



# Coding and noncoding landscape of extracellular RNA released by human glioma stem cells

## Citation

Wei, Z., A. O. Batagov, S. Schinelli, J. Wang, Y. Wang, R. El Fatimy, R. Rabinovsky, et al. 2017. "Coding and noncoding landscape of extracellular RNA released by human glioma stem cells." Nature Communications 8 (1): 1145. doi:10.1038/s41467-017-01196-x. <http://dx.doi.org/10.1038/s41467-017-01196-x>.

## Published Version

doi:10.1038/s41467-017-01196-x

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34492187>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

ARTICLE

DOI: 10.1038/s41467-017-01196-x

OPEN

# Coding and noncoding landscape of extracellular RNA released by human glioma stem cells

Zhiyun Wei<sup>1</sup>, Arsen O. Batagov<sup>2</sup>, Sergio Schinelli<sup>3</sup>, Jintu Wang<sup>4</sup>, Yang Wang<sup>1</sup>, Rachid El Fatimy<sup>1</sup>, Rosalia Rabinovsky<sup>1</sup>, Leonora Balaj<sup>5</sup>, Clark C. Chen<sup>6</sup>, Fred Hochberg<sup>7,8</sup>, Bob Carter<sup>7</sup>, Xandra O. Breakefield<sup>5</sup> & Anna M. Krichevsky<sup>1</sup>

Tumor-released RNA may mediate intercellular communication and serve as biomarkers. Here we develop a protocol enabling quantitative, minimally biased analysis of extracellular RNAs (exRNAs) associated with microvesicles, exosomes (collectively called EVs), and ribonucleoproteins (RNPs). The exRNA complexes isolated from patient-derived glioma stem-like cultures exhibit distinct compositions, with microvesicles most closely reflecting cellular transcriptome. exRNA is enriched in small ncRNAs, such as miRNAs in exosomes, and precisely processed tRNA and Y RNA fragments in EVs and exRNPs. EV-enclosed mRNAs are mostly fragmented, and UTRs enriched; nevertheless, some full-length mRNAs are present. Overall, there is less than one copy of non-rRNA per EV. Our results suggest that massive EV/exRNA uptake would be required to ensure functional impact of transferred RNA on brain recipient cells and predict the most impactful miRNAs in such conditions. This study also provides a catalog of diverse exRNAs useful for biomarker discovery and validates its feasibility on cerebrospinal fluid.

<sup>1</sup>Department of Neurology, Brigham and Women's Hospital and Harvard Medical School, HMS Initiative for RNA Medicine, Boston, MA 02115, USA.

<sup>2</sup>Vishuo Biomedical, #3-33 Teletch Park, 20 Science Park Road, Singapore 117674, Singapore. <sup>3</sup>Department of Drug Sciences, University of Pavia, Pavia 27100, Italy. <sup>4</sup>Beijing Genomics Institute, Shenzhen 518083, China. <sup>5</sup>Department of Neurology and Radiology, Massachusetts General Hospital and Program in Neuroscience, Harvard Medical School, Charlestown, MA 02129, USA. <sup>6</sup>Neurosurgery Department, University of Minnesota, Minneapolis, MN 55455, USA. <sup>7</sup>Department of Neurosurgery, University of California, La Jolla, San Diego, CA 92093, USA. <sup>8</sup>Scintillon Institute, San Diego, CA 92121, USA. Correspondence and requests for materials should be addressed to A.M.K. (email: [akrichevsky@bwh.harvard.edu](mailto:akrichevsky@bwh.harvard.edu))

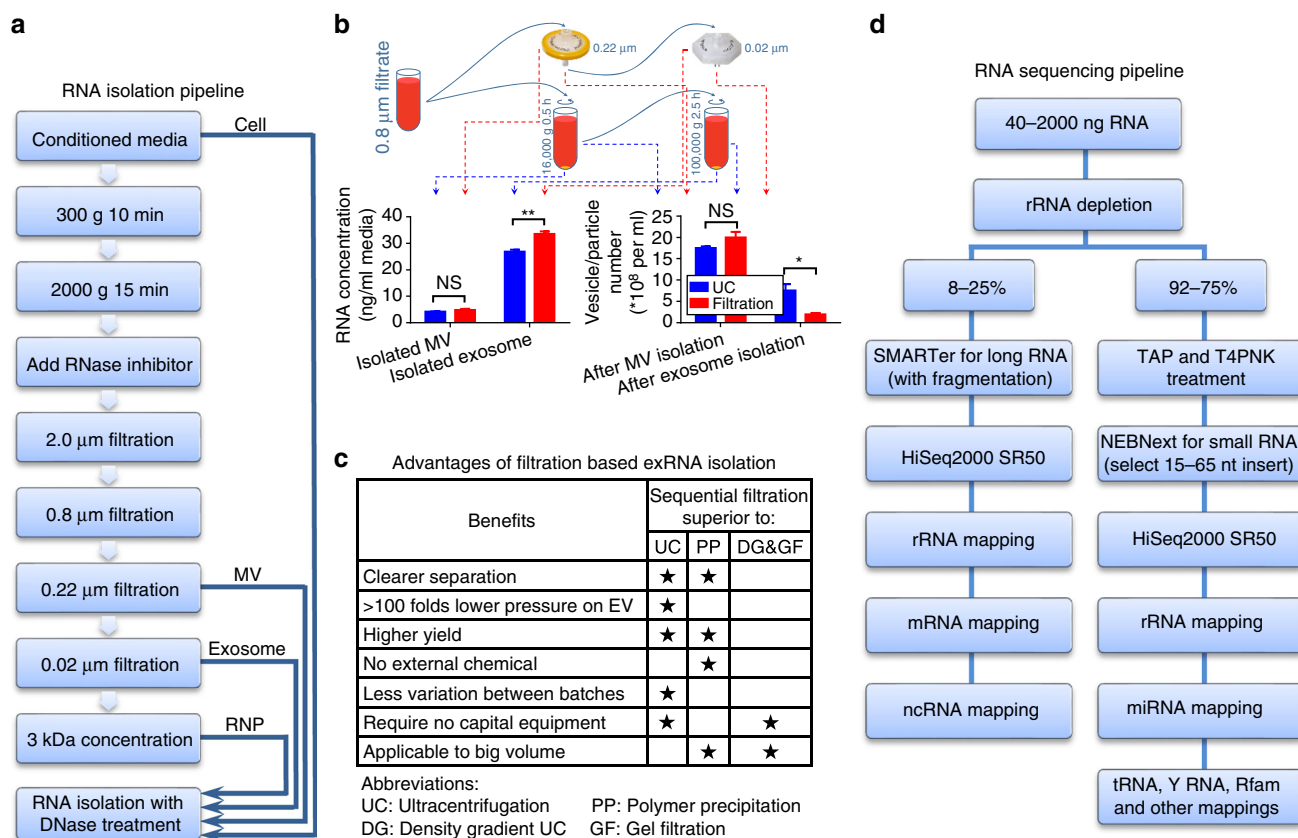
Intercellular communication within complex biological systems, such as cancer and its host microenvironment, via “horizontal” RNA transfer, is an expanding area of research<sup>1</sup>. Extracellular RNAs (exRNAs) are packaged into various extracellular complexes, including microvesicles (MVs), exosomes, and non-vesicular ribonucleoprotein complexes (RNPs)<sup>2,3</sup>. MVs and exosomes, broadly called extracellular vesicles (EVs), are released and taken up by various cells, thereby transferring their content. This process likely plays a role in cancer development and manipulation of its microenvironment<sup>4</sup>. However, methodologies are only beginning to emerge for characterizing the exRNA landscape and monitoring levels of individual coding and regulatory exRNAs. exRNA mostly consists of small RNA species (<200 nt); and the majority of reports to date focus on miRNA<sup>5,6</sup>. As a critical step toward understanding the biological impact of exRNA release and transfer, we investigated the complete spectrum of cancer-derived exRNAs, and the enrichment of specific RNA classes and individual species. By creating cDNA libraries of both small and long exRNA, and reducing the ligation bias favoring miRNAs, we identified a diverse and highly distinct composition of exRNA in MVs, exosomes, and RNP complexes. Furthermore, semi-absolute quantification of RNAseq, performed using RNA spike-in molecules, allowed us to monitor the levels of various RNA classes and species in these extracellular complexes.

This work focused on glioblastoma (GBM), the most common and aggressive brain tumor, as an important model for

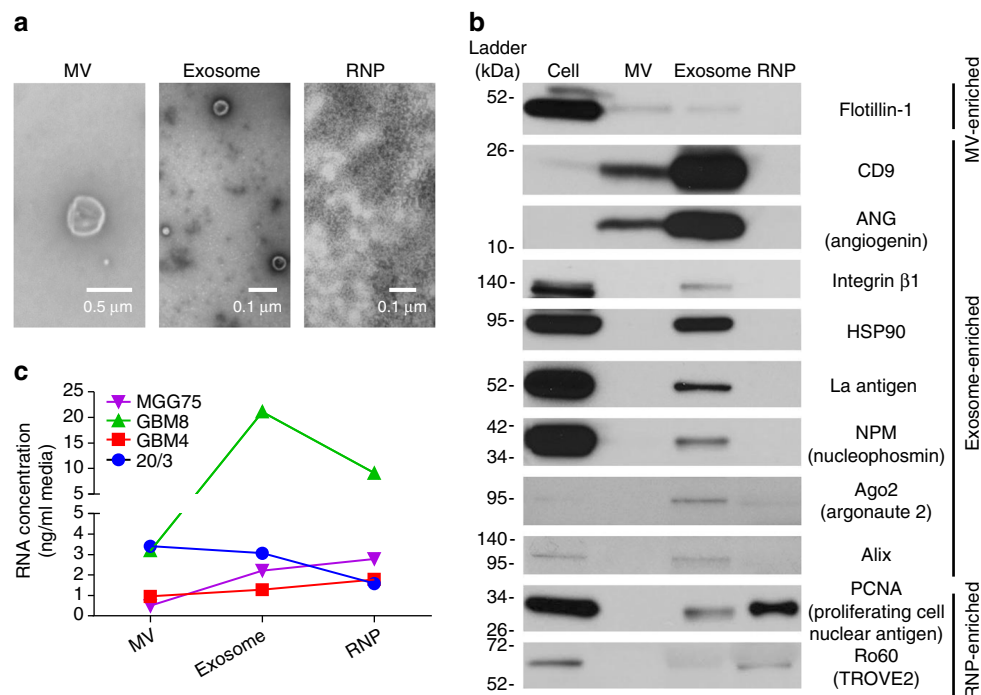
investigation of cancer-derived exRNA. As proliferating and invading GBM cells migrate through brain parenchyma, they interact with the changing landscape of extra-tumoral stimuli and actively modulate this landscape<sup>4</sup>. Communication between tumor cells and diverse normal cells in the brain is nevertheless one of the least investigated aspects of glioma biology. We employed low-passage patient-derived tumorigenic GBM cell cultures that represent the most therapy-resistant stem-like cell population (GSC), and are considered the core cell type within the tumor. Analysis of GSC cellular and extracellular RNA, along with the transcriptome of primary human and mouse cells of the brain microenvironment (neurons, astrocytes, endothelial cells, and microglia) enables us to predict the most impactful miRNAs and expand the repertoire of potentially transferred exRNAs far beyond the classes of miRNAs and mRNAs. We also demonstrate that MVs, large vesicles of 0.2–0.8  $\mu\text{m}$ , most closely mirror the cellular transcriptome and thus present a highly promising but yet poorly explored source of liquid biopsy biomarkers.

