

# SCIENTIFIC REPORTS



OPEN

## Changes in nuclear and cytoplasmic microRNA distribution in response to hypoxic stress

Tiia A. Turunen<sup>1</sup>, Thomas C. Roberts<sup>2,3</sup>, Pia Laitinen<sup>1</sup>, Mari-Anna Väänänen<sup>1</sup>, Paula Korhonen<sup>1</sup>, Tarja Malm<sup>1</sup>, Seppo Ylä-Herttuala <sup>1,4</sup> & Mikko P. Turunen<sup>1</sup>

MicroRNAs (miRNAs) are small non-coding RNAs that have well-characterized roles in cytoplasmic gene regulation, where they act by binding to mRNA transcripts and inhibiting their translation (i.e. post-transcriptional gene silencing, PTGS). However, miRNAs have also been implicated in transcriptional gene regulation and alternative splicing, events that are restricted to the cell nucleus. Here we performed nuclear-cytoplasmic fractionation in a mouse endothelial cell line and characterized the localization of miRNAs in response to hypoxia using small RNA sequencing. A highly diverse population of abundant miRNA species was detected in the nucleus, of which the majority (56%) was found to be preferentially localized in one compartment or the other. Induction of hypoxia resulted in changes in miRNA levels in both nuclear and cytoplasmic compartments, with the majority of changes being restricted to one location and not the other. Notably, the classical hypoxamiR (miR-210-3p) was highly up-regulated in the nuclear compartment after hypoxic stimulus. These findings reveal a previously unappreciated level of molecular complexity in the physiological response occurring in ischemic tissue. Furthermore, widespread differential miRNA expression in the nucleus strongly suggests that these small RNAs are likely to perform extensive nuclear regulatory functions in the general case.

The canonical view of microRNA (miRNA) function is that these small RNA molecules typically repress gene expression by targeting the mRNA 3' untranslated region (3' UTR) in order to induce post-transcriptional gene silencing (PTGS)<sup>1</sup>. A growing body of evidence suggests that miRNA functionality may be much more complex than this canonical paradigm, and recent studies have suggested multiple novel modes of action for small RNAs<sup>2-5</sup>. For example, small RNA-mediated transcriptional gene activation (TGA) has been reported in some cases, whereby synthetic small RNAs or endogenous miRNAs targeted to gene promoters can induce increases in gene expression<sup>6,7</sup>. TGA involves an epigenetic mechanism, although a detailed understanding of this phenomenon has thus far remained elusive<sup>8</sup>.

We have previously shown that short hairpin RNA (shRNA)-mediated transcriptional activation of the *Vegfa* promoter is of therapeutic benefit in murine models of hindlimb ischemia<sup>9</sup> and myocardial infarction<sup>10</sup>. These observations suggested that shRNAs may function by mimicking endogenous miRNAs. To date, intracellular miRNA localization has been largely overlooked, and only a handful of publications have performed global miRNA analysis in nuclear and cytoplasmic fractions. Conversely, the majority of citations focusing on miRNA function have not addressed the issue of subcellular localization. Previous reports have analyzed the subcellular distribution of small RNAs in 5-8F<sup>11</sup> and HTC116 cells<sup>12</sup> and showed that the majority of miRNAs, regardless of their sequence, are shuttled to nucleus. Some sequence motifs have been demonstrated to affect miRNA subcellular distribution<sup>13,14</sup>, but other factors may also direct miRNA localization patterns. For example, Khudayberdiev *et al.*, observed that nuclear accumulation of miRNAs may be associated with global down-regulation of miRNA levels occurring during neuronal development<sup>15</sup>. Importantly, components of the RNA interference machinery have been observed to be functional in the nucleus, although small RNA loading is restricted to the cytoplasm<sup>16</sup>. While

<sup>1</sup>A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Yliopistoranta 1E, 70210, Kuopio, Finland. <sup>2</sup>Department of Paediatrics, University of Oxford, South Parks Road, Oxford, OX1 3QX, UK. <sup>3</sup>Sanford Burnham Prebys Medical Discovery Institute, Development, Aging and Regeneration Program, 10901 North Torrey Pines Road, La Jolla, CA, 92037, USA. <sup>4</sup>Heart Center and Gene Therapy Unit, Kuopio University Hospital, PO Box 100, 70029 KUH, Kuopio, Finland. Thomas C. Roberts and Pia Laitinen contributed equally. Correspondence and requests for materials should be addressed to M.P.T. (email: [mikko.turunen@uef.fi](mailto:mikko.turunen@uef.fi))

several studies have profiled miRNA expression in the nucleus, no study has to date applied this analysis to cells undergoing physiological transitions (i.e. during cellular differentiation, or in response to environmental stimulus).

Hypoxia is a severe pathophysiological condition observed in multiple diseases (e.g. myocardial infarction and cerebral ischemia) where it impairs tissue function, resulting in a rapid and complex response at the molecular and cellular level. In the present study, we selected hypoxia as a prototypical, medically-relevant, pathophysiological stimulus that is known to lead to changes in gene and miRNA expression<sup>17</sup>. To investigate the extent to which mature miRNAs are present in the nucleus, we performed high throughput sequencing of small RNAs (sRNA-seq) in nuclear and cytoplasmic fractions derived from endothelial cells after the induction of hypoxia. This analysis has identified multiple miRNAs that are enriched in either the nucleus or cytoplasm. Furthermore, we show for the first time, that specific miRNAs are differentially expressed in a subcellular location restricted manner in response to hypoxic stimulus. These findings suggest that miRNAs likely execute nuclear functions in addition to their cytoplasmic PTGS activity, and reveal a previously unappreciated level of complexity in the cellular response to hypoxia.

## Results

**Isolation of nuclear and cytoplasmic fractions from endothelial cells.** The murine C166 endothelial cell line was cultured under normal conditions or exposed to hypoxia (1% oxygen) for 24 h. Cells were separated into nuclear and cytoplasmic fractions according to the protocol described by Gagnon *et al.*, and RNA and proteins were extracted from each sample<sup>18</sup> (Supplementary Fig. S1). Fraction purity was confirmed by western blot using antibodies against nuclear (histone H3) and cytoplasmic ( $\beta$ -tubulin) marker proteins (Figs. 1a and S2). Similarly, the nuclear-enriched long non-coding RNAs (lncRNAs) *Malat1* and *Neat1*<sup>19</sup> were found to be 37- and 125-fold enriched in the nucleus respectively, as determined by RT-qPCR (Fig. 1b). Conversely, transfer RNAs (tRNA) tRNA-Lys-TTT and tRNA-Met-CAT were found to be 12- and 17-fold enriched in the cytoplasm, respectively (Fig. 1b), consistent with successful subcellular fractionation.

