



OPEN ACCESS

Edited by:

Paul Takam Kamga,
Université de Versailles
Saint-Quentin-en-Yvelines,
France

Reviewed by:

Xin Lai,
University Hospital
Erlangen, Germany
Paola Campomenosi,
University of Insubria, Italy

***Correspondence:**

Marialuisa Sensi
marialuisa.sensi@istitutotumori.mi.it
Loris De Cecco
loris.dececco@istitutotumori.mi.it

†Present address:

Loris De Cecco,
Molecular Mechanisms Unit,
Department of Research, Fondazione
IRCCS Istituto Nazionale dei Tumori,
Milan, Italy
Andrea Mariancini,
Department of Biomedical Sciences,
Humanitas University, Pieve
Emanuele, Milan, Italy
Arianna Micali,
IRCCS MultiMedica, Department of
Pathological Anatomy, Sesto S.
Giovanni, Milan, Italy
Elisa Mancinelli,
Azienda Ospedaliera Ospedale
Niguarda Ca' Granda, Department of
Services, Clinical Microbiology,
Milan, Italy
Matteo Dugo,
Department of Medical Oncology,
IRCCS Ospedale San Raffaele,
Milan, Italy

Specialty section:

This article was submitted to
Thoracic Oncology,
a section of the journal
Frontiers in Oncology

Received: 02 April 2022

Accepted: 16 June 2022

Published: 19 July 2022

A Cross-Comparison of High-Throughput Platforms for Circulating MicroRNA Quantification, Agreement in Risk Classification, and Biomarker Discovery in Non-Small Cell Lung Cancer

Chiara Gargiuli¹, Loris De Cecco^{1*†}, Andrea Mariancini^{1†}, Maria Federica Iannò¹, Arianna Micali^{1†}, Elisa Mancinelli^{1†}, Mattia Boeri², Gabriella Sozzi², Matteo Dugo^{1†} and Marialuisa Sensi^{1*}

¹ Platform of Integrated Biology Unit, Department of Applied Research and Technology Development, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy, ² Tumor Genomics Unit, Department of Research, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

Background: Circulating microRNAs (ct-miRs) are promising cancer biomarkers. This study focuses on platform comparison to assess performance variability, agreement in the assignment of a miR signature classifier (MSC), and concordance for the identification of cancer-associated miRs in plasma samples from non-small cell lung cancer (NSCLC) patients.

Methods: A plasma cohort of 10 NSCLC patients and 10 healthy donors matched for clinical features and MSC risk level was profiled for miR expression using two sequencing-based and three quantitative reverse transcription PCR (qPCR)-based platforms. Intra- and inter-platform variations were examined by correlation and concordance analysis. The MSC risk levels were compared with those estimated using a reference method. Differentially expressed ct-miRs were identified among NSCLC patients and donors, and the diagnostic value of those dysregulated in patients was assessed by receiver operating characteristic curve analysis. The downregulation of miR-150-5p was verified by qPCR. The Cancer Genome Atlas (TCGA) lung carcinoma dataset was used for validation at the tissue level.

Results: The intra-platform reproducibility was consistent, whereas the highest values of inter-platform correlations were among qPCR-based platforms. MSC classification concordance was >80% for four platforms. The dysregulation and discriminatory power of miR-150-5p and miR-210-3p were documented. Both were significantly dysregulated also on TCGA tissue-originated profiles from lung cell carcinoma in comparison with normal samples.

Conclusion: Overall, our studies provide a large performance analysis between five different platforms for miR quantification, indicate the solidity of MSC classifier, and identify two noninvasive biomarkers for NSCLC.

Keywords: liquid biopsy, circulating microRNAs, high-throughput platforms, lung cancer, microRNA signature classifier, miR-150-5p, miR-210-3p, profiling

INTRODUCTION

MicroRNAs (miRs) are a class of small (18 to 22 nt) non-coding RNAs with known roles in gene regulation (1–3). miRs can be released from cells into the extracellular space and have been detected in all tested biological fluids (1–5). Circulating miRs (ct-miRs) are either stored in particles (exosomes, microvesicles, and apoptotic bodies) or associated with RNA-binding proteins or lipoproteins, which prevent their degradation (3–5). The stability, abundance, and variety of ct-miRs made them attractive candidates as non-invasive biomarkers for diagnosing, predicting, and monitoring diseases like cancer (6–8), and increasing attention is being paid to their role in lung carcinogenesis (9–16).

In our institution, the use of ct-miRs for the early detection of lung cancer has been assessed as a complementary diagnostic tool in the context of low-dose computed tomography (LDCT) screening in large retrospective cohorts (12, 13). These studies led to the development of a plasma miR signature classifier (MSC) based on reciprocal ratios of 24 plasma miRs able to stratify individuals undergoing lung cancer screening into three levels (high, intermediate, and low) according to the risk of developing lethal lung cancer (13, 14). As assessed in samples collected from smokers within the randomized Multicenter Italian Lung Detection trial, a large retrospective validation study, MSC resulted in a sensitivity, specificity, positive predicted value, and negative predictive value of 87, 81, 27, and 99% (13). The utility of the classifier was also recently assessed, thanks to the prospective BioMILD screening trial on 4,119 high-risk volunteers, where MSC-positive participants had a 2-fold higher risk to develop lung cancer within the fourth year of screening than MSC-negative participants, independently of the low-dose computed tomography (LDCT) result (14). The risk level given by MSC reflects microenvironment-related changes associated to lung cancer development and aggressiveness. In detail, the miRs composing the classifier were found to be associated to an immunosuppressive phenotype of specific immune cell subsets, such as neutrophils, macrophages, and lymphocytes (15).

Several high-throughput platforms, based on quantitative reverse transcription PCR (qPCR) or on sequencing (miR-Seq), have been routinely used to quantify miRs in human plasma. However, there is poor consensus on the optimal methodology for the successful clinical application of ct-miR biomarkers (17–21). Pre-analytical and analytical conditions are a major source of variation in results, but many challenges remain in terms of the reliability of ct-miR quantification methods (17–21). In 2014, the “microRNA quality control study” (miRQC)

systematically evaluated 12 available miR platforms across a variety of samples including human universal reference RNA, human brain RNA, and human serum samples (17). The expression level of 196 common miRs was considered. Although no platform was consistently superior to the others, there was substantial variability in performance assessments. Only two miRs (3%) were differentially expressed (DE) by all platforms; about half of the miRs (48%) were concordant for half of the platforms. Since the miRQC study, newer platforms have emerged. Nonetheless, most recent studies report similar findings when comparing the different platforms for profiling low-copy number miRs in human biological fluids (plasma/serum) or extracellular vesicles (18–21).

A few reports have compared ct-miR abundance using multiple high-throughput technologies in defined clinical subgroups. Only one study has reported the use of multiple platforms (Toray 3D Gene System from Toray Systems, nCounter from Nanostring Technologies, and QIAseq from Qiagen) to profile cell-free and extracellular-derived miR fractions from non-small cell lung cancer (NSCLC) patients and healthy donors (20). The patients' cohort was however heterogeneous and not age-matched with the control group, preventing the interpretation of differential expression between NSCLC patients and healthy control samples for different ct-miR fractions and platforms (20). In addition, to the best of our knowledge, none of the previous studies has challenged the ability of different platforms to correctly classify individual samples according to a clinically relevant ct-miR signature.

To address these issues, we determined the miR profile of plasma samples from 10 stage IV NSCLC patients and 10 healthy heavy smokers matched for age, sex, smoking status, and MSC classification assessed with the gold-standard method (13, 14), using five well-established high-throughput methods. Three of them, Taqman OpenArray/Taqman OpenArray Advanced from Thermo Fisher Scientific and miRCURY LNA from Qiagen, were qPCR-based. The remaining two, EdgeSeq from HTG Molecular and QiaSeq miRNA Library from Qiagen, were next generation sequencing (NGS)-based. EdgeSeq allows the assessment of 2,083 human miR transcripts directly from plasma, without extraction, through quantitative nuclease protection, whereas QiaSeq is a true discovery platform enabling the capture of the whole miRNome profile.

The aims of this cross-platform comparison were assessment of intra- and inter-platform reproducibility, agreement in correctly classifying samples according to the MSC classifier, and identification and validation of putative cancer-associated ct-miRs.

MATERIALS AND METHODS

Characteristics of the Participants

Blood was collected from stage IV NSCLC patients and heavy smoker healthy individuals, as controls, with no history of cancer or other diseases. Patients and controls were classified, according to their class of risk, based on the reference MSC test generated from the ratios of 24 plasma miRs (12–14). The test was performed, as previously described, using a Custom RT and Pre-amplification Pools with TaqMan MiR Assays (Thermo Fisher Scientific, Waltham, MA, USA) (12–14). The clinical characteristics and MSC scores of the participants to the study are listed in **Table 1**. Only individuals belonging to high and low risk were included in the study. There was no significant difference in sex, age, smoking history, and nationality between the participants ($p > 0.05$).

Plasma Preparation and RNA Extraction

Blood samples, collected in P100 tubes (BD Bioscience, San Jose, CA, USA), were separated within 2 h of collection into plasma aliquots by two centrifugations of 1,600g for 10 min and stored at -80° until assayed. Total RNA was extracted from 200 μ l of plasma using the automatic nucleic acid extractor Maxwell 48 (Promega, Madison, WI, USA), eluted in nuclease-free water, and stored at -80° C. Exogenous synthetic miRs (ath-miR-159a, cel-miR-39-3p, UniSp2, UniSp4, UniSp5, and UniSp6) (Thermo Fisher Scientific and Qiagen, Hilden, Germany) were added as spike-in controls during sample processing to minimize the loss of the specific RNA template and to monitor the extraction efficiency.

ct-miR Profiling and Quality Controls

The Taqman OpenArray Human microRNA panel (OAC as Open Array “Classic” assay) (Thermo Fisher Scientific) is a fixed-content panel containing validated human TaqMan miR assays derived from Sanger miRBase release v.14. In total, 754 human miRs are amplified in each sample together with 16 replicates each of 4 internal controls (ath-miR159a, RNU48, RNU44, and

U6 rRNA). In brief, according to the manufacturer’s instructions, separate reverse transcription (RT) and pre-amplification reactions were performed on all samples using MegaPlex Pools A (v2.1) and B (v3.0) primer pools, which reverse-transcribe and pre-amplify specific miRs. The pre-amplified products were diluted before mixing with TaqMan OpenArray Real-Time PCR Master Mix and loaded onto a 384-well TaqMan OpenArray loading plate.

The Taqman OpenArray Human Advanced MicroRNA Panel (OAA) (Thermo Fisher Scientific) is also a fixed-content panel containing 754 well-characterized human miR sequences from the Sanger miRBase release v.21. The internal controls are ath-miR-159a and cel-miR-39-3p. Preparation of poly(A) tailing and adapter ligation reactions were performed, according to the manufacturer’s instructions, on all samples before RT and set-up of qPCR in a 384-well TaqMan OpenArray loading plate. The OAC and OAA products were automatically loaded from the 384-well plates onto the OpenArray plates using the AccuFill System (Thermo Fisher Scientific), and the qPCR reactions were carried out on a QuantStudio 12K Flex Real Time PCR system (Thermo Fisher Scientific). Quality controls were performed on raw data to control for batch effects and outliers. The distribution of raw Ct/Crt, AmpScore, and CqConf values of the exogenous spike-in ath-miR-159a was evaluated. Plate images were manually inspected for every sample in every run to control for evaporation, bubbles, or oil leakage. The fluorescence of ROX, a passive dye in the qPCR reagent mix, was controlled to confirm that each well was correctly loaded. Wells with a ROX signal above 1,000 were included.

