Published in final edited form as:

Trends Genet. 2018 March ; 34(3): 165–167. doi:10.1016/j.tig.2017.12.015.

Big Strides in Cellular microRNA Expression

Marc K. Halushka¹, Bastian Fromm², Kevin J. Peterson³, and Matthew N. McCall⁴

¹Department of Pathology, Johns Hopkins University School of Medicine, Baltimore Maryland, USA, 21205

²Department of Tumor Biology, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway, 0424

³Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire, USA, 03755

⁴Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, New York, USA, 14642

Abstract

A lack of knowledge of the cellular origin of microRNAs (miRNAs) has greatly confounded functional and biomarkers studies. Recently, three studies characterized miRNA expression patterns across >78 human cell types. These combined data expand our knowledge of miRNA expression localization and confirm that many miRNAs show cell-type specific expression patterns.

Keywords

microRNA; RNA-seq; cell expression

Since their discovery nearly 25 years ago, much progress has been made on understanding the roles miRNAs play in gene regulatory networks. One of the primary insights into their functionality has been the global characterization of miRNA expression at the tissue/organ level. The general observation has been that some miRNAs have conserved organ-specific expression with other miRNAs having ubiquitous expression. This has led to the idea that many miRNAs have broad expression profiles across cell types, especially in humans. However, tissues and organs are a collection of cell types. Some cell types, such as endothelial cells, red blood cells, and fibroblasts, are found in essentially all organs and the miRNAs they may exclusively express may be misinterpreted as ubiquitous [1–3]. Thus a need for cell specific data to clarify the localization of miRNAs in a cell-by-cell fashion is long overdue.

Corresponding Author: Halushka, M.K., mhalush1@jhmi.edu, <http://labs.pathology.jhu.edu/halushka/>, @Marc_Halushka.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Recently, three groups, the FANTOM5 consortium [4], the Hemmrich-Stanisak laboratory [5] and two of us [6] published complementary papers that provide the first deep look at primary cell miRNA expression using small RNA-seq data. The FANTOM5 consortium performed small RNA-seq on 396 human samples across 118 cell types obtained from different organs [4]. The Hemmrich-Stanisak laboratory performed magnetic activated cell sorting (MACS) separation of 43 buffy coats in addition to serum and whole blood sample collection. This yielded multiple libraries of 7 hematopoietic cell types [5]. Our group grew 34 primary cell types in culture, flow sorted 2 cell types, then scoured the Sequence Read Archive for additional primary and cancer cell data obtaining an additional 226 samples (126 primary cell and 100 cancer cell lines) all with a minimum of 1 million miRNA reads [6]. In combination, the three teams generated small RNA-seq data from over a thousand cell-based samples. The data were both complementary and filled gaps in each other's datasets.

A key point of the two more global surveys is that cellular microRNA expression is generally not ubiquitous. The FANTOM5 group repurposed a "tissue specificity index" [7] into a "cell type specificity index" in which they noted many miRNAs including miR-122-5p, miR-142-5p and miR-302a-5p being enriched in specific cell types. Other miRNAs, such as miR-100-5p and miR-29a-3p could be broadly expressed, but reduced in specific cell types. The McCall paper identified 13 ubiquitously expressed miRNAs, present at a threshold of 100 reads per million miRNA reads (RPM) across 46 cell types, and 94 cell specific miRNAs, present at a level of 1,000 RPM or greater in only one class of cells.

In unison, these works reported the discovery of thousands of additional "novel miRNAs." Interestingly, the studies of McCall and the FANTOM5 consortium in particular express their own concerns that many of their predictions do not fulfill the basic properties of genuine miRNAs, and indeed this miRNA inflation is concerning as humans only share ~450 distinct miRNAs with the macaque [8]. Applying well understood annotation criteria reduces these thousands of novel miRNA loci to, at best, a handful of genes that are truly novel miRNAs, albeit miRNAs that are evolutionary very recent, and thus probably play very minor roles (if any) in translational regulation in human cell types [8]. A serious discussion as to what to name and define these non-canonical RNAs is sorely needed.

