PLOS ONE

RESEARCH ARTICLE

Integrated microRNA and proteome analysis of cancer datasets with MoPC

Marta Lovino[1], Elisa Ficarra[1], Loredana Martignetti[2]*

1 Dipartimento di Ingegneria Enzo Ferrari, University of Modena and Reggio Emilia, Modena, Italy, **2** Institut Curie, INSERM U900, MINES ParisTech, PSL Research University, Paris, France

* loredana.martignetti@curie.fr

Abstract

MicroRNAs (miRNAs) are small molecules that play an essential role in regulating gene expression by post-transcriptional gene silencing. Their study is crucial in revealing the fundamental processes underlying pathologies and, in particular, cancer. To date, most studies on miRNA regulation consider the effect of specific miRNAs on specific target mRNAs, providing wet-lab validation. However, few tools have been developed to explain the miRNA-mediated regulation at the protein level. In this paper, the MoPC computational tool is presented, that relies on the partial correlation between mRNAs and proteins conditioned on the miRNA expression to predict miRNA-target interactions in multi-omic datasets. MoPC returns the list of significant miRNA-target interactions and plot the significant correlations on the heatmap in which the miRNAs and targets are ordered by the chromosomal location. The software was applied on three TCGA/CPTAC datasets (breast, glioblastoma, and lung cancer), returning enriched results in three independent targets databases.

Introduction

MicroRNAs (miRNAs) are small RNA molecules that emerged as important regulators of gene expression at the post-transcriptional level. They are involved in many diverse biological processes [1], and their dysregulation may lead to various diseases, including cancer [2–4]. Accurate prediction of miRNA-target interactions is critical to understanding the function of miRNAs. Although much progress has been made, target identification remains a challenge because of the limited understanding of the molecular basis of miRNA-target coupling, but also due to the context-dependence of post-transcriptional regulation and miRNA mode of action [5–7]. Generally, miRNAs downregulate proteins through a combination of translational inhibition and promotion of mRNA decay [8–10], even though which mechanisms of action of microRNAs are the most dominant remains a matter of debate. The emergence of high-throughput methods in the past decades has allowed researchers to address the question of miRNA action on a global scale. High-throughput experiments have usually been focused on measuring miRNA-mediated changes at the mRNA level (degradation), allowing for the characterization of only a subset of direct targets [11–14]. Further integration of global proteome profiling has allowed the investigation of miRNA-mediated gene expression changes at

both the mRNA and the protein level [15, 16], even though many of these studies are limited
to one or a few miRNAs or proteins. Additionally, few published tools are available to integrate
mRNA and protein profiles to infer miRNA regulation. ProteoMirExpress [17] infers active
miRNAs based on the anticorrelation between miRNAs and their potential targets at either the
mRNA or the protein level, or both. However, pairwise correlation approaches applied to
omics profiles produce very dense networks containing direct and indirect associations that
are very difficult to distinguish.

Here, we propose a novel in silico method, called MoPC (Multi-omics Partial Correlation
analysis), to analyze post-transcriptional regulation by miRNAs by integrating information at
both mRNA and protein level. The MoPC method predicts miRNA-target potential interac-
tions using high-throughput mRNA, protein, and miRNA expression datasets from matched
samples. It computes the conditional dependence (partial correlation) of mRNA and protein
levels, taking into account the effect of a third variable, the expression of a given regulatory
miRNA. The underlying hypothesis assumes that if a miRNA regulates a given gene, then the
partial correlation between the mRNA and the protein expression of that gene conditioned by
the expression of the miRNA will be greater than the bivariate correlation between the mRNA
and protein levels of the gene. Therefore, to identify candidate miRNA targets, MoPC selects
those genes whose partial correlation between the level of RNA and protein levels conditioned
to the expression of a given miRNA is more significant than the bivariate correlation between
the RNA and the protein.

