exonic sequence for each gene divided by transcript length and total number of uniquely mapped reads from a sample). Both the differential translation rate and RNA expression analyses were performed using DESeq (Anders and Huber, 2010) based on the uniquely mapped, CDS ribosome footprint counts and the uniquely mapped exonic RNA counts for each gene computed using HTSeq.

For the human ribosome profiling translation rate analysis depicted in Figure 4D, we show genes for which we measured at least a twofold increase in translation rate in the tumor tissue compared with normal brain for all possible pairwise comparisons of tumor and normal brain specimens.

Gene ontology analysis. Mutual information-based gene ontology analysis was performed using the differential translation rate analysis between RiboTag and normal brain profiles with iPAGE (Goodarzi et al., 2009). Gene identification names from the mm10 reference transcriptome annotation were converted to RefSeq identifiers using Babelomics and filtered based on the annotation provided by iPAGE. We then used log₂-transformed fold-change obtained from DESeq for each statistically significant differentially expressed gene ($p < 0.05$) between the RiboTag profiles and the normal brain profiles as input for iPAGE to obtain over- and under-represented gene ontologies across nine bins of translation rate fold-change.

We identified ~100 genes from our differential translation rate analysis between the homogenate and normal brain profiles with both a significantly higher translation rate in the tumor compared with normal brain and on the consensus list of RiboTag-depleted genes. This list was too short to allow iPAGE analysis, and so we identified enriched gene ontologies using Fisher's exact test as implemented in FuncAssociate 2.0 (Berriz et al., 2009).

Cell-type-specific deconvolution of RNA-Seq. We implemented a modified version of the Population Specific Expression Analysis (PSEA) algorithm for deconvolution of RNA-Seq expression profiles from human and murine proneural glioma tissue specimens (Kuhn et al., 2011). Briefly, we used the relative expression levels of a series of cell-type-specific marker genes (described in the main text) measured by RNA-Seq as a proxy for the fractional composition of each cell type in the tissue. We deconvolved the expression of a given gene by least-squares regression of a system of linear equations (1 equation for each sample) in which the contribution of each cell type to the expression of a gene in a sample was represented as the product of its composition obtained from marker gene expression and the average expression level of the gene in that cell type across all samples. We also included a background (γ -intercept) term. We then used least-squares regression to solve for the average expression level of the gene in each cell type given the equations from each sample. The least-squares problem was constrained by the requirement that the resulting expression values be non-negative, a quadratic programming problem that we solved using the "lsqnonneg" function in MATLAB (MathWorks). Our implementation was distinct from PSEA in that we allowed iterative improvement of the marker gene set used to estimate the cellular composition of each sample. In particular, we included all genes with an $R^2 > 0.9$ in the previous iteration for estimation of cellular composition.

Results

RiboTag glioma mouse model enables tumor-specific ribosome profiling

Previous implementations of tagged ribosomes have relied on cell-type-specific promoters to drive expression of either Cre recombinase or the tagged ribosomal protein itself. Although these strategies can be highly effective in normal tissue, it is nearly impossible to find a promoter used exclusively by transformed cells in a primary, solid tumor. Previous attempts at cell-type-specific profiling in glioma have used the *Olig2*-driven bacTRAP system (Helmy et al., 2012). However, we and others have shown that *Olig2* is highly expressed by untransformed glial progenitors in the brain and in reactive cells recruited to the tumor (Assanah et al., 2006). To address this issue, we deliver Cre recombinase by stereotactic injection of retrovirus to activate RiboTag expression. The RiboTag system uses a conditional knock-in allele such that the tagged ribosomal protein is expressed from its endogenous promoter. This strategy is uniquely synergistic with ribosome profiling because, upon Cre-mediated recombination, the native *Rpl22* locus is deleted and replaced by the tagged version. Thereafter, infected cells will express only the HA-tagged version of Rpl22.

In the RiboTag mouse, exon 4 of the *Rpl22* is flanked by loxP sites followed by an identical exon tagged with three hemagglutinin epitope-coding sequences or HA-tag (Fig. 1; Sanz et al., 2009). We crossed the RiboTag mouse with a *Trp53^{lox/lox}* (Chen et al., 2005) mouse to generate mice that were homozygous for both alleles. Delivery of Cre recombinase and PDGF-B with a replication-incompetent retrovirus simultaneously deletes *Trp53*, activates expression of HA-tagged ribosomes, and drives cellular proliferation (Lei et al., 2011; Sonabend et al., 2014). Upon retrovirus injection, a small number of glial progenitors in subcortical white matter are selectively infected, resulting in a fatal brain tumor within 30 d.

Previous work demonstrated that brain tumors initiated as described arise locally around the injection site from expansion of retrovirus-infected cells (Lei et al., 2011). In the RiboTag glioma model, tagged ribosomes are expressed exclusively in the transformed cells that arise from the retrovirally infected cells. However, these cells can diffusively infiltrate and intermingle with surrounding brain tissue, and therefore the tumor microenvironment contains a complex mixture of cell types including neurons, astrocytes, microglia, and recruited OLIG2-expressing progenitors. To demonstrate the compositional heterogeneity of these tumors and cell-type specificity of RiboTag expression, we used immunofluorescence to stain tissue sections for canonical marker proteins (Fig. 1B). Tumor markers, such as OLIG2 and PDGFRA, colocalize with HA-staining. As expected, there are also OLIG2-expressing cells that do not express the HA-tag, representing a population of recruited oligodendrocyte progenitors that did not arise from the retrovirus-infected cells. In contrast, we do not observe HA in cells expressing GFAP, CD44, RBFOX3, and AIF1, which mark astrocytes, reactive astrocytes, neurons, and microglia, respectively.

After isolating tumor tissue from three end-stage mice, we homogenized frozen tumors, extracted total RNA from a portion of the lysate, and digested the remainder with RNase I for monosome purification on a sucrose gradient. Instead of purifying cell-type-specific polysomal RNA as in previous studies (Heiman et al., 2008; Sanz et al., 2009), we used immunoprecipitated HA-tagged monosomes directly from half of the appropriate sucrose fraction to generate a tumor cell-specific ribosome footprint li-