The miRCURY LNA miRNome PCR Panels (miRCURY) (Qiagen) is a system based on universal RT, followed by qPCR amplification with locked nucleic acid (LNA)-enhanced primers designed for miR detection using SYBR tracking dye. In each sample, a total of 752 unique human miRs based on Sanger miRBase release 21 are profiled using miRNA ready-to-use PCR human panels I and II following the manufacturer’s instruction. The PCR panels also include three small RNA reference genes (U6, SNORD38B, and SNORD49A) and three miR reference

TABLE 1 | Clinicopathological features of the analyzed cohort.

Characteristics	Non-small cell lung cancer (NSCLC) Patients (N = 10)	Healthy Controls (N = 10)
NSCLC types		
Adenocarcinoma	5	0
Sarcomatoid	2	0
Others ^a	3	0
Stage		
IV	10	0
Gender		
Male	7	7
Female	3	3
Age		
>50	10	10
Smoking status		
Current smokers	10	10
MSC status		
High	5	5
Low	5	5

^aOthers: adenocarcinoma (N = 1), poorly differentiated (N = 1), and squamous (N = 1).

genes (miR-103-3p, miR-191-5p, and hsa-miR-423-5p), all found on panel I. Panel I also contains qPCR assays for the 5 synthetic RNAs in the RNA Spike-in Kit (cel-miR-39-3p, UniSp2, UniSp4, UniSp5, and UniSp6). After RT, qPCR reactions were carried out on a QuantStudio 12K Flex Real Time PCR system. To control for run-to-run variations, interplate calibration was performed using the six interplate calibrators, UniSp3 miR, as per the manufacturer's instruction. After the calibration of each plate, the data were merged to obtain a unique data matrix.

The QiaSeq miRNA Library (QiaSeq) (Qiagen) is a discovery platform which captures all small RNA sequences and uses unique molecular indices (UMIs) to enable an unbiased and accurate miRNome-wide quantification of mature miRs by NGS technology. Briefly, the preparation of small RNA libraries was performed according to the manufacturer's procedures. The quality and concentration of libraries were determined using Qubit™ DNA HS Assay Kit on a Qubit fluorometer (Thermo Fisher Scientific), while the library size was assessed using Agilent High Sensitivity D1000 ScreenTape on a 4200 TapeStation, (Agilent Technologies, Santa Clara, CA, USA). The libraries were sequenced on a NextSeq 500 System (Illumina, San Diego, CA, USA). Raw sequences were analyzed using the Qiagen Online Data Analysis Center with default settings, and 1,823 unique miRs were selected for the subsequent analysis.

In the EdgeSeq miR Whole Transcriptome Assay (EdgeSeq) (HTG Molecular Diagnostics, Inc., Tucson, AZ, USA), frozen plasma samples were shipped to HTG to carry out the multiplexed nuclease protection assay, sequencing, quality controls, and primary analysis of the data. The assay, which allows the assessment of miRs directly, without extraction, is based on probes containing sequences complementary to 2,083 specific miRs (miRBase v20) and flanking sequences for downstream amplification. It includes five negative process control probes to the plant gene: "ANT" (Aintegumenta, NM_119937). Probes that successfully hybridize to their cognate miR in the sample are protected from nuclease digestion, amplified with the addition of barcodes, and then sequenced on automated HTG EdgeSeq sequencer system. This study was executed at HTG Molecular in the VERI/O Laboratory following VERI/O processes and procedures. Data are provided as a data table of raw counts, QC raw, and \log_2 CPM (counts per million).

Data Import and Processing

All statistical and bioinformatic analyses were performed using the R statistical program v. 3.6.1. For the three qPCR-based panels, text files were downloaded from the QuantStudio 12K Flex and were imported in R as data tables. The expression matrices in qPCRset format were created for every dataset using the HTqPCR R package (22). Filtering on detection was performed according to the manufacturers' suggested thresholds: $Crt \leq 28$, $AmpScore > 1$, and $CqConf > 0.8$ for OAC and OAA panels; $Ct \leq 35$ and $AmpScore > 1$ and $CqConf > 0.8$ for miRCURY panels I + II. If miRs did not reach the thresholds, they were set to 40 and considered as "undetected".

Since different miRbase versions were used to design the platforms, we downloaded the platform annotations from each manufacturer's website and, using the mature sequence identifier, we converted miR names to miRbase version 21. For qPCR-based platforms, data were normalized using the global median normalization method with the median values of detectable miRs. For QiaSeq, primary analysis was performed with the GeneGlobe online software (<https://geneglobe.qiagen.com/sg/analyze/>). Raw counts were normalized using the trimmed mean of M-value (TMM) method (23) implemented in the edgeR package (24) considering that only the UMI counts had more than 10 counts mapping in at least 30% of samples. For EdgeSeq, raw counts were corrected by background subtraction of the maximum value of the five ANT probes. In addition, control miRs were removed, and miRs with negative counts after the background correction were set to 0 for the subsequent normalization performed using the TMM method (23).

Guanine-Cytosine Content Evaluation

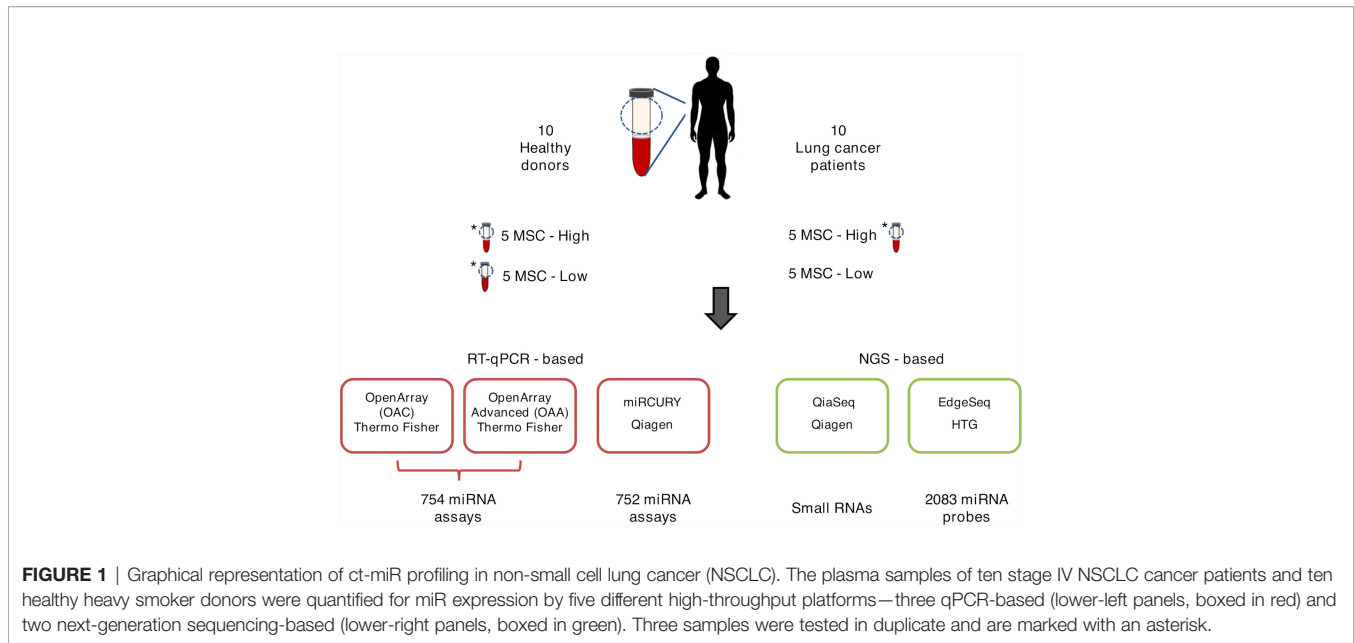
Guanine-cytosine (GC) content was calculated for detected and undetected miRs common to all platforms ($n = 488$). The percentage of GC was calculated as the sum of G and C present in every miR sequence divided by the length of the sequence and multiplied by 100. Differences between detected and undetected miRs in each platform were assessed with Wilcoxon rank-sum test.

Correlation and Concordance Analysis

Three samples deriving from a patient and two healthy subjects were profiled twice each using, depending on the platform, either independent RNA extractions of the same plasma or duplicate aliquots of crude plasma (Figure 1). The concordance and correlation coefficient (CCC) was calculated using the $\Delta Ct/Crt$ and $\log_2(CPM)$ values on pairs of technical replicates for each platform with the epi.ccc function of epiR package (<https://cran.r-project.org/web/packages/epiR/epiR.pdf>). Hierarchical clustering was performed using Euclidean distance and Ward method.

MSC Algorithm

The plasma-based MSC test analyzes the reciprocal levels of 24 ct-miRs (listed in Supplementary Table S1) by qPCR. The expression values of these miRs were determined by gold-standard methodology. Briefly, the Multiplex Pools Protocol on custom-made microfluidic cards (Thermo Fisher Scientific) containing the 24 miRs spotted on duplicates was used as described (14, 15). To remove the batch effect, a ratio-based approach, using the gold-standard methodology as reference array, was first adopted (25). In detail, the normalized data of the 24 miRs from each platform were scaled by the arithmetic mean of the reference array. The fixed MSC algorithm (26) was then applied to the 24 scaled miR profile obtained for each sample in each platform, taking into account the single values. The MSC risk scores were compared with those calculated in the same samples by the gold-standard methodology (Table 1). Cohen's kappa was used to assess the agreement between platforms for MSC classes.



Differential Expression Analysis and Concordance Rate Between Platforms

Differential expression analysis was carried out on normalized data using the linear modeling approach implemented in the limma package (27). Nominal *p*-values were corrected for multiple testing using the Benjamini–Hochberg false discovery rate (FDR). DE ct-miRs were selected according to an FDR <0.1 in all the platforms. We then assessed the pairwise concordance of fold changes (FC) between platforms (platform X vs. platform Y). Four qualitative evaluations were assigned to each comparison: compressed, opposite, overestimate, or concordant (28). When the compared FC were in the same direction but the ratio of X/Y was greater than or equal to 2, a value of “compressed” was assigned. Similarly, if the FC ratio of X/Y is less than or equal to 0.5, the comparison was deemed “overestimate”. FC ratios between these values were named “concordant”. When two FC values were not in the same direction and either of them was greater than 2 or less than 0.5, the comparison was determined to be “opposite”. Concordance rates were calculated by number of miRs with “concordant” and “overestimate” calls divided by the total number of analyzed miRs which were in common and expressed in all the platforms.

