proceedings

MicroRNA Expression Analysis and Biological Pathways in Chemoresistant Non- Small Cell Lung Cancer

Krystallia Gourlia[1], Maria Markaki[2], Chara Papadaki[1], Konstantinos Rounis[3], Eleftherios Vorrias[3], Elena Prokova[2], Eleni Lagoudaki[4], Anastasios Koutsopoulos[4], Ioannis Tsamardinos[2], Dimitrios Mavroudis[1,3], Sofia Agelaki[1,3]*

1. Laboratory of Translational Oncology, School of Medicine, University of Crete, Heraklion, Greece;krysgourlia@gmail.com; chapapadak@uoc.gr; mavroudis@uoc.gr;agelakisofia@gmail.com
2. Department of Computer Science, University of Crete, Heraklion, Crete, Greece; maria.g.markaki@gmail.com; tsamard.it@gmail.com
3. Department of Medical Oncology, University General Hospital, Heraklion, Crete, Greece; rounis@gmail.com; elvorrias@gmail.com
4. Department of Pathology, University General Hospital of Heraklion, School of Medicine, University of Crete, Heraklion, Crete, Greece , elagoudakimd@gmail.com; akoutsop@gmail.com

* Correspondence:Correspondence: agelakisofia@gmail.com; Tel.: +30-281-0392438

Received: date; Accepted: date; Published: date

Abstract: Platinum-based chemotherapy (CT) is a standard treatment for lung cancer, however a variety of chemoresistance mechanisms can impair its efficacy. MicroRNAs (miRNAs) represent potential biomarkers for the prediction of treatment efficacy in non-small cell lung cancer (NSCLC). We herein used a bioinformatics approach to identify differentially expressed (DE) miRNAs associated with response to platinum-based CT in NSCLC. We identified 6 miRNAs targeting signaling molecules participating in biological pathways involved in cancer and drug resistance. In summary, we developed a 6- miRNA signature that potentially predicts the response to cisplatin in NSCLC and warrants further validation in clinical samples.

Keywords: cisplatin resistance; NSCLC; bioinformatics; miRNA signature

1. Introduction

Lung cancer is the leading cause of cancer-related mortality for both sexes, worldwide with non-small cell lung cancer (NSCLC), accounting for 85% of cases [1]. Despite significant advances in systemic therapy with targeted agents and immune checkpoint inhibitors (ICIs), platinum-based chemotherapy remains the cornerstone of treatment for both early and metastatic NSCLC, however inherent or acquired tumor resistance limits its efficacy [2]. Multiple mechanisms and pathways have been proposed in tumor cells that contribute to cisplatin resistance [3]. Mechanisms such as decreased intracellular accumulation of the drug, increased detoxification systems, impaired apoptotic signaling after DNA damage, are involved [2], however, the efficiency of DNA damage response

and DNA repair pathways seem to play a central role in cisplatin resistance. Platinum induces intra-strand and inter-strand crosslinks that pause the replication fork resulting in increased toxicity to proliferating cells [4]. Replication fork prevention activates the DNA damage response, followed by DNA repair through the different repair pathways [5]. Mutations or reduced expression of DNA repair genes are associated with platinum sensitivity [5]. The research for the identification of biomarkers for platinum responsiveness remains a major challenge in order to avoid overtreatment in patients who are not expected to respond.

MicroRNAs (miRNAs) are small non-coding RNA molecules consisting of 20-22 nucleotides that regulate gene expression at the post-transcriptional level [6]. Their expression is deregulated in various types of cancer [7]. Also, several studies suggest that proteins involved in DDR undergo post-transcriptional regulation by miRNAs in response to platinum-induced damage. Conversely, the response of cells to platinum-induced DNA damage causes the transcriptional regulation and modification of miRNA expression [8]. In recent years, miRNAs have attracted particular interest as prognostic and predictive biomarkers in cancer [8].

Importantly, currently available information generated through high-throughput methods can be capitalized by systematic bioinformatics analyses thus allowing the identification of candidate miRNAs that will be validated for their ability to predict response treatment. In the current work we implemented state-of-the-art bioinformatics methods in gene prioritization and regulatory network analysis in order to identify differentially expressed miRNAs associated with the response to platinum-based chemotherapy in NSCLC.