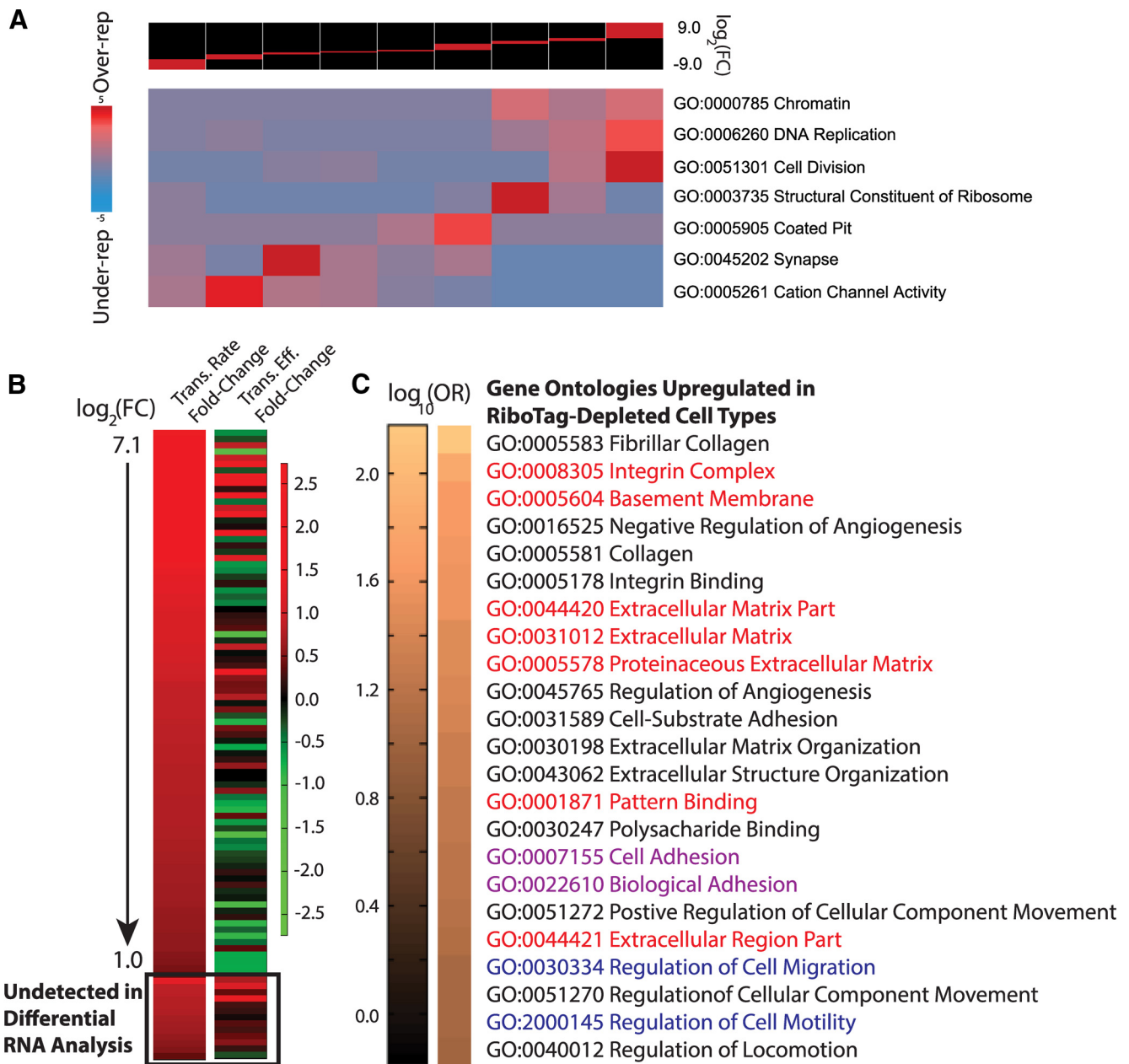


Figure 2. Differential translation rate analysis. **A**, Information theory-based iPAGE analysis of over- and under-represented gene ontologies in genes with statistically significant ($p < 0.05$) high- and low-translation rate fold-changes indicating high translational output in the RiboTag sample and normal brain, respectively. Chromatin, DNA replication, cell division, and ribosomal pathways are over-represented among genes highly translated in the RiboTag sample, whereas coated pit, synapse, and cation-channel activity pathways are over-represented in the normal brain profile. **B**, Heat map displaying the translation rate fold-change and translation efficiency fold-change from differential translation rate analysis between the tumor homogenate and normal brain ribosome profiles. The genes in this heat map show statistically significant increased translation in the tumor homogenate relative to normal brain but are consistently depleted in the RiboTag profile, indicating expression in tumor associated cells. A subset of these genes, all but one of which exhibited higher translation efficiency in tumor tissue, was not found to have a statistically significant change in RNA abundance. **C**, Gene ontology analysis of upregulated, depleted genes from **B** with heat map of the odds ratio. Pathways in red and blue indicate overlap with mesenchymal and classical glioblastoma pathways, respectively. Pathways in purple indicate overlap with both mesenchymal and classical pathways.

brary (Fig. 1A, the RiboTag sample). We used the other half of the monosome fraction to generate a ribosome footprint library from the combination of tagged and untagged monosomes (Fig. 1A, the homogenate sample).

We generated ribosome profiling libraries as described previously (Ingolia et al., 2012) with a few modifications and poly(A)⁺ libraries for RNA-Seq using standard protocols. To demonstrate that the RiboTag immunoprecipitation procedure results in an actual ribosome profile, we computed the power spectrum of the 5'-end mapping positions along the gene body for CDS reads

relative to the annotated start codon (Fig. 1D). As expected, the power spectrum shows a clear peak at ~ 0.33 /nucleotide, consistent with the expected three-base periodicity of codons reflected in the ribosome footprints.

To validate the RiboTag specificity, we computed enrichment scores across all three tumor samples for various tumor and neural lineage markers (Bédard et al., 2007; Cahoy et al., 2008; Verhaak et al., 2010; Lei et al., 2011). We defined the enrichment score for each gene as the ratio of the CDS translation rate in the RiboTag profile to the CDS translation rate in the homogenate.

Table 1. RiboTag-enriched genes with higher translation rates in tumor versus normal

<i>Adam9</i>	<i>FrmD8</i>	<i>Myo9b</i>	<i>Rpl7a</i>
<i>Aebp1</i>	<i>Gng5</i>	<i>Nasp</i>	<i>Rps12</i>
<i>Afap1l2</i>	<i>Gsx1</i>	<i>Nav1</i>	<i>Rps2</i>
<i>Aldh1a3</i>	<i>H2afy2</i>	<i>Nhs1</i>	<i>Rps29</i>
<i>Arfgap3</i>	<i>Igf2bp2</i>	<i>Notch1</i>	<i>Rps9</i>
<i>Atp5g2</i>	<i>Igfbp3</i>	<i>Pde8a</i>	<i>Rrm2</i>
<i>Bag3</i>	<i>Incenp</i>	<i>Pdlim4</i>	<i>Sall3</i>
<i>Birc5</i>	<i>Jam3</i>	<i>Pdlim5</i>	<i>Sapcd2</i>
<i>Cand1</i>	<i>Kdm6b</i>	<i>Plk1</i>	<i>Scrg1</i>
<i>Cd276</i>	<i>Kif20a</i>	<i>Plxn3</i>	<i>Socs3</i>
<i>Cdkn2a</i>	<i>Kif2c</i>	<i>Pnlip</i>	<i>Spc24</i>
<i>Ckap2l</i>	<i>Kif3</i>	<i>Ppfbp1</i>	<i>Spry2</i>
<i>Col11a1</i>	<i>Knstrn</i>	<i>Ppp1r18</i>	<i>Spry4</i>
<i>Creb5</i>	<i>Lima1</i>	<i>Prdx4</i>	<i>Sulf2</i>
<i>Cspg4</i>	<i>Lix1l</i>	<i>Qpct</i>	<i>Tagln2</i>
<i>Dcps</i>	<i>Lmnb1</i>	<i>Rbmx</i>	<i>Tk1</i>
<i>Dnmt1</i>	<i>Lnx1</i>	<i>Rcc1</i>	<i>Tmop</i>
<i>Epn2</i>	<i>Matn4</i>	<i>Rhoc</i>	<i>Trih2</i>
<i>Ezh2</i>	<i>Mcm10</i>	<i>Rpl18</i>	<i>Trim25</i>
<i>Fam101a</i>	<i>Mcm2</i>	<i>Rpl18a</i>	<i>Tubb2b</i>
<i>Fam64a</i>	<i>Metrn</i>	<i>Rpl19</i>	<i>Tubb6</i>
<i>Fam83d</i>	<i>Midn</i>	<i>Rpl21</i>	<i>Vim</i>
<i>Fbl</i>	<i>Morc4</i>	<i>Rpl27a</i>	<i>Ybx1</i>
<i>Fbn2</i>	<i>Moxd1</i>	<i>Rpl28</i>	<i>Zfp488</i>
<i>Fgfr1</i>	<i>Mpz1</i>	<i>Rpl35a</i>	

List of differentially translated genes found to have a statistically significant ($p < 0.05$) higher translation rate in the tumor homogenate compared with normal brain, and that are enriched in the RiboTag profile relative to tumor homogenate.

Table 2. RiboTag-depleted genes with higher translation rates in tumor versus normal

<i>Abca1</i>	<i>Col6a3</i>	<i>Ilgam</i>	<i>Postn</i>
<i>Abcc1</i>	<i>Cpne3</i>	<i>Irgb2</i>	<i>Prc1</i>
<i>Abcg2</i>	<i>Ctsc</i>	<i>Kdr</i>	<i>Psmb8</i>
<i>Acan</i>	<i>Ctss</i>	<i>Klhl5</i>	<i>Ptpn12</i>
<i>Aldh1a2</i>	<i>Dcn</i>	<i>Lama4</i>	<i>Ptprb</i>
<i>Anxa2</i>	<i>Dock1</i>	<i>Lamb1</i>	<i>Rrbp1</i>
<i>Aspm</i>	<i>Ecm1</i>	<i>Lgals3bp</i>	<i>S100a6</i>
<i>B2m</i>	<i>Ednrb</i>	<i>Lig1</i>	<i>Scpep1</i>
<i>C1qtnf6</i>	<i>Eltf1</i>	<i>Lmf2</i>	<i>Sec61a1</i>
<i>Cad</i>	<i>Emp1</i>	<i>Lrg1</i>	<i>Serpine1</i>
<i>Cd248</i>	<i>Enpp1</i>	<i>Lum</i>	<i>Serpinf1</i>
<i>Cd74</i>	<i>Ercc6l</i>	<i>Lyn</i>	<i>Slc16a3</i>
<i>Cd93</i>	<i>Esam</i>	<i>Lyz2</i>	<i>Smc4</i>
<i>Cd97</i>	<i>Fancd2</i>	<i>Mmp2</i>	<i>Snx9</i>
<i>Cdca8</i>	<i>Fat1</i>	<i>Mmp9</i>	<i>Spp1</i>
<i>Cdh5</i>	<i>Flna</i>	<i>Mpeg1</i>	<i>Stab1</i>
<i>Chst11</i>	<i>Fn1</i>	<i>Mtap</i>	<i>Syde1</i>
<i>Clic1</i>	<i>Gene</i>	<i>Ncapd3</i>	<i>Thbs1</i>
<i>Col15a1</i>	<i>Gnb2l1</i>	<i>Nid2</i>	<i>Thbs2</i>
<i>Col19a1</i>	<i>Gnb4</i>	<i>Nup188</i>	<i>Tm4sf1</i>
<i>Col3a1</i>	<i>Gpr56</i>	<i>Nup205</i>	<i>Tnc</i>
<i>Col4a1</i>	<i>Grn</i>	<i>Nup93</i>	<i>Top2a</i>
<i>Col4a2</i>	<i>Igfbp7</i>	<i>Pdgfrb</i>	<i>Usp24</i>
<i>Col5a2</i>	<i>Itga1</i>	<i>Pglyrp1</i>	<i>Vwf</i>
<i>Col5a3</i>	<i>Itga7</i>	<i>Plau</i>	<i>Zfp191</i>

List of differentially translated genes found to have a statistically significant ($p < 0.05$) higher translation rate in the tumor homogenate compared with normal brain, and that are depleted in the RiboTag profile relative to tumor homogenate.