## Results

**Sequential filtration-based exRNA isolation.** To characterize exRNA released by patient-derived GBM cells in various complexes, we assessed several technical approaches. EV and exRNA isolation protocols can be generally categorized into three major groups: based on ultracentrifugation (UC), precipitation using



**Fig. 1** Flowchart of the exRNA fractionation and sequencing. **a** The pipeline of the filtration-based exRNA isolation. Following removal of cells and cellular debris by low speed centrifugation, the supernatants were filtered through a sequence of reduced pore sizes (2.0, 0.8, 0.22, and 0.02  $\mu\text{m}$ ) to separate the extracellular fractions, and a final concentrator with the cutoff of 3 kDa was applied to collect the remaining small particles. **b** The aliquots of conditioned media after 0.8  $\mu\text{m}$  filtration were used for MV and exosome isolation, either by ultracentrifugation (UC) or filtration, and the RNA yield of these fractions compared. The number of remaining vesicles/particles was compared in UC supernatant and filter flow-through.  $N = 4$  aliquots of conditioned media. All bars represent mean  $\pm$  SEM. **c** Comparison between the filtration-based exRNA isolation and other common exRNA isolation methods. The stars mark superior characteristics of sequential filtration over other methods. **d** The optimized pipeline for the broad coverage, minimally biased RNA-sequencing. RNA of 15–65 nt was selected for the small RNA libraries, to reduce the overwhelming levels of tRNAs. NS, not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ;  $t$ -test



**Fig. 2** Quality control of the fraction separation. **a** Transmission electron microscopy of EVs and RNPs isolated using the sequential filtration protocol. TEMs were replicated three times. **b** Protein markers, identified by the western blotting, verified the separation of extracellular fractions. Equal protein input of 50  $\mu$ g per lane was used. Western blots were replicated twice. **c** RNA yields of extracellular fractions from different GSC cultures

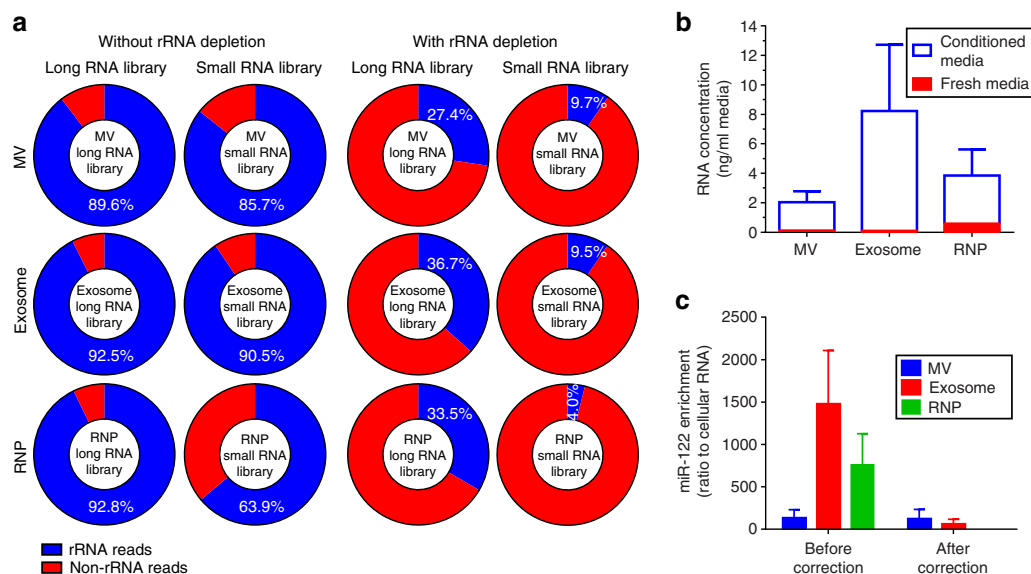
chemical polymers (PP), such as polyethylene glycol, and fractionation, including density gradient UC and gel filtration (DG&GF)<sup>7</sup>. Since specific markers or physical parameters for the various types of EVs and extracellular RNPs are still not clearly defined, UC remains the most commonly used approach to isolate the entire vesiculome<sup>8</sup>. However, based on nanoparticle tracking analysis (NTA; NanoSight) and fluorescent dye-binding quantification (RiboGreen), the yield of EVs and exRNA isolated by this technique is relatively low (20–40%) (Supplementary Fig. 1). Furthermore, this procedure yields a highly heterogeneous mix of EVs and RNP/LNP (liponucleoprotein) particles<sup>9,10</sup>. To separate EVs and RNPs according to their physical size and improve the yield of exRNA isolation, we developed a sequential filtration (SF) protocol (Fig. 1a). This protocol offers several advantages over current methods, including low pressure on EVs, better separation between EVs and RNPs, higher RNA yield, and scalability (summarized in Fig. 1b, c and Supplementary Fig. 1). However, the extended hands-on time of the filtration procedure (Supplementary Fig. 2) and the separation solely on the basis of size are limitations of this method. Also, retrieval of EVs from the filtration membranes is inefficient and could potentially alter their structure; therefore, the utility of this protocol for functional EV analysis needs additional evaluation.

Established glioma cell lines have very limited capacity to reflect GBM biology<sup>11</sup>. We utilized previously characterized low-passage GSC cultures derived from four primary human heterogeneous GBM tumors for the exRNA profiling<sup>12,13</sup>. These cells were grown as neurospheres in serum-free medium, to maintain their initial properties and transcriptional profiles and, therefore, better reflect tumor biology<sup>14</sup>. Transmission electron microscopy confirmed the presence of EVs/particles in the corresponding extracellular fractions isolated from GSC cultures (Fig. 2a). Distinct profiles of several protein markers exhibited by cellular and extracellular fractions served to confirm the purity of fractions and the lack of cellular contamination in the MV and other extracellular fractions (Fig. 2b). GSC cultures derived from

different patients varied in the amount of exRNA released (Fig. 2c), ranging between 5.4 and 38.0 ng/ml accumulated in culture over 7 days. Considering that 3–10  $\mu$ g of cellular RNA was isolated from 1 ml of the corresponding GSC neurosphere cultures (~1.3 million cells), between 0.05 and 0.7% of cellular RNA is accumulated in the extracellular space in 7 days. Of note, total exRNA yield varied ~7-fold among the GSC cultures, and the proportion of exRNA associated with different extracellular complexes also varied between the GSC types, suggesting the variations between cultures reflected intrinsic properties of different tumors. Analytical RNA profiles examined by the Agilent 2100 Bioanalyzer indicated high-quality cellular RNA (RIN > 9.5) with sharp rRNA peaks and no sign of degradation (Supplementary Fig. 3). In contrast, exRNA exhibited mostly short RNA profiles (below 200 nt) with intact rRNA peaks detectable in large MVs, but not exosomes and RNPs.

**Technical considerations for RNAseq.** The protocols commonly utilized for small RNA library construction are based on adaptor ligations to the 5'-phosphate and 3'-hydroxy ends of RNA, the modifications characteristic for miRNA, and thus favor miRNA<sup>15</sup>. In order to characterize RNA content in a minimally biased way, we utilized sequential treatments with tobacco acid pyrophosphatase (TAP) and T4 polynucleotide kinase (T4 PNK) to create more uniform 5' and 3' ends for various types of transcripts, leading to their more accurate representation in the cDNA libraries<sup>16,17</sup>. The caveat is that this end-modifying procedure leads to an overwhelming abundance of rRNA reads in cellular and exRNA samples, and reduces the sequencing depth for other RNA classes. Therefore, we included an rRNA depletion step in the protocol that reduced rRNA reads remarkably (Fig. 3a).

Unlike the established strategies for normalization of cellular RNAseq data sets, which utilize total mapped reads as the normalization factor, there are no adequate standard for

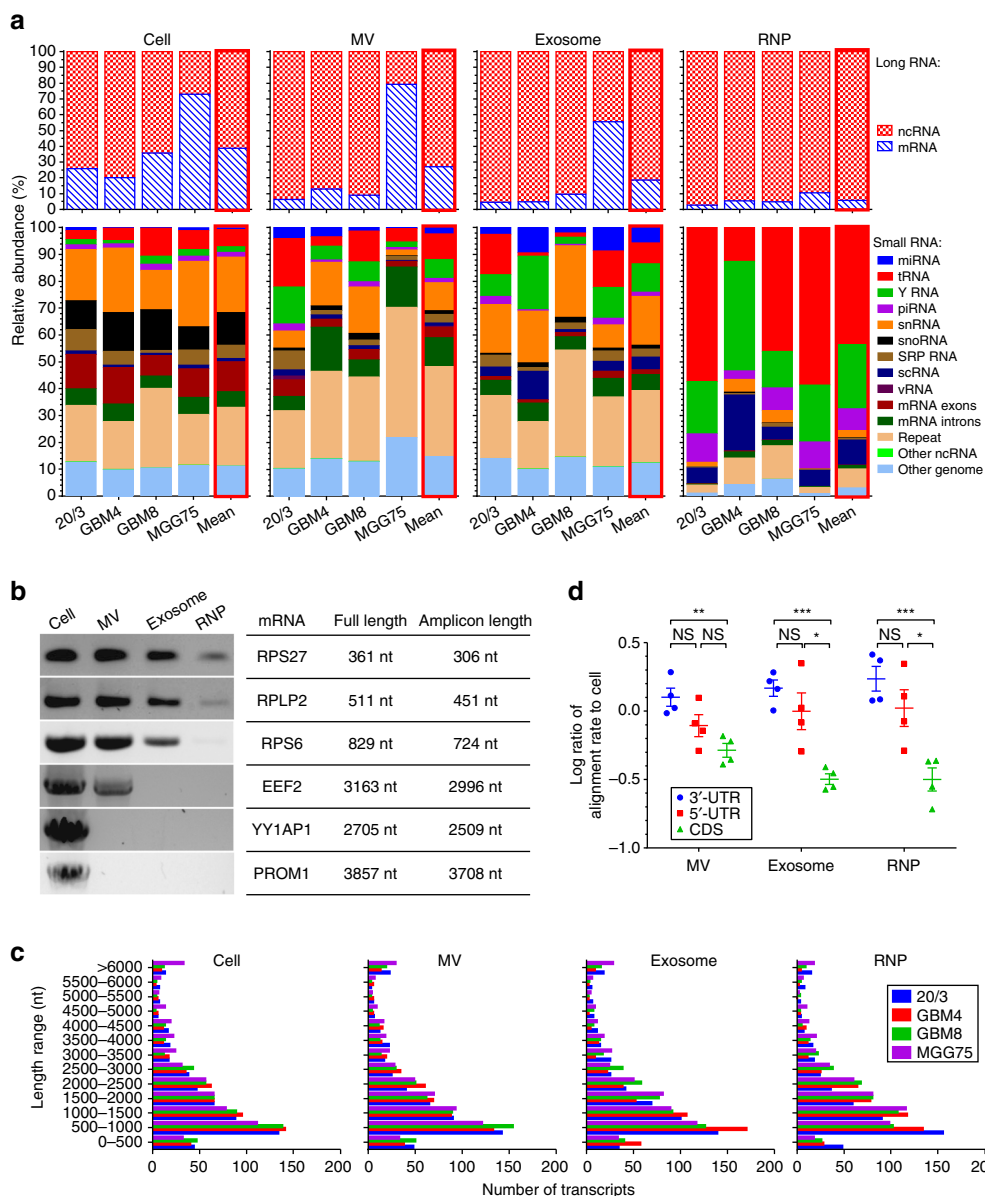


**Fig. 3** rRNA depletion and media correction are warranted for the deep RNAseq analysis. **a** Comparison of the exRNA libraries prepared with and without rRNA depletion. The percent of rRNA reads is shown, indicating that the majority of exRNA reads in non-depleted libraries represent rRNA. Sequencing depth of other exRNA classes is substantially increased by the rRNA depletion. **b** Quantification of total RNA in conditioned and fresh media indicates that the vast majority is cell derived ( $n = 4$  GSC cultures). Nevertheless, quantification of specific RNA species can be skewed by the media, as illustrated in **c**. The levels of miR-122 were assessed in exRNA isolated from the fractions of conditioned and fresh media. miR-122 enrichment in exRNA ( $n = 4$  GSC cultures) was calculated pre- and post-correction to the levels in fresh media, as described in Supplementary Fig. 6. miR-122 was highly enriched in the GSC exRNA, relative to its intracellular level, with up to 1500-fold enrichment in the exosomes before media correction. miR-122 enrichment in GSC exosomes became marginal upon correction. All bars represent mean  $\pm$  SEM

comparative quantitative assessment of cellular vs extracellular RNAs. Since the proportion of total small RNA, and specific types of RNA is vastly different in cellular and extracellular RNA (Fig. 4a and Supplementary Fig. 3), normalization to total mapped reads is not optimal. As recently proposed<sup>18</sup>, we utilized spike-in RNA for normalization, and quantified the abundance of RNA species as fmol per  $\mu$ g of total RNA. Further, since fresh culture media (FM) contain RNA that is co-isolated with cell-derived exRNA<sup>19</sup>, we also assessed the interference of FM RNA with downstream analysis of GSC exRNA. With this goal, we isolated three RNA fractions from the corresponding FM, using the same filtration procedure, and subjected them to RNAseq in parallel with the GSC exRNA. Approximately 1 ng/ml total exRNA was isolated from FM, and it was largely associated with the RNP fraction (Supplementary Fig. 4). Overall, FM contributed 1.3–15% to the exRNA isolated from conditioned media, varying between the fractions (Fig. 3b). Although this amount of contaminating RNA is small, it can still affect the results of exRNA enrichment analysis. For example, miR-122 falsely showed exosomal enrichment (Fig. 3c), consistent with previous reports<sup>20, 21</sup>, due to its abundance in B-27 supplement (Supplementary Fig. 5). Based on this observation, we included FM RNA data set in our RNAseq analysis pipeline, to provide the baseline for GSC-derived exRNA (Supplementary Fig. 6). The results described below were obtained using the media correction.

