

Video Article

# MicroRNA Based Liquid Biopsy: The Experience of the Plasma miRNA Signature Classifier (MSC) for Lung Cancer Screening

Mavis Mensah<sup>1</sup>, Cristina Borzi<sup>1</sup>, Carla Verri<sup>1</sup>, Paola Suatoni<sup>2</sup>, Davide Conte<sup>1</sup>, Ugo Pastorino<sup>2</sup>, Fortunato Orazio<sup>1</sup>, Gabriella Sozzi<sup>1</sup>, Mattia Boeri<sup>1</sup>

<sup>1</sup>Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori

<sup>2</sup>Unit of Thoracic Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori

\*These authors contributed equally

Correspondence to: Fortunato Orazio at [Orazio.fortunato@istitutotumori.mi.it](mailto:Orazio.fortunato@istitutotumori.mi.it)

URL: <https://www.jove.com/video/56326>

DOI: [doi:10.3791/56326](https://doi.org/10.3791/56326)

Keywords: Cancer Research, Issue 128, Lung cancer, early detection, circulating microRNAs, biomarkers, liquid biopsy, Real-Time PCR.

Date Published: 10/26/2017

Citation: Mensah, M., Borzi, C., Verri, C., Suatoni, P., Conte, D., Pastorino, U., Orazio, F., Sozzi, G., Boeri, M. MicroRNA Based Liquid Biopsy: The Experience of the Plasma miRNA Signature Classifier (MSC) for Lung Cancer Screening. *J. Vis. Exp.* (128), e56326, doi:10.3791/56326 (2017).

## Abstract

The development of a minimally invasive test, such as liquid biopsy, for early lung cancer detection in its preclinical phase is crucial to improve the outcome of this deadly disease. MicroRNAs (miRNAs) are tissue specific, small, non-coding RNAs regulating gene expression, which may act as extracellular messengers of biological signals derived from the cross-talk between the tumor and its surrounding microenvironment. They could thus represent ideal candidates for early detection of lung cancer. In this work, a methodological workflow for the prospective validation of a circulating miRNA test using custom made microfluidic cards and quantitative Real-Time PCR in plasma samples of volunteers enrolled in a lung cancer screening trial is proposed. In addition, since the release of hemolysis-related miRNAs and more general technical issues may affect the analysis, the quality control steps included in the standard operating procedures are also presented. The protocol is reproducible and gives reliable quantitative results; however, when using large clinical series, both pre-analytical and analytical features should be cautiously evaluated.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/56326/>

## Introduction

Lung cancer is the most commonly diagnosed cancer worldwide, accounting for about 25% of all cancer diagnoses<sup>1</sup>. Despite the steady reduction in lung cancer mortality rate over the past 2 decades, mainly due to reduced tobacco use, lung cancer remains the leading cause of cancer death in both men and women. In 2017, it is predicted to account for approximately 25% and 20% of the total cancer deaths in the United States and Europe, respectively<sup>1,2</sup>.

Since the symptoms do not usually occur in early disease, the majority of lung cancers are diagnosed at late stages. This leads to a limited possibility for therapeutic intervention and poor efficacy of the treatments<sup>3</sup>. Therefore, many scientific efforts are aimed at identifying an effective test for early lung cancer detection.

A variety of screening trials shows that chest radiography has no value in lung cancer screening, whereas low-dose computed tomography (LDCT) is a better tool compared to radiography for detecting early stage lung cancer<sup>4</sup>. Moreover, it has been shown that screening with the use of LDCT reduced lung cancer mortality up to 20% among current and former heavy smokers<sup>5</sup>. These data support the use of lung cancer screening in a high risk population as defined in the guidelines by several professional societies<sup>4,6</sup>.

Of note, among the major concerns about LDCT screening, there is the high rate of false-positive results and over-diagnosis. That being so, the scientific community is looking for a strategy to overcome these issues and increase the specificity of the LDCT screening test. The development of non-invasive complementary biomarkers could be useful here. To date, there is an extensive list of candidate biomarkers under investigation, such as miRNAs, cell-free DNA, gene methylation, small proteins, and others<sup>7</sup>.