As expected, markers of transformed oligodendrocyte progenitor cells (OPCs) and proneural glioma cells were consistently enriched in libraries generated from affinity-purified footprints, whereas markers associated with other neural cell types were depleted (Fig. 1E).

RiboTag enrichment and differential translation rate analyses reveal dysregulated pathways in tumor and tumor-associated cells

We next sought to determine which biological pathways are over-represented in the transformed cells on the basis of differential translation rate (ribosome density per gene). Using our three biological replicate RiboTag immunoprecipitation and three normal brain profiles, we generated a list of genes with statistically significant differential translation rates ($p < 0.05$). We then used iPAGE, an information theory-based algorithm, to determine which pathways are enriched in different ranges of fold-change (Fig. 2A; Goodarzi et al., 2009). Several factors contribute to the differential translation rate of a particular gene, including cellular composition, transcriptional regulation, and translational regulation. Our analysis associates “synapse” and “cation channel activity” with higher translation rates in normal brain, which contains many neurons, compared with the RiboTag profile, which should not express neuronal genes. The “DNA replication” and “cell division” ontologies were enriched among highly translated genes in the RiboTag profile. Interestingly, we also found significant enrichment in “structural components of the ribosome”, indicating global upregulation of protein synthesis in the transformed cells.

Cell-type-specific ribosome profiling also allowed us to probe the translational status of transcripts in cells within the tumor microenvironment that did not arise from retroviral infection. We first performed a differential translation rate analysis between the tumor homogenate ribosome profile (which reflects both transformed and untransformed cells) and that of normal brain (Table 1). We obtained a set of genes with statistically significant translation rate fold-changes ($p < 0.05$). We then queried this list for genes that were depleted by RiboTag immunoprecipitation. We reasoned that these genes would be predominantly upregulated in tumor-associated cells that did not arise from the retrovirus-infected cells. Interestingly, we found 100 genes (Table 2) that have a higher translation rate in the tumor homogenate, but appear to be expressed mainly in tumor-associated cells. These genes exhibited a broad range of translation efficiencies with both positive and negative fold-changes. Fourteen of these genes did not exhibit a statistically significant change in RNA abundance, which may be explained by upregulation primarily at the level of translation. Indeed, all but one of these 14 genes had a positive fold-change in translation efficiency ($p = 8 \times 10^{-4}$, Fisher’s exact test).

Gene ontology analysis revealed that the highly translated genes attributed to tumor-associated cells are predominantly associated with the extracellular matrix, cell motility, cell adhesion, and regulation of angiogenesis (Fig. 2C). To place this in the context of human disease, we examined the large-scale analysis of gene expression in human GBM by TCGA, where hierarchical clustering delineated four transcriptional subtypes based on 840 classifier genes (Verhaak et al., 2010). We measured expression profiles for 10 mouse gliomas by RNA-Seq and computed the median Spearman correlation coefficient between each of these profiles and each of the subtyped TCGA human microarray profiles for the classifier genes (Sonabend et al., 2013). As expected and shown previously for other PDGF-driven glioma mouse models, our model correlates best with the proneural subtype for all 10 replicates (Lei et al., 2011; Sonabend et al., 2013). Despite this, none of the pathways found in the upregulated, depleted genes expressed by tumor-associated cells are present in the TCGA gene ontology analysis of proneural classifier genes. However, 9 of 23 gene ontologies listed here occur in the mesenchymal

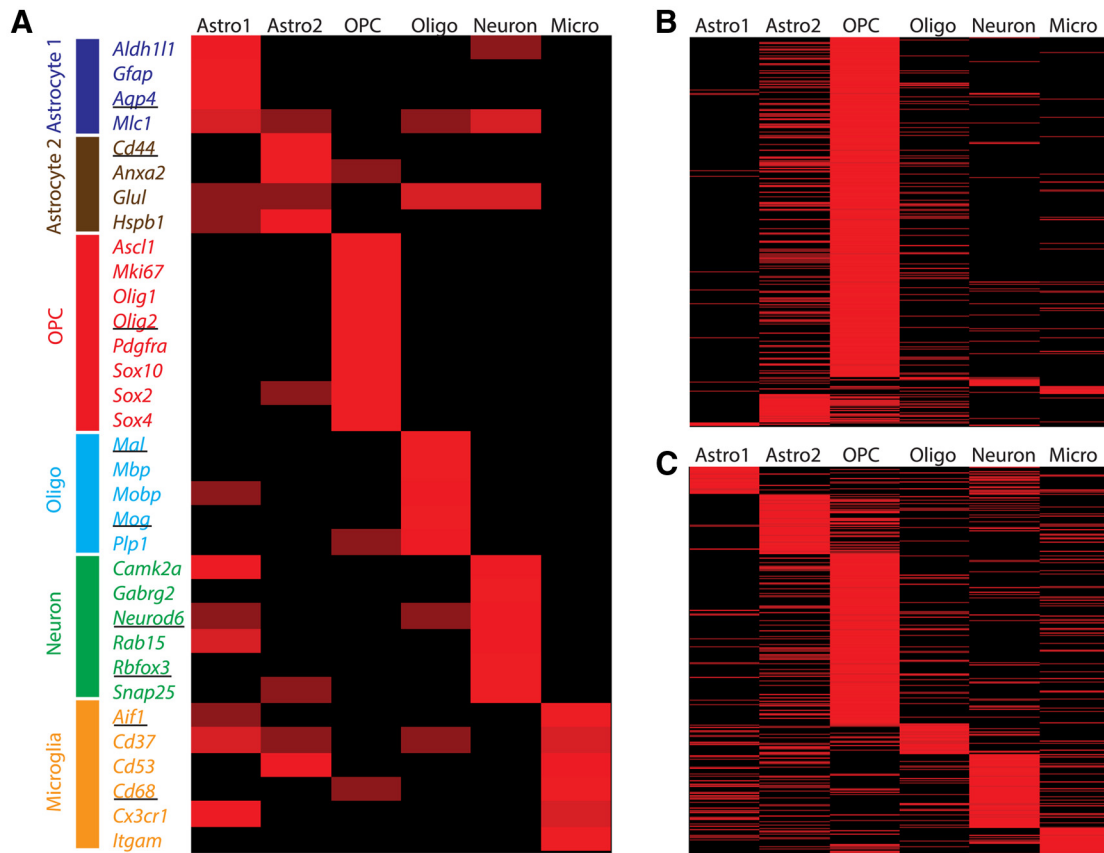


Figure 3. Computational deconvolution of cell-type-specific gene expression. **A**, Cell-type distribution heat map of canonical neural marker genes based on computational deconvolution of 10 RNA-Seq profiles from 10 murine proneural gliomas. The columns are the cell types used for deconvolution and the values are normalized across rows. Genes used to seed the deconvolution algorithm are underlined. **B**, Cell-type distribution heat map of RiboTag-enriched genes based on deconvolution of the profiles in **A** showing a strong enrichment in the OPC lineage, consistent with the glial progenitor origin of the tumor. **C**, Cell-type distribution heat map of RiboTag-depleted genes based on deconvolution of the profiles in **A** showing the expected, significant representation of all six cell types for genes that are not tumor-specific.

signature (indicated in red or purple) and four of the 23 occur in the classical signature (indicated by blue or purple). Remarkably, 13 of 216 mesenchymal classifier genes appear in our list of 100 upregulated, RiboTag-depleted genes ($p = 1.5 \times 10^{-10}$, hypergeometric test). Mesenchymal was the only subtype with significant overlap between its classifier genes and the upregulated, depleted genes, and the overlap was particularly enriched in pathways related to extracellular matrix and cell adhesion.