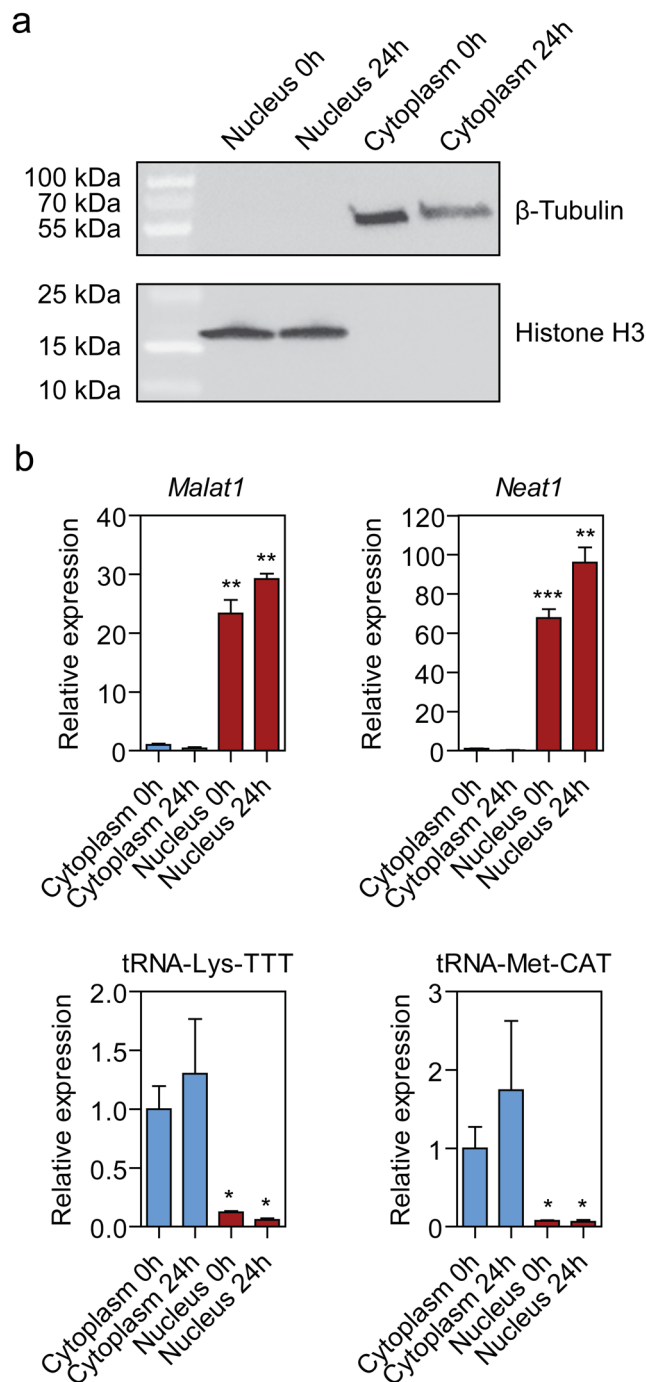
**Identification of nuclear and cytoplasmic enriched miRNAs.** C166 cells were cultured in either normoxic or hypoxic conditions (for 2 h and 24 h) and RNA from nuclear and cytoplasmic fractions analyzed by sRNA-seq ( $n = 2$ ) (Supplementary Fig. S1). 214 million reads were generated in total, and library sizes were found to be comparable between all samples (Supplementary Fig. S3). The data were filtered to remove low abundance miRNAs (<10 reads in each library) after which 350 miRNAs remained. Statistical differences between samples were tested using DESeq2<sup>20</sup> and libraries were visualized by Principal Component Analysis (PCA) (Fig. 2a) which showed tight clustering of biological replicates. Nuclear and cytoplasmic samples were clearly separated in principal component one (which contained 74% of the variation in the dataset), whereas the response to hypoxic stress was reflected in principal component 2 (12% of the variance).

Nuclear and cytoplasmic enrichment was determined by differential expression analysis of all nuclear libraries against all cytoplasmic libraries and visualized by hierarchical clustering (Fig. 2b) and a volcano plot (Fig. 2c). 196 miRNAs (56% of all miRNAs) were significantly different ( $P < 0.01$ ) between nucleus and cytoplasm, suggesting that the majority of miRNAs are preferentially enriched in one of the compartments. Of these, 105 miRNAs were enriched in either compartment by  $\geq 2$ -fold (i.e. 46 enriched in the nucleus, and 59 enriched in the cytoplasm). MA plot analysis showed that differentially enriched miRNAs were found across a broad range of absolute expression (i.e. counts) values (Fig. 2d). The most strongly location-enriched miRNAs were miR-3535 (27.9-fold enriched in the nucleus) and miR-27a-5p (48-fold enriched in the cytoplasm) (Fig. 2e). Differential nuclear-cytoplasmic distribution of these miRNAs was further confirmed by RT-qPCR, where additional replicate cultures were analyzed to increase statistical power and confirm reproducibility of these findings (Fig. 2f). RT-qPCR data were normalized to the levels of miR-186-5p which was determined to be the most stably expressed miRNA across all conditions using the NormFinder method<sup>21</sup>.

**miRNAs are differentially expressed in distinct subcellular compartments.** We next sought to identify miRNAs that were differentially expressed in response to hypoxic stress.

Nuclear and cytoplasmic libraries were considered separately, and differential expression determined between time points. Using a fold-change cut-off of 2, and a stringent statistical cut-off of  $P < 0.01$  (Benjamini-Hochberg adjusted), 54 and 35 significantly different miRNAs were called in the nuclear and cytoplasmic libraries respectively for the 24 h vs 0 h comparison. In contrast, no significant differences were called in the 2 h vs 0 h comparison for either nuclear or cytoplasmic fractions.

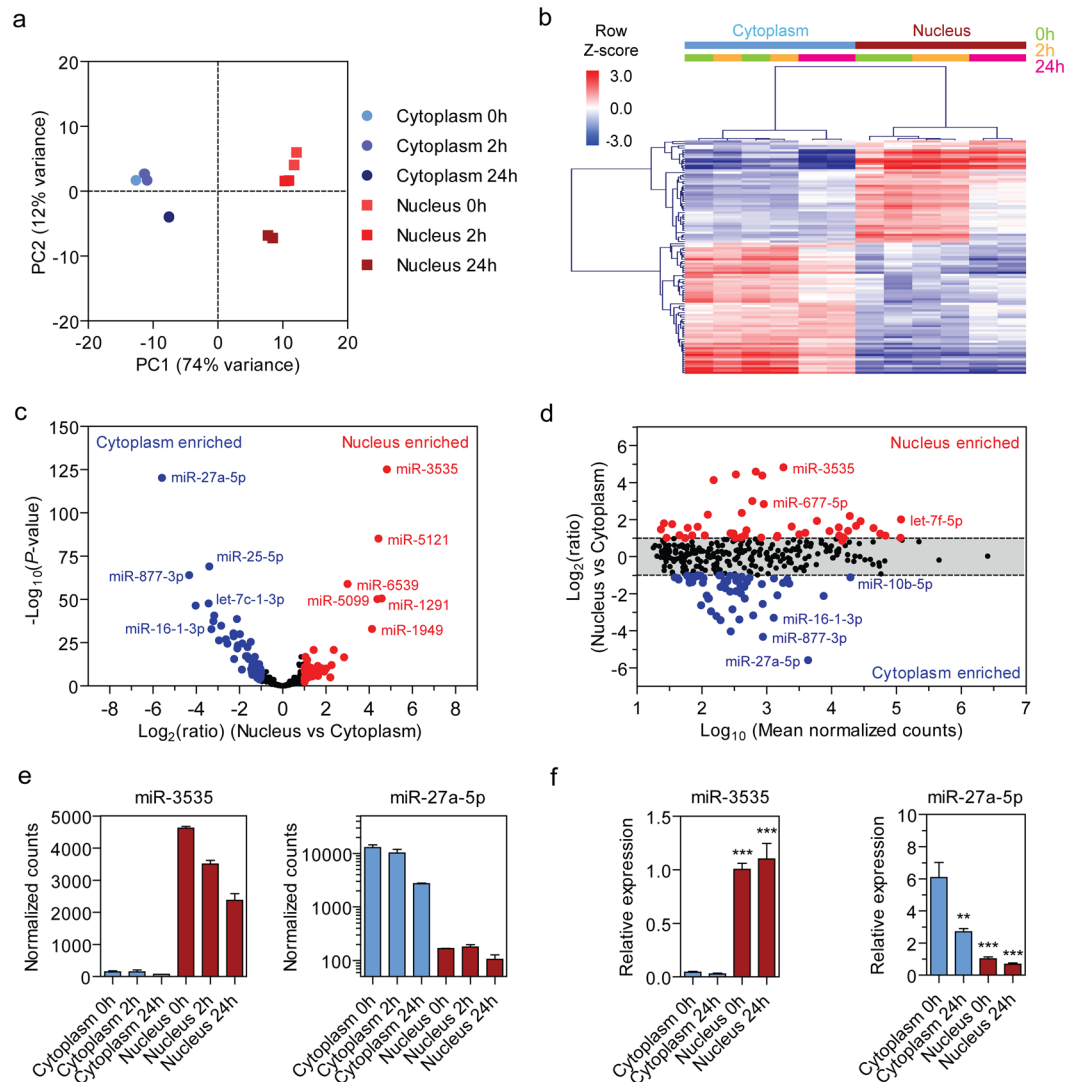
Expression ratios for nuclear and cytoplasmic libraries were not correlated (Pearson  $r = 0.0488$ , Spearman  $r = 0.0912$ ,  $P > 0.05$ ) (Fig. 3a), indicating that differential miRNA expression typically occurred exclusively in one fraction and not the other. Interestingly, the majority of differentially expressed miRNAs were down-regulated under hypoxic conditions. Only 6 miRNAs were identified as exceptions that were differentially expressed in both nucleus and cytoplasm after 24 h of hypoxic stress (Fig. 3b). These included 2 up-regulated miRNAs (miR-210-3p and miR-669c-5p) and 4 down-regulated (let-7a-5p, miR-200c-3p, miR-193a-5p and miR-203-3p). In the cytoplasm, 29 miRNAs were differentially expressed (Fig. 3c) (5 up-regulated and 24 down-regulated). In the nucleus, 48 miRNAs were differentially expressed (Fig. 3d) (13 up-regulated and 35 down-regulated). Surprisingly, miR-210-3p, a well-described hypoxamiR<sup>22</sup>, was observed to be highly up-regulated in the nuclear fraction in the sRNA-seq data (Fig. 3e). This finding was confirmed by RT-qPCR (Fig. 3f). Fluorescent *in situ* hybridization (FISH) was used to analyze miR-210-3p localization (Fig. 4). Detected signal in confocal microscopy for miR-210-3p was punctate and could be observed in the nuclei of C166 cells both in normoxia and hypoxia (24 h). Qualitative assessment of miR-210-3p puncta also correspond with the increase in miR-210 levels observed by RT-qPCR. miR-210-3p is known to have a role in several pathways related to the low cellular oxygen and has been shown to regulate *EFNA3*, *E2F3*, *HOXA3*, *HIP1*, *BDNF*, *KCMF1* and *NDUFA4P1* by canonical PTGS<sup>23,24</sup>.