**EVs and RNPs exhibit distinct RNA composition.** To characterize the repertoire of GSC cellular and extracellular RNA, we sequenced the libraries of small and long RNAs, and first normalized the number of reads for each RNA class to the total number of non-rRNA reads within the library, thereby removing the confounding factor of variable rRNA depletion efficiencies. All reads generated on the long RNA libraries (that generally included transcripts longer than 100 nt), were classified as either

mRNA or non-coding RNA. Although in three out of four GSC cultures, cellular mRNA reads accounted for 20–35% of long RNA libraries, the mRNA proportion in EV fractions was below 10%, and even lower in RNPs (Fig. 4a). The reads obtained from small RNA libraries were first mapped to the most accurately annotated miRNA database (miRBase), and subsequently to other databases with equal mapping priority. In total, all annotated RNA species were categorized into 14 classes (Fig. 4a). Since small RNA libraries were built on 15–65 nt transcripts, the vast majority of the mapped reads represent fragments rather than full-length transcripts, with exception of miRNA and piRNA reads. Despite the remarkable heterogeneity of the GSC cultures, different extracellular fractions exhibited common characteristics of their RNA repertoires. Some of the most distinct features of GSC-derived exRNA are summarized as follows: (1) mRNA exons and snoRNAs are depleted, compared to cellular RNA, in all extracellular fractions; (2) all extracellular fractions, especially non-vesicular RNPs, are highly enriched in specific Y RNA fragments of largely unknown functions; (3) MVs and exosomes differ in their RNA composition, with mRNAs being relatively more enriched in MVs and miRNAs in exosomes; (4) RNP fractions have a highly distinctive RNA repertoire, with tRNA and Y RNA fragments strongly enriched, and snRNA and repeats reduced. The predominance of tRNA and Y RNA fragments in RNP is reflected in the corresponding sharp ~32 nt peak observed in the length distribution profile of reads (Supplementary Fig. 7). The relative abundance of piRNA and scRNA (small cytoplasmic RNA) fragments was also higher in RNP fractions. However, the most abundant individual piRNA and scRNA species were identical or highly homologous to major tRNA and Y RNA fragments, respectively (Supplementary Fig. 8). Whether such identical sequences indeed belong to two functionally distinct classes of transcripts, or tRNA and Y RNA are commonly mis-identified due to the poor quality of databases, is unknown. In addition, recently discovered circular RNAs have been reported as



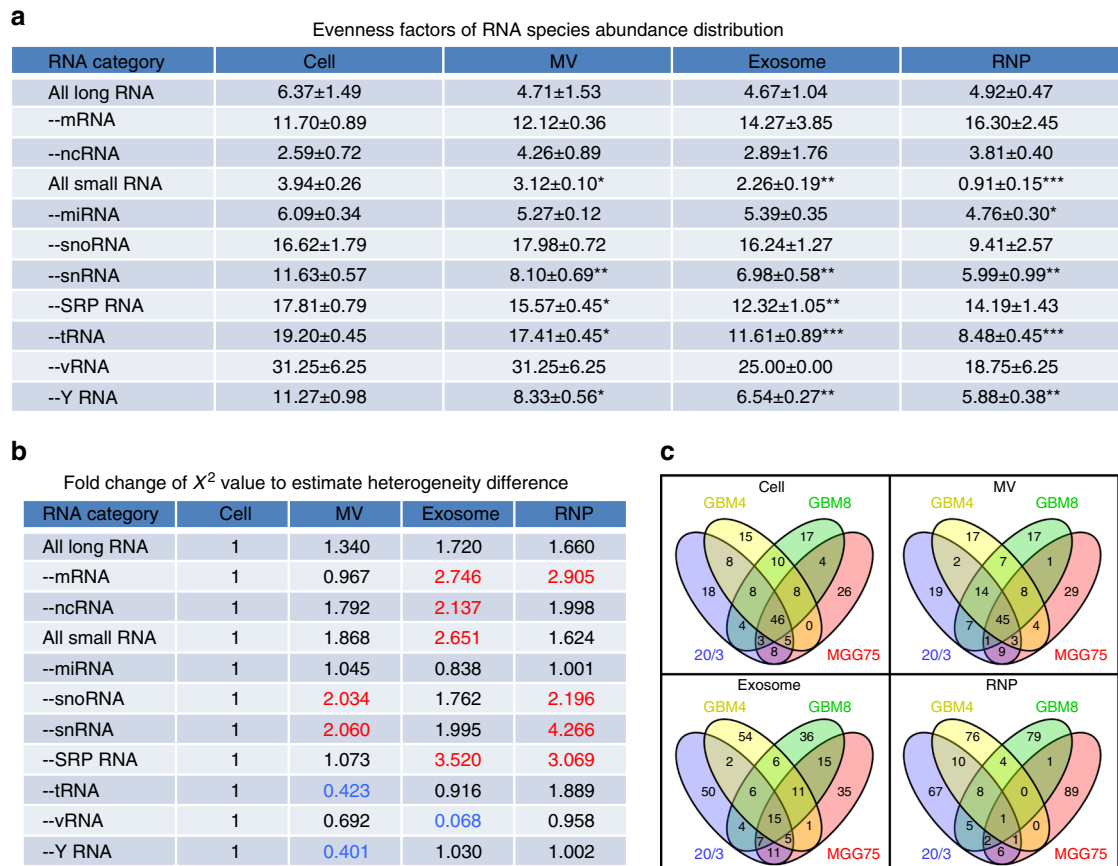
**Fig. 4** Relative composition of diverse RNA classes in cellular and extracellular compartments (MVs, exosomes, and RNPs) in different GSC cultures. **a** The top panels exhibit relative RNA composition in long RNA libraries, and the bottom panels depict the composition in small RNA libraries. The data were normalized to the total number of annotated non-rRNA reads. The bars framed in red represent the mean values of four GSC cultures.

**b** RT-PCR analysis (with equal input of total RNA) of selected mRNAs abundant in exRNA, demonstrates the presence of nearly full-length short, but not long messages in the extracellular fractions. Long RT-PCRs were replicated twice. **c** Long RNA libraries-based analysis of the length distribution of 500 most abundant mRNAs suggests no length preference for shorter parent transcripts in the extracellular fractions. **d** Analysis of the RNAseq reads mapped to mRNAs indicates that UTR regions were more enriched than CDS regions in the extracellular fractions ( $n = 4$  GSC cultures). mRNA reads were aligned to the coding sequences (CDS), 5'-UTRs, and 3'-UTRs separately, and the alignment rates for each extracellular fraction were normalized to the corresponding cellular fraction. The log-transformed ratios of the alignment rates were compared among the three regions. Error bars represent mean  $\pm$  SEM. NS, not significant; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; t-test

enriched in exRNA<sup>22</sup>, and although our RNAseq protocol was not optimized for their quantification, several circular RNA species were detected (Supplementary Table 1).

**UTR fragments are released more than ORFs and intact mRNAs.** The length of intact cellular mRNAs varies between 350 and 12,000 nt, with the average around 2000 nt<sup>23</sup>. Considering the small size of EVs, the maximum length of an mRNA that could be packaged is still an open question. Previous high-throughput studies have not discriminated between full-length mRNAs and

fragments in exRNA<sup>24, 25</sup>. On average, our RNAseq detected 17,148 intracellular mRNAs in GSCs, whereas 17,219, 11,592, and 11,819 mRNAs were detected in MV, exosome, and RNP fractions, respectively. Many of the most abundant extracellular mRNAs corresponded to relatively short transcripts, including mRNAs for various ribosomal proteins (Supplementary Data 1). However, some abundant reads in EVs corresponded to long mRNA transcripts, such as PROM1 and EEF2. To examine whether intact mRNAs are present in extracellular fractions, we designed PCR primers for selected mRNAs to provide their full-length amplification. As shown in Fig. 4b, near-complete short

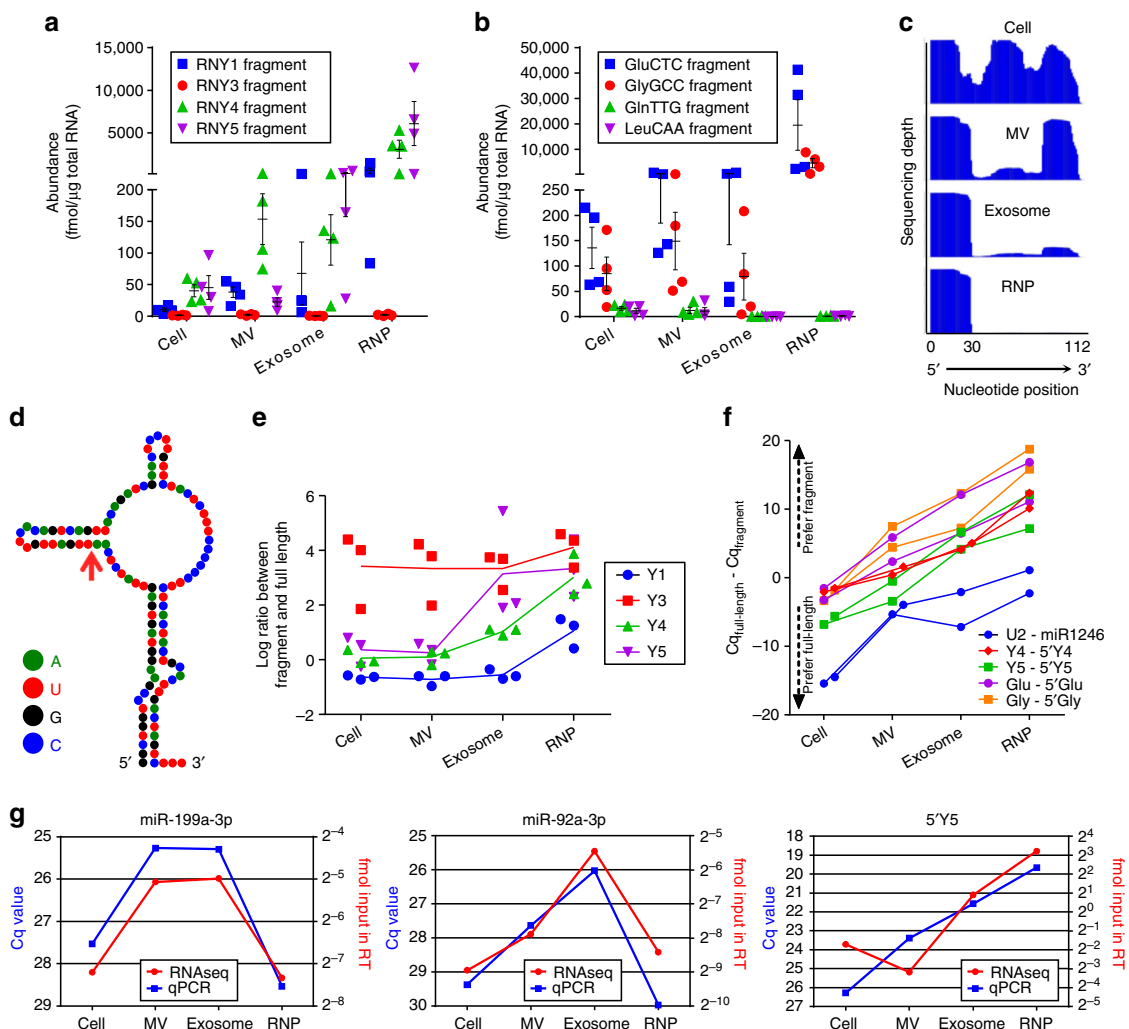


**Fig. 5** Inequality and heterogeneity of the RNA repertoire of extracellular fractions. **a** Evenness factors, reflecting the inequality of abundance' distribution of the indicated RNA classes in various fractions. Higher evenness factors correspond to lower inequality. The classes of tRNA, Y RNA, snRNA, and SRP RNA, but not miRNA, contributed to the increased inequality of exRNA most significantly. Differential evenness factors of cellular and extracellular RNA suggest the selectivity of secretion ( $n = 4$  GSC cultures). **b** For each RNA category, a sum of squared errors ( $\chi^2$  value) was calculated among four GBM cultures, after normalization of each RNA species to the total number of reads in that RNA category. The  $\chi^2$  value of each extracellular fraction was compared to the cellular fraction. Fold change of  $\chi^2$  values higher than 1 reflects the increased heterogeneity. Heterogeneity either increased or decreased more than two-fold is highlighted in red and blue, respectively. **c** Venn diagrams depict the number of common species among 100 top abundant mRNAs, in four GBM cultures and their extracellular fractions, supporting the observation of higher heterogeneity of mRNA composition in exRNA than in cellular RNA. Significant differences between the cellular and extracellular fractions are depicted as following: \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ;  $t$ -test

mRNAs could be detected in cellular and all extracellular fractions, but detection of long mRNAs above 1000 nt was limited to cells and MVs only. These results suggest that either long mRNAs are excluded from packaging into exosomes and RNP complexes, or they are present in these fractions only as fragments. The latter appears to be the case, since the RNAseq demonstrates a similar representation of transcripts of various lengths in the intracellular and extracellular compartments (Fig. 4c). These data suggest that most exRNA reads corresponding to long mRNAs represent fragmented transcripts. Nevertheless, amplifying long RNA from low-input exRNA fractions is technically challenging and cannot be performed in a high-throughput manner, so we cannot exclude that some long full-length mRNAs are present extracellularly. Next, mRNA reads were aligned separately to the coding sequences (CDS), 5'-UTRs, and 3'-UTRs. The UTR regions, and especially 3'-UTRs, were significantly enriched in all extracellular fractions relative to CDS sequences (Fig. 4d), validating the previous observation<sup>26</sup> and suggesting differential release pathways for the protein-coding and regulatory sequences.