Computational deconvolution of murine and human tumors reveals the cellular origin of differentially translated genes

To further assess the cellular origin of the signatures described above in both human and mouse, we applied an algorithm for computational deconvolution of expression profiles into cell-type-specific profiles. We used a modified version of the PSEA algorithm to deconvolve 10 RNA-Seq profiles from our murine tumors and 39 RNA-Seq profiles from human proneural GBM. The algorithm, which has been applied previously to brain tissue, uses expression of cell-type-specific marker genes as a proxy for the fractional composition of each cell type in each sample and linear regression to estimate a single, average expression profile for each cell type based on all input samples. Our implementation of PSEA differs somewhat from the previous report in that we iteratively improve our cellular composition estimate for each sample using the best-fit genes from previous iterations and calculate the absolute expression of each gene in each cell type rather than expression relative to a marker.

We deconvolved the mouse and human expression profiles using six cell types: OPCs, mature oligodendrocytes, neurons, microglia, and two different types of astrocytes. We seeded the algorithm with markers of each population: *Olig2* for OPCs, *Mog/Mal* for oligodendrocytes, *Rbfox3/Neurod6* for neurons, *Aif1/Cd68* for microglia, *Aqp4* for the first astrocyte population (Astro1), and *Cd44* for the second (Astro2). These markers are based largely on previous attempts to deconvolve brain tissue expression and the large-scale study of cell-type-specific expression in the mouse brain (Cahoy et al., 2008). We included a *Cd44*⁺ cells commonly associated with gliomas that resemble the cultured astrocytes from Cahoy et al. (2008) and the expression signature of mesenchymal glioblastoma (Cahoy et al., 2008; Verhaak et al., 2010; Katz et al., 2012). We note that, although *Olig2* is expressed in both OPCs and some mature oligodendrocytes in normal brain tissue, the expression level of *Olig2* in proneural glioma cells and proliferating OPCs is considerably higher than in mature oligodendrocytes, making it essentially cell-type-specific. The expression of *Olig2* across the 10 murine tumors profiled for deconvolution is strongly correlated other OPC and proneural glioma markers ($r = 0.6$ for *Mki67*; $r = 0.7$ for *Pdgfra*, *Sox2*, *Sox11*, and *Sox10*; $r = 0.8$ for *Cspg4*) and anticorrelated with that of canonical oligodendrocyte markers ($r = -0.6$ for *Mog*, *Mal*, *Mobp*, and *Mbp*; $r = -0.5$ for *Plp1*). We found that deconvolution of the mouse tumor expression profiles correctly predicts the cellular distribution of gene expression for most canonical marker genes in mouse tumor tissue (Fig. 3A). For further vali-

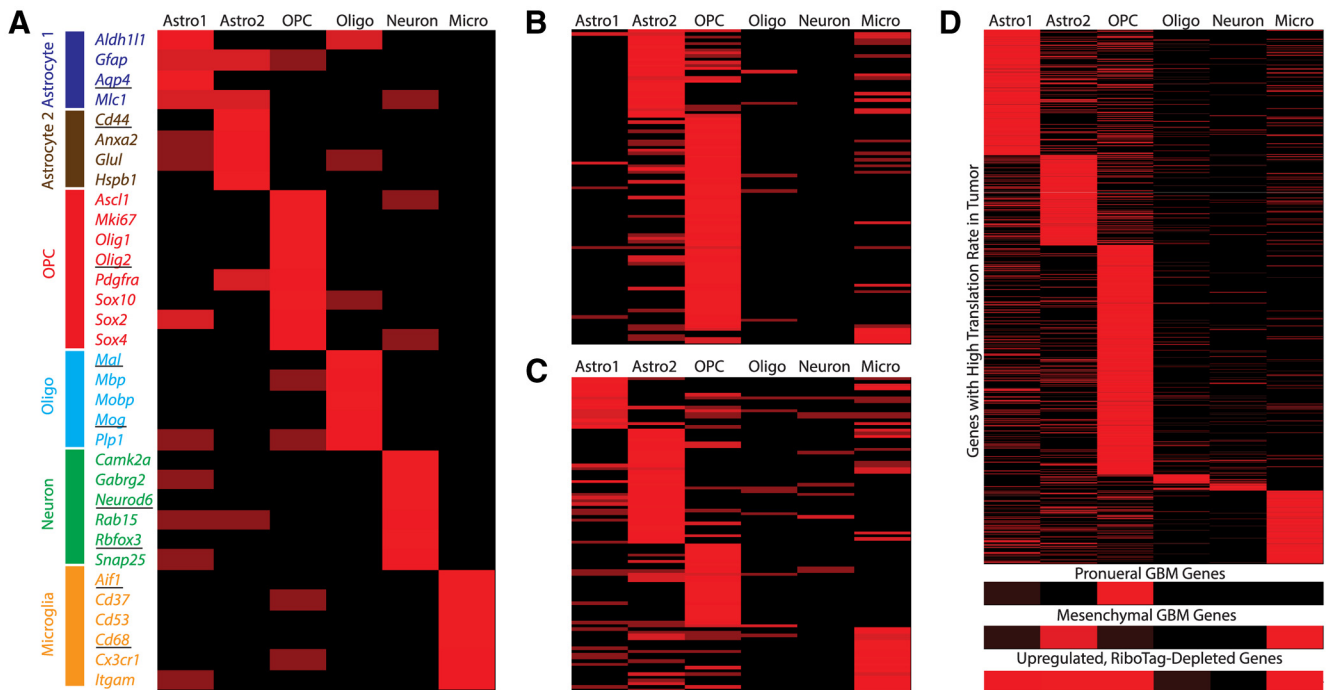


Figure 4. Computational deconvolution in human and murine glioma. **A**, Cell-type distribution heat map of canonical lineage marker genes based on computational deconvolution of 39 RNA-Seq profiles from 39 human proneural gliomas from TCGA. The columns are the cell types used for deconvolution and the values are normalized across rows. Genes used to seed the deconvolution algorithm are underlined. **B**, Cell-type distribution heat map of the upregulated, RiboTag-depleted genes from Figure 2B based on computational deconvolution of the murine RNA-Seq profiles from Figure 3A. **C**, Cell-type distribution heat map of these same genes, as determined by deconvolution of the human RNA-Seq profiles. **D**, Cell-type distribution heat map of genes with high-translation rates (>2 -fold across both specimens) in human proneural gliomas relative to non-neoplastic human brain tissue based on ribosome profiling. Cell type assignments for each gene are based on computational deconvolution of the proneural TCGA profiles (as in **C**). Below this heat map, we show three small heat maps showing the cell-type distribution of proneural classifier genes, mesenchymal classifier genes, and upregulated, RiboTag-depleted genes that also exhibit increased translation rates in human proneural glioma based on ribosome profiling. These small heat maps were constructed by summing the number of genes in each category with increased translation rate in the human glioma tissue that are predominantly deconvolved in each cell type.

dition, we examined the deconvolved profiles of genes that were enriched (Fig. 3B) or depleted (Fig. 3C) by RiboTag immunoprecipitation. Deconvolution predicts that $>87\%$ of the RiboTag-enriched genes are expressed predominantly in the OPC-like, *Olig2*⁺ lineage ($p = 1.3 \times 10^{-111}$, hypergeometric test) and $\sim 94\%$ of these genes have at least some predicted expression in this lineage (Fig. 3B). Conversely, there is considerable representation from all six cellular lineages as predominant cell types in deconvolution of the RiboTag-depleted genes (Fig. 3C). Here, only 44% of genes are predicted to be predominantly expressed in the OPC-like lineage, which is not a statistically significant enrichment ($p = 0.28$, hypergeometric test). Hence, our computational deconvolution algorithm recapitulates the cell type specificity of the RiboTag system.

We repeated this analysis on RNA-Seq profiles from 39 individual proneural human tumor samples (TCGA). Figure 4A shows the deconvolved human profiles across the same 33 canonical CNS-lineage marker genes used in Figure 3A, demonstrating the efficacy of the algorithm in human samples. We used deconvolution to assess the cellular distribution underlying expression of the upregulated, RiboTag-depleted genes from the tumor microenvironment. Figure 4B, C shows the deconvolved profiles of these genes in the mouse and human tumors, respectively. Although there are differences between the two distributions, in both cases we see that neurons and oligodendrocytes are underrepresented compared with the more reactive OPC-like, astrocytic, and microglial lineages.