Individual qPCR Assays

Single qPCR reactions were performed using TaqMan MicroRNA Assays (hsa-miR-150-5p and hsa-miR-93-5p, Thermo Fisher Scientific) according to the manufacturer’s instructions. Briefly, total RNA (3 µl) was reverse-transcribed, and the resulting cDNA was used (2.5 µl) for the pre-amplification reaction. The pre-amplified cDNA was diluted 1:12, and 0.10 µl of the product was used to perform the qPCR amplification reaction using the corresponding miR assay

primers and TaqMan Universal PCR Master Mix no AmpErase UNG, according to the manufacturer’s instructions. The PCR reaction conditions were as follows: enzyme activation at 95°C for 10 min, 40 cycles of denaturation at 95°C for 15 s, and annealing/extension at 60°C for 60 s. The amplification was performed in 384-well plates with QuantStudio 12K Flex Real Time PCR system (Thermo Fisher Scientific) assembled using the Janus automated workstation (PerkinElmer, Waltham, MA) from 96-well plates. Each qPCR analysis was done in triplicate, and data were acquired through QuantStudio 12K Flex v.1.2.3; the obtained mean Ct values were exported for statistical analysis. miR-93-5p was identified as a reference housekeeper by all the platforms using the selectHKgenes function with Vandesompele method (29) of SLqPCR R package (<https://bioconductor.org/packages/release/bioc/html/SLqPCR.html>) calculated on filtered raw data of each platform. The expression levels of miR-150-5p were then normalized according to the DCT method (30) using the Ct mean values of the endogenous control.

External Validation

External validation was performed in The Cancer Genome Atlas (TCGA) dataset. Raw count values for the TCGA miR-seq data of lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) were downloaded from the Genomic Data Commons data portal (<https://portal.gdc.cancer.gov/>). The LUAD project included 519 primary solid tumors, 2 recurrent tumors, and 46 normal samples from adjacent tumor tissues; the LUSC project included 478 primary solid tumors and 45 normal samples (31, 32). TCGA raw count values, samples, and patients’ annotations were obtained using the TCGABiolinks package (33). miRs with less than 10 counts expressed in more than 50% of samples were filtered out. Raw counts were then normalized with the TMM method implemented in the edgeR

package (23, 24). Differential expression between tumor and normal tissue was performed using the limma/voom method (27). Nominal *p*-values were corrected for multiple testing using the Benjamini–Hochberg FDR.

ROC Curves

Receiver operating curves (ROC) with area under the curve (AUC) calculation were used to determine the diagnostic value of miRs in distinguishing between plasma from healthy controls and NSCLC patients (34). ROC curves were obtained by plotting sensitivity against specificity using the ROC function of pROC R package (34). An area greater than 0.5 under the curve suggests the diagnostic potential of each ct-miR candidate.

RESULTS

Study Design and ct-miR Expression Profiling

A total of 20 human specimens were employed for this study, which included plasma from NSCLC patients ($n = 10$) and healthy subjects ($n = 10$) matched for age, sex, smoking status, and MSC risk score (Table 1). Three plasma, one derived from a patient and two from healthy subjects, were in duplicate, bringing the total number of analyzed samples to 23 (Figure 1). RNA derived from these samples was profiled by the following four high-throughput technological platforms: Taqman OpenArray Human miR and Taqman OpenArray Human Advanced miR Panels (Thermo Fisher Scientific), miRCURY LNA miR miRNome PCR Panels (Qiagen), and QiaSeq miRNA Library (Qiagen) (Figure 1). The fifth platform, EdgeSeq miR Whole Transcriptome Assay (HTG Molecular Diagnostics), employed crude blood plasma instead (Figure 1). The starting material for the duplicates was a second aliquot of either crude plasma (for EdgeSeq platform) or RNA independently extracted (for all remaining platforms). Their inclusion was required to assess intra-platform repeatability as described below. The presence and detection of miRs by platform and sample is reported in Supplementary Table S2. The number of common miRs detectable by all platforms was 488 (Supplementary Figure S1A and Supplementary Table S2). For each platform, the average number of ct-miRs detected after normalization and filtering in the different samples ranged from 236 for EdgeSeq platform to 806 for QiaSeq (Supplementary Figure S1B). By considering only the 488 commonly detected miRs, average detection ranged from 120 to 323 (Figure 2A). As shown in Supplementary Figure S1C, the influence of GC content had no or little impact on the detection rate. The 488 common miRs included a list of 26 miRs (named super_core in Supplementary Table S2) highly expressed in all plasma samples as indicated by the empirical cumulative distribution curves of their expression quantiles (Supplementary Figure S1D).

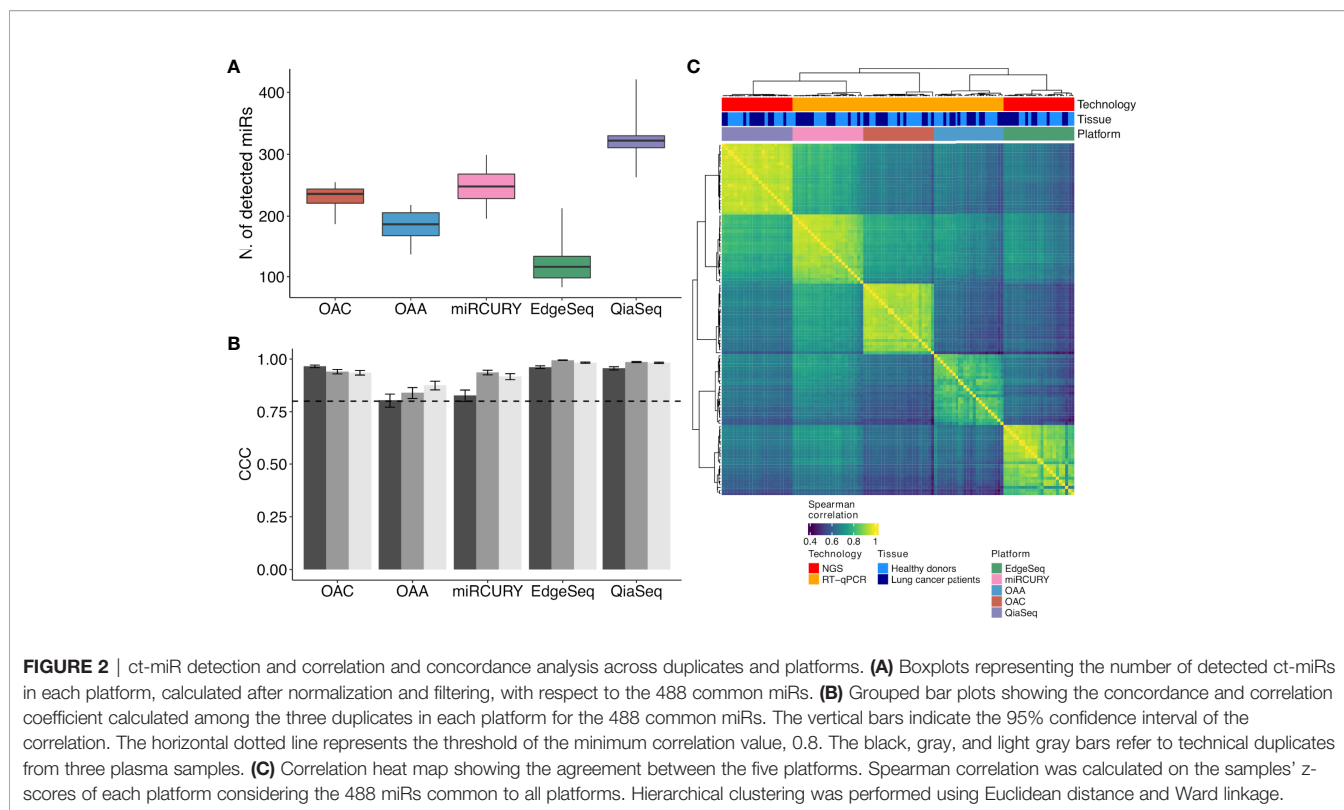
Intra-platform Repeatability and Inter-platform Comparison

To evaluate the intra-platform repeatability, we calculated for each platform the Lin's CCC between the ct-miR profiles of

duplicate samples. CCC between duplicates was >0.8 for all platforms considering either the 488 common miRs (Figure 2B) or the total number of available miRs (Supplementary Figure S2). Pairwise scatterplots for duplicates are displayed in Supplementary Figure S2. These results demonstrate intra-platform consistency and no significant differences among the different technologies. We then calculated Spearman's correlation coefficients between pairs of samples within and between platforms. Hierarchical clustering of the correlation matrices showed that each platform produced very homogenous and highly correlated data (Figure 2C). Within each platform, we did not observe any separate cluster of tumor and normal samples. This suggests that, independently of the platform, most of the ct-miRs are uninformative to distinguish the two groups. We did instead observe clustering according to the profiling platform, indicating that the variability explained by the technological approach is higher than the biological variability. An unsupervised hierarchical clustering algorithm was carried out on Spearman's correlation coefficients calculated between the pair of platforms for each of the 488 common miRs. Four major clusters were identified according to different levels of correlation (Figure 3A). Cluster 1 comprised 17% of miRs displaying the lowest inter-platform correlation for all pairs of platforms. Cluster 3 included 32% of miRs that had low expression levels in all platforms and that were highly correlated when comparing qPCR-based platforms but were negatively correlated between NGS- and qPCR-based platforms. Cluster 2 included 35% of miRs and showed a heterogenous pattern of correlation. A first subset of miRs showed a positive correlation in all comparisons, whereas a second subset showed negative correlations when the comparisons were against OAA, indicating that the expression of these ct-miRs is inconsistent specifically for this platform. Finally, cluster 4 included 15% of miRs that were highly expressed in all platforms and showed a high inter-platform correlation. EdgeSeq did not correlate with any other platforms since many of the 488 miRs showed an expression value of 0 in all samples. miRs belonging to each cluster are reported in Supplementary Table S3. The Spearman correlation coefficients among the six platforms shown in the right boxplot of Figure 3A indicate that the highest inter-platform reproducibility was observed between qPCR-based platforms (miRCURY, OAC, and OAA). For each pair of comparisons between platforms, we counted the number of ct-miRs above increasing correlation cutoffs (Figure 3B). We confirmed that, independently of the correlation cutoff, the comparison between qPCR-based platforms returned the highest number of correlated ct-miRs, especially for OAC vs. miRCURY. Comparisons including EdgeSeq showed the lowest number of correlated ct-miRs due to the lower detection rate of EdgeSeq compared with the other platforms.

Cross-Platform Concordance in the Assignment of a Clinical Validated miR Risk Score

Our cohort consisted of subjects equally distributed within high and low risk (Table 1) in both classes (NSCLC and controls) as



previously assessed by the gold-standard methodology. MSC algorithm was adopted to classify each sample according to the expression profiles of the 24 ct-miRs (**Supplementary Table S1**) determined in each platform. The classification of each sample (including duplicates from two subjects) for each platform is displayed in **Supplementary Table S4A**. All qPCR-based platforms and QiaSeq displayed a classification highly concordant to the original assessment by the gold-standard method (**Figure 3C** and **Supplementary Table S4A**). The same results were obtained by computing Cohen's kappa statistics as pairwise measure of similarity when each platform was confronted to the reference (**Supplementary Table S4B**). A lower fidelity was displayed by EdgeSeq when compared with the reference (**Supplementary Tables S4A, B**). Except for EdgeSeq, all other platforms correctly classified all samples from MSC-low individuals, whereas the situation was more heterogeneous for MSC-high individuals (**Supplementary Table S4A**). Overall, OAA, miRCURY, and QiaSeq were the three platforms with 91% of correctly classified samples, followed by OAC (82%) and EdgeSeq (59%) (**Figure 3C** and **Supplementary Table S4A**). These results demonstrate that the classification obtained by the standard protocol could be replicated with a good agreement using at least two qPCR-based technologies and one sequencing technology.