**Inequality of RNA representation increases in exRNA.** A relatively small number of the most abundant miRNA species

usually account for the majority of total miRNA in a given cellular context<sup>27</sup>. Consistent with this observation, 31 miRNA species accounted for 80% of the total miRNA in GBM8 cells, indicative of a diverse range of miRNA expression (or “inequality” of miRNA levels). Even fewer, 19 miRNA species, accounted for 80% of the total miRNome in GBM8 exosomes, suggesting a higher inequality of miRNAs in exRNA. Such comparison, however, relies on a randomly selected cutoff (80% used above). To compare the inequality of intracellular and extracellular transcripts more objectively, we developed two alternative strategies. The first one is an improved version of traditional evaluation named the evenness factor ( $\epsilon$ ), which defines that  $\epsilon\%$  of RNA species can account for  $(100-\epsilon)\%$  of total abundance. The second strategy is based on the Gini coefficient, commonly used for evaluation of inequality in economics. Of note, although both are objective and use no preset cutoff, higher  $\epsilon$  and lower Gini coefficient correspond to a more equal representation (or lower inequality). The detailed illustration of the concepts behind these evaluation systems can be found in Supplementary Fig. 9. Taking advantage of these parameters, we compared transcript inequality in cellular and extracellular fractions, either for the whole RNA library or a specific RNA class. The inequality of RNA levels was similar for cellular and extracellular long RNA libraries, but



**Fig. 6** Y RNA and tRNA fragments are abundant and enriched in exRNA. **a, b** Specific Y RNA and tRNA species are among the most abundant in small RNA libraries ( $n = 4$  GSC cultures). The reads corresponding to Y1, Y4, and Y5 are highly enriched in extracellular fractions, especially in RNP (**a**). The reads corresponding to specific tRNAs, such as GluCTC and GlyGCC, are highly enriched in exRNA, while others (e.g. GlnTTG and LeuCAA) are not (**b**). **c** Mapping coverage of Y1 reads indicates that the Y1 is precisely processed, and mostly its 5' fragment is present in exosomes and RNPs, as evidenced by the steep peaks corresponding to the 5'-end 30 nt. These profiles are distinct from the more uniform full-length coverage observed for the cellular RNA, and to a lesser extent MV RNA. Similar analysis for other Y RNAs and tRNAs is presented in Supplementary Fig. 12. **d** The predicted secondary structure of Y1 RNA<sup>30</sup>, Copyright (1993) National Academy of Sciences, USA, and the position of its cleavage (indicated by the arrow) that produces the 5' fragment which is highly abundant in exRNA. **e** Quantification of Y RNA reads in long and small RNA libraries, demonstrates different fragment to full-length ratios in the cell and exRNA fractions ( $n = 3$  GSC cultures; MGG75 was excluded due to very low abundance of the full-length Y RNA). The ratios are increased in extracellular fractions. **f** qRT-PCR analysis with primers specific to either full-lengths or fragments of several RNA species validates the enrichment of 5' tRNA and Y RNA fragments in extracellular fractions. For each specific transcript examined, two lines represent GBM8 and 20/3 cells, respectively. **g** qRT-PCR analysis of selected transcripts confirms the quantitative character of the RNAseq pipeline. The blue dots represent qRT-PCR Cq values, while the red dots represent the results of RNAseq quantification in fmol. Both analyses were performed on the same set of RNA samples. Error bars represent mean  $\pm$  SEM

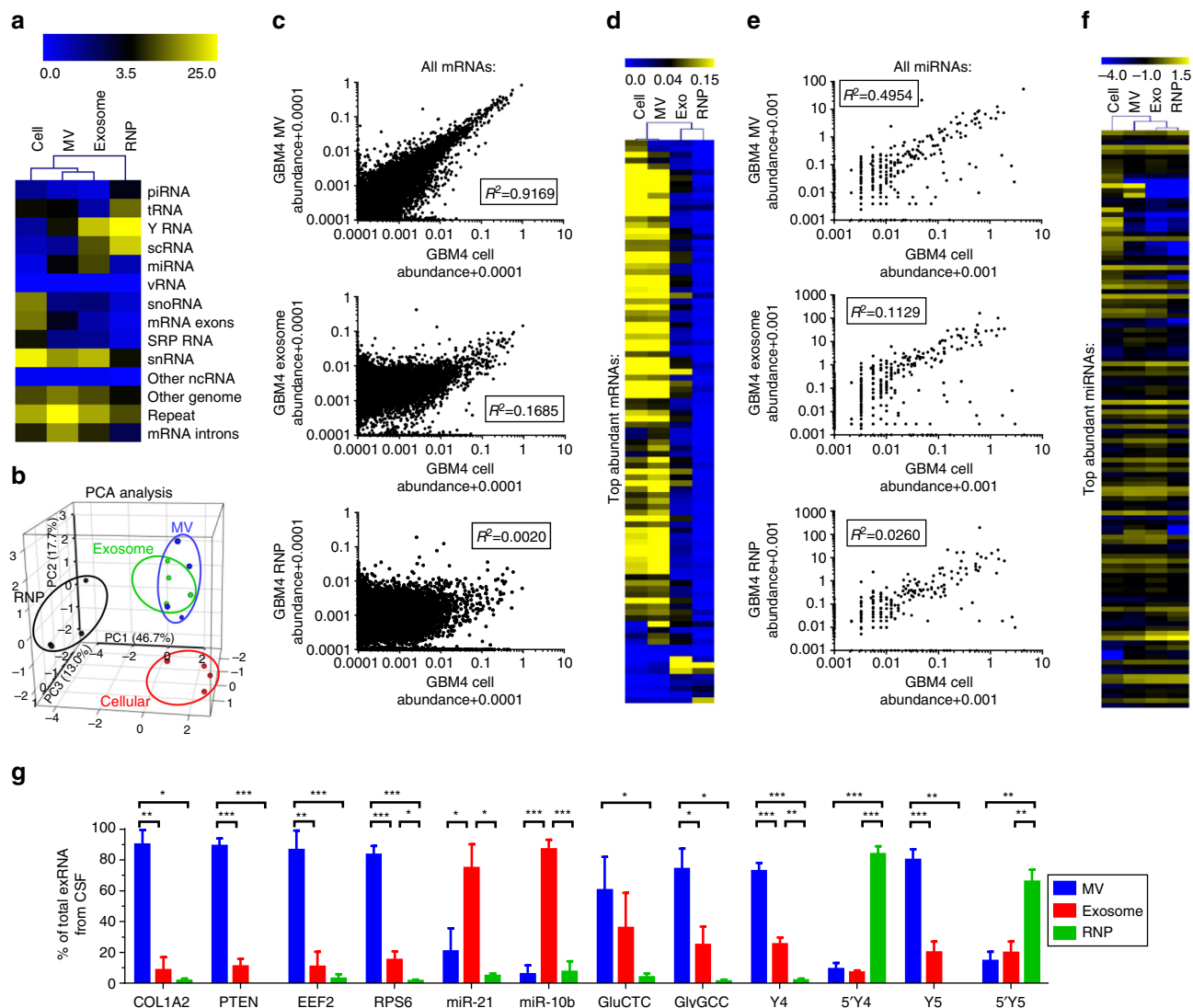
increased significantly in small exRNA libraries (Fig. 5a and Supplementary Table 2). These results provide support for the idea of selective RNA incorporation into different exRNA fractions, that is still highly debated in the field.

**Heterogeneity of RNA repertoire increases in exRNA fractions.**

Despite the genetic diversity of GSCs, cellular RNA class composition was similar among the four different GSC cultures analyzed, with the exception of MGG75 cells that expressed more mRNA than long ncRNA (Fig. 4a). However, the RNA repertoire of extracellular fractions was more heterogeneous among GSC cultures than that of cells. To estimate the heterogeneity of

cellular and extracellular RNA species across the GSC cultures, we first normalized the reads to the total read number within an individual RNA category, and then evaluated the sum of squared errors ( $\chi^2$  value). The higher  $\chi^2$  value reflects the higher diversity/heterogeneity of GSC cultures in terms of RNA composition. As shown in Fig. 5b, relative to cellular RNA, the extracellular fractions were, overall, more heterogeneous in their composition in both long and small RNA libraries, as well as for the majority of specific RNA classes. This phenomenon was not caused by technical irreproducibility of the exRNA analysis, because independently analyzed exosomes produced by different passages of GBM8 cells were much more concordant than the pairwise-compared exosomes released by GSC cultures established from



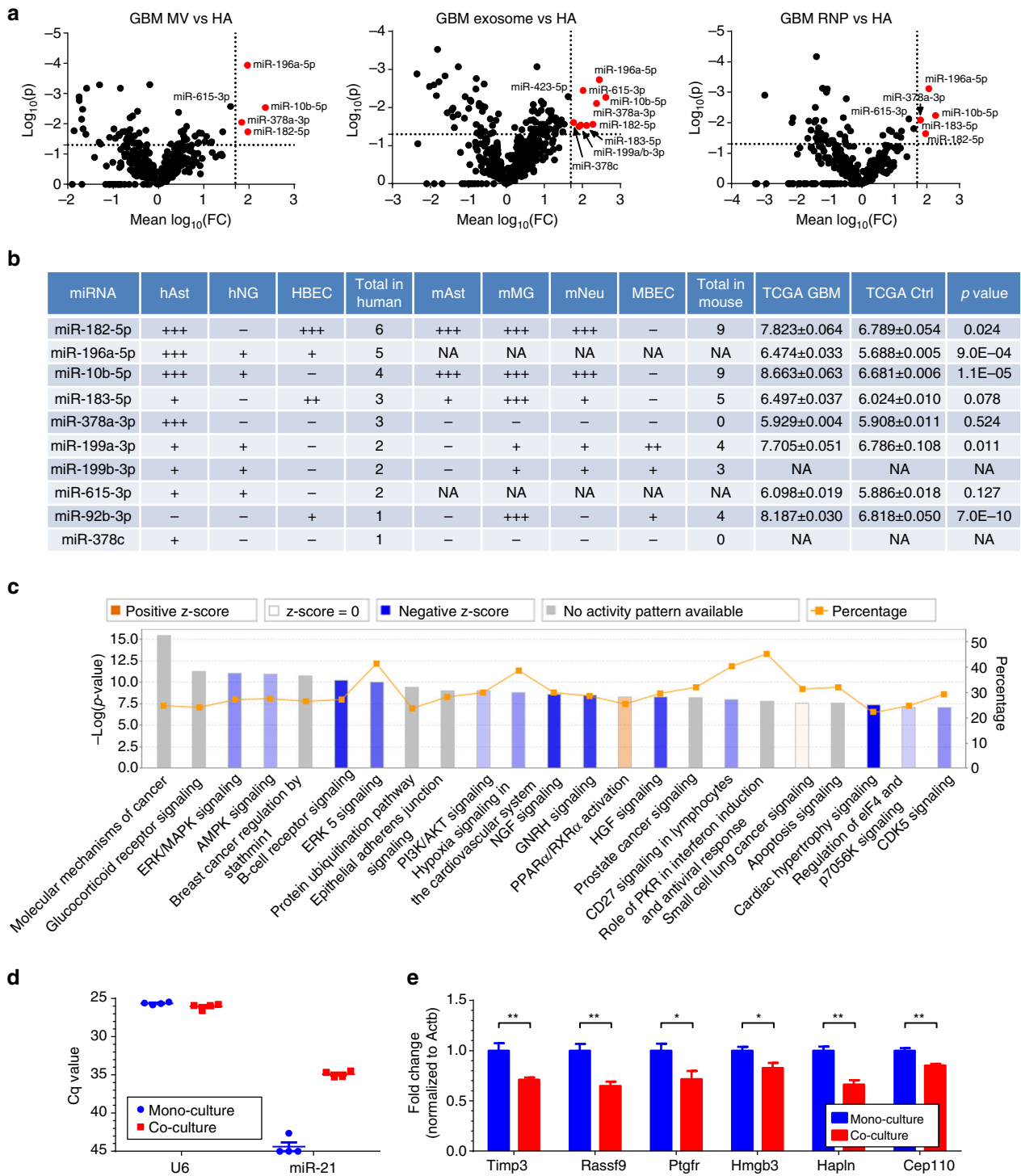


**Fig. 7** RNA repertoire of MV most closely reflects cellular RNA composition. **a** Heat map cluster analysis of RNA classes indicates relative similarity of the composition of MVs and the source cells. The scale bar represents the percentage of non-rRNA annotated reads. **b** PCA analysis of RNA classes. Different fractions are marked in different colors. Within each fraction, four dots represent four GSC cultures. **c, d** GBM4 MV represents the cellular mRNA content closely, and much better than exosomes or RNPs, based on the correlation analysis of all mRNA species (**c**), and cluster analysis of the top abundant mRNAs (**d**). The scale bar represents mRNA abundance. **e, f** Extracellular miRNA composition, in general, is less reflective of the cellular miRNome; nevertheless, MV fraction still remains the best simulator, based on the correlation analysis of all miRNA species (**e**), and cluster analysis of the top abundant miRNAs (**f**). The scale bar represents the log-transformed miRNA abundance. Similar analyses of other cell cultures can be found in Supplementary Figs. 14–17. **g** qRT-PCR analysis of exRNA fractions isolated from the CSF of GBM patients, using the same filtration-based procedure, indicates its applicability to clinical samples. The data verify preferential association of selected RNA species with different exRNA fractions in the human biofluid ( $n = 4$  CSF samples). All bars represent mean  $\pm$  SEM. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ;  $t$ -test

different patients (Supplementary Fig. 10). Consistently, Venn diagram analysis showed less commonalities (or increased heterogeneity) for mRNAs in exosomes and RNPs, relative to cellular RNA (Fig. 5c). The heterogeneity of miRNA class, however, was similar between GSC cellular and exRNA compartments (Supplementary Fig. 11). Combined analyses of inequality and heterogeneity suggest that GSC cultures may utilize various sorting mechanisms for exRNA release.