2. Results

2.1. Dataset selection and expression profiling data analysis

Two datasets retrieved from Gene Expression Omnibus Datasets (GEO) included 69 patients with NSCLC treated with platinum-based CT, from whom, 33 were responders and 36 were non-responders. Analysis in limma package in R revealed 1614 DE miRNAs from the two datasets and the number reduced to 1150 miRNAs, due to miRNAs with non-zero values that were excluded from further analysis. After meta- analysis, 72 and 40- DE miRNAs were consistently up- and down-regulated, respectively in both datasets (Figure 1). Furthermore, out of the 112- DE miRNAs, 24 were up- or down- regulated with a *p*-value<0.05 and $|\log FC| \geq 0.5$.

2.2. Survival Analysis by KM Plotter

We used KM plotter database as a validation step to evaluate the association of the 24-DE miRNAs, with overall survival in patients with NSCLC. We checked the concordance of the log fold change (logFC, as reported by limma) of these 24 miRNAs with the corresponding Hazard Ratio (HR, as reported by KM plotter), meaning consistent up- or down regulation of respective miRNAs in NSCLC patients. We also required a statistical significance of *p*- value < 0.05 in KM plotter (Table 1). Therefore, after the last integration, a 6-miRNA signature was revealed consisting of hsa-miR-497, hsa-miR-29c, hsa-miR-26a, hsa-miR-34a, hsa-miR-30e-5p and hsa-miR-30e-3p.

Table 1. The logFC values and p. values after the limma and KM Plotter analysis for the selected miRNAs .

DE miRNAs	GSE56036		GSE56264		Hazard Ratio (Adeno)		Hazard Ratio (Squamous)	
	logFC	*P*- Value	logFC	*P*- Value	logFC	*P*-Value	logFC	*P*-Value
hsa-miR-497	-0.99268	0.010906	-0.82107	0.019803	0.52	0.0009	1.17	0.29
hsa-miR-29c	-1.0811	0.013295	-0.82175	0.003344	0.54	0.012	0.8	0.15
hsa-miR-26a	-1.4049	0.016655	-0.52735	0.023449	0.63	0.038	0.74	0.033
hsa-miR-30e	-1.19279	0.024696			0.56	8.9e-0.5	0.75	0.048
hsa-miR-30e*			-0.53971	0.043327	0.56	8.9e-0.5	0.75	0.048
hsa-miR-34a	-1.31859	0.01225	-0.45893	0.041946	0.71	0.062	1.21	0.2

2.4. Validation of the predicted 6-miRNA signature by pathway enrichment analysis

Pathway enrichment analysis for the 6- DE miRNAs was performed by DIANA Tools. The analysis revealed significantly enriched pathways related to cancer such as pathways in cancer, NSCLC, apoptosis and proliferation, p53 pathway, Hippo and proteoglycans in cancer. Among those the p53, Hippo and Proteoglycans in cancer have been previously correlated with DDR. All the selected miRNAs, except hsa-miR-29c, are significantly involved in at least two of the above pathways (Figure 1).

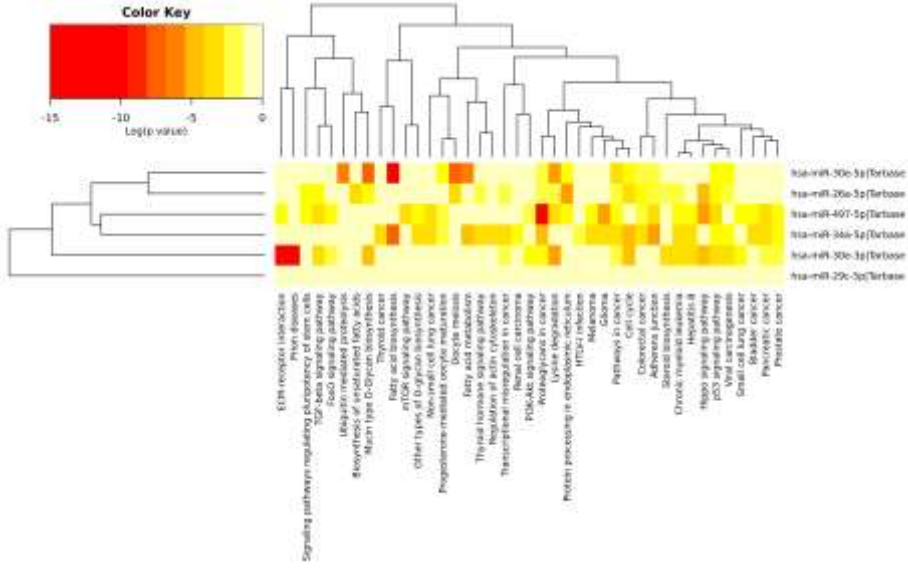

Figure 1. The heatmap exhibits significantly enriched pathways targeted from the 6- DE miRNAs. The colour key depicts the statistical significance expressed by the log (*p*-value); as the colour gets darker, the more statistically significant is the involvement of the miRNA in the specific pathway.