Blood-based biomarkers, including immune-response biomarkers and circulating miRNAs, are those that reach the more advanced validation phase<sup>8</sup>. The immune-response biomarkers are based on the knowledge that immune responses against both intracellular and surface tumor antigens are built up in lung cancer patients<sup>9</sup>. For instance, Ajona *et al.* evaluated the role of C4d, a degradation product of the classical complement pathway, in the diagnosis and prognosis of lung cancer<sup>10</sup>. Otherwise, a commercially available autoantibodies serum test examining seven tumor-related antigens has been validated in non-small cell lung cancer patients and showed 36% of sensibility and 91% of specificity<sup>11</sup>. Immune-response biomarkers are interesting, but further validation studies are needed for a potential clinical application.

miRNAs are endogenous small non-coding RNAs with a conserved mechanism and large functional significance all over the animal and plant kingdoms. The role of miRNAs is to regulate protein translation: they repress the expression of a large set of target genes<sup>12</sup>. It has been well documented that alterations of miRNAs' expression contribute to the pathogenesis of most human cancer<sup>12,13,14</sup>. In addition, both tumor

and stromal cells release miRNAs, stabilized by their incorporation into microvesicles or through association to RNA-binding proteins, into body fluids such as plasma and serum<sup>15,16</sup>, urine<sup>17</sup>, and saliva<sup>18</sup>. Circulating miRNAs may thus reflect the host response and the dynamic interaction between the tumor and its microenvironment<sup>19</sup>. Taken together these observations have made circulating miRNAs a promising class of biomarkers for the detection of human cancer.

Circulating miRNAs can be detected by using different methods: microarray platforms<sup>20</sup>, quantitative reverse transcription PCR (RT-qPCR)<sup>21</sup>, and next generation sequencing (NGS)<sup>14,22</sup>. Various expression profile studies based on the aforementioned technologies have been developed in the last years. Microarray is a high-throughput technology, able to analyze up to thousands of miRNAs in one single assay. However, it has lower dynamic range and specificity compared to RT-qPCR or NGS. In addition, microarray is a non-quantitative method, therefore further experimental validation is required. Sequence-based methods can allow the identification of unknown miRNAs and are suitable especially in a discovery phase. On the other hand, NGS methods are still expensive and necessitate special equipment and expert bioinformaticians, thus limiting their use in large validation cohorts and in clinical settings<sup>23</sup>. At the moment, RT-qPCR is the most adopted platform for circulating miRNA detection in a diagnostic/clinical context<sup>24</sup>. There are different RT-qPCR technologies available for miRNAs analysis, but the stem-loop RT followed by TaqMan PCR is the most widely used<sup>25</sup>. This technique is extremely sensitive and accurate (single nucleotide discrimination) and has a high dynamic range; however, it allows the detection of known miRNAs only.

The microRNA quality control study (miRQC study) by Mestdagh *et al.* extensively investigated the characteristics of some commercially available platforms for miRNAs profiling based on the three aforesaid technologies<sup>26</sup>. By analyzing 196 miRNAs in tissues and serum samples, the performance of the different platforms was assessed in terms of reproducibility, sensitivity, accuracy, specificity, and concordance of differential expression. The results have shown that platforms based on NGS and microarray technologies had a higher reproducibility and specificity, but RT-qPCR platforms were more accurate, sensitive, and had a higher detection rate, especially for low input RNA samples, such as body fluids. Thus RT-qPCR seems to be the most suitable method for a potential application in the routine clinical diagnostics.