**Y RNA and tRNA fragments are abundant and enriched in exRNA.** The reads corresponding to tRNA and Y RNA species constitute a significant proportion of the rRNA-depleted small exRNA libraries (8.5% in cells, 13.5% in EVs, and 67.5% in RNPs, Fig. 4a). All four human Y RNA (Y1, Y3, Y4, and Y5) and some

tRNA species are highly abundant in extracellular fractions, especially the RNP, reaching the quantities of up to dozens of pmol per  $\mu\text{g}$  of total RNA (Fig. 6a, b). Relative to let-7b-5p, one of the most abundant miRNAs in exRNA, those molecules are at least hundred times more abundant. Since human Y RNAs are 84–113 nt and tRNAs are 68–176 nt, and our libraries were constructed from 15 to 65 nt RNA, we reasoned that the reads represent fragments rather than full-length transcripts. Notably, the coverage analysis of the reads mapped to Y RNAs and tRNAs suggested the presence of specifically processed fragments (Fig. 6c and Supplementary Fig. 12), with the processing sites located within the loop domains that are known to bind several proteins<sup>28–30</sup> (Fig. 6d and Supplementary Fig. 12). Further integration of the small and long RNA data sets revealed that the ratios of fragment to full-length Y RNAs differed significantly among



**Fig. 8** Comparative analysis of GSC-secreted miRNAs and cellular miRNomes of the normal cells of the brain predicts the most impactful GBM miRNAs in tumor-to-microenvironment communication. **a** Extracellular fractions of GSC cultures were compared to primary human astrocytes (HA), based on the corresponding RNAseq data sets. The fold-changes in miRNA levels were log transformed and a *t*-test was applied to examine the significance of difference. MiRNAs with log-fold changes higher than 1.7, which corresponded to 50 times higher levels in the GSC-derived exRNA fractions relative to the recipient cells, and  $p < 0.05$  (*t*-test), were defined as potentially impactful (colored in red). The horizontal axis of Volcano plots shows the log-fold difference, and the vertical axis shows the statistical significance. Similar analyses of primary human neuroglial and endothelial cells are shown in Supplementary Fig. 18. **b** A full list of most impactful GSC miRNAs for human and mouse astrocytes, neurons, microglia, or brain endothelial cells. The number of “+” symbols reflects the number of extracellular fractions in which an miRNA meets the indicated criteria as in **a**. Most of these miRNAs are also upregulated in the GBM tumors compared to non-neoplastic brain tissues in the TCGA microarray data set, as indicated in the three right columns ( $n = 496$  GBM vs 10 control). **c** Top enriched IPA pathways for the validated mRNA targets of the most impactful miRNAs. Predicted activation and inhibition of pathways are labeled as orange and blue bars, respectively. The yellow line shows the percent of genes in each pathway that are validated targets. **d, e** Co-cultured with GBM8 neurospheres, primary miR-21-null astrocytes exhibit steady miR-21 levels (**d**) and downregulation of validated miR-21 targets (**e**). Cq value of miR-21 in mono-cultures was defined as 45 (undetectable expression).  $N = 4$  wells in 24-well plate. All bars represent mean  $\pm$  SEM. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; *t*-test

the Y RNA species (Fig. 6e). Overall, the fragment to full-length ratio was significantly higher in extracellular vs cellular fractions for all four Y RNA species (Fig. 6e), indicating that whereas full-length Y RNAs are highly abundant in the cells and MVs, specific fragments of them are enriched in the exosomes and RNPs. These data were further confirmed by qRT-PCRs (Fig. 6f). Moreover, such preference for fragments in the exosomes and RNPs was not limited to Y RNA. For example, full-length U2 snRNA (RNU2-1), which also serves as a precursor for miR-1246<sup>31</sup>, and specific tRNAs are also released to a lesser extent in extracellular fractions than their corresponding processed products (Fig. 6f). As indicated by the experiment with unfractionated exRNA, fragmentation is a feature of exRNA rather than an artifact of the filtration procedure (Supplementary Fig. 13). Of note, qRT-PCRs quantification of the selected transcripts correlated well with the RNAseq results (Fig. 6g), supporting the accuracy of our RNAseq analysis.

**MVs most closely reflect cellular RNA composition.** Cancer-derived exRNA may serve as clinical biomarkers for disease diagnostics, prognostics and monitoring. An extracellular fraction that closely mirrors the deranged cellular transcriptome would be the most valuable for such applications. To date, mostly unfractionated, precipitated or 100,000 g-pelleted EVs have been explored as a source for potential biomarkers. To examine what type of extracellular complexes (MVs, exosomes, or RNPs) might serve as the closest proxy of the GSC transcriptome, we performed correlation and clustering analysis of their RNA composition. Overall, clustering analysis based on small RNA libraries demonstrated the highest similarity between the cellular transcriptome and MV content, and to a lesser extent to exosome content, while the RNP fraction had a highly distinct RNA composition (shown for GBM4 in Fig. 7a). This was further confirmed by principal component analysis (PCA) of all four GSC cultures (Fig. 7b). Similar analyses of mRNA and miRNA classes, and the most abundant transcripts within them, further supported this conclusion (Fig. 7c–f and Supplementary Figs. 14–17). Consistent with this idea, there were fewer miRNA species, and overall fewer different RNA species significantly enriched in MVs than in exosomes and RNPs (Supplementary Data 2). The observation that MVs provide a more accurate peripheral read-out of the source cell content is in line with recent extracellular proteomic analysis<sup>32</sup>. Therefore, MVs which include large vesicles (200–800 nm) appear to be a good source for RNA biomarker discovery, and have the potential to outperform the more studied smaller exosomes.

**Analysis of exRNA complexes in cerebrospinal fluid.** To investigate the differential potential of exRNA fractions for biomarker discovery, we tested our sequential filtration method on cerebrospinal fluid (CSF) samples obtained from four GBM patients. As shown in Fig. 7g, sufficient amounts of total RNA for qRT-PCR analysis were isolated from each CSF fraction (MVs, exosomes, and RNPs), and distinct profiles observed for different RNA classes. Although we could not examine the resemblance of CSF exRNA profiles to those of parental tumor and neural cells due to unavailability of tumor material from the corresponding patients, the data indicated that, indeed, MVs and exosomes should be analyzed separately in biomarker discovery studies. Similarly to the results obtained on cultured GSC-derived exRNA, the CSF mRNAs were preferentially associated with MVs and miRNAs with exosomes, suggesting that MV enrichment is warranted for mRNA biomarkers, while exosomes represent a superior source for miRNA biomarkers. Of note, several mRNAs frequently mutated in GBM, including PTEN and COL1A2, were

detected in CSF MVs. The preference for full-length Y RNAs in CSF MVs, and their fragments in RNPs was also consistent with the observations made on GSC-derived exRNA.

**Exosomal miRNAs with high functional potential upon transfer.** A common, but still highly debated, hypothesis is that EVs and EV-associated exRNA are taken up by recipient cells and mediate intercellular communication. MiRNA transferred from GBM to cells in their microenvironment could have significant impact on the transcriptome of the peritumoral cells. We reasoned that the greatest impact would be mediated by those miRNAs that are most abundant in EVs/exRNA, while not or minimally expressed in recipient cells, such that their transfer would significantly alter their levels in recipient cells and thus affect the regulation of mRNA targets. To identify such miRNAs, we compared relative levels of all miRNAs in exRNA fractions from GSCs with those in the major types of human and mouse brain cells, including astrocytes, mature neurons, microglia, and brain-derived endothelial cells. Comparison of three GSC extracellular fractions with primary astrocytes is shown in Fig. 8a, and similar analysis for other cell types is presented in Supplementary Fig. 18. Summary of this analysis, performed on three extracellular fractions and four types of recipient cells, revealed a list of the GBM-derived miRNAs that might potentially have the strongest impact on the normal cells of brain microenvironment (Fig. 8b). In agreement with these data, most of the listed miRNAs were also found to be elevated in GBMs relative to the control brain tissues in The Cancer Genome Atlas (TCGA) data set (Fig. 8b). To predict the downstream effects of the 10 most impactful miRNAs on the recipient cells, we analyzed their direct mRNA targets previously validated by at least three supporting CLIP-Seq data sets in the starBase database<sup>33</sup>. In total, 2267 mRNA species interact with at least one of the impactful miRNAs. Based on the ingenuity pathway analysis (IPA), these targets are significantly enriched in many canonical cancer-related pathways and bioterms, including the molecular mechanism of cancer, ERK/MAPK signaling, PI3K/AKT signaling, and NGF signaling (Fig. 8c).

**Less than one copy per EV on average for most RNA species.** A substantial amount of transferred exRNA might be required to exert a functional effect in a recipient cell; however, the quantitative data for the levels of specific RNA classes and individual transcripts in exRNAs is very limited thus far. To address this issue, we performed a stoichiometric analysis of the EV-associated RNA. On average, GSC EVs (MVs and exosomes collectively) provided ~8.9 ng total RNA per ml of conditioned media. Considering the concentration  $2 \times 10^9$  EVs per ml, measured by NTA, one EV contained ~4.45 ag total RNA, or ~0.445 ag non-rRNA, which corresponded to ~836 ribonucleotides. Similar analysis was performed for individual RNA classes and species. As shown in Table 1, rRNA and snRNA are present in more than one copy per EV on average. RNA classes represented by approximately one copy per EV include Y RNA and Y RNA fragments, tRNA fragments, lncRNA, and miRNA. One copy of mRNA or mRNA fragment can be found in ~10 EVs. Of note, quantification of individual RNA species revealed that only several specific molecules are present at the level of one copy per EV (e.g., RNY1, Y5 fragments, GluCTC tRNA fragment, U1 and U2 snRNA fragments). The most abundant miRNA species (e.g., let-7b, miR-21) were present at the level of one copy per 10 EVs approximately, consistent with the previous report<sup>34</sup>. The most abundant individual mRNA species were present at one copy per 1000 EVs approximately (e.g., TMSB10), and the levels of mRNAs most commonly mutated in GBM (e.g., EGFR, IDH1,

TP53, PTEN, and COL1A2) were not higher than one copy per 100,000 EVs approximately. Based on these data, continuous barrage, massive and/or highly selective uptake of EVs might be

required to affect signaling or alter the phenotype of the recipient cells via their exRNA content.

To examine the functionality of transferred GSC-derived miRNAs, we established 3D transwell co-cultures of GBM8 neurospheres with primary mouse astrocytes, a system more physiological than the commonly used exposure of recipient cultures to super-concentrated EVs isolated from donor cell conditioned medium. In line with the low copy number of individual miRNAs per EV, we were unable to detect significant elevation of either miR-21 or miR-10b (top GBM-promoting miRNAs) in astrocytes co-cultured with GBM8. However, miR-21 became readily detectable in primary astrocytes established from miR-21 knockout mice upon their co-culture with GBM8 (Fig. 8d). Furthermore, this latter miR-21 transfer led to functional effects in these cells, i.e., repression of its previously validated mRNA targets (Fig. 8e)<sup>35</sup>. These data support intercellular transfer of miRNA via exRNA and suggest that, despite the apparently low levels in EVs, miRNAs may exert regulatory functions in recipient cells, albeit with specific conditions and highly sensitive methodologies required for detection.