Cell-type-specific translation rates in human proneural glioma

To further our cross-species comparison, we generated ribosome profiling libraries from two human proneural glioma specimens and three human non-neoplastic brain tissue specimens procured during surgical resection for epilepsy. Figure 4D shows the deconvolution of genes with consistently higher translation rates in the glioma tissue samples compared with the non-neoplastic brain samples. As expected, the *Olig2*⁺ OPC-like lineage predominates along with astrocytes and microglia, whereas there is little contribution for the mature oligodendrocytic and neuronal lineages.

We then asked which cell types are most enriched in the proneural and mesenchymal classifier genes with high translation rates, as measured by ribosome profiling of the human glioma samples (Fig. 4D). Although the *Olig2*⁺ OPC-like lineage is significantly enriched in the proneural classifier genes ($p = 2.4 \times 10^{-12}$, hypergeometric test), only the *Cd44*⁺ reactive astrocyte ($p = 0.03$, hypergeometric test) and microglial ($p = 8 \times 10^{-13}$, hypergeometric test) lineages are significantly enriched for mesenchymal classifier genes. Furthermore, the genes that are upregulated and RiboTag-depleted in the mouse proneural gliomas, and also have a high translation rate in the human proneural gliomas are predominately distributed in astrocytes, *Olig2*⁺ cells, and microglia in our human tumor specimens, however, they are only statistically enriched in the *Cd44*⁺ reactive astrocyte lineage.

Unlike the RiboTag glioma mouse model, where we can clearly delineate which cells received the initiating genetic lesion,

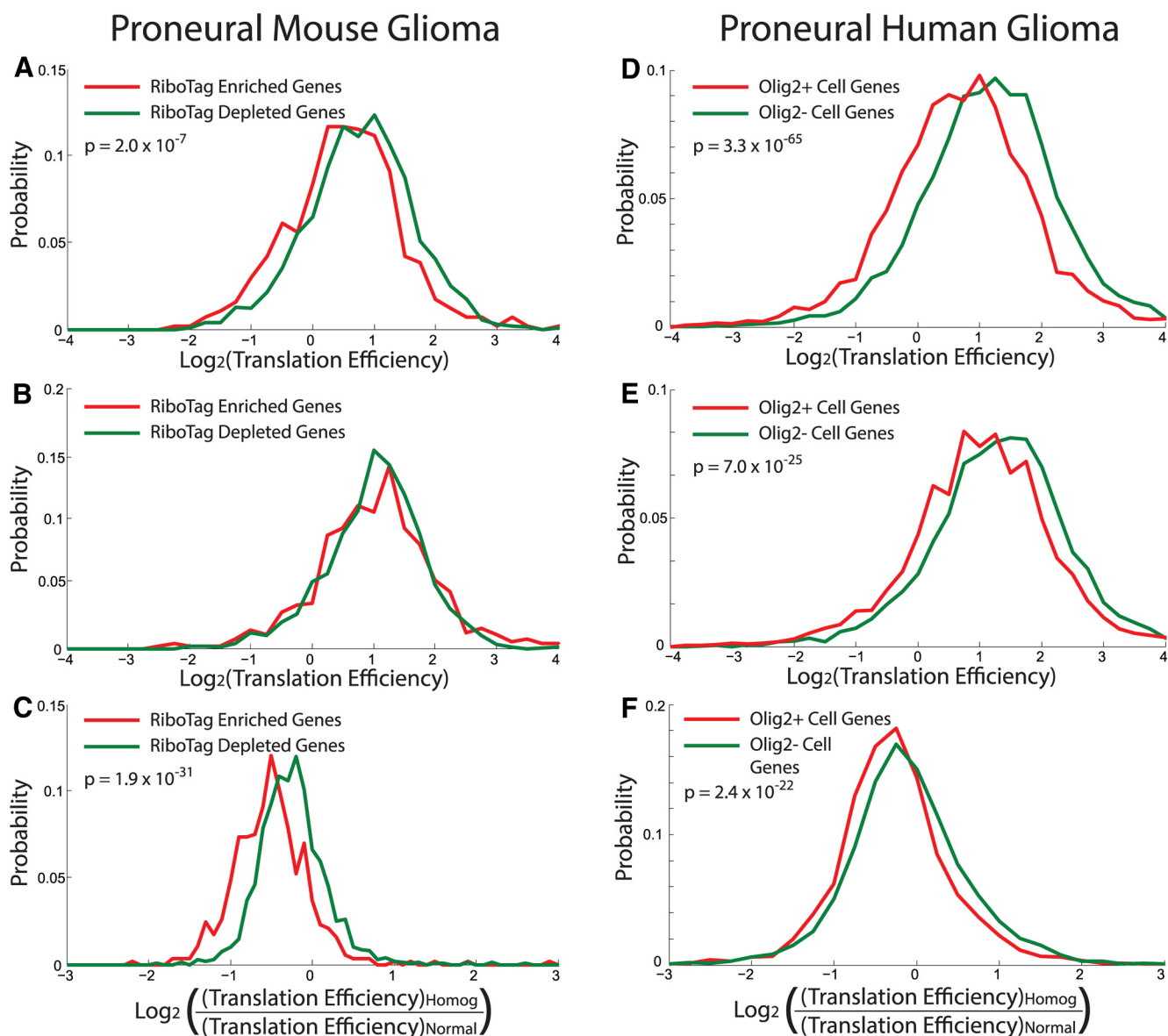


Figure 5. Translation efficiency analysis from mouse and human ribosome profiling. **A**, Histograms of the mean translation efficiency for genes that are either enriched (red) or depleted (green) computed from ribosome profiles and RNA-Seq of the homogenate murine tumor samples. The tumor-specific, RiboTag-enriched genes show a statistically significant tendency toward lower translation efficiency. **B**, Histograms of the mean translation efficiency for genes that are either enriched (red) or depleted (green) computed from the ribosome profiles and RNA-Seq of the murine normal brain samples. There is no statistically significant difference between the two gene sets in normal brain. **C**, Histogram of the translation efficiency fold-change between tumor homogenate and normal brain for the RiboTag-enriched and RiboTag-depleted genes showing that $\sim 90\%$ of RiboTag-enriched genes are translationally downregulated in the murine tumors. **D**, Histograms of the mean translation efficiency for genes that are predominantly expressed in *Olig2*⁺ (red) or *Olig2*[−] (green) cells (based on proneural TCGA deconvolution) computed from ribosome profiles and RNA-Seq of the human proneural tumor samples. Genes deconvolved in *Olig2*⁺ cells show a statistically significant tendency toward lower translation efficiency, similar to the RiboTag-enriched genes in the mouse model. **E**, Histograms of the mean translation efficiency for genes that are predominantly expressed in *Olig2*⁺ (red) or *Olig2*[−] (green) cells (same genes as in **D**) computed from the ribosome profiles and RNA-Seq of the human non-neoplastic brain tissue. There is a significantly smaller shift in translation efficiency compared with the tumor tissue in **D**. **F**, Histogram of the translation efficiency fold-change between tumor and non-neoplastic brain tissue for genes deconvolved in the *Olig2*⁺ (red) or *Olig2*[−] (green) cells showing that $>70\%$ of genes associated with *Olig2*⁺ cells are translationally downregulated in the human proneural tumors.

we cannot make this distinction in human tumors. However, given the overall resemblance of proneural glioma to glial progenitors and the characteristic expansion of *Olig2*⁺ cells, the *Cd44*⁺ reactive astrocyte lineage and the microglial lineage most likely represents non-neoplastic or reactive cells in the tumor microenvironment. Hence, we draw an analogy between our finding that mesenchymal, RiboTag-depleted genes are upregulated primary in reactive astrocytes and microglia in our mouse glioma model and these results showing upregulation of mesenchymal classifier genes in *Olig2*[−], reactive astrocytes and microglia in human proneural tumors.