3. Discussion

Platinum-based chemotherapy still has a major role both in the early and metastatic NSCLC, however primary or acquired resistance is encountered in most of the cases. Refining prognostics and personalized treatment approaches represents a significant challenge in the effort to improve patient outcomes. In the last few years miRNAs have

emerged as promising biomarkers for diagnosis, prognosis and prediction to treatment outcomes. We herein, applied a multistep bioinformatics approach to identify miRNAs that are differentially expressed among responders and non-responders patients with NSCLC treated with CT. We analyzed miRNA expression data from two GEO datasets, by employing strict thresholds for both logFC and *p*-value to select the DE miRNAs. After a multistep analysis the number of DE expressed miRNAs was reduced to 24. We next validated the prognostic significance of the 24-DE miRNAs using KM plotter database. We finally extracted a 6-miRNA signature consisting of miR-497, miR-29c, miR-26a, miR-34a, miR-30e-5p and miR-30e-3p that were down-regulated in non-responders. Interestingly, the role of the above miRNAs in platinum resistance is supported by bibliographic data. Indicative, miR-34a and miR-26a have been found to sensitize lung cancer cells to cisplatin [9] [10]. To uncover potential pathways that could be regulated by the 6-miRNA signature we performed pathway enrichment analysis, revealing significant associations of a number of pathways related to cancer, apoptosis, proliferation, p53 pathway and Hippo. Among those the p53 [11] and Hippo [12] have been previously reported to be correlated with DDR. Additional analysis is required to reveal the regulatory components of the above pathways and the possible regulatory relationships between the 6-miRNAs of the signature and their mRNA targets.

In summary, we developed a 6-miRNA signature which potentially predicts the response of NSCLC patients to cisplatin. These miRNAs found to be down-regulated in non-responders and bibliographic data exist to support their role in cisplatin sensitivity. Further studies are needed to evaluate the prognostic significance of the above signature in clinical samples.

4. Materials and Methods

4.1. Dataset Collection

Two miRNA expression microarray datasets (GSE56036 [13], GSE56264 [14]) published in 2015, were retrieved from GEO. According to the inclusion criteria, the datasets had to include samples from patients with NSCLC treated with chemotherapy (CT) and also had records on patients' response to CT. In total, the two datasets included 69 samples, 33 of them corresponded to responders and 36 to non-responders.

4.2. Dataset Statistical Analysis

The Linear Models for Microarray Analysis (limma) package in R software environment for statistical computing and graphics was applied to compare the expression of miRNAs between responders and non-responders included in the two datasets. Using the limma package, we were able to extract the *p*-value and log fold change (logFC) expression for each microRNA from each individual dataset.

4.3. Meta- analysis

We conducted random- effects meta-analysis integrating the two miRNA datasets and used a *p*- value <0.05 as significant and logFC for each individual miRNA greater or equal

than 0.5 was adjusted as a filter to find DE miRNAs. More specifically, miRNAs which were consistently up/down- regulated and had |logFC|≥0.5 were chosen.

4.4. Survival analysis with KM Plotter

Kaplan-Meier plotter (KM-Plotter; http://kmplot.com/lung), an online tool that performs univariate Cox regression analysis with data from TCGA and GEO, was used to further validate the prognostic significance of the candidate miRNAs. KM plotter contained survival data for 513 adenocarcinoma (AD) and 478 squamous carcinoma (SQ) non- small cell lung cancer patients including stage information. We determined the hazard ratio (HR) and long rank *p*- value for the 24 DE miRNAs that were revealed from meta-analysis. In our study, we chose the DE miRNAs which their logFC was in concordance with the HR. Specifically, a HR greater than 1 is expected when logFC value is positive and a HR lower than one is expected when logFC value is negative, respectively. Furthermore, the statistical significance was set as *p*- value< 0.05.

4.5. Pathway enrichment analysis by DIANA Tools

DIANA-mirPath is a miRNA pathway analysis web-server, providing accurate statistics, while being able to accommodate advanced pipelines. mirPath can utilize predicted miRNA targets (in CDS or 3′-UTR regions) provided by the DIANA-microT-CDS algorithm or even experimentally validated miRNA interactions derived from DIANA-TarBase. Tools database was utilized to determine in which biological pathways the miRNAs of interest are involved into.