**Figure 1.** Isolation of nuclear and cytoplasmic fractions. **(a)** Western blot analysis of cytoplasmic ( $\beta$ -tubulin) and nuclear (histone H3) marker protein expression. **(b)** RT-qPCR analysis of nuclear-enriched lncRNAs (*Malat1* and *Neat1*) or cytoplasm-enriched tRNAs (tRNA-Lys-TTT and tRNA-Met-CAT). Values are mean  $\pm$  SEM,  $n = 3-4$ , \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  ( $t$ -test with Welch correction for unequal variance, comparisons are to the matched group-matched Cytoplasmic fraction).

Furthermore, miR-210-3p was previously shown to interact with *XIST*, a lncRNA located in the nucleus which is responsible for the initiation of X chromosome inactivation<sup>25</sup>. Together, these data suggest that miR-210-3p likely regulates nuclear processes, in addition to well-established cytoplasmic PTGS mechanisms.

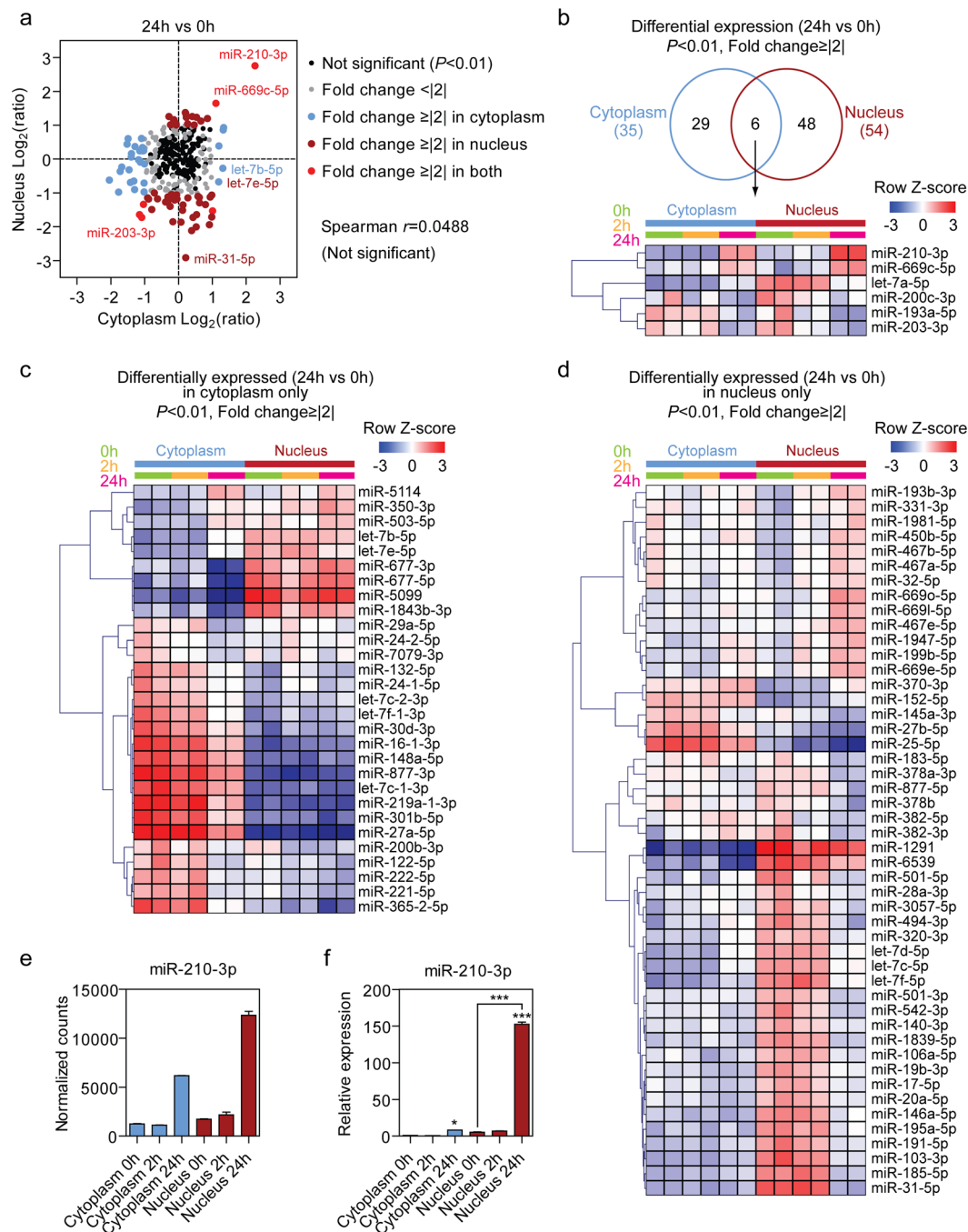
To complement the analysis in C166 cells, we validated the most interesting miRNAs in other mouse cell lines by RT-qPCR. The expression of the miRNAs with nuclear and cytoplasmic localization, miR-3535 and miR-1291 (nuclear) and miR-27a-5p (cytoplasmic), was also observed to be similar in the cell lines MS1 (pancreatic endothelial), MOVAS (smooth muscle) and C2C12 (skeletal muscle) (Supplementary Fig. S4). In addition, miR-210-3p was upregulated in hypoxia in the cytoplasmic and nuclear fractions of these cells in a similar fashion as to that observed in C166 cells. Overall, these results demonstrate that the nuclear-cytoplasmic distribution of certain



**Figure 2.** Nuclear and cytoplasmic small RNAs. **(a)** Principal component analysis (PCA) shows tight clustering of biological replicates. Principal component one (PC1) shows the separation between nuclear and cytoplasmic samples (74% of variation), whereas PC2 depicts the response to hypoxic stress (12% of variation). **(b)** Hierarchical clustering of differential expression between nuclear and cytoplasmic samples (adjusted  $P < 0.01$ , fold change  $\geq |2|$ ). Scale bars show mean-centered  $\log_2$  normalized counts (row Z-score) where red and blue indicate higher and lower than mean abundance respectively. **(c)** Volcano plot depicting nuclear and cytoplasmic miRNA enrichment. **(d)** MA plot analysis of location-enriched miRNAs showing that differentially nuclear/cytoplasm-enriched miRNAs were observed over a broad range of absolute expression values. **(e)** The two most location-enriched miRNAs across the dataset were miR-3535 (in the nucleus) and miR-27a-5p (in the cytoplasm). **(f)** miRNA-expression levels of nuclear (miR-3535) and cytoplasmic (miR-27a-5p) miRNAs identified by miRNA-seq were validated by RT-qPCR. miRNA levels were normalized to miR-186-5p and the mean value of the Nucleus 0 h group scaled to a value of one. Values are mean  $\pm$  SEM,  $n = 2$  (sRNA-seq),  $n = 3$  (RT-qPCR), \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (one-way ANOVA with Bonferroni *post hoc* test, comparisons are to the Cytoplasm 0 h group).