MoPC has been applied to TCGA/CPTAC omics datasets of breast, glioblastoma, and lung
cancer, predicting key miRNAs involved in these diseases. To assess the reliability of the inter-
actions predicted by the MoPC, they have been tested for the enrichment in predicted targets
by three independent databases, namely miRDB [18, 19], TargetScan [20], and miRTarBase
[21]. The targets predicted by MoPC in breast and lung cancer show significant overlap with
those predicted by the three independent databases. In contrast, the targets predicted in glio-
blastoma significantly overlap with targets predicted in TargetScan and miRTarBase. Finally,
MoPC provides the user with a heatmap to visualize results and simultaneously explore the
conditional correlation map of all genes and miRNAs included in the analysis. The user can
visualize the interactions between multiple miRNAs and multiple targets by considering hori-
zontal or vertical bands in the heatmap.

Materials and methods

MoPC method

The MoPC workflow is reported in Fig 1.

To predict miRNA-target interactions by integrating information at both mRNA and pro-
tein level, MoPC calculates the partial correlation of the mRNA and protein levels of a gene con-
ditioned by miRNA expression. Input requires three matrices of mRNA, protein, and miRNA
expression denoted respectively as M, P, and MI with the same number of columns n, corre-
sponding to the samples. The M and P matrices must have the same rows IDs corresponding to
the genes for which the mRNA and protein levels were measured. Partial correlation $r_{yx.z}$
between mRNA x and protein expression y while accounting for the effect of a given miRNA z
is computed as in Eq 1 (r_{yx} stands for the bivariate correlation between variables y and x):

$$r_{yx.z} = \frac{r_{yx} - (r_{yz})(r_{xz})}{\sqrt{1 - r_{yz}^2}\sqrt{1 - r_{xz}^2}}$$
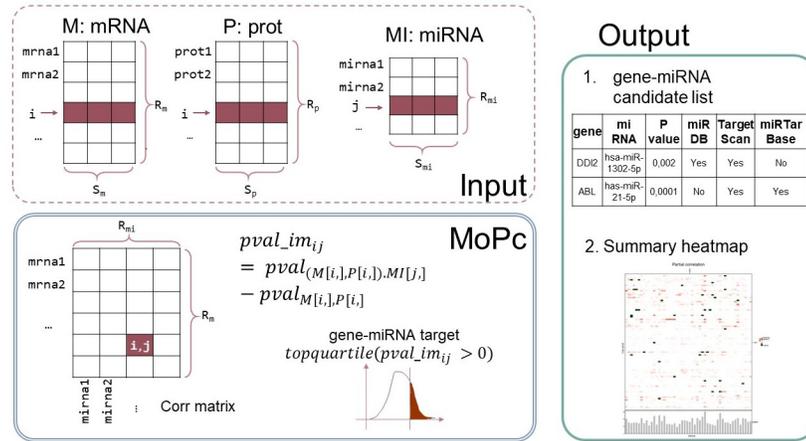
(1)

Fig 1. MoPC method in a glance. Input data consist of M, P, and MI matrices representing mRNA, proteomics, and miRNA expression matrices. First, the partial correlation matrix is computed. Note that this matrix has R_m rows corresponding to the number of rows in M and P (genes) and R_{mi} columns equal to the number of MI rows (miRNAs). Next, the improvement matrix is computed as the difference between the partial correlation p-value and the bivariate correlation p-value. Finally, relevant gene-miRNA interactions are chosen as the top quartile of positive improvement values.

The bivariate Pearson correlation r_{yx} between mRNA expression x and protein expression y is also computed. Relevant miRNA-target interactions are selected as those for which a lower p-value is observed with partial correlation than that obtained with bivariate correlation. To ensure a large increase in the significance of the p-value, only interactions with an improvement in the p-value in the upper quartile of the distribution are selected.

The MoPC returns two outputs:

1. **miRNA-target significant interactions**. This list reports all miRNA-target candidate interactions based on the partial correlation measure. Additional columns contain the validation information (e.g., if the miRNA-target interaction was previously known in the literature according to three databases: miRDB, TargetScan, and miRTarBase).