In 2011, Sozzi *et al.* analyzed the miRNA profile of plasma samples collected from volunteers enrolled in a first LDCT screening trial (INT-IEO trial)<sup>27</sup> and four miRNA signatures composed by reciprocal ratios among 24 circulating miRNAs were generated. These signatures were able to discriminate disease free individuals from patients with any or lethal lung cancer at the time or up to 2 years before LDCT disease detection<sup>28</sup>. In a further paper, the same group described for the first time the miRNA signature classifier (MSC), obtained by combining the four miRNA signatures in order to stratify patients in three levels of risk: high, intermediate, and low. Its clinical utility was validated in a large retrospective set of more than 1,000 individuals enrolled in the MILD screening trial, a randomized study comparing annual or biennial LDCT arms with an observation arm<sup>29</sup>. Indeed, the combination of MSC and LDCT reduced the LDCT false-positive rate by approximately 5 times and the three MSC risk groups were associated with overall survival.

Later on, in 2013 at the "Fondazione IRCCS Istituto Nazionale dei Tumori" (Milan, IT) a prospective screening trial (BioMILD) implementing LDCT with the plasma based MSC test was launched. Volunteers enrolled in the BioMILD trial undergo LDCT and blood withdrawal that is immediately processed to separate plasma for miRNA profiling using custom made microfluidic cards. The combination of the LDCT and MSC results defines the specific screening algorithm<sup>24</sup>.

In the present work, the entire methodological protocol used in the BioMILD trial is described, starting from blood sample collection, to plasma isolation, RNA extraction, and the 24 miRNAs expression profiling using custom made RT-qPCR microfluidic cards. Given the high consistency and reproducibility, these procedures can also be used for the development of miRNA-based liquid biopsies in other diseases.

## Protocol

The protocol was approved by the research ethics committee of our institution.

### 1. Plasma Sample Collection

1. Collect 10 mL of whole blood sample in Vacutainer tubes with spray-coated K<sub>2</sub>EDTA and store at room temperature.  
Note: To minimize the hemolysis, separate plasma within 2 h. Do not store the whole blood at low temperature (*i.e.*, 4 °C) to avoid thermal shock and cell lysis that lead to an unspecific miRNA release.
2. Within 1 h, separate the plasma by a first centrifugation step at 1,258 x g and 4 °C for 10 min.
3. Transfer the plasma supernatant into a 15 mL tube, carefully avoiding contact with the lymphocytic ring.
4. Centrifuge the plasma a second time at 1,258 x g and 4 °C for 10 min.
5. Aliquot 1 mL of plasma into 1.5 mL cryovials, avoiding collecting the plasma fraction at the base of the tube.
6. Store all the aliquots at -80 °C, except one for the evaluation of the hemolysis. The molecular analysis should be performed within 5 weeks.

### 2. Evaluation of Hemolysis by Spectrophotometric Measurement

1. Immediately after the plasma separation step, in a cuvette for spectrophotometry, make a 1:10 dilution of the plasma sample in 1X PBS (*e.g.*, 100 µL of plasma in 900 µL of 1X PBS) and mix to homogenize it.  
Note: The plasma sample should not be frozen before the spectrophotometric measurements, as the freeze-thawing of the sample (and in particular in lipemic samples), could form flocculates that interfere with the spectrophotometric analysis.
2. Read the absorbance at 375 nm, 414 nm, 541 nm, and 576 nm using a UV-Vis spectrophotometer, with a baseline correction settled at 750 nm and a path length of 10 mm. Perform a blank measurement using 1 mL of 1x PBS.
3. Calculate the ratio between the absorbance at 414 nm and 375 nm; if higher than 1.4, consider the sample hemolyzed and repeat the blood withdrawal (QC1 in **Table 1**).