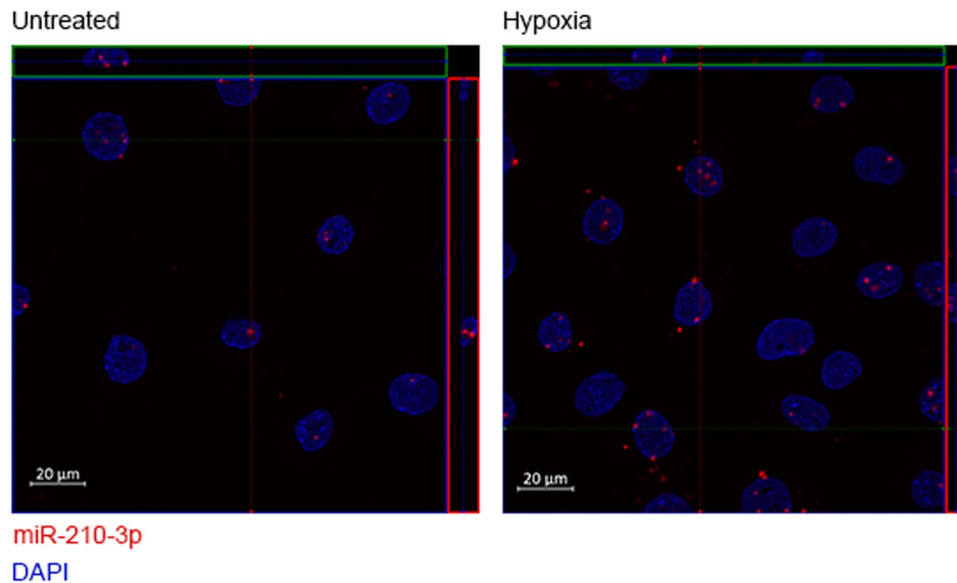
location-enriched miRNAs, and the nuclear up-regulation of miR-210-3p are similar between different cell lines derived from distinct lineages.

sRNA-seq analysis was performed using equal amounts of input RNA for each sample, and the RT-qPCR analysis performed above normalized according to a stable reference gene identified from the sequencing dataset. These experiments therefore assume that the amount of RNA in the nuclear and cytoplasmic fractions is equal. To address this potential confounding factor, RT-qPCR validation experiments were repeated using new C166 cultures in which a synthetic spike oligonucleotide (the *C. elegans* miRNA: cel-miR-39) was added to the nuclear and cytoplasmic fractions at the phenol extraction phase. This technical modification allowed for assumption-free RT-qPCR normalization independent of any difference in total RNA or total miRNA content between the nuclear and cytoplasmic fractions. Results for the miRNAs-of-interest (miR-27a-5p, miR-3535, miR-210-3p and miR-1291) were found to be very similar irrespective of whether the data were normalized to the exogenous spike (cel-miR-39)



**Figure 3.** Differential expression of nuclear and cytoplasmic miRNAs upon hypoxic stimulus. **(a)** Comparison of expression ratios for nuclear and cytoplasmic libraries (0 h vs 24 h hypoxia). No significant correlation was observed between the nuclear and cytoplasmic expression ratios. **(b)** Heatmap of common differentially expressed miRNAs upon hypoxia in both nuclear and cytoplasmic samples. **(c)** Heatmap of the differentially expressed miRNAs in cytoplasmic fraction only. **(d)** Heatmap of the differentially expressed miRNAs in nuclear fraction only. Scale bars show mean-centered  $\log_2$  normalized counts (row Z-score) where red and blue indicate higher and lower than mean abundance respectively. **(e)** miR-210-3p was identified to be highly up-regulated upon hypoxia in the nuclear fraction. **(f)** miR-210-3p expression in nuclear and cytoplasmic extracts was validated with RT-qPCR. miRNA-levels were normalized to miR-186-5p levels and the mean value of the Cytoplasm 0 h group scaled to a value of one. Values are mean  $\pm$  SEM,  $n = 2$  (sRNA-seq),  $n = 3$  (RT-qPCR),  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  (one-way ANOVA and Bonferroni *post hoc* test, statistical comparisons are to the Cytoplasm 0 h group unless otherwise indicated).

or the endogenous reference (miR-186-5p) (Supplemental Fig. S5). miRNA expression levels and changes upon hypoxia in cytoplasmic and nuclear compartments were similar with both normalization methods used, indicating that the expressional changes observed are not related to possible different total RNA amount of different samples.



**Figure 4.** miR-210-3p localization upon hypoxic stimulus. miR-210-3p localization in C166 cells in normoxia and hypoxia (24 h) using fluorescent *in situ* hybridization (FISH). miR-210-3p is found in nucleus in normoxia and hypoxia. DAPI was used to stain nuclei of the cells. miR-210-3p is stained in red. To better display the co-localization of nuclear and miR-210-3p staining, areas within green and red rectangles show the view to the confocal microscopy image stack from the side.

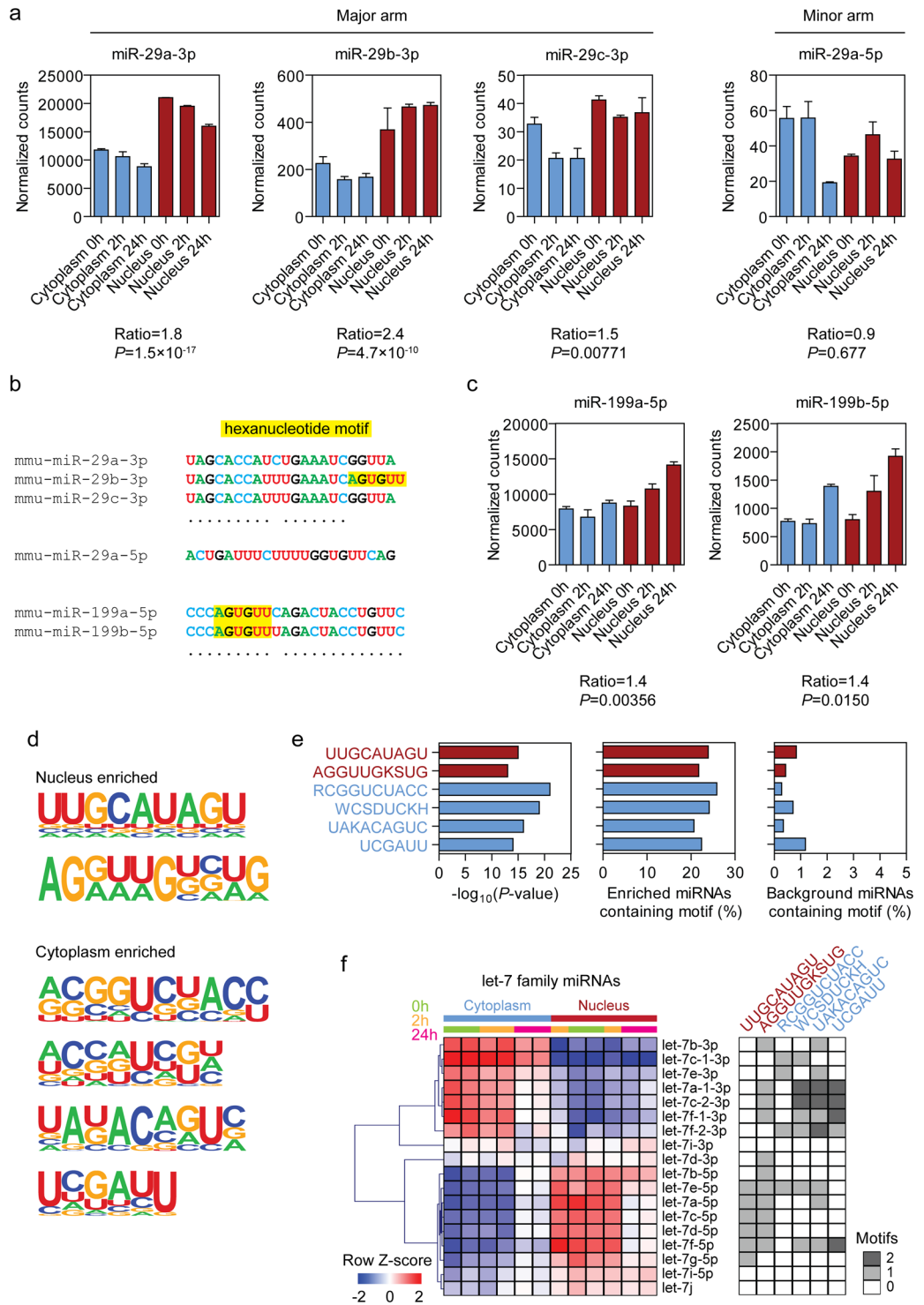
#### Analysis of sequence motifs associated with enrichment in nuclear and cytoplasmic fractions.