2. **summary heatmap**. The heatmap reports genes on the rows and miRNAs on the column ordered according to their chromosomal location. Significant miRNA-targets are reported in red (the more intense the color, the higher the partial correlation estimate). MiRNA-targets both significant and validated in at least one of the three databases (miRDB, TargetScan, and miRTarBase) are reported in dark green. The user can identify miRNAs targeting multiple genes by visually inspecting horizontal or vertical bands in the heatmap.

The method is freely available as an R package at the following link: https://github.com/martalovino/MoPc.

In addition, we investigated the miRNA-targets found by MoPC in all three datasets (breast, lung, and glioblastoma) to search if they contain the seed region of the corresponding miRNA. Since the MoPC method does not exploit at all gene and miRNA sequences in the analysis, the presence of the miRNA seed in the target mRNA is of interest. The gene sequences were obtained by processing the human reference sequence and the annotation files (fasta and gtf format, respectively) from the ENSEMBL GRCh38, version 109, database. Each gene sequence has been selected as the sequence (from the start to the end coordinates) of the longest

Table 1. Data dimensions as provided in input to MoPC. For each dataset, the number of samples, quantified mRNA, proteins and miRNAs are reported after the adequate preprocessing.

Dataset name	# of samples	# of genes	# of proteins	# of miRNAs
Breast cancer	77	5263	5263	332
Glioblastoma	106	4355	4355	316
Lung cancer	201	4279	4279	252

# of samples column refers to samples for which mRNA, protein and miRNAs quantifications are available.

transcript. The miRNA sequences were obtained from miRDB in fasta format, and the miRNA seed region was identified from base 2 to base 8, counting from the miRNA 5p end. To verify the presence of the seed region in the target gene, we enforced an exact match.

Data pre-processing

The MoPC method was tested on three TCGA/CPTAC datasets from breast, glioblastoma, and lung cancer studies [22–24].

The breast cancer dataset was obtained from the work of Meryins et al., published in 2016 [22]. The published matrices contain the quantifications of 17814 genes, 7665 proteins, and 332 miRNAs for 77 samples, respectively. The glioblastoma dataset was obtained from the publication of Wang et al. of 2021 [23], including mRNA, protein, and miRNA expression values for 106 tumor samples. The input matrices contain, respectively, 45914 genes, 10998 proteins, and 2883 miRNAs. The lung squamous cell tumor dataset comes from the 2021 publication of Satpathy et al., in which mRNA, proteins, and miRNA expression are quantified [24]. More than 200 samples have been analyzed, of which 201 have all three omics quantifications of our interest. Overall, 21792 genes, 11111 proteins, and 2585 miRNAs are reported.

The three datasets were processed according to the following procedure. These datasets contain missing values, particularly in protein and miRNA expression matrices. For our purposes, proteins and miRNAs with less than 10% of missing values across samples were kept. The remaining missing values were imputed by imposing the median value. Finally, the mRNA e miRNA matrices were transformed by applying the log2 function.

In all datasets, only the top 50% most highly variable proteins were included, since MoPC requires a certain variance in the expression profiles.

Table 1 shows the datasets size supplied to MoPC.

Results

MoPC detects biologically relevant regulatory miRNAs in cancer datasets

MoPC was applied to three omic datasets from TCGA/CPTAC projects, namely breast, glioblastoma, and lung cancer. MoPC output returns 1) a table of miRNAs and predicted target genes and 2) a heatmap to visualize the chromosomal position of miRNAs and target genes. In particular, the table listing miRNAs and target genes contains both the estimate value and p-value returned from the partial correlation (the higher the estimate and the lower the FDR adjusted p-value, the more significant the interaction between that gene and that miRNA). In addition, it returns if the gene-miRNA target interaction is already annotated in the three external databases (miRDB, TargetScan, and miRTarBase).

Computational validation

In order to test the reliability of MoPC predictions, the enrichment in miRDB, TargetScan, and miRTarBase predicted targets was assessed for each of the three cancer datasets, using hypergeometric test. In addition, the same enrichment test was performed for miRNA-target predictions based on the bivariate correlation between mRNA and miRNA expression and between protein and miRNA expression. Pearson correlation with FDR corrected p-values was computed and the miRNA-target pairs with a p-value lower than alpha 0.05 were selected. The same conditions of the partial correlation were used to compute the bivariate correlations to perform a fair comparison. Tables 2–4 report for each of the three cancer datasets the percentage of miRNA-targets selected as significant, and the p-value of the hypergeometric test in each of the three databases (miRDB, TargetScan, and miRTarBase) both for the MoPC method and the mRNA-miRNA and protein-miRNA bivariate correlation.