### 3. Plasma Total RNA Extraction

1. Prepare a 1-Thioglycerol/Homogenization (OMG) solution with 20  $\mu\text{L}$  of 1-Thioglycerol per mL of Homogenization Solution. Since 1-Thioglycerol is viscous, carefully pipette for accurate measurement. Chill the OMG Solution on ice or at 2-10  $^{\circ}\text{C}$  before using it.  
NOTE: The working solution is stable at 2-10  $^{\circ}\text{C}$  for 1 month.
2. Suspend the lyophilized DNase I by adding 275  $\mu\text{L}$  of nuclease-free water into the vial and gently mix (do not vortex). As a visual aid, add 5  $\mu\text{L}$  of blue dye to the reconstituted DNase I and dispense the solution into single-use aliquots in nuclease-free tubes. Store reconstituted DNase I at -30  $^{\circ}\text{C}$  to -10  $^{\circ}\text{C}$ .
3. Starting from 200  $\mu\text{L}$  of plasma, add 200  $\mu\text{L}$  of the chilled OMG solution.
4. Vortex 15-30 s to ensure a complete homogenization. If foaming occurs, let the sample settle on ice.
5. Add 200  $\mu\text{L}$  of Lysis Buffer and 25  $\mu\text{L}$  of Proteinase K to the homogenized sample and vortex for 20 s.
6. Incubate the samples for 15 min in a thermomixer preheated at 37  $^{\circ}\text{C}$ .
7. Load a cartridge for each sample on the deck tray of the instrument and place the plugger in the proper position.
8. Transfer the lysate to the appropriate position in the instrument cartridge.
9. Add 5  $\mu\text{L}$  of DNase I solution to the proper position in the cartridge.
10. Add 60  $\mu\text{L}$  of nuclease-free water to the base of each elution tube.
11. Select the "RSC miRNA Tissue" method and begin the automated purification run. The total RNA samples can be stored at -80  $^{\circ}\text{C}$ .

### 4. Taq-based RT

1. Use the Taq RT Primer Pool with the miRNAs of interest to convert the RNA to cDNA with the Taq MicroRNA Reverse Transcription Kit.
2. Prepare 12  $\mu\text{L}$  of RT reaction mix on ice in a 1.5 mL microcentrifuge tube according to the kit instructions, using 6  $\mu\text{L}$  of the custom Taq RT Primer Pool.
3. Add 3  $\mu\text{L}$  of total RNA for a final volume of 15  $\mu\text{L}$  and incubate in ice for 5 min.
4. Load the RT reaction on a thermo-cycler configured as follows: 16  $^{\circ}\text{C}$  for 30 min, 42  $^{\circ}\text{C}$  for 30 min, 85  $^{\circ}\text{C}$  for 5 min, and hold at 4  $^{\circ}\text{C}$ .  
Note: The cDNA can be store at -15 to -20  $^{\circ}\text{C}$  for at least one week.

### 5. Pre-amplification

1. Mix 2.5  $\mu\text{L}$  of each RT product with 12.5  $\mu\text{L}$  of Taq Master Mix 2x, 3.75  $\mu\text{L}$  Custom Taq pool, and 6.25  $\mu\text{L}$  nuclease-free water, for a total volume of 25  $\mu\text{L}$ .
2. Perform the pre-amplification reaction using a thermo-cycler according to the following thermal profile: 95  $^{\circ}\text{C}$  for 10 min, 55  $^{\circ}\text{C}$  for 2 min, 72  $^{\circ}\text{C}$  for 2 min, 12 cycles of 95  $^{\circ}\text{C}$  for 15 s, and 60  $^{\circ}\text{C}$  for 4 min, then hold at 99.9  $^{\circ}\text{C}$  for 10 min and at 4  $^{\circ}\text{C}$ .
3. Dilute the product by adding 175  $\mu\text{L}$  of 0.1X TE, pH 8.0.  
Note: The diluted product can be store at -15 to -20  $^{\circ}\text{C}$  for at least one week.