Differential ct-miR Modulation in NSCLC Patients Compared to the Healthy Control Group

To evaluate the differential expression concordance among platforms, we identified DE ct-miRs between NSCLC patients

and healthy donors for each platform. The number of miRs that passed the detection filter and were available for the contrast differed among platforms: 689 for QiaSeq, 337 for miRCURY, 305 for OAC, 269 for OAA, and 246 for EdgeSeq. Among the 488 miRs measured by all platforms, those commonly detected were over 80% for qPCR technologies, were 50% for EdgeSeq, and dropped to 44% for QiaSeq. In total, 100 miRs were altogether detected by all platforms, 164 by all but EdgeSeq, which presented the lowest number of ct-miRs passing the detection filters. The results of the DE analysis for all platforms are presented in **Supplementary Table S5**. For each platform, we evaluated the number of DE ct-miRs at varying FDR thresholds, ranging from 0.25 to 0.01 (**Supplementary Figure S3A**). On average, the miRCURY platform gave the highest number of DE ct-miRs, followed by OAC, QiaSeq, and EdgeSeq. No DE ct-miRs were identified for OAA at any FDR threshold. At the usual FDR <0.05, the miRCURY platform gave 43 DE ct-miRs, followed by QiaSeq ($n = 5$) and EdgeSeq ($n = 1$). No DE ct-miRs were found for OAC and OAA at an FDR <0.05. Considering a stringent FDR of 0.01, only QiaSeq identified two DE ct-miRs. We next evaluated the intersection between the lists of DE ct-miRs identified for each platform at different FDR thresholds (**Supplementary Figure S3B**). At FDR <0.01, no DE ct-miRs were shared between two or more platforms. At FDR <0.05, one ct-miR was identified by three platforms and two by two platforms. At increasing FDR, the number of shared DE ct-miRs across platforms increased. Since the selection of DE ct-miRs by different FDR cutoffs influences the comparison of the platforms, we evaluated the correlation of the t-statistics to

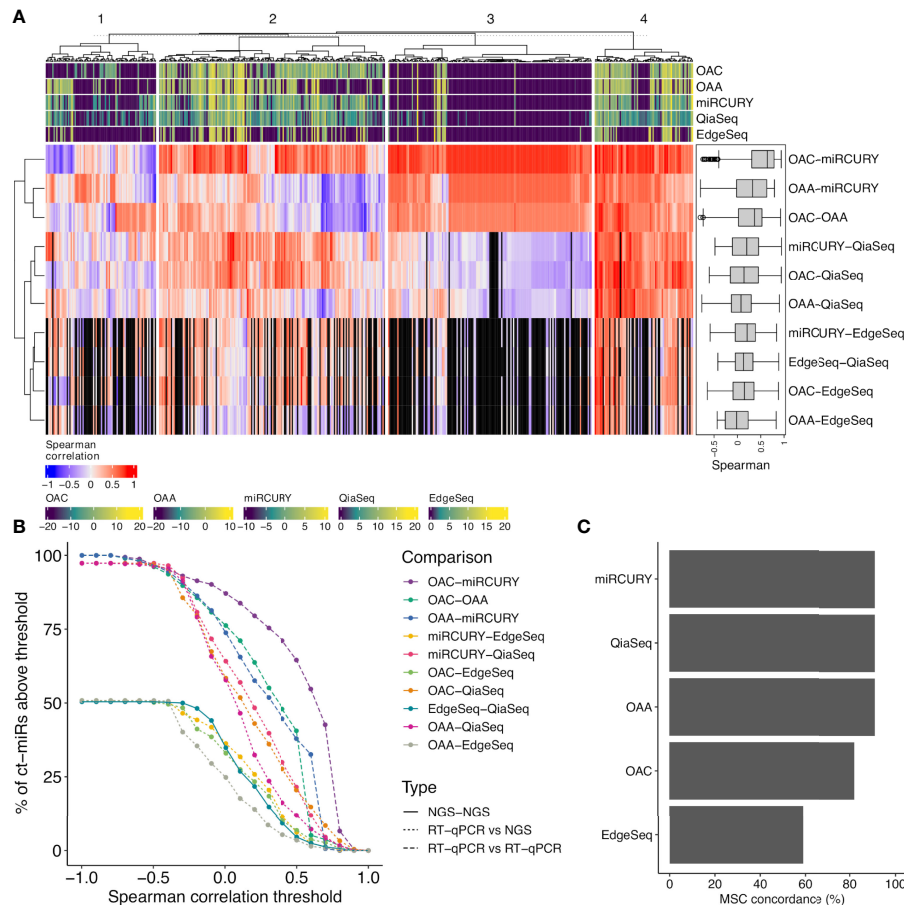


FIGURE 3 | Correlation analysis between each pair of platform and concordance assessment for the miR signature classifier (MSC). **(A)** The correlation heat map shows how the different platforms correlate with respect to the expression values of the 488 common miRs. The median pairwise Spearman's correlation values are shown also as boxplots in a black box (right corner). The colored bars on the top and bottom of the heat map (violet- to yellow-colored gradient) define the median-normalized expression values of each platform. Four functional groups are identified and defined according to different levels of correlation (1, scarce; 2, intermediate; 3, both positive (red) and negative correlation (blue); 4, high). The black vertical bars represent the miRs with an expression value of 0 in EdgeSeq platform that do not correlate. **(B)** Curves showing the number of ct-miRs correlated above increasing correlation cutoffs for each pairwise comparison between platforms. **(C)** Bar plots displaying the percentage of concordance in assigning the label miR risk classifier MSC—high or MSC—low compared with the reference platform (Custom-made Microfluidic Cards, Thermo Fisher) used to calculate the clinical validated score.

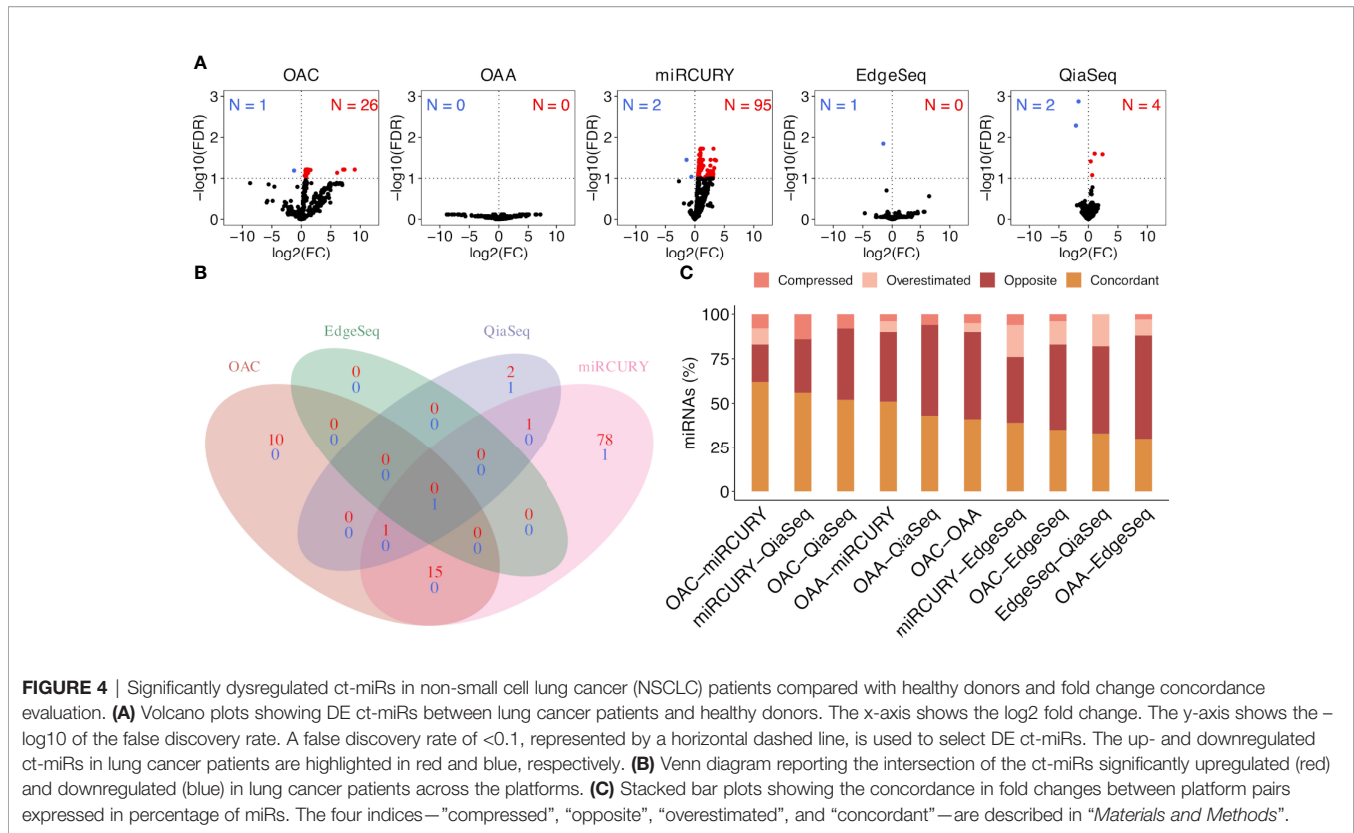
assess whether at least the direction of the modulation was concordant across platforms (**Supplementary Figure S3C**). All pairwise comparisons between platforms showed positive correlation values, indicating that, on average, the trend of modulation of ct-miRs between lung cancer patients and healthy donors was similar between platforms. However, only OAC vs. miRCURY and QiaSeq vs. miRCURY had correlation values higher than 0.5.

To select ct-miRs DE in at least four platforms, we therefore applied an FDR cutoff of 0.1. Volcano plots representing the results of the DE analysis between lung cancer patients and healthy controls at an FDR <0.1 are shown in **Figure 4A**. At a threshold of FDR <0.1, we detected 27 DE ct-miRs on OAC, 6 on QiaSeq, 97 on miRCURY, 1 on EdgeSeq, and none on OAA, corresponding to 4.3, 0.5, 5.2, 0.8, 0.4, and 0% of miRs available for the contrast. A Venn diagram displaying the intersection between the lists of significantly up- or downregulated ct-miRs in

each platform is shown in **Figure 4B**. Among upregulated ct-miRs, at FDR <0.1, one was common to OAC, miRCURY, and QiaSeq platforms, whereas 16 were commonly detected on two of them. The downregulated ct-miRs included 1 miR shared by miRCURY and QiaSeq and 1 common to the four platforms (miR-150-5p, FDR <0.05 in miRCURY, QiaSeq, and EdgeSeq; FDR <0.1 in OAC).