MiRNA-targets visualization map

MoPC returns a heatmap where miRNAs and predicted targets are arranged according to their chromosomal position. The heatmap represents in red the miRNA-targets significantly predicted according to the MoPC partial correlation analysis and in dark green, the miRNA-targets statistically predicted and validated in at least one of the three databases (miRDB, TargetScan, and miRTarBase). The user can visually explore the presence of miRNAs that regulate multiple genes and vice-versa (the same gene regulated by multiple miRNAs). S1–S3 Figs show breast, glioblastoma, and lung cancer heatmaps, respectively.

Discussion

The goal of MoPC is to predict relevant miRNA-targets integrating information from miRNA, mRNA and proteome profiles. The predicted targets are validated with three reference databases (miRDB, TargetScan, and miRTarBase), both for the MoPC method and for the bivariate correlation between mRNA-miRNA and protein-miRNA. As reported in Tables 2–4, MoPC predictions are enriched in independent database targets for the three cancer datasets analyzed. Except for the enrichment of glioblastoma in miRDB, all enrichments are statistically significant. The miRNA-target pairs predicted from the bivariate mRNA-miRNA and protein-miRNA correlation are also statistically enriched for each cancer dataset and database, unless for breast cancer in the bivariate protein-miRNA correlation in the mirDB database.

An essential aspect to compare is the total number of miRNA-target interactions predicted by MoPC and bivariate mRNA-miRNA and protein-miRNA correlation. Analyses of glioblastoma and lung cancer datasets benefit from the results obtained through MoPC. Indeed, MoPC returns 17.20%, 18.26%, and 18.49% of miRNA-target interactions for breast, glioblastoma, and lung cancer. Although the bivariate mRNA-miRNA and protein-miRNA

Table 2. Validation of results on breast cancer dataset. The first two columns contain the results related to the bivariate correlation between mRNA-miRNA and protein-miRNA. The last column contains the MoPC results. Pearson correlation with FDR corrected p-values was computed and the miRNA-target pairs with a p-value lower than alpha 0.05 were selected. The same conditions of the partial correlation were used to compute the bivariate correlations to perform a fair comparison.

	Bivariate correlation mRNA-miRNA	Bivariate correlation protein-miRNA	MoPC based on partial correlation
% of predicted miRNA-targets	**2,59%**	0,68%	17,20%
miRDB enrichment	2.66e-05	0.607	**3.73e-05**
TargetScan enrichment	0.038	0.094	**0.0003**
miRTarBase enrichment	0.0094	0.023	**2.24e-11**

Table 3. Validation of results on glioblastoma dataset. The first two columns contain the results related to the bivariate correlation between mRNA-miRNA and protein-miRNA. The last column contains the MoPC results. Pearson correlation with FDR corrected p-values was computed and the miRNA-target pairs with a p-value lower than alpha 0.05 were selected. The same conditions of the partial correlation were used to compute the bivariate correlations to perform a fair comparison.

	Bivariate correlation mRNA-miRNA	Bivariate correlation protein-miRNA	MoPC based on partial correlation
% of predicted miRNA-targets	36,69%	41,70%	**18,26%**
miRDB enrichment	<e-70	<e-70	0.999
TargetScan enrichment	<e-70	<e-70	9.48e-23
miRTarBase enrichment	<e-70	<e-70	1.57e-63

Table 4. Validation of results on lung cancer dataset. The first two columns contain the results related to the bivariate correlation between mRNA-miRNA and protein-miRNA. The last column contains the MoPC results. Pearson correlation with FDR corrected p-values was computed and the miRNA-target pairs with a p-value lower than alpha 0.05 were selected. The same conditions of the partial correlation were used to compute the bivariate correlations to perform a fair comparison.