### 6. RT-qPCR Reaction on Custom Taq Array MicroRNA Cards

1. Use the 384-well microfluidic Custom Taq Array MicroRNA Card to measure plasma levels of the 24 specific miRNAs (spotted in duplicate) on 8 samples simultaneously. MiRBase ID (v21) for the 24 miRNAs are: hsa-miR-101-3p, hsa-miR-106a-5p, hsa-miR-126-3p, hsa-miR-133a-3p, hsa-miR-140-3p, hsa-miR-140-5p, hsa-miR-142-3p, hsa-miR-145-5p, hsa-miR-148b-3p, hsa-miR-15b-5p, hsa-miR-16-5p, hsa-miR-17-5p, hsa-miR-197-3p, hsa-miR-19b-3p, hsa-miR-21-5p, hsa-miR-221-3p, hsa-miR-28-3p, hsa-miR-30b-5p, hsa-miR-30c-5p, hsa-miR-320a, hsa-miR-451a, hsa-miR-486-5p, hsa-miR-660-5p, and hsa-miR-92a-3p.
2. Mix 1.13  $\mu\text{L}$  of the diluted PreAmp sample with 56.25  $\mu\text{L}$  of 2X Taq Universal Master Mix and 55.69  $\mu\text{L}$  of nuclease-free water in a 0.5 mL tube.
3. Load up to 8 PCR reaction mix on the Custom Taq Array MicroRNA Cards.
4. Centrifuge at 311 x g for 2 min.
5. Seal the custom card with the array card sealer.
6. Run the Real-Time reaction using a Real-Time PCR System modifying the cycling parameters as followed: 94.5  $^{\circ}\text{C}$  for 10 min, 40 cycles of 97  $^{\circ}\text{C}$  for 30 s, and 59.7  $^{\circ}\text{C}$  for 60 s.

### 7. Data Extrapolation and Ratios Generation

1. Use the software to obtain raw Ct values, setting an automatic baseline to remove the background signals and a fixed threshold of 0.15 for all assays and samples. Remove spots with poor amplification curves and/or poor passive reference signals.
2. Export raw Ct values in ".xls" format and use the Ct Mean value of the two duplicates for further analysis.
3. Correlate miRNAs' expression levels to historic measurements on clinical study samples (Table 2). If at least 50% of samples on each custom made microfluidic card result in a Pearson's correlation < 97.5, repeat the card (QC2 in Table 1).
4. Calculate the  $-\Delta\text{Ct}$ s between all miRNAs, equivalent to the log<sub>2</sub> values of the miRNA ratios.

### 8. Evaluation of Hemolysis by the Related miRNA Signature

1. Generate the hemolysis-related miRNA signature using the following miRNA ratios with respective cut-offs (log<sub>2</sub> values) for short storage plasma samples (1-5 weeks at -80  $^{\circ}\text{C}$ ): miR-126/451 < -0.07, 15b/451 < -3.67, 221/451 < -3.18, 30b/451 < -1.1, 126/486-5p < -0.33, 15b/486-5p < -3.86, 221/486-5p < -3.17, 30b/486-5p < -1.42, 126/92a < 1.8, 15b/92a < -1.8, 221/92a < -1.04, 30b/92a < 0.87, 126/16 < -2.85, 15b/16 < -6.33, 221/16 < -5.9, 30b/16 < -3.68.

- Classify as hemolyzed samples where at least 50% of ratios (8 out of 16) exceed respective cut-off (QC3 in **Table 1**).