Fidelity of Fold Change Across Platforms and Experimental Validation of miR-150-5p

We selected all ct-miRs identified as DE in at least one platform and evaluated the fold change concordance between platforms as defined in the “Materials and Methods” section. As shown in **Figure 4C**, the highest rate of concordant miRs was found between OAC and miRCURY, followed by either OAC or miRCURY compared with QiaSeq. The percentage of miRs displaying fold changes in the opposite direction increased



when the comparisons were done against EdgeSeq and OAA. miR-142-3p was the concordant upregulated ct-miR across OAC (FDR <0.1), miRCURY (FDR <0.05) and QiaSeq (FDR <0.1) in the plasma of NSCLC patients compared with healthy subjects. The ct-miRs concordantly and significantly upregulated in at least two platforms were as follows: miR-590-3p, miR-766-3p, miR-103a-3p, miR-98-5p, miR-296-5p, miR-191-5p, and miR-24-3p (OAC and miRCURY) and let-7f-5p (QiaSeq and miRCURY). Except for miR-142-3p and miR-98-5p, absent from EdgeSeq, all others belonged to the 488 miRs present on all platforms. The majority was included in cluster 2, whereas miR-590-3p and miR-191-5p were in cluster 4 (**Figure 3A** and **Supplementary Table S3**). The only ct-miR downregulated in cancer patients compared with donors, miR-150-5p (cluster 4, **Figure 3A** and **Supplementary Table S3**), was significantly DE in four platforms (**Figure 5A**). A trend toward significance (nominal p -value = 0.008) was also observed in the fifth platform (OAA) (**Figure 5A**). To further investigate the robustness of differences in the abundance of miR-150-5p, individual qPCR assays were performed. By ranking the 26 ct-miRs detected in all samples and platforms according to their average rank across platforms (**Supplementary Table S6**), it was shown that hsa-miR-93-5p is the most stable ct-miR in the cohort and was selected as the normalizer for the single assay. The results confirmed that the relative normalized expression of miR-150-5p in the plasma of NSCLC patients was significantly lower than in healthy donors (**Figure 5A**).

Discrimination of NSCLC Patients and Controls by Receiver Operating Characteristic Curves

To assess the translation of differential expression into diagnostic power, we evaluated the ability of ct-miRs to discriminate NSCLC patients from controls using a ROC curve analysis. Overall, OAA and OAC showed the highest number of potentially diagnostic ct-miRs, followed by QiaSeq, miRCURY, and EdgeSeq (**Supplementary Figure S4A**). The performance of the platforms in identifying diagnostic ct-miRs varied according to the AUC cutoff selected but, in general, decreased rapidly at increasing values of AUC. At AUC >0.8, miRCURY was the best-performing platform, followed by OAC, QiaSeq, OAA, and EdgeSeq. At AUC >0.9, miRCURY was again the top-ranking platform with 7 diagnostic ct-miRs, followed by QiaSeq, OAC, EdgeSeq, and OAA. We next compared the lists of ct-miRs with AUC above a certain threshold (**Supplementary Figure S4B**). For AUC >0.8, only one ct-miR was shared by at least four platforms. Upon increasing the AUC to 0.9, no shared ct-miRs were found for four and five platforms and only one for at least two or three platforms. The correlation of AUC values showed a poor consistency between platforms, with correlation values ranging from a minimum of -0.267 for OAA vs. QiaSeq to a maximum of 0.407 for miRCURY vs. QiaSeq (**Supplementary Figure S4C**). The intersections between miRs with an area under ROC curve (AUC) value above or equal to 0.7 are shown with a Venn diagram in **Figure 5B**. Two ct-miRs, miR-150-5p and

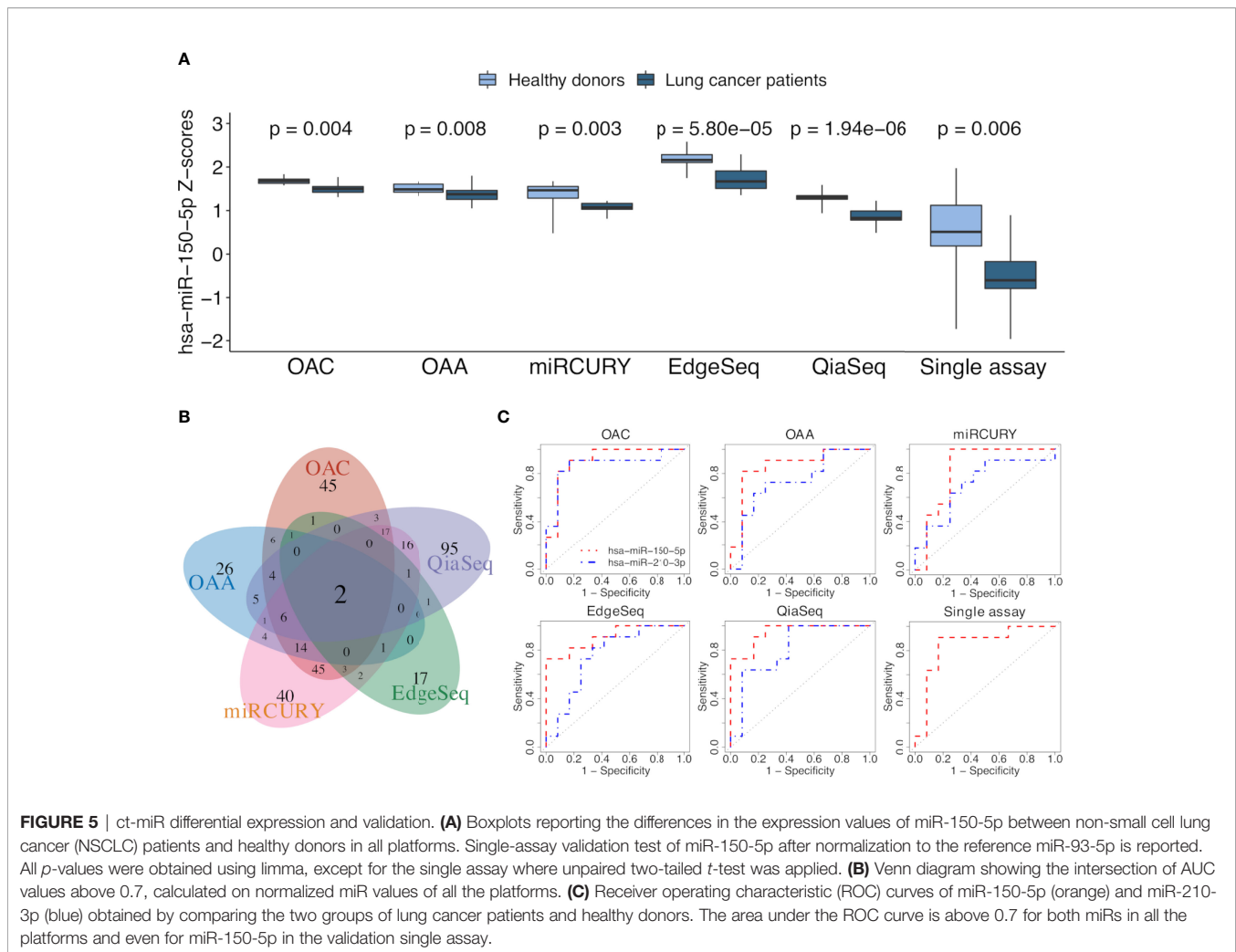


FIGURE 5 | ct-miR differential expression and validation. **(A)** Boxplots reporting the differences in the expression values of miR-150-5p between non-small cell lung cancer (NSCLC) patients and healthy donors in all platforms. Single-assay validation test of miR-150-5p after normalization to the reference miR-93-5p is reported. All p -values were obtained using limma, except for the single assay where unpaired two-tailed t -test was applied. **(B)** Venn diagram showing the intersection of AUC values above 0.7, calculated on normalized miR values of all the platforms. **(C)** Receiver operating characteristic (ROC) curves of miR-150-5p (orange) and miR-210-3p (blue) obtained by comparing the two groups of lung cancer patients and healthy donors. The area under the ROC curve is above 0.7 for both miRs in all the platforms and even for miR-150-5p in the validation single assay.

miR-210-3p, were in common to all the platforms. Although miR-210-3p upregulation in NSCLC was statistically significant only in QiaSeq (FDR <0.1), it displayed the same trend of modulation in the other platforms (data not shown). As shown in **Figure 5C** and **Supplementary Table S7**, the AUC value of miR-150-5p, including the single-assay qPCR results, ranged from 0.95 for QiaSeq [95% confidence interval (CI): 0.87–1] to 0.83 for miRCURY (95% CI: 0.64–1). The AUC for miR-210-3p ranged from 0.87 for OAC (95% CI: 0.7–1) to 0.71 for miRCURY (95% CI: 0.49–0.94). These results indicate that all platforms can detect the discriminatory power between NSCLC patients and healthy donors of these two ct-miRs, even if the accuracy is platform dependent.

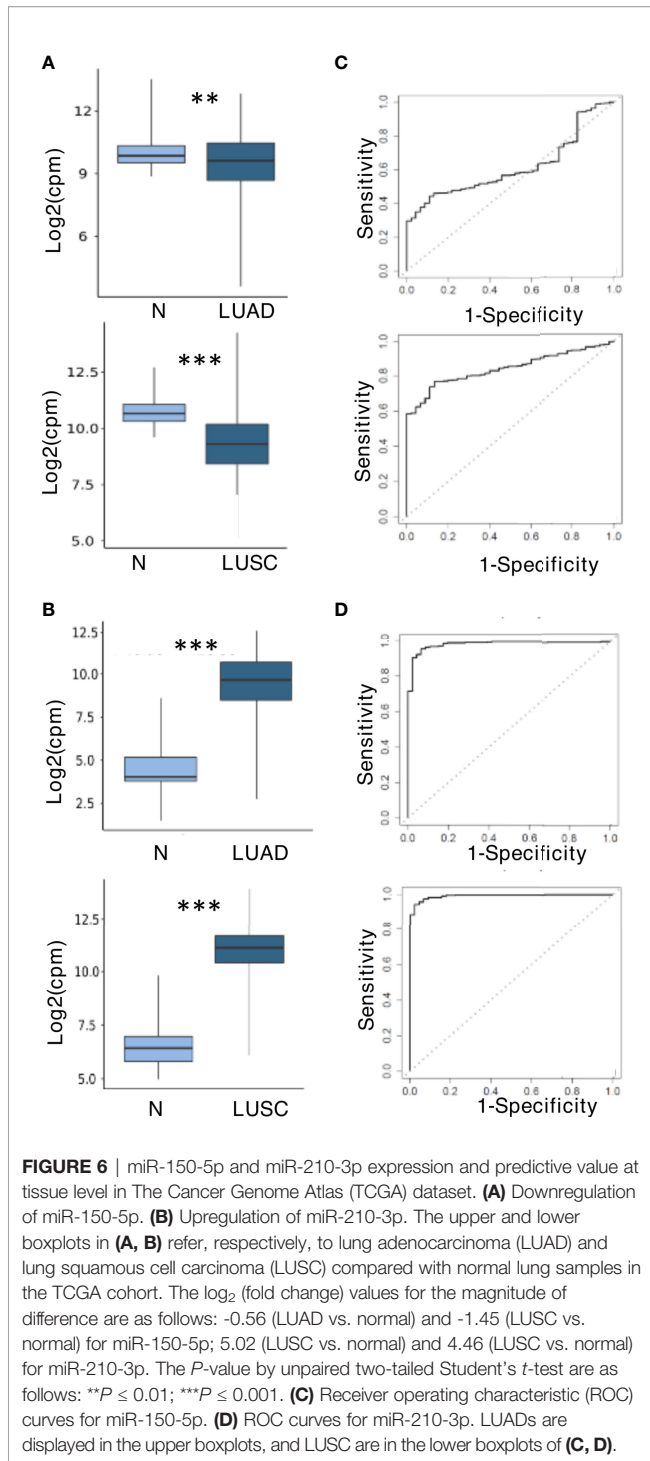
Validation of miR-150-5p and miR-210-3p as Potential Biomarkers in Tissues

To further explore the role of miR-150-5p and miR-210-3p as potential biomarkers even for NSCLC tissues, the TCGA miR sequencing data for tumors and normal tissues of patients affected by LUAD and LUSC were analyzed. The results shown in **Figures 6A, B** indicated that the trend of dysregulation of

these two miRs at the tissue level agreed to that observed in plasma. In comparison with normal tissues, the downregulation of miR-150-5p was however higher for LUSC than LUAD (**Figure 6A**). In contrast, miR-210-3p was significantly upregulated in both histologies (**Figure 6B**). ROC curve analysis was performed to evaluate the diagnostic value of the two miRs at the tissue level. As shown in **Figures 6C, D**, they appeared to represent valuable diagnostic markers. The miR-210-3p AUC values were high in both LUAD and LUSC cohorts at 0.98 and 0.99, respectively (**Figure 6D**), whereas those for miR-150-5p had higher AUC in LUSC (0.84) than in LUAD (0.61) (**Figure 6C**). The related data corresponding to AUCs are summarized in **Supplementary Table S7**.