Translation efficiency is reduced in transformed cells in proneural glioma

We computed the mean homogenate translation efficiency for genes that were enriched or depleted by the RiboTag across all three mice (Fig. 5A). As observed in other mammalian cell types, this distribution is broad, highlighting the role of translational regulation in determining protein output (Ingolia et al., 2011). Notably, we found that the median translation efficiency of depleted genes was $\sim 25\%$ higher than that of the enriched genes, and that there was a significant difference in their translation efficiency distributions ($p = 2 \times 10^{-7}$, two-sample Kolmogorov–

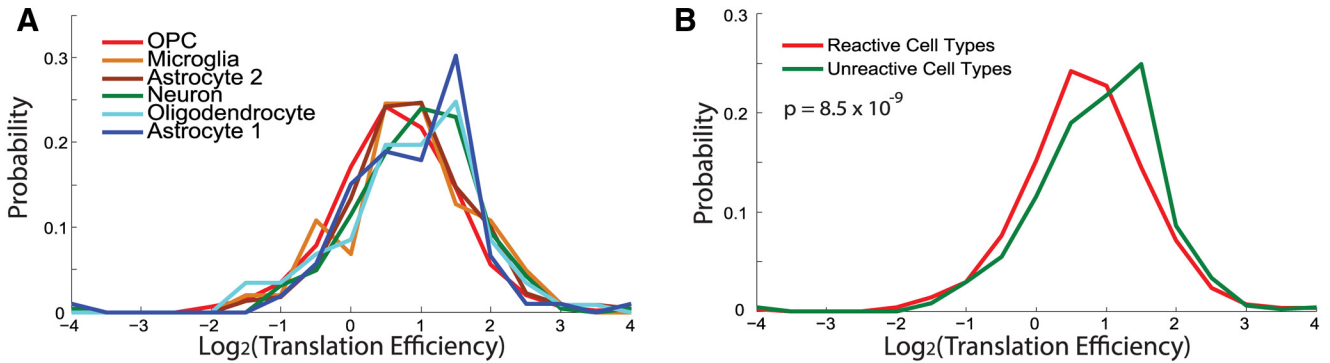


Figure 6. Translation efficiency in reactive cell types in the mouse model. **A**, Cell-type-specific translation efficiency histograms for all RiboTag-depleted genes based on the homogenate ribosome profiling and deconvolution. **B**, Grouping the cell types into reactive (astro2, OPC, microglia) and unreactive (astro1, oligodendrocyte, neuron) reveals a significant shift toward lower translation efficiency for reactive cells.

Smirnov, KS, test). To determine whether this difference was specific to the tumor tissue, we performed ribosome profiling on three normal brain samples. We found no significant difference between the translation efficiency distributions of these two gene sets in normal brain (Fig. 5B). Furthermore, we found that the translation efficiency was lower in tumor homogenate than in normal brain for >88% of RiboTag-enriched genes (Fig. 5C), implying a global difference in the translational status of transformed cells.

We repeated this analysis in our human proneural glioma ribosome profiles where *Olig2*⁺ cells comprise the majority of the transformed population. We used computational deconvolution to compare translation efficiency in the *Olig2*⁺ lineage to the other CNS lineages in the tumor. Like the RiboTag-enriched genes in the mouse model, we found that genes associated with the *Olig2*⁺ lineage exhibited ~40% lower translation efficiencies than genes expressed predominantly by other CNS lineages (Fig. 5D). Although still statistically significant, this difference was reduced upon examination of the same genes in non-neoplastic brain tissue (Fig. 5E). Finally, we found that >70% of genes associated with the *Olig2*⁺ lineage were translated with lower efficiency in human glioma tissue than in non-neoplastic human brain tissue (Fig. 5F). Hence, the qualitative aspects of our translation efficiency phenotype described in the RiboTag mouse model are also present in the human disease.

The RiboTag system allows us to clearly separate cells that received the tumor-initiating genetic alteration from tumor-associated cells, which would be a more complex task in human tumors. The deconvolution suggests that, in the mouse tumors, the OPC, microglial, and *Cd44*⁺ astrocytes recruited to the tumor microenvironment are responsible for the mesenchymal-like signature observed among RiboTag-depleted genes. Given the reactive nature of these cell types, we next asked whether the translational status of these cells is globally altered as we observed for transformed cells. We used the RNA-Seq deconvolution to assign a predominant cell type for each RiboTag-depleted gene. We then calculated the translation efficiency distributions from the homogenate ribosome profiling data for genes assigned to each cell type (Fig. 6A). Here, we see a clear shift toward lower translation efficiency for genes expressed by OPCs, microglia, and *Cd44*⁺ astrocytes compared with *Cd44*⁻ astrocytes, oligodendrocytes, and neurons. Figure 6B shows the same distribution after grouping genes into reactive (OPC, microglial, Astro2) and unreactive (Astro1, neuronal, oligodendrocytic) lineages and a significant shift ($p = 8.5 \times 10^{-9}$, two-sample KS test). Hence, the low-efficiency translation phenotype seen in the transformed

cells (Fig. 5A–C) is also seen in the recruited or reactive cell types in the microenvironment.

Sequence-dependent differential regulation of 5'-leader ribosomal occupancy

We detected ribosomal density in the 5'-leader sequences of hundreds of genes in the profiles obtained from RiboTag immunoprecipitation. Ribosomal density in 5'-leader sequences, which are typically annotated as noncoding 5'-UTRs, has been implicated in translational control in multiple contexts (Barbosa et al., 2013; Somers et al., 2013), and our results suggest that it could play a role for specific genes expressed by transformed cells in the tumor. Previous genome-wide studies of translation in yeast have implied that the presence of upstream AUG start codons (uAUGs) can interfere with translation initiation at the downstream coding sequence by recruiting ribosomes to the 5' leader (Brar et al., 2012; Arribere and Gilbert, 2013). To explore the potential implications of 5'-leader density on protein synthesis, we investigated the sequence content of 5'-leader sequences in the mouse transcriptome. Figure 7A shows the CDS translation efficiency distributions from the tumor homogenate for genes with 5'-leader density in the RiboTag profile that either contain or lack uAUG. Despite ribosomal density in the 5'-leader sequence, genes without uAUG are translated nearly twofold more efficiently on average than genes with uAUG in the tumor ($p = 4.1 \times 10^{-13}$, two-sample KS test). We observed a similar trend in normal brain tissue (Fig. 7B). Hence, the phenomenon is general to healthy and neoplastic tissue.

The observed 5'-leader activity and the significant difference in translation efficiency between transcripts containing and lacking uAUG led us to investigate whether 5'-leader ribosomal occupancy is differentially regulated. We first calculated the 5'-leader efficiency based on the tumor homogenate profiles for genes with 5'-leader ribosomal density in the RiboTag profiles. Unlike the translation efficiency of the CDS, which we found to be broadly reduced in the tumor relative to normal brain for RiboTag-enriched genes (and for ~90% of RiboTag-enriched genes with 5'-leader density), the 5'-leader efficiency was not affected to the same extent (only ~60% showed reduced 5'-leader efficiency). Figure 7C shows histograms of the ratio of the 5'-leader efficiency fold-change relative to the CDS translation efficiency fold-change for transcripts containing or lacking uAUGs for RiboTag-enriched genes with 5'-leader density. For uAUG-containing transcripts, the 5'-leader and CDS translation efficiency fold-changes relative to normal brain are similar, and