Surprisingly, there is a tremendous amount of variability in the RPM values between multiple cultures of the same cell type. Some of this may be related to the use of RPM as a normalization tool [3] or technical variation/batch effects [9]. But it may also suggest fluctuating biology in cells. For example, in three donors of brachiocephalic smooth muscle cells from the FANTOM5 data, the RPM values for miR-100-5p ranged from 40,000 to nearly 340,000. This nearly 9-fold difference in one of the most abundant miRNAs among these cells is surprising and would impact the RPM value of other miRNAs [3]. To estimate the variability in miRNA expression within cell types, it is crucial to include data that capture both biological and technical sources of variation. For example, by merging the data from these three studies, we now have access to 46 fibroblast samples from 4 different studies, allowing us to assess both within and between study variation. This is essential for downstream analyses such as tissue deconvolution with cell-specific miRNA signatures for disease studies (Figure).

As understanding the cell type specific expression of miRNA signal is so useful, all three groups have made their data publically available. The FANTOM5 group built an excellent webserver that is searchable by miRNA, cell type or promoter location for both human and mouse (http://fantom.gsc.riken.jp/5/suppl/De_Rie_et_al_2017/vis_viewer/). The Hemmrich-Stanisak laboratory has a webtool for miRNA comparisons across the blood cells (<http://134.245.63.235/ikmb-tools/bloodmiRs/>). The McCall group made their data available as an R package (<http://bioconductor.org/packages/microRNAome>) and combined the data from the three manuscripts (analyzing all data with miRge [10]) into a barChart track at the UCSC Genome Browser (<http://www.genome.ucsc.edu/cgi-bin/hgHubConnect>). These tools nicely complement a tissue-level miRNA atlas (<https://ccb-web.cs.uni-saarland.de/tissueatlas/patterns>) by the Keller group [7]. Finally, MirGeneDB.org has been created to serve as a repository for hand-curated, high-confident miRNA genes [8].

As much as this cell specific miRNA expression data will clear up confusion in the literature, it also points to some gaps in our collective understanding of miRNA expressions. For example, among the sampled cell types, miR-1 is only found in skeletal and cardiac myocytes. However, a consistent, low miR-1-3p signal can be detected in organs including the prostate, breast, and colon which are devoid of skeletal muscle cells. This could suggest the appropriate miR-1-3p expressing cell is not present in the combined atlases. Another interpretation is that miR-1 is expressed in a cell type that was covered, but that the miR-1 expression was lost in culture. Support for the later contention comes from the Levenon group [11] who compared the miRNA RNA-seq expression signatures of human umbilical vein endothelial cells (HUVECs) from the time of isolation through 3 passages. They specifically noted a loss of miR-126-3p from ~130,000 transcripts per million (TPM) down to ~55,000 TPM in passaging. Conversely, miRNAs of the ~17–92 cluster, thought to be involved in cellular proliferation, were 3–6 fold higher in passage 3 cells [11]. miR-126 is abundant and cell enriched in endothelial cells. Its lower levels in culture may indicate that many cell-restricted miRNAs that have specific tasks in matured cells, are reduced while cells are actively proliferating and in a less mature phenotypic state in cell culture. Indeed, from our data [6], the levels of miR-200c, an epithelial-specific miRNA, is higher in bladder (~60,000 RPM) than in urothelial cells (~5,000 RPM; the epithelial cell of the bladder wall). No other cell in bladder can account for the discrepancy, supporting this idea of reduced maturity-related miRNA expression.

Unlike the cell culture data of the FANTOM5 and McCall reports, the Hemmrich-Stanisak miRNA expression data was built upon MACS-sorted cells [5]. This data would not have the same miRNA alterations observed from cell culture and should hew closer to accurate *in vivo* data. However, isolation methods (flow sorting vs positive/negative bead selection) greatly impact miRNA expression variability [12]. In our experience, the extended time periods in buffers from similar activities (flow sorting) reduces the miRNA yield relative to other small RNAs. Indeed, for many cell types, the relative abundance of miRNAs in their samples was low. For example, of 40 neutrophil samples with at least 1 million total reads, only 3 had >500,000 miRNA reads (median 96,000).