DISCUSSION

We here analyzed the miR profiles of the plasma fluids of 10 NSCLC lung cancer patients and 10 healthy donors by using five different high-throughput platforms that are among the most commonly used commercially available technologies.



Each platform was assessed for performance parameters (intra-platform reproducibility, detection rate, and inter-platform correlation), for MSC classification concordance, and for the ability to detect differences between biological groups (e.g., healthy individuals and patients). The ct-miR detection rate was more similar across qPCR technologies. QiaSeq exhibited the highest miR counts in all sample groups, indicating that it is a

true discovery technology that greatly expands miR repertoire detection and allows the identification of novel miRs. In contrast, EdgeSeq, directly performed on crude human plasma specimens without RNA extraction, displayed the lowest sensitivity. The intra-platform reliability, assessed by calculating pairwise concordance correlation coefficients between duplicates, was very high for all platforms except for OAA which had slightly lower CCC values. The results from the inter-platform reproducibility are consistent with those of previous studies, indicating that the overlap between different technologies is small (17–21). Our clustering analysis demonstrated that the correlation between ct-miRs depends on the platform and on the expression level of the miRs. The highest inter-platform reproducibility was observed between the qPCR-based platforms miRCURY, OAC, and OAA. EdgeSeq, which displayed a very high number of miRs with an expression value of 0 after background correction, had the lowest number of miRs, showing a low inter-platform correlation. It is the only technology that performs direct miR-targeted sequencing without RNA extraction procedures, and the results probably reflect the lower sensitivity for the quantification of low-abundance miRs as already reported (19). However, for specific highly expressed ct-miRs such those in cluster 4, it showed an inter-platform correlation comparable with the other platforms. Our study pinpoints the challenges inherent to the choice of a downstream detection technology for ct-miR profiling in a clinical setting and advises the use of a dual-platform approach to overcome the limitations of single platforms. If cost will prevent this approach, the aim of the experiment should be considered. At a discovery stage, unbiased high-throughput screens of miRs like that offered by QiaSeq small-RNA sequencing could be recommended. High-throughput qPCR technique by miRCURY or OAC could be also a good option for discovery as well as for more focused studies. Regardless of the platform used, putative biologically relevant miR biomarkers should be further validated by an independent technology. Except for EdgeSeq, concordance of MSC classification to the gold-standard assay was high for all other platforms, in particular, for OAA, miRCURY, and QiaSeq, establishing that the classifier could be reproducibly implemented in other multiplexed platforms.

Despite the fact that many studies investigating plasma miRs in patients with NSCLC provide evidence of the potential value of ct-miRs as non-invasive biomarkers, uncertainties remain regarding the clinical validity and utility of dysregulated ct-miRs for lung cancer diagnosis, prognosis, and prediction of response to treatment (35, 36). There are many reasons underlying the variability among published studies, including the use of different technologies and platforms, as also shown here, and heterogeneity of clinical cohorts. Indeed a recent multicentric study in the context of the EU network CANCER-ID reported low concordance among the miR results obtained by comparing two hybridizations (Toray 3D and nCounter), one sequencing (QiaSeq), and two qPCR (miRCURY and two-tailed qPCR) on biological samples composed of cell-free and extracellular-derived miR fractions

from NSCLC patients ($n = 27$) and healthy control samples ($N = 20$) (20). In addition, no common DE ct-miRs among cancer patients and donors were detected by the different quantification technologies. This result could be related to the composition of the NSCLC cohort which included different stages of the disease before and during systemic treatment or radiotherapy as well as before and after surgery and to the control cohort not age-matched with that of the patients (20).

Though smaller, our NSCLC cohort was properly matched with healthy heavy smoker donors and allowed the detection of differential ct-miR expression between cancer patients and controls. The ability to detect statistically significant DE ct-miRs was platform dependent. However, when we disregarded the statistical significance and we focused on the direction of the modulation, we observed that most of the ct-miRs were concordant, except for EdgeSeq and OAA that displayed a high number of discordant DE ct-miRs. Finally, all platforms identified miR-150-5p and miR-210-3p as the best circulating biomarkers able to discriminate NSCLC patients from healthy donors. Of note is the fact that since these two groups were matched for MSC test results, we can speculate that miR-150-5p and miR-210-3p are diagnostic markers independent of the MSC test result.

They were also confirmed at the tissue level, where the same trend of significant differences was observed in comparison with healthy tissue. miR-150-5p was also validated in the same plasma samples using single-assay qPCR, which is considered the gold-standard method for expression quantification.

miR-150-5p plays a critical role in the development of lymphoid and myeloid lineages in both mice and humans and has been observed to be dysregulated in solid and hematological malignancies where, depending on the context, it can exert oncogenic or oncosuppressor functions (37, 38). Several lines of evidence point to its downregulation in different human cancers, like head and neck squamous carcinoma, cholangiocarcinoma, prostate, and hepatocellular carcinoma (39–42), supporting the tumor suppressor role of miR-150-5p. In addition, it was found downregulated in non-neoplastic diseases like advanced heart failure, critical illness, and sepsis (43, 44). The results on the expression and role of miR-150-5p in NSCLC are however conflicting since both oncogenic and tumor suppressor functions have been reported (37, 45–48). Its expression level, as detected in tissues by *in situ* hybridization, negatively correlates with metastasis, including lymph node and distant metastasis, at the time of diagnosis (45). The follow-up data indicate that patients with a low expression of miR-150-5p have a poor progression-free survival rate and a poor overall survival rate compared with those with high miR-150-5p expression (45). Conversely, as assessed by qPCR, the expression of miR-150-5p was found at levels significantly more elevated in NSCLC in comparison with that in non-tumor tissues (46). At the circulating level, once again, either up- and downregulation in plasma or serum of different cohorts of NSCLC patients in comparison with healthy donors was reported (49, 50). By profiling blood plasma miRs in NSCLC patients and healthy individuals using the miRCURY platform with the LNA qPCR Serum/Plasma Panel, the upregulation of miR-210 and the

downregulation of miR-150-5p were observed for both pre-miR and mature miR levels (49). Our results agree with the above-described studies but contradict the finding indicating that the plasma levels of miR-150 and miR-210, among a panel of 12 candidate miRs, were both significantly upregulated in the plasma of NSCLC patients compared with healthy controls (50). As previously mentioned, several parameters like differences in research design, populations and specimens, and experimental methods can be relevant for inconsistencies from study to study. In addition, normalization of expression is a common challenge of miR studies in biological fluids in the absence of stable normalizers. Therefore, the function of miR-150-5p in NSCLC warrants further investigations. Nonetheless, it is worth to point out that, in our study, by applying distinct normalization strategies for data derived from small RNA sequencing, high-throughput qPCR methods, and individual qPCR assay, miR-150-5p was found to be coherently downregulated in plasma samples from NSCLC patients by five different miR profiling platforms, starting from different materials (RNA and crude plasma) and further validated by a single assay.

The role of miR-150-5p downregulation in the early diagnosis of lung tumor development is further supported by recent findings in chronic obstructive pulmonary disease, often associated with comorbidities and an increased risk of cancer, in a large-scale collection of samples from patients without cancer at baseline but with follow-up data (51–54). An increasing number of new strategies for therapeutic miR approaches are currently being pursued to restore the level of downregulated miRs and regain their tumor suppressor function (55). miR-150-5p activity as tumor suppressor has been related to its ability to inhibit wingless (Wnt)- β -catenin signaling pathway, closely associated with NSCLC progression, by targeting known activators like glycogen synthase kinase 3 beta interacting protein, β -catenin, and high mobility group AT-hook 2 (45, 56) as well as to reduce the matrix metalloproteinase 14 (MMP14) levels, whose overexpression correlates with a poor prognosis in NSCLC patients (47–57). The regulation of miR-150-5p is complex, and several long noncoding RNAs or circular RNAs can promote NSCLC cell growth and metastasis through sponging miR-150-5p (58–62).

At difference to miR-150-5p, miR-210-3p has been unambiguously described as a promising biomarker for NSCLC lung cancer due to its upregulation at the tissue, plasma, and serum levels and to its discriminatory accuracy in patients *versus* healthy controls (10, 12, 63–71). Investigations into the effects of miR-210 on lung cancer cell behavior as well as into the specific mechanisms underlying the role of miR-210 in the pathogenesis of NSCLC have been performed. It has been shown to regulate proliferation and apoptosis by targeting the transcriptional regulator SIN3A (69), a tumor suppressor gene for NSCLC cells (70). In addition, exosomal miR-210-3p derived by cancer stem cells targets fibroblast growth factor receptor-like 1 to elicit a pro-metastatic phenotype (71).