	Bivariate correlation mRNA-miRNA	Bivariate correlation protein-miRNA	MoPC based on partial correlation
% of predicted miRNA-targets	61,13%	59,60%	**18,49%**
miRDB enrichment	<e-70	<e-70	2.891e-11
TargetScan enrichment	<e-70	<e-70	2.973e-06
miRTarBase enrichment	<e-70	<e-70	6.68e-27

correlation are generally more enriched in the three external target databases, they return a higher percentage of miRNA-target correlated pairs, typically greater than 36% for both mRNA-miRNA and protein-miRNA. This aspect is of uttermost importance, as the number of investigable gene-miRNA pairs is not clinically useful if it is close to 30–60% of all gene-miRNA pairs in the input dataset. MoPC proves to be an effective method in identifying key miRNAs within a dataset, reducing the number of targets while maintaining the statistical significance of the enrichment.

After the MoPC computational validation in the three external databases, we focused on specific miRNAs known in the literature to be associated with specific diseases. Indeed, the enrichment in miRDB, TargetScan, and miRTarBase verifies that the miRNA-target interaction is known or predicted in the literature. However, such enrichment does not consider whether the predicted miRNAs are key in the specific disease (breast, glioblastoma, lung cancer). Therefore we compared MoPC miRNA-target pairs with disease-associated miRNA in the literature [25–27] (see Tables 5–7 for validation information in breast, gliobastoma and lung cancer, respectively.

Regarding breast cancer, MoPC detects the main miRNAs involved in the cancer genesis, including miR-126–5p and miR-140–5p (with anti-angiogenesis and anti-tumorigenesis function) and miR-200b-5p and miR-200c-5p (promoting metastasis and invasion). Furthermore, among the miRNAs known in glioblastoma, MoPC successfully identifies the tumor suppressors miR-7–5p and miR-128–5p and the oncomiRs miR-21–5p and miR-93–5p, responsible for survival, proliferation, apoptosis, and drug resistance. For lung cancer, the miRNAs known and identified by MoPC are miR-7–5p and miR-21–5p (chemosensitive and chemoresistant, respectively) and miR-200c-5p and miR-223–5p with EGFR-TKI and ALK-TKI sensitivity. The Tables below show the top 5 percentile of miRNAs with the highest significant number of targets found by MoPC and their associated disease.

Table 8 shows for each dataset the percentage of miRNA-targets in which the seed was found in any region of the gene, the percentage in which the seed region is located at 5p and 3p of the gene, and the percentage in which the seed region is located in a coding region (CDS region). The last row shows the percentage in which the seed region is either on the 5p UTR,

Table 5. Breast cancer validation information.

miRNA	gene/pathways	known effect	Reference
hsa-miR-3200–3p			
hsa-miR-197–3p			
hsa-miR-505–5p			
hsa-miR-505–3p			
hsa-miR-18a-5p			
hsa-miR-17–5p			
hsa-miR-577–5p			
hsa-miR-135b-5p	LATS2, CDK2, p-YAP	Promotion of cell proliferation and S–G2/M cell cycle progression	[25]
hsa-miR-93–3p			
hsa-miR-93–5p			
hsa-miR-20a-3p			
hsa-miR-301a-3p			
hsa-miR-127–3p			
hsa-miR-29b-1–5p	Akt3, VEGF, c-MYC	Anti-angiogenesis and anti-tumorigenesis	[25]
hsa-miR-15b-3p	Cyclin E1, E2F7	Anti-proliferative and G1–S cell cycle arrest	[25]
hsa-miR-17–3p			

Table 6. Glioblastoma validation information.