**Table 1 Quantification of selected RNA classes and individual species in copy number per EV**

Copy level	RNA class or individual species	
More than 1 copy per EV	RNA class	rRNA fragment <sup>a</sup> Repeat fragment snRNA fragment tRNA fragment
One copy per 1 EV approximately	RNA class	Intron fragment Y RNA fragment SRP RNA fragment miRNA Exon fragment lncRNA Y RNA snoRNA fragment U2 snRNA fragment
	Individual RNA species	U1 snRNA fragment GluCTC tRNA fragment RNY1 RNY5 fragment mRNA
One copy per 10 EVs approximately	RNA class	vRNA vRNA fragment RNY4 fragment
	Individual RNA species	RNY4 RNY1 fragment let-7b-5p GlyCCC tRNA fragment miR-21-5p vRNA1-2 vRNA1-2 fragment U3 snoRNA fragment
One copy per 100 EVs approximately	Individual RNA species	miR-1246 miR-10b-5p RNY5 vRNA1-1 fragment vRNA1-3 fragment vRNA1-3 RNY3 fragment
One copy per 1000 EVs approximately	Individual RNA species	TMSB10 mRNA (top 1 mRNA) miR-93-5p ACTB mRNA
One copy per 10,000 EVs approximately	Individual RNA species	GAPDH mRNA miR-132-3p COL1A2 mRNA
One copy per 100,000 EVs approximately	Individual RNA species	RNY3 miR-34a-5p IDH1 mRNA EGFR mRNA
1 copy per 1,000,000 EVs approximately	Individual RNA species	TP53 mRNA PTEN mRNA

<sup>a</sup>The transcripts were defined as fragments if present and quantified in small RNA libraries

## Discussion

exRNA studies have expanded rapidly and multidirectionally in this decade. Revealing the composition of exRNA complexes released by defined cell types remains one of the most fundamental milestones toward understanding the role of exRNA in intercellular communication, as well as discovery of RNA biomarkers for disease. This study provides the first minimally biased quantitative analysis of the exRNA released by tumor-derived cells from GBM patients. Despite the high heterogeneity of GSC cultures established from different tumors, and even higher heterogeneity of the exRNA they release, extracellular complexes share key characteristics. The RNA profiles of MVs, exosomes, and RNPs are highly distinct. They all display selectivity in their RNA loading compared with the cells, with the MVs being the most like the cells and the RNPs the least. Small and fragmented RNAs account for the majority of exRNA in all three types of complexes. Between 64 and 93% of all exRNA consists of fragmented rRNA (Fig. 3a), although there is little-to-no intact rRNA (Supplementary Fig. 3). miRNA, the most studied class of exRNA, constitutes <10% of non-rRNA exRNA (Fig. 4a). Consistent with some previous publications<sup>36</sup>, and in contrast to others<sup>20, 37–41</sup>, we found that miRNA species are relatively enriched in exosomes, but not in MVs or RNPs. Of note, the isolation approaches, the library preparation strategies, and the normalization methods utilized in these different studies were diverse, making the results not directly comparable. Other non-coding RNA species are more abundant, and some of them are enriched in EVs or RNPs. Notable among them are precisely processed tRNA and Y RNA fragments, associated with both EVs and extravesicular RNPs, and supported by observations of other cultured cells<sup>20, 37</sup> and body fluids<sup>42</sup>.

The most common extracellular 30–32 nt-long 5'-tRNA fragments (5'-tRFs), also called 5'-tRNA halves or tiRNAs, are evolutionarily conserved molecules<sup>43</sup> produced by angiogenin (ANG), a member of ribonuclease A family<sup>44</sup>. This multifunctional ribonuclease regulates angiogenesis, cell proliferation and viability of cancer cells, as well as neuronal survival and stress response<sup>43</sup>. Specific 5'-tRFs are known to perform crucial functions, often associated with regulation of gene expression in stress response. They repress protein translation by displacing eukaryotic translation initiation factors eIF4E and eIF4G<sup>45</sup>, and modulate stress response by inducing formation of stress granules—cytoplasmic foci where untranslated mRNAs are transiently stored<sup>44</sup>. Importantly, specific 5'-tRFs (e.g., 5'Ala,

5'His, and 5'Cys) selectively regulate translation of subsets of mRNAs, both capped and uncapped, and therefore reprogram protein synthesis<sup>45</sup>. They may also function in a miRNA-like manner<sup>46</sup>. ANG is upregulated in GBM<sup>47, 48</sup> and is one of the proteins most highly secreted by glioma cells<sup>49</sup>. Despite its high abundance, functions of ANG, including its extracellular activity and its exRNA products remain to be investigated. Remarkable enrichment of both ANG and 5'-tRFs observed in GSC-derived exosomes (Figs. 2b and 6f) suggests that tRNA cleavage may occur in exosomes, outside of the cells.

Another highly abundant but poorly studied classes of non-coding exRNAs are Y RNAs and their specific 5' fragments. The majority of information about Y RNA has come from studies of bacteria and invertebrates. In vertebrates, Y RNAs are expressed in all tissues and cells and have been proposed to participate in many important cellular functions<sup>50, 51</sup>. Their lower stem domain, which recruits chaperone Ro60 and exoribonuclease PNPase, may play a role in RNA quality control and degradation of misfolded RNAs<sup>50</sup>. The upper stem domain may participate in initiation of chromosomal DNA replication<sup>52</sup>. With biogenesis and activity independent of Dicer and Ago2<sup>53</sup>, Y RNA fragments do not appear to silence gene expression in a miRNA-like manner<sup>54</sup>. Recent studies reported that Y RNA fragments may be involved in histone mRNA processing<sup>55</sup> and cell damage<sup>56</sup>. Interestingly, Y5 fragment, the most abundant in exRNA, is proposed to specialize in surveillance of ribosomal RNA<sup>57</sup>. Studying the functions of these precisely processed, highly enriched extracellular transcripts represents an exciting new avenue in RNA biology.

In addition to non-coding RNA, we detected low levels of mRNA reads with UTRs relatively enriched compared to the coding regions in exRNA. The mechanism underlying this enrichment and its biological impact has yet to be investigated. Methylation and other RNA modifications of UTRs<sup>58, 59</sup> might cause differential sorting or stability, leading to enrichment in exRNAs. UTRs and their fragments might function as molecular sponges for various regulatory molecules, including miRNA, translation factors (e.g., eIF4F and ribosomal complexes), other RNA-binding proteins, and thereby exert functions in recipient cells. Altogether, our data indicate the strong preference of processed fragments for multiple classes of RNA, protein-coding and non-coding, in exRNA (Fig. 6). The co-packaging of processed RNA with various RNA-binding proteins, including RNases (e.g., ANG) and effector complexes (e.g., Ago2), and the significant depletion of full-length RNA in exosomes, suggests that they may function as an exRNA processing machinery with specialized autonomous functions.

Although several studies have explored the RNA bound to specific secreted or circulating proteins, such as high-density lipoprotein (HDL)<sup>3</sup> and Ago2<sup>60</sup>, overall little attention has been paid thus far to the non-vesicular exRNA complexes that account for the significant proportion of exRNA (Fig. 2c). We present data indicating that extracellular non-vesicular RNPs exhibit a RNA signature readily distinguishable from EVs. Notably, ncRNA accounts for nearly entire RNA population in RNPs, with a large prevalence of tRNA, Y RNA and their products, and depletion of miRNA species (Figs. 4 and 6), consistent with the analysis of breast cancer extracellular RNPs<sup>20</sup>. Although extracellular RNPs might be highly heterogeneous, this collective data suggest the key pathways involved in the biogenesis. Of note, we found that massive levels of albumin, an abundant component in B-27 supplement, hamper the protein characterization of extracellular RNPs, and markers of RNPs have yet to be defined. Benefiting from an efficient (>95%) albumin-depletion protocol, we were able to compare protein content in MVs, exosomes, and non-vesicular RNPs using immunoblots. Interestingly, two

nuclear proteins PCNA and Ro60 appeared enriched in RNPs but not in EVs. While the co-localization of Ro60 with its binding Y RNA partners in RNPs may point to the functionality of this extracellular complex, the role of extracellular PCNA, a protein involved in DNA replication, remains to be elucidated. Although the possibility of direct non-vesicular, RNP-mediated RNA uptake and function in recipient cells is currently unexplored, it is supported by the utility of RNPs for RNA and drug delivery<sup>61, 62</sup>. Regardless of the potential biological function of RNPs, their associated RNAs expand the repertoire of potential biomarkers that should be further explored in body fluids, in parallel to EV transcripts.

The key outcome of our work, overall, is an expansion of the repertoire of exRNAs released by GSCs in different vehicles, with functional and biomarker potential, far beyond the class of miRNA. This conclusion challenges the commonly assumed predominant role of miRNA in exRNA-mediated intercellular communication. It further points to the need for in-depth investigation of other classes of exRNAs and their impact on the physiology of recipient cells and use as biomarkers. The future should bring the development of novel experimental techniques and computational resources for integrating complex expression data sets into comprehensive biologic networks and biomarker discovery.

## Methods

**GBM stem cell cultures.** Human low-passage (below 20) GBM cells (kind gift from Dr. Hiroaki Wakamoto, MGH) were cultured as neurospheres in Neurobasal medium (Gibco) supplemented with 3 mM GlutaMAX (Gibco), 1× B-27 supplement (Gibco), 0.5× N-2 (Gibco), 20 ng/ml EGF (R&D systems, MN), 20 ng/ml FGF (PEPROTECH, NJ) and 0.5% Antibiotic-Antimycotic Solution (Corning), and passaged by NeuroCult Chemical Dissociation Kit (Mouse) (Stemcell Technologies, Canada) following the manual. Approximately  $5 \times 10^5$  dissociated cells were seeded per 10 cm dish (Corning) in 10 ml fresh media, and 1/3 volume of fresh medium was added every 3 days. Mature neurospheres, typically formed in 7–10 days, were dissociated and replated. All cells were tested for mycoplasma. Human cells were used in accordance with the policies of institutional review boards at Brigham and Women's Hospital.

**Primary cultures of normal brain cells.** Brain cortices of E18 and P1 C57BL/6 mice were dissected for primary cultures of neurons and glial cells, respectively. The tissues were dissociated with 0.25% Trypsin (Gibco) and 0.1 mg/ml DNaseI (Roche) for 15 min at 37 °C. The cells were plated in poly-D-lysine-coated T25 flasks or 24-well plates at ~80,000 cells per cm<sup>2</sup>, in the seeding medium consisting of DMEM-F12 (Corning), 10% FBS (Gibco), and 1% Antibiotic-Antimycotic Solution (Corning). For neuronal cultures, the media was exchanged to Neurobasal (Gibco), 2% B-27 (Gibco), 1% Antibiotic-Antimycotic Solution (Corning) and 0.5 mM Glutamax (Gibco) 1 day after plating. Mature neurons at 21 days in culture have been utilized for the RNAseq. For glial cultures, the flasks were shaken (200 rpm at 37 °C) three times overnight to remove microglia, and astrocytes trypsinized and further cultured in 24-well plates. For microglia cultures, the media was supplemented with the recombinant M-CSF mouse protein (10 ng/ml; Gibco). Floating microglial cells were collected from the conditioned media by gentle spin (300×g, 10 min), and re-plated in 24-well poly-D-lysine-coated plates at ~100,000 cells per cm<sup>2</sup>. For human primary neuroglial cultures, fetal cortical tissues (gestational age 18 weeks) were provided by Advanced Bioscience Resources, Inc. (Alameda, CA), dissociated with papain (12 U/ml; Worthington), seeded with neuronal media plus 2% FBS (Gibco), and cultured as mouse neurons<sup>63</sup>. Neuronal cultures at 30 days in vitro have been utilized for the RNAseq. Human (HBEC) and mouse (MBEC) primary brain microvascular endothelial cells were purchased from Cell Biologicals, IL (Catalog# H-6023; Lot# 021514F14 and Catalog# C57-6023; Lot# 070613T2MP, respectively) and cultured accordingly to the manufacturer. All animal experiments have been approved by the Harvard Medical Area Standing Committee on Animals.

**Fractionation and RNA isolation.** Approximately 100 ml conditioned medium was used as input for exRNA isolation. Conditioned media was centrifuged at 300×g, 4 °C for 10 min, following by the additional centrifugation at 2000×g, 4 °C for 15 min, to remove cells and cell debris. To monitor EVs, the samples were diluted in DPBS and examined using the Nanoparticle Tracking Analysis system (NanoSight LM10; Malvern Instruments, UK), and EV concentrations were quantified within the optimal linear range (2–10 × 10<sup>8</sup> particles per 1 ml).