Author Contributions: Conceptualization, S.A.; methodology, K.G., M.M. and C.P.; software, K.G. and M.M.; formal analysis, K.G., M.M., C.P. and S.A..; investigation, K.G., M.M., L.V., K.R., E.P., A.K. and CP.; resources, S.A, D.M and IT.; data curation, K.G. and M.M..; writing—original draft preparation, K.G, M.M, C.P. and S.A..; writing—review and editing, K.G., M.M, C.P, L.V., K.R., E.P., E.L., A.K., D.M, I.T. and S.A.; visualization, K.G., M.M., C.P..; supervision, M.M, C.P. and S.A.; project administration, C.P., M.M. and S.A..; funding acquisition, D.M. and I.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partly supported by Anticancer Research Support Association (ARSA) and Gnosis Data Analysis (Gnosis DA)

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] J. R. Molina, P. Yang, S. D. Cassivi, S. E. Schild, and A. A. Adjei, "Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship," *Mayo Clin. Proc.*, vol. 83, no. 5, pp. 584–594, 2008, doi: 10.4065/83.5.584.

[2] L. Galluzzi *et al.*, "Molecular mechanisms of cisplatin resistance," *Oncogene*, no. 31, pp. 1869–1883, 2012, doi: 10.1038/onc.2011.384.

[3] S. O'Grady, S. P. Finn, S. Cuffe, D. J. Richard, K. J. O'Byrne, and M. P. Barr, "The role of DNA repair pathways in cisplatin resistant lung cancer," *Cancer Treat. Rev.*, vol. 40, no. 10, pp. 1161–1170, 2014, doi: 10.1016/j.ctrv.2014.10.003.

[4] A. Manuscript, "Cisplatin in cancer therapy : molecular mechanisms of action," *Eur J Pharmacol.*, pp. 364–378, 2015, doi: 10.1016/j.ejphar.2014.07.025.Cisplatin.

[5] T. Taniguchi *et al.*, "Disruption of the Fanconi anemia-BRCA pathway in cisplatin-sensitive ovarian tumors," *Nat. Med.*, vol. 9, no. 5, pp. 568–574, 2003, doi: 10.1038/nm852.

[6] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, Jan. 2009, doi: 10.1016/j.cell.2009.01.002.

[7] A. A. Svoronos, D. M. Engelman, and F. J. Slack, "OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer," *Cancer Res.*, vol. 76, no. 13, pp. 3666–3670, Jul. 2016, doi: 10.1158/0008-5472.CAN-16-0359.

[8] I. Fadejeva, H. Olschewski, and A. Hrzenjak, "MicroRNAs as regulators of cisplatin-resistance in non-small cell lung carcinomas," *Oncotarget*, vol. 8, no. 70, pp. 115754–115773, 2017, doi: 10.18632/oncotarget.22975.

[9] C. Song, P. Lu, G. Sun, L. Yang, Z. Wang, and Z. Wang, "miR-34a sensitizes lung cancer cells to cisplatin via p53/miR-34a/MYCN axis," *Biochem. Biophys. Res. Commun.*, vol. 482, no. 1, pp. 22–27, 2017, doi: 10.1016/j.bbrc.2016.11.037.

[10] Y. Yang, P. Zhang, Y. Zhao, J. Yang, G. Jiang, and J. Fan, "Decreased MicroRNA-26a expression causes cisplatin resistance in human non-small cell lung cancer," *Cancer Biol. Ther.*, vol. 17, no. 5, pp. 515–525, 2016, doi: 10.1080/15384047.2015.1095405.

[11] H. C. Reinhardt and B. Schumacher, "The p53 network: Cellular and systemic DNA damage responses in aging and cancer," *Trends Genet.*, vol. 28, no. 3, pp. 128–136, 2012, doi: 10.1016/j.tig.2011.12.002.

[12] D. E. Pefani and E. O'Neill, "Hippo pathway and protection of genome stability in response to DNA damage," *FEBS J.*, vol. 283, no. 8, pp. 1392–1403, 2016, doi: 10.1111/febs.13604.

[13] Y. Fujita *et al.*, "The clinical relevance of the miR-197/CKS1B/STAT3-mediated PD-L1 network in chemoresistant non-small-cell lung cancer," *Mol. Ther.*, vol. 23, no. 4, pp. 717–727, 2015, doi: 10.1038/mt.2015.10.

[14] M. Saito *et al.*, "A Three-microRNA signature predicts responses to platinum-based doublet chemotherapy in patients with lung adenocarcinoma," *Clin. Cancer Res.*, vol. 20, no. 18, pp. 4784–4793, 2014, doi: 10.1158/1078-0432.CCR-14-1096.