miRNA	gene/pathways	known effect	Reference
hsa-miR-128–3p	Tumor suppressor	Proliferation, apoptosis, angiogenesis, stemness, radioresistance	[28]
hsa-miR-139–5p	In vitro, human	ELTD1, cyclin A, cyclin D1	[28]
hsa-miR-21–5p	oncomiR	Survival, proliferation, apoptosis, migration, invasion, chemoresistance	[28]
hsa-miR-21–3p	oncomiR	Survival, proliferation, apoptosis, migration, invasion, chemoresistance	[28]
hsa-miR-7–5p	Tumor suppressor	Survival, proliferation, apoptosis, invasion, angiogenesis	[28]
hsa-miR-1249–3p			
hsa-miR-155–5p	In vitro, human	MAPK13, MAPK14	[28]
hsa-miR-124–3p	Cdk6, PPPR13L, SNAI2/GJIC	miR-124–3p inhibits glioblastoma cell growth in part by diminishing the protein expression of cdk6, leading to cell cycle arrest at the G0/G1 phase.	[28]
hsa-miR-874–3p			
hsa-miR-15b-5p	In vitro, human	miR-15b inhibited the proliferation, cell cycle arrest, and invasion.	[28]
hsa-miR-769–5p			
hsa-miR-132–5p	In vitro, human	miR-132 promoted the proliferation and sphere formation of cells.	[28]

3p UTR, or CDS (the seed could be found in multiple locations, so the last row could not be the mathematical sum of the three previous ones).

The seed is always present in the miRNA-targets selected by MoPC. In addition, although MoPC does not profit from any information related to gene or miRNA sequences, more than 22% of the total miRNA-target interaction is found in 5p UTR, 3p UTR, or coding regions.

Table 7. Lung cancer validation information.

miRNA	gene/pathways	known effect	Reference
hsa-miR-182–5p			
hsa-miR-26b-5p			
hsa-let-7a-5p	LIN28	Chemo sensitive	[27]
hsa-miR-26a-5p			
hsa-miR-183–5p			
hsa-let-7f-5p	LIN28	Chemo sensitive	[27]
hsa-miR-30a-3p	PI3K-SIAH2	EGFR TKI-resistant	[27]
hsa-miR-125a-5p			
hsa-miR-205–5p	PTEN, Mcl-1 and Survivin	Chemo resistant	[27]
hsa-miR-126–3p	Akt and ERK pathways	Chemo sensitiveEGFR TKI-sensitive	[27]
hsa-miR-7–5p	EGFR	Chemo sensitive	[27]
hsa-miR-9–3p			

https://doi.org/10.1371/journal.pone.0289699.t007

The MoPC method could be also exploited to identify significant targets of miRNA isoforms. Indeed, it is sufficient to supply the expression matrix of the miRNA isoforms instead of the miRNA expression matrix, and MoPC will return the miRNA-target isoforms calculated through partial correlation.

Conclusion

Identifying key miRNAs is of fundamental importance for the early diagnosis and treatment of many diseases, including cancer. However, in the literature, various tools deal with providing a list of relevant miRNAs without exploiting the information coming from the expression of proteins. This paper presents MoPC, a novel method for predicting key miRNA-target pairs by exploiting the computation of the partial correlation between mRNA and protein expression conditioned by miRNA expression. The method is applied to three TCGA/CPTAC datasets: breast, glioblastoma, and lung cancer. The results show enrichment for each dataset in the three databases, miRDB, TargetScan, and miRTarBase, except for breast cancer in miRDB. Furthermore, MoPC successfully identifies key miRNAs in each cancer type. Finally, the user can inspect the results both through the list of significant miRNA-gene pairs and graphically through the generated heatmap, in which genes and miRNAs are chromosomally ordered. In the future, this approach can be extended to other types of omics data, given that the underlying biological phenomenon can be detected through partial correlation.

Table 8. Presence of miRNA seed in miRNA-targets selected by MoPC. Each cell contains the percentage of miRNA-targets detected by MoPC for the specific tissue (Breast, Lung, and Glioblastoma) in which the target mRNA contains the miRNA seed region (anywhere in the gene, in 5pUTR, 3pUTR, coding region). The last row is not the sum of the previous rows since the miRNA seed can be present in multiple gene locations.