Given that it has been previously reported that a hexanucleotide motif (5'-AGUGUU-3') located at the 3' terminus of miR-29b-3p was capable for directing this miRNA to the nucleus<sup>14</sup> we sought to determine the nucleocytoplasmic distribution of miR-29 family members in our sRNA-seq data. miR-29b-3p was found to be significantly enriched in the nucleus by 2.4-fold (Fig. 5a). However, miR-29a-3p and miR-29c-3p (which lack the hexanucleotide motif) were similarly found to be significantly nuclear-enriched, although to a lesser extent; 1.8-fold and 1.5-fold respectively. The minor arm of miR-29a (i.e. miR-29a-5p) was also detected in our sequencing data, and was not significantly enriched in either fraction (Fig. 5a). We next searched for other detected miRNAs containing the hexanucleotide motif. Only two such miRNAs were identified; miR-199a-5p and miR-199b-5p, in which the motif was located towards the 5' end, starting at position 4 (Fig. 5b). These miRNAs were modestly nuclear-enriched by 1.4-fold and 1.5-fold respectively, but dynamically changed their nucleocytoplasmic distribution upon hypoxic stimulus. Furthermore, in the unperturbed normoxic state (time 0 h) miR-199a-5p and miR-199b-5p were expressed at equivalent levels in both compartments (Fig. 5c). Taken together, these findings suggest that the miR-29b-3p hexanucleotide motif is not a *bone fide* nuclear-enrichment signal, at least in this context. Notably, several other studies have failed to detect enrichment of this motif in nuclear sRNA libraries, consistent with the findings reported here<sup>11,12</sup>.

We next used *de novo* motif analysis to identify potential signals that might confer enrichment in either the nucleus or the cytoplasm. Two motifs were significantly enriched in nuclear-localized miRNAs, whereas four motifs were enriched in cytoplasm-localized miRNAs (Fig. 5d,e). Upon inspection of motif annotation results, it became apparent that in many cases, motifs were commonly identified in related miRNA family members. This effect was most strongly apparent for the let-7 family (Fig. 5f). As such, the empirically-determined nuclear enrichment signals 5'-UUGCAUAGU-3' and 5'-AGGUUGKSUG-3' (these motifs also have the potential to overlap) were found to be present in almost all of the let-7 family members that exhibited nuclear enrichment. Interestingly, nuclear enrichment was observed only for the 5' arm (i.e. the dominant arm) of the let-7 family members. In contrast, 3' arm let-7 family members (i.e. the minor arm that is expressed with ~10 times fewer read counts than the dominant arm) were almost exclusively enriched in the cytoplasm, and contained many of the empirically-identified cytoplasmic enrichment motifs. These findings demonstrate that the 5' and 3' arms of let-7 family miRNAs exhibit opposite nucleocytoplasmic localization. The detection of multiple related miRNAs is therefore likely to have skewed the motif enrichment analysis. It is consequently non-trivial to determine if the identified nuclear enrichment motifs are directing their associated miRNAs to distinct subcellular locations, or if they are identified as a consequence of multiple co-regulated miRNAs with high sequence similarity biasing the motif enrichment analysis. Specific details of the *de novo* motif enrichment analysis are described in Supplementary File S1.

#### Discussion

In this study we have profiled the subcellular localization of miRNAs in endothelial cells, and identified a set of miRNAs which exhibit nuclear enrichment. Furthermore, induction of hypoxia resulted in differential miRNA expression changes that were spatially restricted to either the nucleus or cytoplasm, or a small number of changes that occurred simultaneously in both compartments. This is the first study to profile subcellular



**Figure 5.** Analysis of putative miRNA sequence motifs affecting subcellular localization. **(a)** sRNA-seq normalized counts values for C166 nuclear and cytoplasmic fractions, with and without hypoxic stress, for the major arms of miR-29 family, and the minor arm of miR-29a-5p. **(b)** Alignments of all detected miR-29 miRNA family members, and miR-199a/b-5p. The positions of the 5'-AGUGUU-3' hexanucleotide motif are indicated in yellow. **(c)** Normalized sRNA-seq counts for miR-199a/b-5p. **(d)** Nuclear and cytoplasmic enrichment sequences identified by *de novo* motif enrichment. **(e)** Output statistics for *de novo* motif discovery; enrichment *P*-values, percentage of enriched miRNAs in containing each motif, and percentage of non-enriched miRNAs (i.e. background) containing each motif. (For ambiguous nucleotides; K = U or G, S = G or C, R = A or G, W = A or U, D = A, G, or U, and H = A, C, or U). **(f)** Heatmap of sRNA-seq data for let-7 family members. Scale bars show mean-centered log<sub>2</sub> normalized counts (row Z-score) where red and blue indicate higher and lower than mean abundance respectively. Occurrences of each enrichment motif in each family member are indicated in the greyscale heatmap. All values are mean ± SEM, *n* = 2 (sRNA-seq).

compartment-specific differential miRNA expression, and provides a strong rationale for the investigation of nuclear and cytoplasmic miRNA levels in other biological contexts.

Hypoxia-associated changes in miRNA levels in the nucleus are highly suggestive of functional significance. However, the mechanisms-of-action for nuclear miRNAs are at present incompletely understood. Studies of TGS and TGA mediated by exogenous small RNAs provided early clues that endogenous miRNAs might regulate transcription in a similar manner. Indeed, there have now been multiple studies that have demonstrated miRNA-mediated transcriptional regulation<sup>7,26–33</sup>. It is generally thought that in these instances, the miRNA directs an Argonaute-containing protein complex to non-coding, promoter-associated transcripts<sup>26,27,29</sup>. Exogenous small RNAs have also been shown to regulate pre-mRNA splicing decisions<sup>34–36</sup>, although such effects have yet to be demonstrated for endogenous miRNAs. The observation that siRNAs are capable of silencing nuclear-localized target transcripts implied a similar activity for miRNAs<sup>16,37,38</sup>. Indeed, a recent study by Sarshad *et al.*, showed that AGO2 protein is highly expressed in the nuclei of stem cells, where it mediates the silencing of nuclear transcripts, such as pre-mRNAs<sup>39</sup>. Aside from the regulation of gene expression, the miRNA processing machinery has also been implicated in the cellular response to DNA damage<sup>40,41</sup>, the regulation of 3-dimensional chromatin structure<sup>42</sup>, and regulation of the production of intron-included, truncated proteins as a consequence of direct interactions between the miRNA and genomic DNA<sup>43</sup>.