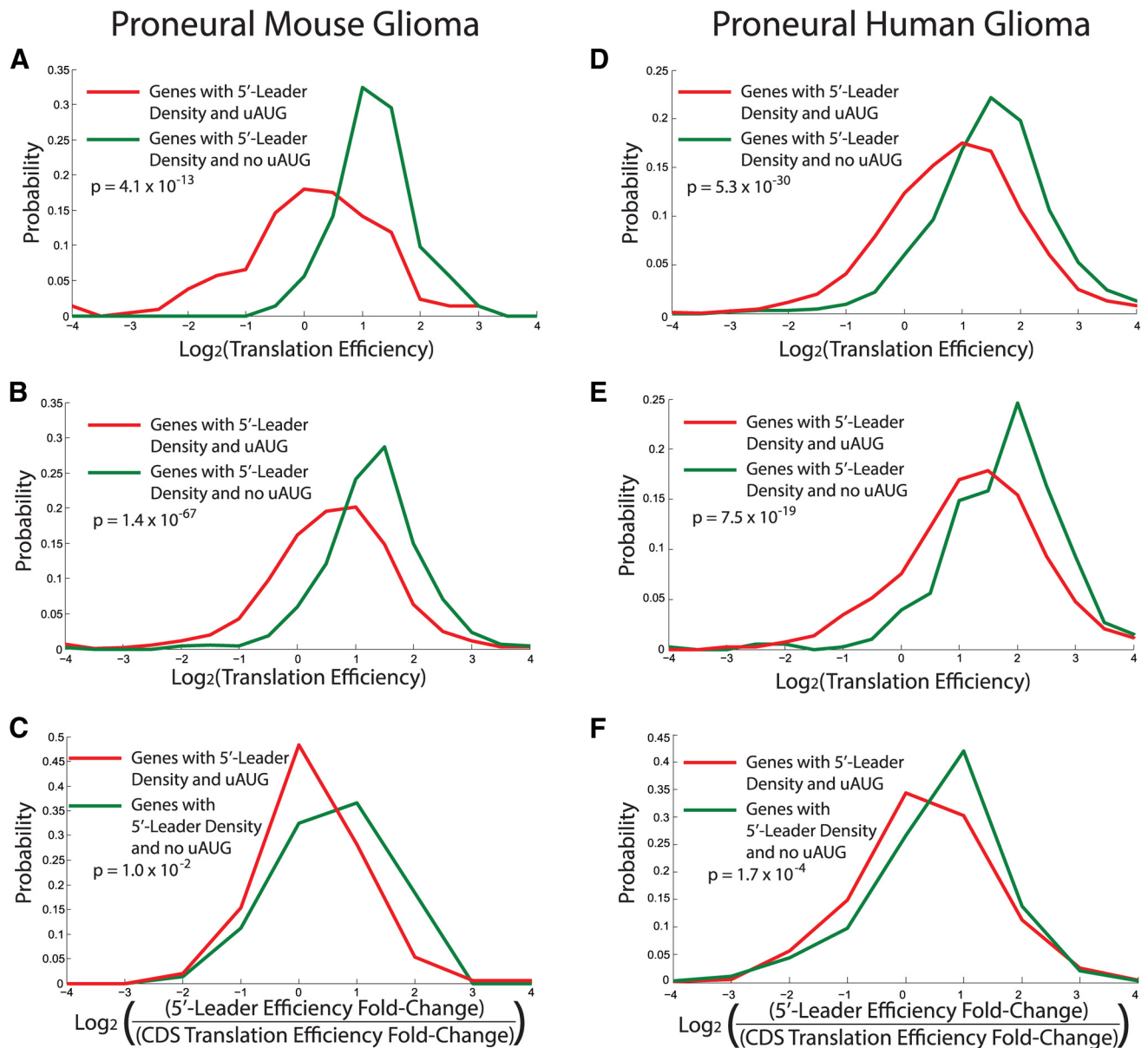


Figure 7. Analysis of noncanonical translation in murine and human glioma. **A**, Histograms of the translation efficiency of genes with 5'-leader density across all three RiboTag profiles that either contain (red) or lack (green) uAUG. **B**, Histograms of the translation efficiency of genes with 5'-leader density across all three normal brain profiles that either contain (red) or lack (green) uAUG. **C**, Histogram of the ratio of 5'-leader efficiency fold-change to CDS translation efficiency fold-change for genes with 5'-leader density in all three RiboTag mice and either contain (red) or lack (green) uAUG. Fold-change in 5'-leader and CDS translation efficiencies are calculated between the tumor homogenate and normal brain samples. **D**, Histograms of the translation efficiency of genes with 5'-leader density across both human proneural glioma ribosome profiles that either contain (red) or lack (green) uAUG. **E**, Histograms of the translation efficiency of genes with 5'-leader density across all three human non-neoplastic brain ribosome profiles that either contain (red) or lack (green) uAUG. **F**, Histogram of the ratio of 5'-leader efficiency fold-change to CDS translation efficiency fold-change for genes with 5'-leader density in both human proneural glioma ribosome profiles and either contain (red) or lack (green) uAUG. Fold-change in 5'-leader and CDS translation efficiencies are calculated between the human tumor and non-neoplastic brain samples.

the log-scale histogram peaks at zero. However, for transcripts lacking uAUG, the 5'-leader and CDS translation efficiencies appear differentially regulated to a larger extent, and the ratio of their fold-changes is greater than one for ~73% of these genes. The difference in these two distributions is statistically significant ($p = 0.01$, two-sample KS test). This can be improved to $p = 4.5 \times 10^{-16}$ by including genes with 5'-leader density in any of the three RiboTag mice in the analysis. This result raises the possibility that translation initiation on non-AUG start codons and canonical initiation are regulated by different factors and that those factors are differentially regulated in the transformed cells compared with normal brain.

Finally, we repeated all three analyses of 5'-leader ribosomal occupancy on our human proneural and non-neoplastic ribosome profiles. Just as in the mouse, we find a distinctive shift toward higher CDS translation efficiency for genes with 5'-leader density that lack uAUG in their 5'-leader sequence in both tumor and non-neoplastic tissue (Figs. 7D,E). We also find that differential regulation of 5'-leader density and CDS translation efficiency in comparing tumor and non-neoplastic ribosome profiles (Fig. 7F). The median fold-change ratio between 5'-leader and CDS efficiency is close to one for genes with ribosomal density in uAUG-containing 5'-leaders and shifted to higher values for genes without uAUG ($p = 1.7 \times 10^{-4}$). Hence, differ-

ential regulation of translation between neoplastic and non-neoplastic tissue is dependent on 5'-leader sequence content for proneural glioma in both human and mouse.

Discussion

The brain is compositionally heterogeneous and uniquely interconnected, while the translational machinery is highly dynamic. Together, these features demand a minimally perturbative, cell-type-specific approach to quantifying translation. Our strategy combines the RiboTag system for cell-type specificity with ribosome profiling to achieve a comprehensive, quantitative picture of translation in glioma. Dysregulation of protein synthesis plays a role in many diseases of the CNS, and we have made several key observations about translation in glioma.

Our results reveal a program of non-cell autonomous upregulation of cell adhesion and extracellular matrix-related genes in tumor-associated cells. These pathways are known to be upregulated in gliomas and play an important role in tumor cell invasion. Despite the fact that our glioma mouse model globally recapitulates the proneural subtype of human glioblastoma (Lei et al., 2011; Sonabend et al., 2014), the upregulated pathways found in tumor-associated cells are most closely associated with the mesenchymal subtype. Deconvolution of the upregulated subset of RiboTag-depleted genes in both mouse and human showed little contribution from oligodendrocytes and neurons. Instead, these genes were found in more reactive cell types such as the OPC-like, astrocytic, and microglial lineages. Hence, despite the resemblance of our animal model to the proneural subtype and the glial progenitor origin of these tumors, multiple cell types in the tumor microenvironment contribute a mesenchymal signature.

We found that the translation efficiency of the RiboTag-depleted genes is cell-type dependent. First, genes expressed in cells that arise from the retrovirus-infected lineage tend to have lower translation efficiencies than genes predominantly expressed in other CNS lineages. Although the total translational output of these genes is higher in the tumor than in normal brain and ribosomal components comprise an enriched ontology in the RiboTag profile, they are almost universally translated more efficiently in normal brain. Thus, glial transformation is associated with a significant reduction in global translation efficiency. Similarly, in human proneural gliomas, we find that genes expressed by *Olig2*⁺ cells exhibit significantly lower translation efficiency than genes expressed in other CNS lineages. There are multiple potential explanations for this phenomenon. Although gene-specific mechanisms of translational regulation could be responsible, it is also possible that the translational machinery is globally saturated with mRNA in the transformed cells. Another possibility is that tight translational control is essential to maintenance of an immature cell state and avoiding differentiation as demonstrated in other contexts (Signer et al., 2014; Tahmasebi et al., 2014).

We used deconvolution to divide the RiboTag-depleted genes into reactive (*Cd44*⁺ astrocytes, microglia, and OPCs) and unreactive (*Cd44*⁻ astrocytes, neurons, and oligodendrocytes) cells and found that reactive cells share the low translation efficiency phenotype with the transformed cells. Hence, protein expression in the reactive and transformed cell types may be driven through common, potentially dysregulated signaling pathways.

Finally, we detected ribosome footprint density in the 5'-leader sequences of many genes in the RiboTag profile and in human proneural glioma ribosome profiles. Although specific cases have been described, the precise role of 5'-leader density is

poorly understood on a gene-by-gene basis. For example, 5'-leader-bound ribosomes could be translationally inactive and mainly affect RNA stability or translation of the downstream CDS. Alternatively, these ribosomes could be actively translating upstream open reading frames (uORFs) to generate short peptides, fusion proteins, or full-length proteins from an unannotated ORF.

The downstream effects of 5'-leader ribosomal density on translation efficiency are strongly sequence-dependent. In normal brain, transcripts lacking the conventional start codon upstream of the annotated start of the CDS (uAUG) exhibit higher CDS translation efficiencies than those containing uAUG. We see that this regulatory framework is intact in the tumor, and the 5'-leader sequence dependence leads to differential regulation of translation relative to normal brain. For most genes expressed in the RiboTag cells, the CDS translation efficiency is lower in the tumor than in normal brain, but the same is not true for the 5'-leader efficiency. In particular, we discovered that the fold-changes in 5'-leader and CDS translation efficiencies for a given transcript relative to normal brain diverge for transcripts lacking uAUG. For these genes, translation initiation in a uORF and the annotated ORF must initiate from different start codons, and these two processes may be regulated by distinct factors or pathways that are differentially regulated in the tumor.