	Breast	Lung	Glioblastoma
Perc. with seed region	100%	100%	100%
Perc. seed region in 5pUTR	1.82%	1.94%	1.94%
Perc. seed region in 3pUTR	12.36%	11.93%	11.93%
Perc. seed region in coding region	11.50%	11.39%	11.39%
Perc. seed region in 5pUTR or 3pUTR or coding region	22.64%	22.04%	22.04%

https://doi.org/10.1371/journal.pone.0289699.t008

Supporting information

S1 Fig. Breast cancer summary heatmap. The heatmap reports in red the miRNA-genes pairs resulting significantly with the MoPC analysis. In dark green, the miRNA-genes pairs are statistically significant in the input dataset and validated in at least one of the three databases (miRDB, TargetScan, and miRTarBase). Genes are reported on the rows, and miRNAs on the columns. Both genes and miRNAs are chromosomally ordered. Only the 5p' most expressed miRNAs in breast cancer, according to isomiRTar [6], have been reported in the columns to provide more understandable results.
(TIF)

S2 Fig. Glioblastoma summary heatmap. The heatmap reports in red the miRNA-genes pairs resulting significantly with the MoPC analysis. In dark green, the miRNA-genes pairs are statistically significant in the input dataset and validated in at least one of the three databases (miRDB, TargetScan, and miRTarBase). Genes are reported on the rows, and miRNAs on the columns. Both genes and miRNAs are chromosomally ordered. Only the 25% most expressed miRNAs in this glioblastoma dataset have been reported in the columns to provide more understandable results.
(TIF)

S3 Fig. Lung cancer summary heatmap. The heatmap reports in red the miRNA-genes pairs resulting significantly with the MoPC analysis. In dark green, the miRNA-genes pairs are statistically significant in the input dataset and validated in at least one of the three databases (miRDB, TargetScan, and miRTarBase). Genes are reported on the rows, and miRNAs on the columns. Both genes and miRNAs are chromosomally ordered. Only the 5p' most expressed miRNAs in lung cancer, according to isomiRTar [6], have been reported in the columns to provide more understandable results.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Marta Lovino, Elisa Ficarra, Loredana Martignetti.

Formal analysis: Marta Lovino, Loredana Martignetti.

Funding acquisition: Elisa Ficarra, Loredana Martignetti.

Project administration: Elisa Ficarra, Loredana Martignetti.

Resources: Elisa Ficarra, Loredana Martignetti.

Software: Marta Lovino.

Supervision: Loredana Martignetti.

Writing – original draft: Marta Lovino, Loredana Martignetti.

References

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. cell. 2004; 116(2):281–297. https://doi.org/10.1016/S0092-8674(04)00045-5 PMID: 14744438

2. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, et al. A microRNA polycistron as a potential human oncogene. nature. 2005; 435(7043):828–833. https://doi.org/10.1038/nature03552 PMID: 15944707

3. Sheedy P, Medarova Z. The fundamental role of miR-10b in metastatic cancer. American journal of cancer research. 2018; 8(9):1674. PMID: 30323962

4. Lovino M, Montemurro M, Barrese VS, Ficarra E. Identifying the oncogenic potential of gene fusions exploiting miRNAs. Journal of Biomedical Informatics. 2022; 129:104057. https://doi.org/10.1016/j.jbi.2022.104057 PMID: 35339665

5. Bartel DP. MicroRNAs: target recognition and regulatory functions. cell. 2009; 136(2):215–233. https://doi.org/10.1016/j.cell.2009.01.002 PMID: 19167326

6. Nersisyan S, Gorbonos A, Makhonin A, Zhiyanov A, Shkurnikov M, Tonevitsky A. isomiRTar: a comprehensive portal of pan-cancer 5′-isomiR targeting. PeerJ. 2022; 10:e14205. https://doi.org/10.7717/peerj.14205 PMID: 36275459

7. Reyes-Herrera PH, Ficarra E. Computational methods for CLIP-seq data processing. Bioinformatics and Biology insights. 2014; 8:BBI–S16803. https://doi.org/10.4137/BBI.S16803 PMID: 25336930

8. Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. Nature reviews genetics. 2015; 16(7):421–433. https://doi.org/10.1038/nrg3965 PMID: 26077373