For RNA preparation, 5 µl of SUPERase RNase Inhibitor (Ambion) was added to the supernatants per 10 ml media. The media was then filtered

sequentially through the 2  $\mu\text{m}$  filter (GE Healthcare, UK), 0.8  $\mu\text{m}$  filter (EMD Millipore, MA), and 0.22  $\mu\text{m}$  filter (EMD Millipore), with no/minimal pressure applied. The filtrate was split to 15 ml per sample and further filtered through the 0.02  $\mu\text{m}$  filter (GE Healthcare) with up to 75 psi pressure applied. To facilitate the 0.02  $\mu\text{m}$  filtration, a mechanical syringe pump was designed and manufactured (Supplementary Fig. 2). Upon filtration, each filter was washed with 1 ml DPBS (Corning), and the corresponding fractions were lysed with 600  $\mu\text{l}$  lysis solution of the miRCURY RNA Isolation Kit—Cell & Plant (Exiqon, Denmark). The fractions collected on 0.02  $\mu\text{m}$  filters were lysed with 900  $\mu\text{l}$  lysis solution. The last flow-through fractions of 0.02  $\mu\text{m}$  filters were pooled together (up to 30 ml) and concentrated  $\sim$ 60 times using 3 kDa Amicon Ultra Centrifugal Filters (EMD Millipore) at 4000 $\times$ g, 4  $^{\circ}\text{C}$ , for 60 min. The concentrates were collected and lysed with six volumes of the same lysis solution (Exiqon). Total RNA was then isolated from all fractions as recommended by miRCURY protocol, with on-column DNase treatment (Qiagen, Germany). The corresponding 1.2 ml of the source neurospheres were span down at 300 $\times$ g, 4  $^{\circ}\text{C}$  for 5 min, and total cellular RNA was isolated from them and analyzed in parallel. The same protocol was carried out for RNA isolation from fresh media, with 500 ml media input. For RNA isolation from primary cells cultured in 24-well plates, the cells were lysed with 350  $\mu\text{l}$  lysis solution per well. The concentrations of cellular and extracellular RNA were determined using the NanoDrop 2000 Spectrophotometer and Quant-iT RiboGreen RNA Assay Kit (Thermo Fisher Scientific), respectively. The RNA quality was examined using Agilent 2100 Bioanalyzer (Agilent, CA) and the RNA Integrity Number (RIN) estimated.

**Transmission electron microscopy.** The material collected on the filters was resuspended using DPBS (Corning), and further pelleted by 100,000 $\times$ g UC for 80 min at 4  $^{\circ}\text{C}$ . The material diluted in DPBS was added to a glow-discharged carbon-coated grid. The grids were washed with distilled water, stained with 0.75% uranyl formate, examined using Tecnai G<sup>2</sup> Spirit BioTWIN microscopy (FEI, OR), and images recorded by the AMT 2k CCD camera (Advanced Microscopy Techniques, MA) at the Harvard Medical School EM Facility.

**RNA sequencing.** Two sets of spike-in RNAs were added to the samples prior to library preparation: ERCC RNA Spike-In Control Mixes (Ambion; 0.02  $\mu\text{l}$  per 1  $\mu\text{g}$  total RNA), and miRCURY Spike-in kit, part 1, with UniSp2's final concentrations of 1.25 and 5.0 fmol per 1  $\mu\text{g}$  of total RNA for cellular and extracellular RNA, respectively. Total RNA, either 40–200 ng of exRNA, or 2  $\mu\text{g}$  of cellular RNA, was rRNA-depleted using the Ribo-Zero rRNA Removal Kits (Illumina, CA). One quarter of the rRNA-depleted RNA was fragmented to 100–500 nt using the 5 $\times$  First-Strand Buffer (Clontech, CA), and utilized for the long RNA library construction by SMARTer Stranded RNA-Seq Kit (Clontech). The remaining 75% of the rRNA-depleted RNA was treated sequentially with Tobacco Acid Pyrophosphatase (TAP; Illumina) and T4 Polynucleotide Kinase (T4PNK; New England Biolabs, MA) to create more uniform 5'- and 3'-ends for various classes of transcripts. The RNA was then used as input for the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs), with size selection of 15–65 nt inserts for small RNA libraries. The quality of libraries was examined using the Agilent DNA 1000 kit at the Agilent 2100 Bioanalyzer instrument, and cDNA quantified by qRT-PCR. The libraries were sequenced on HiSeq 2000 (Illumina) with single read 50 cycles by Beijing Genomics Institute (BGI, China).

**Reads annotation.** Sequencing reads were treated using the BGI pipeline that included multiple filtering steps, as follows: (1) removing reads with adapters; (2) removing reads with >10% of unknown bases; and (3) removing low-quality reads (sequencing quality <10). After filtering, the remaining clean reads were subjected to the bioinformatics analysis. Generally, 20 and 10 M clean reads were generated per long RNA library and small RNA library, respectively. The clean reads produced from the long RNA libraries were first mapped to the Homo sapiens rRNA database using the SOAPaligner/SOAP2 short read alignment software, to remove the remaining rRNA reads. The non-rRNA reads were used to perform the transcriptome assembling and quantification. First, non-rRNA reads were mapped to human reference genome hg19 using an improved version of TopHat2, which aligns reads across splice junction without relying on gene annotation. Default -g/--max-multihits was used to allow up to 20 multimapping. Next, the reads mapped to genome were assembled using Cufflinks. Reference Annotation Based Transcript (RABT) assembly was performed with the reference gene annotation to compensate incompletely assembled transcripts caused by read coverage gaps in the regions of reference gene. The set of transfrags generated was then compared with the reference transcripts to remove transfrags that were approximately equivalent to the whole or a portion of a reference transcript. After the assembling, the whole parsimonious set of transcripts was obtained. These transcripts were blasted with the NONCODE database using the filter set (identity >0.9 and coverage >0.8) to identify known long non-coding RNA. Then, the rest of assembled transcripts were aligned to the reference annotation utilizing Cuffcompare. Thereafter, Cuffmerge was utilized to merge several assemblies from different samples together, which automatically filtered out a number of transfrags that probably were artifacts and produced a single annotation file for downstream gene expression analysis. Two mismatches were allowed for annotation generally, and only one mismatch was allowed for the ERCC spike-ins. mRNA and lncRNA expression analyses were

performed by Cuffdiff with parameter -u/--multi-read-correct and normalized based on the ERCC spike-in controls using the cyclic loess robust local regression.

Clean reads produced from the small RNA libraries were first aligned to carrier RNA (Enterobacteria phage MS2) of miRCURY Spike-in kit with bowtie (parameter: -v 2 -l 7 --all). The unmapped reads were aligned against human rRNA sequences with bowtie (parameter: -v 2 -l 7 --all) to remove the remaining rRNA reads. The remaining reads were mapped to human miRNA precursors (miRBase V19) with bowtie (parameter: -v 1 --all), and the miRNA precursors reads were separated according to the mapping position. The mapped reads with more than 10% mismatched nucleotides were excluded. The remaining reads were mapped to spike-ins (UniSp2, UniSp4 and UniSp5; parameter: -v 1 --all), and other non-coding RNA classes including tRNAs (gtrNadb), piRNAs (RNadb), snoRNA (snoRNA-LBME-db), scRNAs (Genbank), and others (Rfam database) (parameter: -v 2 --all). Finally, the remaining reads were mapped to the hg19 human genome, and the reads mapped to exons, introns, intergenic regions, and repeats (parameter: -v 0 --all) identified. The SAMtools was used to calculate the reads depth for each base position and R package barplot was used to draw the corresponding bar plot.

**Data analysis.** Based on the known amounts of spike-in RNAs, the read count for each RNA species was first normalized to spike-ins' reads, to quantify the absolute amounts in fmol per  $\mu\text{g}$  of total RNA. Next, the obtained values were corrected to the corresponding fresh media as the blank control, using the equation shown in Supplementary Fig. 6. For the analysis of class composition, the total abundance of corrected non-rRNA was used for normalization between the samples. To compare heterogeneity of the samples, the sum of squared errors ( $\chi^2$  value) of species composition was calculated using MS Excel Macro (available in Supplementary Data 3). To estimate the inequality of abundance among all RNA species in one class, the evenness factor, Gini coefficient, and traditional pre-set evaluations were calculated using a MS Excel Macro (available in Supplementary Data 4). The hierarchical clustering analysis was performed using the MultiExperiment Viewer (Dana-Farber Cancer Institute, MA). The principle component analysis (PCA) was performed and visualized with R package rgl. miRNA targets were determined based on starBase v2.0, with at least three supporting CLIP-Seq experiments<sup>33</sup>. Pathway analysis of target mRNAs was performed using the Ingenuity Pathway Analysis (IPA; Qiagen). Venn diagrams were plotted using Venny 2.1.

**Long RNA reverse transcription PCR.** Maxima Reverse Transcriptase (100U; Thermo Fisher Scientific) was used to reverse transcribe 20 ng RNA with both oligo (dT) and random hexamers in a 10  $\mu\text{l}$  reaction system. Next, 0.5  $\mu\text{l}$  cDNA was used in 10  $\mu\text{l}$  PCR reactions based on Phire Hot Start II PCR Master Mix (Thermo Fisher Scientific) and primers (0.5  $\mu\text{M}$  each), and amplified using a touchdown program with either 15 or 70 s extension step for short or long amplification, respectively. The PCR products were examined on 0.8–1.2% agarose gels (Thermo Fisher Scientific). Primer sequences are provided in Supplementary Table 3.

**Quantitative reverse transcription PCR.** Generally, for small RNA qRT-PCR, 10 ng of total RNA was used in 10  $\mu\text{l}$  reverse transcription reaction with Universal cDNA Synthesis kit II (Exiqon). The cDNA was diluted 80 times, and 4  $\mu\text{l}$  was used in 10  $\mu\text{l}$  qPCR reactions using the ExiLent SYBR Green master mix and custom-designed LNA primers (Exiqon). For mRNA qRT-PCR, 10 ng of total RNA was used in 10  $\mu\text{l}$  reverse transcription reaction with PrimeScript RT Master Mix (Takara, Japan). One microliter cDNA was used in 10  $\mu\text{l}$  qPCR reactions using the ExiLent SYBR Green master mix and synthesized primers (Supplementary Table 4; Eton Bioscience, CA). The qPCR reactions were run on a Viia 7 instrument (Thermo Fisher Scientific) in duplicates. The specificity of qPCR products was verified by the presence of a single peak at the melting curves.

**Immunoblotting.** For protein analysis, cellular and extracellular fractions were lysed using the modified RIPA buffer containing 2% SDS, 1% sodium deoxycholate and 3 M urea. The RNP fractions were BSA-depleted using the Aurum Affi-Gel Blue Mini Columns (Bio-Rad, CA), and concentrated using Pierce SDS-PAGE Sample Prep Kit (Pierce, MA). Protein concentrations were quantified using the Micro BCA Protein Assay Kit (Pierce). Equal amounts of the total protein (50  $\mu\text{g}$ ) were loaded per lane in Novex WedgeWell 14% Tris-Glycine Mini Gel (Thermo Fisher Scientific). Proteins were transferred to 0.45  $\mu\text{m}$  PVDF membrane (Thermo Fisher Scientific). After blocking with 5% (wt/vol) fat-free milk in Tris-buffered saline with 0.075% Tween-20 (TBST), membranes were incubated with 1:1000 diluted primary antibodies (Flotillin-1 #18634S, CD9 #13174S, Integrin  $\beta$ 1 #9699S, HSP90 #4874S, La #5034S, NPM #3542S, Ago2 #2897S, Alix #2171S, and PCNA #13110S from Cell Signaling Technology, MA; Ro60 #AV40534 from Sigma-Aldrich; ANG #sc-74528 from Santa Cruz Biotechnology, TX) overnight at 4  $^{\circ}\text{C}$ . The membranes were washed and incubated with horseradish peroxidase-conjugated secondary antibodies (#7074S and #7076S from Cell Signaling Technology, 1:2000 dilution) for 1 h at room temperature. The blots were developed by the Amersham ECL Reagent (GE Healthcare) and, if required, stripped using One Minute Western Blot Stripping Buffer (GM Biosciences, CT). Uncropped scans of blots are shown in Supplementary Fig. 19.

**Transwell cocultures of GSC and astrocytes.** Mouse cortical astrocytes established from WT or miR-21 KO P1 pups, passage 1 or 2, were cultured in 24-well plates. Small GBM8 neurospheres grown in GSC conditions were transferred to the upper chamber of the Millicell hanging insert with 1.0  $\mu\text{m}$  pore size (Millipore), and co-cultured with the pre-established astrocytes. The co-culture media was adjusted to contain DMEM-F12 (Corning), 1 $\times$  B-27 supplement (Gibco), 0.5 $\times$  N-2 (Gibco), and 1% Antibiotic-Antimycotic Solution (Corning), enabling co-culturing of healthy astrocytes and GSC neurospheres over a period of at least 6 days. The corresponding control “mono-cultures” of astrocytes were cultured in the identical media under the same conditions. Three days later, astrocytes were washed with cold DPBS (Corning) three times, and cellular RNA isolated as in other experiments.

**Fractionation of cerebrospinal fluid.** Cerebrospinal fluid (CSF) samples from four patients diagnosed with primary GBM were collected during surgery, with informed consent obtained according to the appropriate protocol, approved by the UC San Diego Institutional Review Board. The samples were centrifuged at 1500 $\times$ g at room temperature for 10 min, immediately after collection to remove cells. The supernatants were filtered through the 0.8  $\mu\text{m}$  filters (Millipore), aliquoted into cryotubes (Thermo Fisher Scientific), and stored at  $-80^\circ\text{C}$  for 12–24 months. The aliquots of 3 ml were thawed on ice, and used to isolate MV, exosomes, and RNP fractions by sequential filtration, as described above.