In conclusion, our study provides a comparison of ct-miRs, relevant in NSCLC, using widely used high-throughput platforms. We could show that the correlation between ct-miRs depends on both the type of platform and the miRs expression levels. Indeed a high inter-platform correlation was observed for ct-miRs profiled in qPCR-based platforms and, for

all platforms, within highly expressed ct-miRs. Concordance of MSC classification among most miR detection technologies with the “gold-standard” method established that the classifier could be successfully implemented in other multiplex platforms. Finally, we here demonstrate, for the first time, that the decreased abundance of miR-150-5p and the increased abundance of miR-210-3p in the plasma of lung cancer patients is independent of the detection technology. Both miRs display promising attributes and constitute attractive circulating biomarkers for NSCLC cancer detection. Larger and prospective studies composed of patients with different NSCLC histological cancer subtypes and at different stages of the disease are needed to confirm their significance.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE204951. Normalized data are available in **Supplementary Table S8**.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MS, LDC, GS, MB, and MD conceived and designed the study. MB and GS provided samples. LDC, AM, MFI, MB, AM, and EM analyzed the cases. CG, MD, and MB conducted the computational analyses. MS, MD, and CG contributed to the preparation of the original draft. All authors contributed to the article and approved the submitted version

REFERENCES

- Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* (2004) 116:281–97. doi: 10.1016/S0092-8674(04)00045-5
- He L, Hannon GJ. MicroRNAs. Small RNAs With a Big Role in Gene Regulation. *Nat Rev Genet* (2004) 5:522–31. doi: 10.1038/nrg1379
- Ameres SL, Zamore PD. Diversifying microRNA Sequence and Function. *Nat Rev Mol Cell Biol* (2013) 14:475–88. doi: 10.1038/nrm3611
- Nik Mohamed Kamal NNSB, Shahidan WNS. Non-Exosomal and Exosomal Circulatory MicroRNAs. Which Are More Valid as Biomarkers? *Front Pharmacol* (2020) 10:1500. doi: 10.3389/fphar.2019.01500
- Anfossi S, Babayan A, Pantel K, Calin GA. Clinical Utility of Circulating non-Coding RNAs - an Update. *Nat Rev Clin Oncol* (2018) 15:541–63. doi: 10.1038/s41571-018-0035-x
- Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogossova-Agadjanian EL, et al. Circulating microRNAs as Stable Blood-Based

FUNDING

This research was supported by grant 12162 (Special Program “Innovative Tools for Cancer Risk Assessment and Early Diagnosis” 5 × 1000) from the Italian Association for Cancer Research (to GS as PI and MS as GL) and by 5 × 1000 Funds (Italian Ministry of Health 2014—institutional grant BRI2017) from Fondazione IRCCS Istituto Nazionale dei Tumori (to LDC as PI) and by Italian Ministry of Health (Ricerca Corrente 2022 Funds).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.911613/full#supplementary-material>

Supplementary Figure 1 | (A) Venn diagram reporting the intersection of the miRs present on all platforms, highlighting the 488 in common. **(B)** Boxplots representing the number of detected ct-miRs in each platform with respect to those present in each platform. **(C)** Boxplots reporting the percentage of GC in detected (pink) and undetected (blue) ct-miRs for each platform. *P*-value calculation by Wilcoxon rank-sum test. **(D)** Comparison of the empirical cumulative density function for each platform of the expression quantiles of the 26 ct-miRs detected in all samples and in all five platforms.

Supplementary Figure 2 | Scatterplot of duplicate profiles derived from three plasma samples (#1, NSCLC patient; #2 and #3, donors) calculated on normalized and filtered data in each platform. The platforms from top to bottom are OAC, OAA, miRCURY, EdgeSeq, and QiaSeq.

Supplementary Figure 3 | Comparison of differential expression results across platforms. **(A)** Number of DE ct-miRs in each platform according to different false discovery rate (FDR) thresholds. **(B)** Number of DE ct-miRs shared by at least 2, 3, 4, and 5 platforms according to different FDR thresholds. **(C)** Pairwise scatter plots of the *t*-statistic values obtained from the differential expression analysis performed for each platform. Pearson’s correlation coefficients and significance are reported. ****p*-value is <0.001, ***p*-value is <0.01, and **p*-value is <0.05.

Supplementary Figure 4 | Comparison of area under the curve (AUC) values of ct-miRs across platforms. **(A)** Number of ct-miRs in each platform with AUC values above increasing cutoffs. **(B)** Number of ct-miRs shared by at least 2, 3, 4, and 5 platforms according to increasing AUC thresholds. **(C)** Pairwise scatter plots of AUC values for the classification of lung cancer patients and healthy donors. Pearson’s correlation coefficients and significance are reported. ****p*-value is <0.001, ***p*-value is <0.01, and **p*-value is <0.05.

- Markers for Cancer Detection. *Proc Natl Acad Sci U.S.A.* (2008) 105:10513–18. doi: 10.1073/pnas.0804549105
- Kosaka N, Iguchi H, Ochiya T. Circulating microRNA in Body Fluid. A New Potential Biomarker for Cancer Diagnosis and Prognosis. *Cancer Sci* (2010) 101:2087–92. doi: 10.1111/j.1349-7006.2010.01650.x
 - Valihrach L, Androvic P, Kubista M. Circulating miRNA Analysis for Cancer Diagnostics and Therapy. *Mol Aspects Med* (2020) 72:100825. doi: 10.1016/j.mam.2019.10.002
 - Lampignano R, Kloten V, Krahn T, Schlange T. Integrating Circulating miRNA Analysis in the Clinical Management of Lung Cancer. *Present future? Mol Aspects Med* (2020) 72:100844. doi: 10.1016/j.mam.2020.100844
 - Song Y, Yu X, Zang Z, Zhao G. Circulating or Tissue microRNAs and Extracellular Vesicles as Potential Lung Cancer Biomarkers: A Systematic Review. *Int J Biol Markers* (2018) 33:3–9. doi: 10.5301/ijbm.5000307
 - Boeri M, Pastorino U, Sozzi G. Role of microRNAs in Lung Cancer. microRNA Signatures in Cancer Prognosis. *Cancer J* (2012) 18:268–74. doi: 10.1097/PPO.0b013e318258b743

12. Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, et al. microRNA Signatures in Tissues and Plasma Predict Development and Prognosis of Computed Tomography Detected Lung Cancer. *Proc Natl Acad Sci U.S.A.* (2011) 108:3713–18. doi: 10.1073/pnas.1100048108
13. Sozzi G, Boeri M, Rossi M, Verri C, Suatoni P, Bravi F, et al. Clinical Utility of a Plasma-Based miRNA Signature Classifier Within Computed Tomography Lung Cancer Screening. A Correlative MILD Trial Study. *J Clin Oncol* (2014) 32:768–73. doi: 10.1200/JCO.2013.50.4357
14. Pastorino U, Boeri M, Sestini S, Sabia F, Milanese G, Silva M, et al. Baseline Computed Tomography Screening and Blood microRNA Predict Lung Cancer Risk and Define Adequate Intervals in the BioMILD Trial. *Ann Oncol* (2022) 33:395–405. doi: 10.1016/j.annonc.2022.01.008
15. Fortunato O, Borzi C, Milione M, Centonze G, Conte D, Boeri M, et al. Circulating Mir-320a Promotes Immunosuppressive Macrophages M2 Phenotype Associated With Lung Cancer Risk. *Int J Cancer* (2019) 144:2746–61. doi: 10.1002/ijc.31988
16. Fehlmann T, Kahraman M, Ludwig N, Backes C, Galata V, Keller V, et al. Evaluating the Use of Circulating microRNA Profiles for Lung Cancer Detection in Symptomatic Patients. *JAMA Oncol* (2020) 6:714–23. doi: 10.1001/jamaoncol.2020.0001
17. Mestdagh P, Hartmann N, Baeriswyl L, Andreasen D, Bernard N, Chen C, et al. Evaluation of Quantitative miRNA Expression Platforms in the microRNA Quality Control (miRQC) Study. *Nat Methods* (2014) 11:809–15. doi: 10.1038/nmeth.3014
18. Farr RJ, Januszewski AS, Joglekar MV, Liang H, McAulley AK, Hewitt AW, et al. A Comparative Analysis of High-Throughput Platforms for Validation of a Circulating microRNA Signature in Diabetic Retinopathy. *Sci Rep* (2015) 5:10375. doi: 10.1038/srep10375
19. Godoy PM, Barczak AJ, DeHoff P, Srinivasan S, Etheridge A, Galas D, et al. Comparison of Reproducibility, Accuracy, Sensitivity, and Specificity of miRNA Quantification Platforms. *Cell Rep* (2019) 29:4212–22.e5. doi: 10.1016/j.celrep.2019.11.078
20. Babayan A, Neumann MHD, Herdean A, Shaffer JM, Janning M, Kobus F, et al. Multicenter Evaluation of Independent High-Throughput and RT-qPCR Technologies for the Development of Analytical Workflows for Circulating miRNA Analysis. *Cancers (Basel)* (2020) 12:1166. doi: 10.3390/cancers12051166
21. Hong LZ, Zhou L, Zou R, Khoo CM, Chew ALS, Chin CL, et al. Systematic Evaluation of Multiple qPCR Platforms, NanoString and miRNA-Seq for microRNA Biomarker Discovery in Human Biofluids. *Sci Rep* (2021) 11:4435. doi: 10.1038/s41598-021-83365-z
22. Dvinge H, Bertone P. HTqPCR. High-Throughput Analysis and Visualization of Quantitative Real-Time PCR Data in R. *Bioinformatics* (2009) 25:3325. doi: 10.1093/bioinformatics/btp578
23. Robinson MD, Oshlack A. A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data. *Genome Biol* (2010) 11:R25. doi: 10.1186/gb-2010-11-3-r25
24. Robinson MD, McCarthy DJ, Smyth GK. EdgeR. A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* (2010) 26:139–40. doi: 10.1093/bioinformatics/btp616
25. Luo J, Schumacher M, Scherer A, Sanoudou D, Megherbi D, Davison T, et al. A Comparison of Batch Effect Removal Methods for Enhancement of Prediction Performance Using MAQC-II Microarray Gene Expression Data. *Pharmacogenomics J* (2010) 10:278–91. doi: 10.1038/tpj.2010.57
26. Mensah M, Borzi C, Verri C, Suatoni P, Conte D, Pastorino U, et al. microRNA Based Liquid Biopsy. The Experience of the Plasma miRNA Signature Classifier (MSC) for Lung Cancer Screening. *J Vis Exp* (2017) 26:6326. doi: 10.3791/56326
27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007
28. Yu J, Cliften PF, Juehne TI, Sinnwell TM, Sawyer CS, Sharma M, et al. Multi-Platform Assessment of Transcriptional Profiling Technologies Utilizing a Precise Probe Mapping Methodology. *BMC Genomics* (2015) 16:710. doi: 10.1186/s12864-015-1913-6
29. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paep A, et al. Accurate Normalization of Real-Time Quantitative RT-PCR Data by Geometric Averaging of Multiple Internal Control Genes. *Genome Biol* (2002) 3:RESEARCH0034. doi: 10.1186/gb-2002-3-7-research0034
30. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2⁻ΔΔCT Method. *Methods* (2001) 25:402–8. doi: 10.1006/meth.2001.1262
31. Collisson E, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, et al. Cancer Genome Atlas Research Network. Comprehensive Molecular Profiling of Lung Adenocarcinoma. *Nature* (2014) 511:543–50. doi: 10.1038/nature13385
32. Hammerman PS, Lawrence MS, Voet D, Jing R, Cibulskis K, Sivachenko A, et al. Cancer Genome Atlas Research Network. Comprehensive Genomic Characterization of Squamous Cell Lung Cancers. *Nature* (2012) 489:519–25. doi: 10.1038/nature11404
33. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAAbioliinks. An R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res* (2016) 44:e71. doi: 10.1093/nar/gkv1507
34. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC. @ an Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77
35. Goyosso-Gómez LV, Ortiz-Quintero B. Circulating MicroRNAs in Blood and Other Body Fluids as Biomarkers for Diagnosis, Prognosis, and Therapy Response in Lung Cancer. *Diagnostics (Basel)* (2021) 11:421. doi: 10.3390/diagnostics11030421
36. Vösa U, Vooder T, Kolde R, Vilo J, Metspalu A, Annilo T. Meta-Analysis of microRNA Expression in Lung Cancer. *Int J Cancer* (2013) 132:2884–93. doi: 10.1002/ijc.27981
37. Wang F, Ren X, Zhang X. Role of microRNA-150 in Solid Tumors. *Oncol Lett* (2015) 10:11–6. doi: 10.3892/ol.2015.3170
38. He Y, Jiang X, Chen J. The Role of miR-150 in Normal and Malignant Hematopoiesis. *Oncogene* (2014) 33:3887–93. doi: 10.1038/onc.2013.346
39. Koshizuka K, Nohata N, Hanazawa T, Kikkawa N, Arai T, Okato A, et al. Deep Sequencing-Based microRNA Expression Signatures in Head and Neck Squamous Cell Carcinoma. Dual Strands of pre-miR-150 as Antitumor miRNAs. *Oncotarget* (2017) 8:30288–304. doi: 10.18632/oncotarget.16327
40. Okato A, Arai T, Kojima S, Koshizuka K, Osako Y, Idichi T, et al. Dual Strands of Pre-Mir-150 (Mir-150-5p and Mir-150-3p) Act as Antitumor miRNAs Targeting SPOCK1 in Naïve and Castration-Resistant Prostate Cancer. *Int J Oncol* (2017) 51:245–56. doi: 10.3892/ijo.2017.4008
41. Wu X, Xia M, Chen D, Wu F, Lv Z, Zhan Q, et al. Profiling of Downregulated Blood-Circulating miR-150-5p as a Novel Tumor Marker for Cholangiocarcinoma. *Tumour Biol* (2016) 37:15019–29. doi: 10.1007/s13277-016-5313-6
42. Shi KQ, Lin Z, Chen XJ, Song M, Wang YQ, Cai YJ, et al. Hepatocellular Carcinoma Associated microRNA Expression Signature. Integrated Bioinformatics Analysis, Experimental Validation and Clinical Significance. *Oncotarget* (2015) 6:25093–108. doi: 10.18632/oncotarget.4437
43. Scrutinio D, Conserva F, Passantino A, Iacoviello M, Lagiolo R, Gesualdo L. Circulating microRNA-150-5p as a Novel Biomarker for Advanced Heart Failure. A Genome-Wide Prospective Study. *J Heart Lung Transplant* (2017) 36:616–24. doi: 10.1016/j.healun.2017.02.008
44. Roderburg C, Luedde M, Vargas Cardenas D, Vucur M, Scholten D, Frey N, et al. Circulating microRNA-150 Serum Levels Predict Survival in Patients With Critical Illness and Sepsis. *PLoS One* (2013) 8:e54612. doi: 10.1371/journal.pone.0054612
45. Dai FQ, Li CR, Fan XQ, Tan L, Wang RT, Jin H. miR-150-5p Inhibits Non-Small-Cell Lung Cancer Metastasis and Recurrence by Targeting HMGA2 and β-Catenin Signaling. *Mol Ther Nucleic Acids* (2019) 16:675–85. doi: 10.1016/j.omtn.2019.04.017
46. Wu Z, Li W, Li J, Zhang Y, Zhang X, Xu Y, et al. Higher Expression of miR-150-5p Promotes Tumorigenesis by Suppressing LKB1 in non-Small Cell Lung Cancer. *Pathol Res Pract* (2020) 216:153145. doi: 10.1016/j.prp.2020.153145
47. Suetsugu T, Koshizuka K, Seki N, Mizuno K, Okato A, Arai T, et al. Downregulation of Matrix Metalloproteinase 14 by the Antitumor miRNA, miR-150-5p, Inhibits the Aggressiveness of Lung Squamous Cell Carcinoma Cells. *Int J Oncol* (2018) 52:913–24. doi: 10.3892/ijo.2017.4232
48. Cao M, Hou D, Liang H, Gong F, Wang Y, Yan X, et al. miR-150 Promotes the Proliferation and Migration of Lung Cancer Cells by Targeting SRC Kinase Signalling Inhibitor 1. *Eur J Cancer* (2014) 50:1013–24. doi: 10.1016/j.ejca.2013.12.024