Relative to the enormous research effort directed towards canonical miRNA activity, the nuclear targets and functions of specific miRNAs have so far been largely neglected. Further studies are needed to elucidate the mechanisms through which nuclear miRNAs execute their functions in different cell types and physiological conditions. Given the diversity of functional possibilities this presents a significant task. Analogous to the manner by which the post-transcriptional regulation of mRNA expression and translation can be fine-tuned by multiple miRNAs binding on multiple co-acting sites<sup>44</sup>, multiple non-canonical mechanisms mediated by multiple miRNA may be acting on a gene in the nucleus at any one time. (For example, miRNAs also have the potential to target mRNA coding regions<sup>45</sup> and intronic sequences in newly transcribed pre-mRNA<sup>39</sup> simultaneously). One approach is to predict miRNA target sequences in regions that are considered non-canonical targets<sup>46</sup>. However, this strategy does not take into account the transcriptional landscape of the cell of interest, which is expected to be highly lineage specific (e.g. a gene expressed at very high levels might not be susceptible to gene activation). Predictions are also expected to have a high false positive rate, and so extensive experimental validation is required to identify *bona fide* miRNA-target interactions. Alternatively, crosslinking immunoprecipitation techniques can be used to empirically determine miRNA binding events<sup>47,48</sup>. These techniques suffer from the limitation of having relatively low resolution, and of being technically challenging.

miRNAs constitute a promising class of therapeutic target molecules that can be inhibited using antisense oligonucleotides or mimicked using synthetic siRNAs, shRNAs or artificial miRNAs<sup>49</sup>. Our group has previously demonstrated transcriptional regulation of *Vegfa* expression by promoter-targeted shRNAs *in vitro* and *in vivo*<sup>9,10</sup>. Depending on their target site, these shRNAs are able to either up-regulate (TGA) or down-regulate (TGS) *Vegfa* expression<sup>9</sup>. These shRNAs were shown to exert positive therapeutic effects in both hindlimb ischemia and myocardial infarction models<sup>9,10</sup>. Similarly, an siRNA which activates CEBPA expression<sup>50</sup> is being explored as a treatment for hepatocellular carcinoma with patients currently being enrolled for a phase I clinical trial sponsored by MiNA Therapeutics<sup>51</sup>. As such, nuclear miRNA functions offer new possibilities for the manipulation of gene expression, or may be novel targets in their own right.

Given that altered miRNA expression in hypoxia gives rise to multiple gene expression changes associated with canonical miRNA function<sup>17</sup>, our data suggest that nuclear miRNAs may similarly regulate a wide variety of hypoxia-associated genes through non-canonical mechanisms. This is exemplified by the well-established hypoxamiR miR-210-3p, which was significantly increased in the nucleus of the cells upon hypoxic stimulus, and therefore likely regulates a plethora of nuclear targets. This notion is supported by a previous report showing the interaction between miR-210-3p and the nuclear lncRNA *XIST*<sup>25</sup>. Indeed, interaction between small RNAs and lncRNAs is a common theme. For example, miR-9 has been shown to downregulate *MALAT1* in the nucleus<sup>52</sup>, and promoter-associated transcripts have been shown to regulate transcription of the *VEGFA* locus by acting as targets for TGA-inducing small RNA<sup>53</sup>.

Notably, hypoxic conditions can vary in different physiological contexts with corresponding differences in the cellular gene expression response. This was demonstrated in a recent study in the context of the tumor microenvironment where the commonly-used, continuous, long-term hypoxia condition was compared to a shorter cyclic hypoxia protocol<sup>54</sup>. Previously-described hypoxamiRs were identified in the long-term hypoxia model, but cyclic hypoxic conditions resulted in a different expression pattern, with only 31 common miRNAs between these groups (including miR-210-3p). These observations illustrate the complexity of ncRNA regulation in hypoxia, an issue compounded when considering the possible functions of miRNAs in the nucleus.

Components of the RNAi machinery such as AGO2, Dicer, and GW182 (TRNC6A) have been shown to be present in the nucleus<sup>16</sup>. Given that Dicer processing and RISC loading are generally considered to be restricted to the cytoplasm<sup>16</sup>, this would necessarily require mechanisms for the re-import of miRNAs to the nucleus. The karyopherins exportin-1 (XPO1)<sup>55</sup> and importin-8 (IPO8)<sup>56</sup> have been implicated in the transport of miRNA shuttling across the nuclear envelope. Recently, it was also shown the stress-induced response complex (SIRC) (consisting of AGO1, AGO2, YB1, CTCF, FUS, SMAD1, SMAD3, and SMAD4 proteins) is involved in nucleocytoplasmic trafficking of miRNAs in a cellular stress-dependent manner<sup>57</sup>. As such, RNA binding proteins and complexes like SIRC may regulate the subcellular distribution of miRNAs, as we observe also in response to hypoxic stimulus. Some studies have suggested that short sequence motifs might act as nuclear localization signals, as in the case of the 3' terminal hexanucleotide motif in miR-29b-3p, presumably by recognizing specific RNA binding proteins such as ANT2 (ADP/ATP translocase 2)<sup>58</sup>. However, while we did observe a modest nuclear enrichment of miR-29b-3p, the related miRNAs miR-29a-3p and miR-29c-3p which lack the hexanucleotide motif were similarly nuclear-enriched. We have previously suggested an alternative hypothesis whereby



miRNAs are non-specifically shuttled between nucleus and cytoplasm and then become enriched in the compartment where their target transcripts are most concentrated<sup>2,59,60</sup>.

In conclusion, this study demonstrates that miRNAs are differentially expressed in the nucleus and cytoplasm in response to hypoxic stress. These findings offer new insights into the molecular response to hypoxia, and for the function of miRNAs in general. Studies of ncRNA networks in gene regulation by our group and others significantly expand our understanding of the role of miRNAs beyond their canonical PTGS functions, and suggest that a vibrant world of ncRNA biology and regulatory potential resides within the nucleus.

## Materials and Methods

**Cell culture.** C166 (yolk-sac-derived mouse endothelial cells, ATCC:CRL-2581), C2C12 (skeletal myoblast-derived mouse muscle cells, ATCC:CRL-1772), MS1 (pancreas/islet of Langerhans-derived mouse endothelial cells, ATCC:CRL-2279) or MOVAS (aorta-derived smooth muscle cells, ATCC:CRL-2797) cells were maintained under normal conditions (37 °C, 5% CO<sub>2</sub>). Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich) containing 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (PS). For hypoxia experiments, cells were cultured in a hypoxia chamber with 1% O<sub>2</sub>, 5% CO<sub>2</sub> (Baker Ruskinn).

**Western blot.** Protein samples were extracted from nuclear and cytoplasmic fractions according to protocol by Gagnon *et al.*,<sup>18</sup> and with TRI reagent (Molecular Research Center) according to the manufacturer's instructions. For western blot, equal volumes of extracted protein were loaded on precast gels (Mini-Protean TGX Stain-free Precast Gel, 4–20%, Bio-Rad) and transferred to nitrocellulose membranes (Trans-Blot Turbo Bio-Rad Midi, 0.2 µm nitrocellulose, Bio-Rad). The membranes were blocked with 5% milk for 1.5 hours at room temperature, washed with TBST (0.15 M sodium chloride, 0.050 M TRIS-HCl buffer; 0.05% Tween 20; pH 7.6) and incubated with antibodies against a known nuclear protein (anti-trimethyl-histone H3 (Lys27), Millipore, 1:2,500) and a cytoplasmic protein (anti-β-tubulin, Sigma Aldrich, 1:1,000) overnight at 4 °C. The membranes were washed with TBST and incubated with secondary antibodies (goat anti-rabbit IgG (H + L), HRP-conjugated, Invitrogen; anti-mouse IgG, HRP-conjugated, R&D systems; both 1:5,000) for 1 hour at room temperature. Membranes were analyzed using ECL Plus Western Blotting Substrate (Pierce) and imaged with the ChemiDoc Imaging System (Bio-Rad).