**Statistics.** Values are given as mean  $\pm$  SEM. Numbers of experimental replicates are given in the figure legends. When two groups were compared, significance was determined using an unpaired two-sided *t*-test. Before performing *t*-test, normal distribution was verified by one-sample Kolmogorov–Smirnov test and equality of variances was verified by Levene’s test using SPSS (IBM, NY). A *p*-value  $< 0.05$  is considered as statistical significance. All bars represent mean  $\pm$  SEM.

**Data availability.** RNAseq fq files and processed data are available through GEO under GSE93143. MS Excel Macros used to calculate inequality and heterogeneity are available in Supplementary Data 3 and 4. All relevant data are available within the Article and Supplementary Files, or available from the authors upon request.

Received: 5 January 2017 Accepted: 25 August 2017

Published online: 26 October 2017

## References

1. Tkach, M. & Thery, C. Communication by extracellular vesicles: where we are and where we need to go. *Cell* **164**, 1226–1232 (2016).
2. Valadi, H. et al. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
3. Vickers, K. C., Palmisano, B. T., Shoucri, B. M., Shamburek, R. D. & Remaley, A. T. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* **13**, 423–433 (2011).
4. D’Asti, E., Chennakrishnaiah, S., Lee, T. H. & Rak, J. Extracellular vesicles in brain tumor progression. *Cell. Mol. Neurobiol.* **36**, 383–407 (2016).
5. Sato-Kuwabara, Y., Melo, S. A., Soares, F. A. & Calin, G. A. The fusion of two worlds: non-coding RNAs and extracellular vesicles—diagnostic and therapeutic implications (review). *Int. J. Oncol.* **46**, 17–27 (2014).
6. Chen, X., Liang, H., Zhang, J., Zen, K. & Zhang, C. Y. Secreted microRNAs: a new form of intercellular communication. *Trends Cell Biol.* **22**, 125–132 (2012).
7. Laurent, L. C. et al. Meeting report: discussions and preliminary findings on extracellular RNA measurement methods from laboratories in the NIH Extracellular RNA Communication Consortium. *J. Extracell. Vesicles* **4**, 26533 (2015).
8. Nakano, I., Garnier, D., Minata, M. & Rak, J. Extracellular vesicles in the biology of brain tumour stem cells—implications for inter-cellular communication, therapy and biomarker development. *Semin. Cell. Dev. Biol.* **40**, 17–26 (2015).
9. Gyorgy, B. et al. Detection and isolation of cell-derived microparticles are compromised by protein complexes resulting from shared biophysical parameters. *Blood* **117**, e39–e48 (2011).
10. Cvjetkovic, A., Lotvall, J. & Lasser, C. The influence of rotor type and centrifugation time on the yield and purity of extracellular vesicles. *J. Extracell. Vesicles* **3**, 23111 (2014).
11. Jacobs, V. L., Valdes, P. A., Hickey, W. F. & De Leo, J. A. Current review of in vivo GBM rodent models: emphasis on the CNS-1 tumour model. *ASN Neuro* **3**, e00063 (2011).
12. Wakimoto, H. et al. Human glioblastoma-derived cancer stem cells: establishment of invasive glioma models and treatment with oncolytic herpes simplex virus vectors. *Cancer Res.* **69**, 3472–3481 (2009).
13. Wakimoto, H. et al. Maintenance of primary tumor phenotype and genotype in glioblastoma stem cells. *Neuro Oncol.* **14**, 132–144 (2012).
14. Teplyuk, N. M. et al. Therapeutic potential of targeting microRNA-10b in established intracranial glioblastoma: first steps toward the clinic. *EMBO Mol. Med.* **8**, 268–287 (2016).
15. Raabe, C. A., Tang, T. H., Brosius, J. & Rozhdzhevskiy, T. S. Biases in small RNA deep sequencing data. *Nucleic Acids Res.* **42**, 1414–1426 (2014).
16. Zhuang, F., Fuchs, R. T. & Robb, G. B. Small RNA expression profiling by high-throughput sequencing: implications of enzymatic manipulation. *J. Nucleic Acids* **2012**, 360358 (2012).
17. Ando, Y., Burroughs, A. M., Kawano, M., Hoon, M. J. L. d. & Hayashizaki, Y. in *Regulatory RNAs* (eds Mallick, B. & Ghosh, Z.) 253–271 (Springer, 2012).
18. Roberts, T. C., Coenen-Stass, A. M. & Wood, M. J. Assessment of RT-qPCR normalization strategies for accurate quantification of extracellular microRNAs in murine serum. *PLoS ONE* **9**, e89237 (2014).
19. Wei, Z., Batagov, A. O., Carter, D. R. & Krichevsky, A. M. Fetal bovine serum RNA interferes with the cell culture derived extracellular RNA. *Sci. Rep.* **6**, 31175 (2016).
20. Tosar, J. P. et al. Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res.* **43**, 5601–5616 (2015).
21. Montecalvo, A. et al. Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. *Blood* **119**, 756–766 (2012).
22. Li, Y. et al. Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res.* **25**, 981–984 (2015).
23. Ravasi, T. et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.* **16**, 11–19 (2006).
24. Jenjaroenpun, P. et al. Characterization of RNA in exosomes secreted by human breast cancer cell lines using next-generation sequencing. *PeerJ* **1**, e201 (2013).
25. Li, J. et al. Identification and characterization of 293T cell-derived exosomes by profiling the protein, mRNA and microRNA components. *PLoS ONE* **11**, e0163043 (2016).
26. Batagov, A. O. & Kurochkin, I. V. Exosomes secreted by human cells transport largely mRNA fragments that are enriched in the 3’-untranslated regions. *Biol. Direct.* **8**, 12 (2013).
27. Landgraf, P. et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401–1414 (2007).
28. Kowalski, M. P. & Krude, T. Functional roles of non-coding Y RNAs. *Int. J. Biochem. Cell Biol.* **66**, 20–29 (2015).
29. Kohn, M., Pazaitis, N. & Huttelmaier, S. Why YRNAs? About versatile RNAs and their functions. *Biomolecules* **3**, 143–156 (2013).
30. O’Brien, C. A., Margelot, K. & Wolin, S. L. Xenopus Ro ribonucleoproteins: members of an evolutionarily conserved class of cytoplasmic ribonucleoproteins. *Proc. Natl Acad Sci USA* **90**, 7250–7254 (1993).
31. Mazieres, J. et al. Alternative processing of the U2 small nuclear RNA produces a 19–22nt fragment with relevance for the detection of non-small cell lung cancer in human serum. *PLoS ONE* **8**, e60134 (2013).
32. Haraszti, R. A. et al. High-resolution proteomic and lipidomic analysis of exosomes and microvesicles from different cell sources. *J. Extracell. Vesicles* **5**, 32570 (2016).
33. Li, J. H., Liu, S., Zhou, H., Qu, L. H. & Yang, J. H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **42**, D92–D97 (2014).
34. Chevillet, J. R. et al. Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl Acad. Sci. USA* **111**, 14888–14893 (2014).
35. Gabriely, G. et al. MicroRNA 21 promotes glioma invasion by targeting matrix metalloproteinase regulators. *Mol. Cell Biol.* **28**, 5369–5380 (2008).
36. Bang, C. et al. Cardiac fibroblast-derived microRNA passenger strand-enriched exosomes mediate cardiomyocyte hypertrophy. *J. Clin. Invest.* **124**, 2136–2146 (2014).
37. Baglio, S. R. et al. Human bone marrow- and adipose-mesenchymal stem cells secrete exosomes enriched in distinctive miRNA and tRNA species. *Stem Cell Res. Ther.* **6**, 127 (2015).
38. van Balkom, B. W., Eisele, A. S., Pegtel, D. M., Bervoets, S. & Verhaar, M. C. Quantitative and qualitative analysis of small RNAs in human endothelial cells and exosomes provides insights into localized RNA processing, degradation and sorting. *J. Extracell. Vesicles* **4**, 26760 (2015).
39. Liao, J., Liu, R., Yin, L. & Pu, Y. Expression profiling of exosomal miRNAs derived from human esophageal cancer cells by solexa high-throughput sequencing. *Int. J. Mol. Sci.* **15**, 15530–15551 (2014).
40. Koppers-Lalic, D. et al. Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. *Cell Rep.* **8**, 1649–1658 (2014).
41. Li, C. C. et al. Glioma microvesicles carry selectively packaged coding and non-coding RNAs which alter gene expression in recipient cells. *RNA Biol.* **10**, 1333–1344 (2013).
42. Vojtech, L. et al. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Res.* **42**, 7290–7304 (2014).

43. Anderson, P. & Ivanov, P. tRNA fragments in human health and disease. *FEBS Lett.* **588**, 4297–4304 (2014).
44. Emara, M. M. et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J. Biol. Chem.* **285**, 10959–10968 (2010).
45. Ivanov, P., Emara, M. M., Villen, J., Gygi, S. P. & Anderson, P. Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol. Cell* **43**, 613–623 (2011).
46. Kumar, P., Anaya, J., Mudunuri, S. B. & Dutta, A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol.* **12**, 78 (2014).
47. Eberle, K. et al. The expression of angiogenin in tissue samples of different brain tumours and cultured glioma cells. *Anticancer. Res.* **20**, 1679–1684 (2000).
48. Xia, W. et al. Angiogenin promotes U87MG cell proliferation by activating NF- $\kappa$ B signaling pathway and downregulating its binding partner FHL3. *PLoS ONE* **10**, e0116983 (2015).
49. Skog, J. et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **10**, 1470–1476 (2008).
50. Chen, X. et al. An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. *Cell* **153**, 166–177 (2013).
51. Christov, C. P., Gardiner, T. J., Szuts, D. & Krude, T. Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. *Mol. Cell Biol.* **26**, 6993–7004 (2006).
52. Krude, T., Christov, C. P., Hyrien, O. & Marheineke, K. Y RNA functions at the initiation step of mammalian chromosomal DNA replication. *J. Cell Sci.* **122**, 2836–2845 (2009).
53. Nicolas, F. E., Hall, A. E., Csorba, T., Turnbull, C. & Dalmay, T. Biogenesis of Y RNA-derived small RNAs is independent of the microRNA pathway. *FEBS Lett.* **586**, 1226–1230 (2012).
54. Meiri, E. et al. Discovery of microRNAs and other small RNAs in solid tumors. *Nucleic Acids Res.* **38**, 6234–6246 (2010).
55. Kohn, M., Ihling, C., Sinz, A., Krohn, K. & Huttelmaier, S. The Y3\*\* ncRNA promotes the 3' end processing of histone mRNAs. *Genes Dev.* **29**, 1998–2003 (2015).
56. Chakraborty, S. K., Prakash, A., Nechooshtan, G., Hearn, S. & Gingeras, T. R. Extracellular vesicle-mediated transfer of processed and functional RNY5 RNA. *RNA* **21**, 1966–1979 (2015).
57. Hogg, J. R. & Collins, K. Human Y5 RNA specializes a Ro ribonucleoprotein for 5S ribosomal RNA quality control. *Genes Dev.* **21**, 3067–3072 (2007).
58. Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
59. Ke, S. et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* **29**, 2037–2053 (2015).
60. Arroyo, J. D. et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl Acad. Sci. USA* **108**, 5003–5008 (2011).
61. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).
62. Zhang, F. et al. Reconstituted high density lipoprotein mediated targeted co-delivery of HZ08 and paclitaxel enhances the efficacy of paclitaxel in multidrug-resistant MCF-7 breast cancer cells. *Eur. J. Pharm. Sci.* **92**, 11–21 (2016).
63. Wang, Y. et al. Downregulation of miR-132/212 impairs S-nitrosylation balance and induces tau phosphorylation in Alzheimer's disease. *Neurobiol. Aging* **51**, 156–166 (2017).

## Acknowledgements

This work was supported by the U19 CA179563 grant and U19 administrative supplement through the NIH Common Fund and the Office of Strategic Coordination/Office of the NIH Director (A.M.K. and X.O.B.), R21 NS098051 (A.M.K.), and CA069246 (X.O.B.) grants. We thank Robert Kitchen, Sai Lakshmi Subramanian, Alain Charest, Steven Gould, Thorsten Mempel, and the RNaseq working group of the Extracellular RNA Communication Consortium for valuable insights on the data analysis. We thank Maria Ericsson for TEM, Hongjun Wang for designing and manufacturing syringe pumps, Pavel Ivanov for sharing ANG antibody, Oleg Butovsky for sharing miR-21 knockout mice, and Erik J. Uhlmann for reviewing the manuscript. We thank members of Krichewsky and Breakefield laboratories for helpful discussions.

## Author contributions

A.M.K. conceived the study; Z.W. and A.M.K. designed experiments; Z.W. performed experiments; S.S., R.E.F., Y.W., R.R. and L.B. assisted with experiments; A.B. and J.W. performed bioinformatics analysis; X.O.B. contributed to data analysis; F.H., C.C. and B.C. provided CSF samples; Z.W. and A.M.K. analyzed the data and wrote the paper. All authors revised and approved the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-01196-x.

**Competing interests:** X.O.B. is a consulting member of the Scientific Advisory Board of Evox Therapeutics, Ltd. and Exocyte Therapeutics, Ltd. The remaining authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017