9. Bartel DP. Metazoan micrornas. Cell. 2018; 173(1):20–51. https://doi.org/10.1016/j.cell.2018.03.006 PMID: 29570994

10. Biasini A, Abdulkarim B, de Pretis S, Tan JY, Arora R, Wischnewski H, et al. Translation is required for miRNA-dependent decay of endogenous transcripts. The EMBO journal. 2021; 40(3):e104569. https://doi.org/10.15252/embj.2020104569 PMID: 33300180

11. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005; 433(7027):769–773. https://doi.org/10.1038/nature03315 PMID: 15685193

12. Huang JC, Babak T, Corson TW, Chua G, Khan S, Gallie BL, et al. Using expression profiling data to identify human microRNA targets. Nature methods. 2007; 4(12):1045–1049. https://doi.org/10.1038/nmeth1130 PMID: 18026111

13. Volinia S, Visone R, Galasso M, Rossi E, Croce CM. Identification of microRNA activity by targets' reverse EXpression. Bioinformatics. 2010; 26(1):91–97. https://doi.org/10.1093/bioinformatics/btp598 PMID: 19897564

14. Wang YP, Li KB. Correlation of expression profiles between microRNAs and mRNA targets using NCI-60 data. BMC genomics. 2009; 10(1):1–13. https://doi.org/10.1186/1471-2164-10-218 PMID: 19435500

15. Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. Nature. 2008; 455(7209):64–71. https://doi.org/10.1038/nature07242 PMID: 18668037

16. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. nature. 2008; 455(7209):58–63. https://doi.org/10.1038/nature07228 PMID: 18668040

17. Liu Q, Halvey PJ, Shyr Y, Slebos RJ, Liebler DC, Zhang B. Integrative omics analysis reveals the importance and scope of translational repression in microRNA-mediated regulation. Molecular & Cellular Proteomics. 2013; 12(7):1900–1911. https://doi.org/10.1074/mcp.M112.025783 PMID: 23550052

18. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome biology. 2019; 20(1):1–10. https://doi.org/10.1186/s13059-019-1629-z PMID: 30670076

19. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic acids research. 2020; 48(D1):D127–D131. https://doi.org/10.1093/nar/gkz757 PMID: 31504780

20. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. elife. 2015; 4:e05005. https://doi.org/10.7554/eLife.05005 PMID: 26267216

21. Huang HY, Lin YCD, Cui S, Huang Y, Tang Y, Xu J, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions. Nucleic acids research. 2022; 50(D1):D222–D230. https://doi.org/10.1093/nar/gkab1079 PMID: 34850920

22. Mertins P, Mani D, Ruggles KV, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016; 534(7605):55–62. https://doi.org/10.1038/nature18003 PMID: 27251275

23. Wang LB, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y, et al. Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell. 2021; 39(4):509–528. https://doi.org/10.1016/j.ccell.2021.01.006 PMID: 33577785

24. Satpathy S, Krug K, Beltran PMJ, Savage SR, Petralia F, Kumar-Sinha C, et al. A proteogenomic portrait of lung squamous cell carcinoma. Cell. 2021; 184(16):4348–4371. https://doi.org/10.1016/j.cell.2021.07.016 PMID: 34358469

25. Loh HY, Norman BP, Lai KS, Rahman NMANA, Alitheen NBM, Osman MA. The regulatory role of microRNAs in breast cancer. International journal of molecular sciences. 2019; 20(19):4940. https://doi.org/10.3390/ijms20194940 PMID: 31590453

26. Chen M, Medarova Z, Moore A. Role of microRNAs in glioblastoma. Oncotarget. 2021; 12(17):1707. https://doi.org/10.18632/oncotarget.28039 PMID: 34434499

27. Uddin A, Chakraborty S. Role of miRNAs in lung cancer. Journal of Cellular Physiology. 2018;. https://doi.org/10.1002/jcp.26607 PMID: 29676470

28. Rezaei O, Honarmand K, Nateghinia S, Taheri M, Ghafouri-Fard S. miRNA signature in glioblastoma: Potential biomarkers and therapeutic targets. Experimental and Molecular Pathology. 2020; 117:104550. https://doi.org/10.1016/j.yexmp.2020.104550 PMID: 33010295