Both the RiboTag glioma mouse model and computational deconvolution are powerful tools for studying transcriptional and translational regulation in brain tumors, which are composed of a complex mixture of different cell types. Previous studies have shown that paracrine signaling between different cell types play an important role in glioma growth and induces non-cell autonomous alterations (Assanah et al., 2006; Hoelzinger et al., 2007; Bonavia et al., 2011; Pong and Gutmann, 2011). The proneural subtype of human glioblastoma, which is effectively modeled here, is driven by growth factor signaling, particularly through the PI3K/AKT pathway, which directly impacts translation through mTOR and other downstream signaling molecules. Our detailed characterization of glioma's altered translational landscape raises the possibility that the translational machinery is selectively vulnerable in transformed and reactive cells. The analytical tools described here will be essential to assessing and predicting the effects of therapeutic strategies that target protein synthesis.

References

- Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11:R106. [CrossRef Medline](#)
- Arribere JA, Gilbert WV (2013) Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res* 23:977–987. [CrossRef Medline](#)
- Assanah M, Lochhead R, Ogden A, Bruce J, Goldman J, Canoll P (2006) Glial progenitors in adult white matter are driven to form malignant gliomas by platelet-derived growth factor-expressing retroviruses. *J Neurosci* 26:6781–6790. [CrossRef Medline](#)
- Barbosa C, Peixeiro I, Romão L (2013) Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet* 9:e1003529. [CrossRef Medline](#)
- Bédard A, Tremblay P, Chernomoretz A, Vallières L (2007) Identification of genes preferentially expressed by microglia and upregulated during cuprizone-induced inflammation. *Glia* 55:777–789. [CrossRef Medline](#)
- Berriz GF, Beaver JE, Cenik C, Tasan M, Roth FP (2009) Next generation software for functional trend analysis. *Bioinformatics* 25:3043–3044. [CrossRef Medline](#)
- Bonavia R, Inda MM, Cavenee WK, Furnari FB (2011) Heterogeneity maintenance in glioblastoma: a social network. *Cancer Res* 71:4055–4060. [CrossRef Medline](#)
- Brar GA, Yassour M, Friedman N, Regev A, Ingolia NT, Weissman JS (2012)

- High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* 335:552–557. [CrossRef Medline](#)
- Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, Thompson WJ, Barres BA (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* 28:264–278. [CrossRef Medline](#)
- Chen Z, Trotman LC, Shaffer D, Lin HK, Dotan ZA, Niki M, Koutcher JA, Scher HI, Ludwig T, Gerald W, Cordon-Cardo C, Pandolfi PP (2005) Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* 436:725–730. [CrossRef Medline](#)
- Doyle JP, Dougherty JD, Heiman M, Schmidt EF, Stevens TR, Ma G, Bupp S, Shrestha P, Shah RD, Doughty ML, Gong S, Greengard P, Heintz N (2008) Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* 135:749–762. [CrossRef Medline](#)
- Fan QW, Cheng CK, Nicolaidis TP, Hackett CS, Knight ZA, Shokat KM, Weiss WA (2007) A dual phosphoinositide-3-kinase/alpha/mTOR inhibitor cooperates with blockade of epidermal growth factor receptor in PTEN-mutant glioma. *Cancer Res* 67:7960–7965. [CrossRef Medline](#)
- Goodarzi H, Elemento O, Tavazoie S (2009) Revealing global regulatory perturbations across human cancers. *Mol Cell* 36:900–911. [CrossRef Medline](#)
- Heiman M, Schaefer A, Gong S, Peterson JD, Day M, Ramsey KE, Suárez-Farinas M, Schwarz C, Stephan DA, Surmeier DJ, Greengard P, Heintz N (2008) A translational profiling approach for the molecular characterization of CNS cell types. *Cell* 135:738–748. [CrossRef Medline](#)
- Helmy K, Halliday J, Fomchenko E, Setty M, Pitter K, Hafemeister C, Holland EC (2012) Identification of global alteration of translational regulation in glioma in vivo. *PLoS One* 7:e46965. [CrossRef Medline](#)
- Hoelzinger DB, Demuth T, Berens ME (2007) Autocrine factors that sustain glioma invasion and paracrine biology in the brain microenvironment. *J Natl Cancer Inst* 99:1583–1593. [CrossRef Medline](#)
- Hsieh AC, Liu Y, Edlind MP, Ingolia NT, Janes MR, Sher A, Shi EY, Stumpf CR, Christensen C, Bonham MJ, Wang S, Ren P, Martin M, Jessen K, Feldman ME, Weissman JS, Shokat KM, Rommel C, Ruggiero D (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* 485:55–61. [CrossRef Medline](#)
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324:218–223. [CrossRef Medline](#)
- Ingolia NT, Lareau LF, Weissman JS (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147:789–802. [CrossRef Medline](#)
- Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 7:1534–1550. [CrossRef Medline](#)
- Jiang BH, Liu LZ (2009) PI3K/PTEN signaling in angiogenesis and tumorigenesis. *Adv Cancer Res* 102:19–65. [CrossRef Medline](#)
- Katz AM, Amankulor NM, Pitter K, Helmy K, Squatrito M, Holland EC (2012) Astrocyte-specific expression patterns associated with the PDGF-induced glioma microenvironment. *PLoS One* 7:e32453. [CrossRef Medline](#)
- Kim D, Perlea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36. [CrossRef Medline](#)
- Kuhn A, Thu D, Waldvogel HJ, Faull RL, Luthi-Carter R (2011) Population-specific expression analysis (PSEA) reveals molecular changes in diseased brain. *Nat Methods* 8:945–947. [CrossRef Medline](#)
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* 9:357–359. [CrossRef Medline](#)
- Lei L, Sonabend AM, Guarnieri P, Soderquist C, Ludwig T, Rosenfeld S, Bruce JN, Canoll P (2011) Glioblastoma models reveal the connection between adult glial progenitors and the proneural phenotype. *PLoS One* 6:e20041. [CrossRef Medline](#)
- Liu C, Sage JC, Miller MR, Verhaak RG, Hippenmeyer S, Vogel H, Foreman O, Bronson RT, Nishiyama A, Luo L, Zong H (2011) Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* 146:209–221. [CrossRef Medline](#)
- Parsa AT, Holland EC (2004) Cooperative translational control of gene expression by Ras and Akt in cancer. *Trends Mol Med* 10:607–613. [CrossRef Medline](#)
- Pong WW, Gutmann DH (2011) The ecology of brain tumors: lessons learned from neurofibromatosis-1. *Oncogene* 30:1135–1146. [CrossRef Medline](#)
- Rajasekhar VK, Viale A, Socci ND, Wiedmann M, Hu X, Holland EC (2003) Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol Cell* 12:889–901. [CrossRef Medline](#)
- Sanz E, Yang L, Su T, Morris DR, McKnight GS, Amieux PS (2009) Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci U S A* 106:13939–13944. [CrossRef Medline](#)
- Signer RA, Magee JA, Salic A, Morrison SJ (2014) Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* 509:49–54. [CrossRef Medline](#)
- Somers J, Pöyry T, Willis AE (2013) A perspective on mammalian upstream open reading frame function. *Int J Biochem Cell Biol* 45:1690–1700. [CrossRef Medline](#)
- Sonabend AM, Yun J, Lei L, Leung R, Soderquist C, Crisman C, Gill BJ, Carminucci A, Sisti J, Castelli M, Sims PA, Bruce JN, Canoll P (2013) Murine cell line model of proneural glioma for evaluation of anti-tumor therapies. *J Neurooncol* 112:375–382. [CrossRef Medline](#)
- Sonabend AM, Bansal M, Guarnieri P, Lei L, Amendolara B, Soderquist C, Leung R, Yun J, Kennedy B, Sisti J, Bruce S, Bruce R, Shakya R, Ludwig T, Rosenfeld S, Sims PA, Bruce JN, Califano A, Canoll P (2014) The transcriptional regulatory network of proneural glioma determines the genetic alterations selected during tumor progression. *Cancer Res* 74:1440–1451. [CrossRef Medline](#)
- Tahmasebi S, Alain T, Rajasekhar VK, Zhang JP, Prager-Khoutorsky M, Khoutorsky A, Dogan Y, Gkogkas CG, Petroulakis E, Sylvestre A, Ghorbani M, Assadian S, Yamanaka Y, Vinagolu-Baur JR, Teodoro JG, Kim K, Yang XJ, Sonenberg N (2014) Multifaceted regulation of somatic cell reprogramming by mRNA translational control. *Cell Stem Cell* 14:606–616. [CrossRef Medline](#)
- Takeuchi H, Kondo Y, Fujiwara K, Kanzawa T, Aoki H, Mills GB, Kondo S (2005) Synergistic augmentation of rapamycin-induced autophagy in malignant glioma cells by phosphatidylinositol 3-kinase/protein kinase B inhibitors. *Cancer Res* 65:3336–3346. [CrossRef Medline](#)
- Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O’Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, et al. (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17:98–110. [CrossRef Medline](#)