Thus, while the curtain has been pulled back on cellular miRNA expression by the evidence of these papers, another curtain appears behind it. To pull this second curtain back, we must

develop strategies to isolate cells directly from tissues. What this successful technology will be is not yet clear, but without this information, our understanding of the roles miRNAs play in cellular differentiation and homeostasis remains incomplete. Accurate miRNA expression estimates are even more important as we learn about the importance of the relative abundance of miRNAs to their targets and sponges [13].

All together, these insights and resources greatly advance miRNA research.

Acknowledgments

M.K.H. and M.N.M are supported by grant 1R01HL137811 from the NIH. M.K.H. is also supported from an American Heart Association Grant-in-Aid (17GRNT33670405). M.N.M. is also supported by the University of Rochester CTSA award number UL1 TR002001 from the National Center for Advancing Translational Sciences of the National Institutes of Health. B.F. is supported by the South-Eastern Norway Regional Health Authority (grant number 2014041). K.J.P. is supported by NASA-Ames.

References

1. Kent OA, et al. Lessons from miR-143/145: the importance of cell-type localization of miRNAs. *Nucleic Acids Res.* 2014; 42(12):7528–38. [PubMed: 24875473]
2. McCall MN, et al. MicroRNA profiling of diverse endothelial cell types. *BMC Med Genomics.* 2011; 4:78. [PubMed: 22047531]
3. Witwer KW, Halushka MK. Toward the promise of microRNAs - Enhancing reproducibility and rigor in microRNA research. *RNA Biol.* 2016; 13(11):1103–1116. [PubMed: 27645402]
4. de Rie D, et al. An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat Biotechnol.* 2017; 35(9):872–878. [PubMed: 28829439]
5. Juzenas S, et al. A comprehensive, cell specific microRNA catalogue of human peripheral blood. *Nucleic Acids Res.* 2017; 45(16):9290–9301. [PubMed: 28934507]
6. McCall MN, et al. Toward the human cellular microRNAome. *Genome Res.* 2017; 27(10):1769–1781. [PubMed: 28877962]
7. Ludwig N, et al. Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* 2016; 44(8):3865–77. [PubMed: 26921406]
8. Fromm B, et al. A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu Rev Genet.* 2015; 49:213–42. [PubMed: 26473382]
9. Pritchard CC, et al. MicroRNA profiling: approaches and considerations. *Nat Rev Genet.* 2012; 13(5):358–69. [PubMed: 22510765]
10. Baras AS, et al. miRge - A Multiplexed Method of Processing Small RNA-Seq Data to Determine MicroRNA Entropy. *PLoS One.* 2015; 10(11):e0143066. [PubMed: 26571139]
11. Kuosmanen SM, et al. MicroRNA Profiling Reveals Distinct Profiles for Tissue-Derived and Cultured Endothelial Cells. *Sci Rep.* 2017; 7(1):10943. [PubMed: 28887500]
12. Schwarz EC, et al. Deep characterization of blood cell miRNomes by NGS. *Cell Mol Life Sci.* 2016; 73(16):3169–81. [PubMed: 26874686]
13. Pinzon N, et al. microRNA target prediction programs predict many false positives. *Genome Res.* 2017; 27(2):234–245. [PubMed: 28148562]

Figure.

Robust cell-specific miRNA expression data can be used to improve the analysis of tissue miRNA expression data. This hypothetical example shows how one could use such data to estimate the composition of 5 tissue samples. Specifically, the input data is computationally compared to the cell-specific miRNA expression distributions of constitutive cell types to estimate the cellular composition of each tissue sample. Such deconvolution methods rely on a vast amount of cell-specific expression data to assess cell-specificity. Adjusting for changes in tissue composition is important when comparing normal and diseased tissues where the composition, due to inflammation or fibrosis, is altered.