49. Zaporozhchenko IA, Morozkin ES, Ponomaryova AA, Rykova EY, Cherdynseva NV, Zheravin AA. Profiling of 179 miRNA Expression in Blood Plasma of Lung Cancer Patients and Cancer-Free Individuals. *Sci Rep* (2018) 8:6348. doi: 10.1038/s41598-018-24769-2
50. Jiang HG, Dai CH, Xu YP, Jiang Q, Xia XB, Shu Y, et al. Four Plasma miRNAs Act as Biomarkers for Diagnosis and Prognosis of non-Small Cell Lung Cancer. *Oncol Lett* (2021) 22:792. doi: 10.3892/ol.2021.13053
51. Keller A, Fehlmann T, Ludwig N, Kahraman M, Laufer T, Backes C, et al. Genome-Wide MicroRNA Expression Profiles in COPD. Early Predictors for Cancer Development. *Genomics Proteomics Bioinf* (2018) 16:162–71. doi: 10.1016/j.gpb.2018.06.001
52. Keller A, Ludwig N, Fehlmann T, Kahraman M, Backes C, Kern F, et al. Low miR-150-5p and miR-320b Expression Predicts Reduced Survival of COPD Patients. *Cells* (2019) 8:1162. doi: 10.3390/cells8101162
53. Decramer M, Janssens W, Miravittles M. Chronic Obstructive Pulmonary Disease. *Lancet* (2012) 379:1341–51. doi: 10.1016/S0140-6736(11)60968-9
54. Durham AL, Adcock IM. The Relationship Between COPD and Lung Cancer. *Lung Cancer* (2015) 90:121–7. doi: 10.1016/j.lungcan.2015.08.017
55. Diener C, Keller A, Meese E. Emerging Concepts of miRNA Therapeutics. From Cells to Clinic. *Trends Genet* (2022) 38:613–26. doi: 10.1016/j.tig.2022.02.006
56. Stewart DJ. Wnt Signaling Pathway in non-Small Cell Lung Cancer. *J Natl Cancer Inst* (2014) 106:djt356. doi: 10.1093/jnci/djt356
57. Wang YZ, Wu KP, Wu AB, Yang ZC, Li JM, Mo YL, et al. MMP-14 Overexpression Correlates With Poor Prognosis in non-Small Cell Lung Cancer. *Tumour Biol* (2014) 35:9815–21. doi: 10.1007/s13277-014-2237-x
58. Lu W, Zhang H, Niu Y, Wu Y, Sun W, Li H, et al. Long non-Coding RNA Linc00673 Regulated non-Small Cell Lung Cancer Proliferation, Migration, Invasion and Epithelial Mesenchymal Transition by Sponging miR-150-5p. *Mol Cancer* (2017) 16:118. doi: 10.1186/s12943-017-0685-9
59. Jin M, Shi C, Yang C, Liu J, Huang G. Upregulated circRNA ARHGAP10 Predicts an Unfavorable Prognosis in NSCLC Through Regulation of the miR-150-5p/GLUT-1 Axis. *Mol Ther Nucleic Acids* (2019) 18:219–31. doi: 10.1016/j.omtn.2019.08.016
60. Zeng Z, Zhao G, Rao C, Hua G, Yang M, Miao X, et al. Knockdown of lncRNA ZFAS1-Suppressed non-Small Cell Lung Cancer Progression via Targeting the miR-150-5p/HMGA2 Signaling. *J Cell Biochem* (2019) 121:3814–24. doi: 10.1002/jcb.29542
61. Wu Y, Li L, Wang Q, Zhang L, He C, Wang X, et al. LINC00511 Promotes Lung Squamous Cell Carcinoma Proliferation and Migration via Inhibiting miR-150-5p and Activating TADA1. *Transl Lung Cancer Res* (2020) 9:1138–48. doi: 10.21037/tlcr-19-701
62. Xiao G, Wang P, Zheng X, Liu D, Sun X. FAM83A-AS1 Promotes Lung Adenocarcinoma Cell Migration and Invasion by Targeting miR-150-5p and Modifying MMP14. *Cell Cycle* (2019) 18:2972–85. doi: 10.1080/15384101.2019.1664225
63. Świtlik WZ, Karbownik MS, Suwalski M, Kozak J, Szemraj J. Serum miR-210-3p as a Potential Noninvasive Biomarker of Lung Adenocarcinoma. A Preliminary Study. *Genet Test Mol Biomarkers* (2019) 23:353–8. doi: 10.1089/gtmb.2018.0275
64. Leng Q, Wang Y, Jiang F. A Direct Plasma miRNA Assay for Early Detection and Histological Classification of Lung Cancer. *Transl Oncol* (2018) 11:883–8. doi: 10.1016/j.tranon.2018.05.001
65. Jiang M, Li X, Quan X, Li X, Zhou B. Clinically Correlated microRNAs in the Diagnosis of non-Small Cell Lung Cancer. A Systematic Review and Meta-Analysis. *BioMed Res Int* (2018) 5930951:1–14. doi: 10.1155/2018/5930951
66. Li C, Yin Y, Liu X, Xi X, Xue W, Qu Y. Non-Small Cell Lung Cancer Associated microRNA Expression Signature. Integrated Bioinformatics Analysis, Validation and Clinical Significance. *Oncotarget* (2017) 8:24564–78. doi: 10.18632/oncotarget.15596
67. Zhu W, Zhou K, Zha Y, Chen D, He J, Ma H, et al. Diagnostic Value of Serum miR-182, miR-183, miR-210, and miR-126 Levels in Patients With Early-Stage non-Small Cell Lung Cancer. *PLoS One* (2016) 11:e0153046. doi: 10.1371/journal.pone.0153046
68. Daugaard I, Venø MT, Yan Y, Kjeldsen TE, Lamy P, Hager H, et al. RNA Sequencing Reveals Metastasis-Related microRNAs in Lung Adenocarcinoma. *Oncotarget* (2017) 8:27047–61. doi: 10.18632/oncotarget.15968
69. Ren J, Li X, Dong H, Suo L, Zhang J, Zhang L, et al. miR-210-3p Regulates the Proliferation and Apoptosis of non-Small Cell Lung Cancer Cells by Targeting SIN3A. *Exp Ther Med* (2019) 18:2565–73. doi: 10.3892/etm.2019.7867
70. Suzuki H, Ouchida M, Yamamoto H, Yano M, Toyooka S, Aoe M, et al. Decreased Expression of the SIN3A Gene, a Candidate Tumor Suppressor Located at the Prevalent Allelic Loss Region 15q23 in non-Small Cell Lung Cancer. *Lung Cancer* (2008) 59:24–31. doi: 10.1016/j.lungcan.2007.08.002
71. Wang L, He J, Hu H, Tu L, Sun Z, Liu Y, et al. Lung CSC-Derived Exosomal miR-210-3p Contributes to a Pro-Metastatic Phenotype in Lung Cancer by Targeting FGFR1. *J Cell Mol Med* (2020) 24:6324–39. doi: 10.1111/jcmm.15274

Conflict of Interest: GS and MB are coinventors for three patent applications licensed to Gensignia Life Sciences and regarding the miR MSC signature used in this article.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Citation: Gargiuli C, De Cecco L, Mariancini A, Iannò MF, Micali A, Mancinelli E, Boeri M, Sozzi G, Dugo M and Sensi M (2022) A Cross-Comparison of High-Throughput Platforms for Circulating MicroRNA Quantification, Agreement in Risk Classification, and Biomarker Discovery in Non-Small Cell Lung Cancer. *Front. Oncol.* 12:911613. doi: 10.3389/fonc.2022.911613

Copyright © 2022 Gargiuli, De Cecco, Mariancini, Iannò, Micali, Mancinelli, Boeri, Sozzi, Dugo and Sensi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.