**RNA extraction and small RNA-sequencing.** For sRNA-sequencing samples, C166 cells were cultured in hypoxic conditions for 0, 2 or 24 hours and cells were separated into nuclear and cytoplasmic fractions, as described previously<sup>18</sup>. Briefly, the cells were collected from 15 cm plates by scraping and washed once with cold PBS. Cells were pelleted by centrifugation at 700 g for 5 minutes and lysed with hypotonic lysis buffer (10 mM Tris (pH 7.5), 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.3% (vol/vol) NP-40, 10% (vol/vol) glycerol) to collect the cytoplasmic fraction. Cytoplasmic RNA was obtained by ethanol precipitation overnight at –20 °C followed by re-extraction using TRI reagent. The remaining nuclear pellet was washed three times with the hypotonic lysis buffer, followed by extraction with TRI reagent according to the manufacturer's instructions. Libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina kit (New England Biolabs). Equimolar quantities of each library were pooled and small RNA sequencing performed as a service by Exiqon A/S (Vedbaek, Denmark). FASTQ files were trimmed using Trimmomatic<sup>61</sup> and aligned to reference genome (miRBase release 20) using bowtie2 (v2.1.0)<sup>62</sup>. Differential expression was tested using DESeq2<sup>20</sup> which utilizes the Wald test for determining statistical significance. (Benjamini-Hochberg adjusted *P*-values are reported). Counts data for the full dataset are provided in Supplementary Data S1. sRNA-seq data were analyzed using the NormFinder method<sup>21</sup> which identified miR-186-5p as being stably expressed between all experimental groups. This miRNAs was subsequently selected as a reference gene for RT-qPCR data normalization. Alternatively, 15 fmol of cel-miR-39 ssRNA was added to the samples during phenol extraction phase and was used as spike-in to calculate relative expression of miRNAs.

**RT-qPCR.** Total RNA samples were treated with DNase I, RNase-free (Thermo Fisher Scientific) in order to eliminate genomic DNA contamination. For miRNA analysis, cDNA synthesis was performed using the TaqMan MicroRNA Reverse Transcription Kit (Thermo Fisher Scientific) according to manufacturer's protocol and analyzed using miRNA-specific TaqMan assays (mmu-miR-27a-5p ID: 002445; mmu-miR-3535 ID: CTEPR23; mmu-miR-186-5p, ID: 002285, mmu-miR-210-3p ID: 000512; mmu-miR-1291, ID: 466942\_mat; cel-miR-39-3p, ID: 000200; Thermo Fisher Scientific).

For expression analysis of lncRNAs and tRNAs, cDNA was synthesized using RevertAid Reverse Transcriptase (Thermo Scientific) and random hexamer primers (lncRNA) or gene-specific primers (tRNA) (reverse primers, Supplemental Table 1) and quantified using Maxima SYBR Green/ROX qPCR Master Mix (2 ×) (Thermo Fisher Scientific). Thermal cycling was performed using a LightCycler480 (Roche) with the following program: 10 min at 95 °C, followed by 50 cycles of 15 s at 95 °C and 60 s at 60 °C. Primers for lncRNA and tRNA analysis are shown in Supplemental Table 1. RT-qPCR data were analyzed using the  $\Delta\Delta C_q$  method where normalization was available, or the  $\Delta C_q$  method for un-normalized data (validation of nuclear-cytoplasmic fractionation only). Statistical significance was assessed by one-way ANOVA with Bonferroni *post hoc* correction or *t*-test with Welch correction as appropriate (GraphPad Prism 5 (GraphPad Software, La Jolla, CA, USA)).

**miRNA fluorescent *in situ* hybridization (FISH).** For miRNA FISH, C166 cells were seeded on 8-well chamber slides and grown in normoxia or hypoxia (24 h). ViewRNA miRNA ISH Cell Assay Kit (Thermo Fisher Scientific) was used for the hybridizations according to the manufacturer's protocol. ViewRNA Cell Plus Probe Set (Thermo Fisher Scientific) was used to detect miR-210-3p (assay ID: VM1-10263-VCP, detection label Alexa

Fluor 546). Cell nuclei were visualized using DAPI stain. Pictures were taken using ZEISS LSM700 confocal microscope using 40× oil objective and analyzed with ZEN lite blue 2.6 software (ZEISS).

**de novo motif enrichment.** Sequence motifs were identified in lists of miRNA sequences (enriched in either nucleus or cytoplasm) using HOMER (v4.9.1)<sup>63</sup>. Motifs were restricted to between 6 and 10 nucleotides, and analysis run in strand-specific mode using the findMotifsGenome.pl function. miRNA sequences were subsequently annotated to identify those which contained one or more HOMER-identified motifs using the annotatePeaks.pl function.

## Data Availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus<sup>64</sup> and are accessible through GEO Series accession number GSE125390 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125390>).

## References

- Filipowicz, W., Jaskiewicz, L., Kolb, F. A. & Pillai, R. S. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol.* **15**, 331–41 (2005).
- Roberts, T. C. The MicroRNA Biology of the Mammalian Nucleus. *Mol Ther Nucleic Acids* **3**, e188 (2014).
- Leung, A. K. L. The Whereabouts of microRNA Actions: Cytoplasm and Beyond. *Trends Cell Biol.* **25**, 601–610 (2015).
- Zhang, X. *et al.* MicroRNA Directly Enhances Mitochondrial Translation during Muscle Differentiation. *Cell* **158**, 607–619 (2014).
- Roberts, T. C. & Wood, M. J. A. Non-Canonical MicroRNA Biogenesis and Function in *Applied RNAi: From Fundamental Research to Therapeutic Applications* (eds Arbutnot, P. & Weinberg, M. S.) 19–42 (Caister Academic Press, 2014).
- Li, L.-C. *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. USA* **103**, 17337–17342 (2006).
- Place, R. F., Li, L.-C., Pookot, D., Noonan, E. J. & Dahiya, R. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc. Natl. Acad. Sci. USA* **105**, 1608–1613 (2008).
- Laham-Karam, N., Laitinen, P., Turunen, T. A. & Ylä-Herttuala, S. Activating the chromatin by non-coding RNAs. *Antioxid. Redox Signal.* **29**, 813–831 (2017).
- Turunen, M. P. *et al.* Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism a novel example of epigenotherapy. *Circ. Res.* **105**, 604–609 (2009).
- Turunen, M. P. *et al.* Epigenetic upregulation of endogenous VEGF-A reduces myocardial infarct size in mice. *PLoS One* **9** (2014).
- Liao, J. Y. *et al.* Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of mirnas and tRNA 3' trailers. *PLoS One* **5** (2010).
- Park, C. W., Zeng, Y., Zhang, X., Subramanian, S. & Steer, C. J. Mature microRNAs identified in highly purified nuclei from HCT116 colon cancer cells. *RNA Biol.* **7**, 606–614 (2010).
- Jeffries, C. D., Fried, H. M. & Perkins, D. O. Nuclear and cytoplasmic localization of neural stem cell microRNAs. *RNA* **17**, 675–686 (2011).
- Hwang, H.-W., Wentzel, E. A. & Mendell, J. T. A hexanucleotide element directs microRNA nuclear import. *Science*. **315**, 97–101 (2007).
- Khudayberdiev, S. A., Zampa, F., Rajman, M. & Schratt, G. A comprehensive characterization of the nuclear microRNA repertoire of post-mitotic neurons. *Front. Mol. Neurosci.* **6**, 1–19 (2013).
- Gagnon, K. T., Li, L., Chu, Y., Janowski, B. A. & Corey, D. R. RNAi Factors Are Present and Active in Human Cell Nuclei. *Cell Rep.* **6**, 211–221 (2014).
- Nallamshetty, S., Chan, S. Y. & Loscalzo, J. Hypoxia: A master regulator of microRNA biogenesis and activity. *Free Radic. Biol. Med.* **64**, 20–30 (2013).
- Gagnon, K. T., Li, L., Janowski, B. A. & Corey, D. R. Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. *Nat Protoc* **9**, 2045–60 (2014).
- Hutchinson, J. N. *et al.* A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**, 39 (2007).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **64**, 5245–5250 (2004).
- Chan, Y. C., Banerjee, J., Choi, S. Y. & Sen, C. K. miR-210: the master hypoxamir. *Microcirculation* **19**, 215–223 (2012).
- Pulkkinen, K., Malm, T., Turunen, M., Koistinaho, J. & Ylä-Herttuala, S. Hypoxia induces microRNA miR-210 *in vitro* and *in vivo*. Ephrin-A3 and neuronal pentraxin 1 are potentially regulated by miR-210. *FEBS Letters* **582**, 2397–2401 (2008).
- Giannakakis, A. *et al.* miR-210 links hypoxia with cell cycle regulation and is deleted in human epithelial ovarian cancer. *Cancer Biol. Ther.* **7**, 255–264 (2008).
- Fasanaro, P. *et al.* An integrated approach for experimental target identification of hypoxia-induced miR-210. *J. Biol. Chem.* **284**, 35134–35143 (2009).
- Kim, D. H., Saetrom, P., Snoe, O. & Rossi, J. J. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci.* **105**, 16230–16235 (2008).
- Tan, Y. *et al.* Transcriptional inhibition of Hoxd4 expression by miRNA-10a in human breast cancer cells. *BMC molecular biology* **10**, 12 (2009).
- Majid, S. *et al.* MicroRNA-205 directed transcriptional activation of tumor suppressor genes in prostate cancer. *Cancer* **116**, 5637–5649 (2010).
- Benhamed, M., Herbig, U., Ye, T., Dejean, A. & Bischof, O. Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells. *Nat. Cell Biol.* **14**, 266–275 (2012).
- Zardo, G. *et al.* Polycombs and regulate human granulopoiesis by transcriptional control of target gene expression. *Blood* **119**, 4034–4047 (2012).
- Adilakshmi, T., Sudol, I. & Tapinos, N. Combinatorial action of mirnas regulates transcriptional and post-transcriptional gene silencing following *in vivo* pns injury. *PLoS One* **7** (2012).
- Matsui, M. *et al.* Promoter RNA links transcriptional regulation of inflammatory pathway genes. *Nucleic Acids Res.* **41**, 10086–10109 (2013).
- Wang, C. *et al.* Up-regulation of p21WAF1/CIP1 by miRNAs and its implications in bladder cancer cells. *FEBS Lett.* **588**, 4654–4664 (2014).
- Alló, M. *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **16**, 717–724 (2009).

35. Liu, J., Hu, J. & Corey, D. R. Expanding the action of duplex RNAs into the nucleus: Redirecting alternative splicing. *Nucleic Acids Res.* **40**, 1240–1250 (2012).
36. Kornblihtt, A. R. Chromatin, transcript elongation and alternative splicing. *Nature structural & molecular biology* **13**, 5–7 (2006).
37. Robb, G. B., Brown, K. M., Khurana, J. & Rana, T. M. Specific and potent RNAi in the nucleus of human cells. *Nat. Struct. Mol. Biol.* **12**, 133–137 (2005).
38. Langlois, M. A. *et al.* Cytoplasmic and nuclear retained DMPK mRNAs are targets for RNA interference in myotonic dystrophy cells. *J. Biol. Chem.* **280**, 16949–16954 (2005).
39. Sarshad, A. A. *et al.* Argonaute-miRNA Complexes Silence Target mRNAs in the Nucleus of Mammalian Stem Cells. *Mol. Cell* **71**, 1040–1050.e8 (2018).
40. Tang, K. F. *et al.* Decreased Dicer expression elicits DNA damage and up-regulation of MICA and MICB. *J. Cell Biol.* **182**, 233–239 (2008).
41. Francia, S. *et al.* Site-specific DICER and DROSHA RNA products control the DNA damage response. *Nature* **488**, 231–235 (2012).
42. Shuaib, M. *et al.* AGO1 in association with NEAT1 lncRNA contributes to nuclear and 3D chromatin architecture in human cells. *bioRxiv*, <https://doi.org/10.1101/525527> (2019).
43. von Brandenstein, M. *et al.* Beyond the 3'UTR binding-microRNA-induced protein truncation via DNA binding. *Oncotarget* **9**, 32855–32867 (2018).
44. Fang, Z. & Rajewsky, N. The impact of miRNA target sites in coding sequences and in 3'UTRs. *PLoS One* **6**, 1–6 (2011).
45. Hausser, J., Syed, A. P., Bilen, B. & Zavolan, M. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.* **23**, 604–615 (2013).
46. Younger, S. T., Pertsemliadis, A. & Corey, D. R. Predicting potential miRNA target sites within gene promoters. *Bioorganic Med. Chem. Lett.* **19**, 3791–3794 (2009).
47. Zhang, X., Shen, B. & Cui, Y. Ago HITS-CLIP expands microRNA-mRNA interactions in nucleus and cytoplasm of gastric cancer cells. *BMC Cancer* **19**, 29 (2019).
48. Chi, S. W., Zang, J. B., Mele, A. & Darnell, R. B. Ago HITS-CLIP decodes miRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
49. Roberts, T. C. & Wood, M. J. A. Therapeutic targeting of non-coding RNAs. *Essays Biochem.* **54**, 127–145 (2013).
50. Voutila, J. *et al.* Development and Mechanism of Small Activating RNA Targeting CEBPA, a Novel Therapeutic in Clinical Trials for Liver Cancer. *Mol. Ther.* **25**, 2705–2714 (2017).
51. MiNA Therapeutics. MiNA Therapeutics Announces Enrolment of Patients in Expansion of Phase Ib Trial Evaluating MTL-CEBPA in Combination with Sorafenib. *Press release*, <https://www.minatx.com/wp-content/uploads/2018/12/> (2018).
52. Leucci, E. *et al.* microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Sci. Rep.* **3**, 1–6 (2013).
53. Lopez, P., Wagner, K.-D., Hofman, P. & Van Obberghen, E. RNA Activation of the Vascular Endothelial Growth Factor Gene (VEGF) Promoter by Double-Stranded RNA and Hypoxia: Role of Noncoding VEGF promoter transcripts. *Mol. Cell. Biol.* **36**, 1480–93 (2016).
54. Kochan-Jamroz, K., Króliczewski, J., Moszyńska, A., Collawn, J. F. & Bartoszewski, R. miRNA networks modulate human endothelial cell adaptation to cyclic hypoxia. *Cell. Signal.* **54**, 150–160 (2019).
55. Castanotto, D., Lingeman, R., Riggs, A. D. & Rossi, J. J. CRM1 mediates nuclear-cytoplasmic shuttling of mature microRNAs. *Proc. Natl. Acad. Sci. USA* **106**, 21655–9 (2009).
56. Wei, Y., Li, L., Wang, D., Zhang, C. Y. & Zen, K. Importin 8 regulates the transport of mature microRNAs into the cell nucleus. *J. Biol. Chem.* **289**, 10270–10275 (2014).
57. Castanotto, D. *et al.* A stress-induced response complex (SIRC) shuttles miRNAs, siRNAs, and oligonucleotides to the nucleus. *Proc. Natl. Acad. Sci.* **115**, E5756–E5765 (2018).
58. Kriegel, A. J. *et al.* Isomer-specific effect of microRNA miR-29b on nuclear morphology. *J. Biol. Chem.* **293**, 14080–14088 (2018).
59. Berezina, S. Y., Supekova, L., Supek, F., Schultz, P. G. & Deniz, A. A. siRNA in human cells selectively localizes to target RNA sites. *Proc. Natl. Acad. Sci. USA* **103**, 7682–7687 (2006).
60. Ahlenstiel, C. L. *et al.* Direct evidence of nuclear Argonaute distribution during transcriptional silencing links the actin cytoskeleton to nuclear RNAi machinery in human cells. *Nucleic Acids Res.* **40**, 1579–1595 (2012).
61. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
63. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
64. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).

## Acknowledgements

This work was funded by Academy of Finland Center of Excellence and project funding #268376, The Finnish Foundation for Cardiovascular Research, Instrumentarium Science Foundation and ERC advanced grant (CleverGenes).

## Author Contributions

M.P.T. and T.C.R. designed the study. T.A.T., P.L., M.-A.V. and P.K. performed the experiments. T.C.R. analyzed sequencing data. T.A.T., T.C.R., S.Y.-H. and M.P.T. wrote the paper. T.M., M.P.T. and S.Y.-H. acquired funding.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-46841-1>.

**Competing Interests:** The authors declare no competing interests.

**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) 2019