## 9. Definition of the Level of Risk: High, Intermediate, and Low

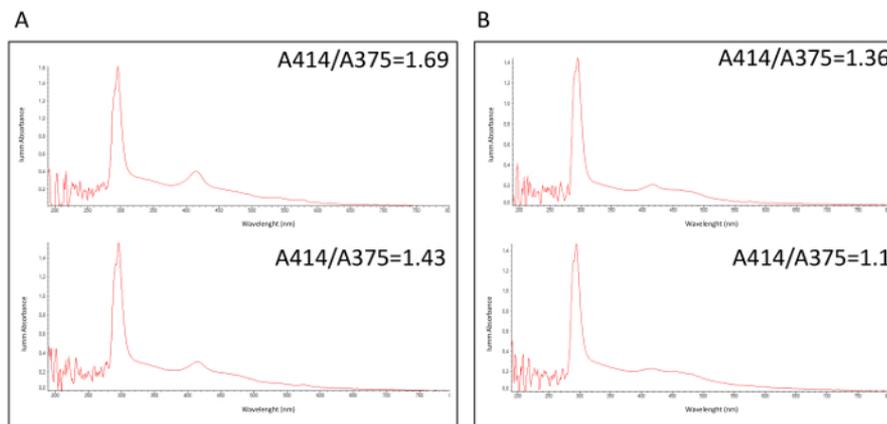
- Define the four signatures of miRNA ratios composing the MSC as reported in **Figure 3**: risk of disease (RD), risk of aggressive disease (RAD), presence of disease (PD), and presence of aggressive disease (PAD).
- For each signature, define the ratios exceeding the respective cut-off value for short storage plasma samples (1-5 weeks at  $-80^{\circ}\text{C}$ ). The number of exceeding ratios needed to be considered positive is 9 out of 27 for RD and PD, and 14 out of 28 for RAD and PAD (**Figure 3A**).
- Attribute to each sample the respective MSC risk level as follows: low risk if  $\text{RD}^{\text{neg}} \cap \text{PD}^{\text{neg}} \cap \text{RAD}^{\text{neg}} \cap \text{PAD}^{\text{neg}}$ ; intermediate risk if  $\text{RD}^{\text{pos}} \cap \text{PD}^{\text{pos}} \cap \text{RAD}^{\text{neg}} \cap \text{PAD}^{\text{neg}}$ ; or high risk if  $\text{RAD}^{\text{pos}} \cap \text{PAD}^{\text{pos}}$  (**Figure 3B**).

## Representative Results

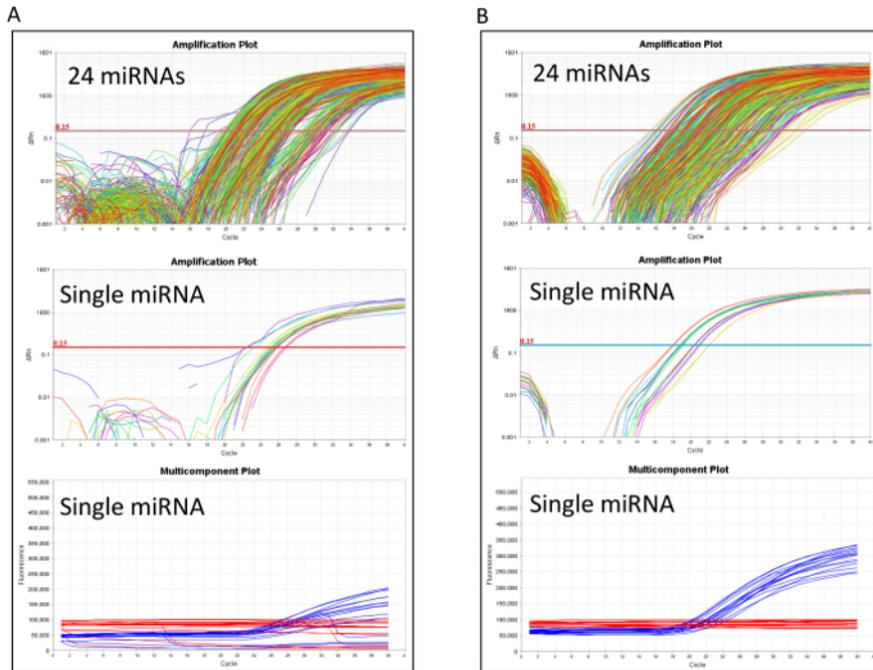
The BioMILD trial is a prospective study for early detection of lung cancer with the purpose of testing the efficiency of a combined LDCT-MSCT approach as first-line screening tests in a large cohort of heavy smoker individuals. A class of risk is attributed to each volunteer on the basis of the miRNA test, which evaluates ratios between 24 plasma circulating miRNAs. As largely reported<sup>30,31,32</sup>, hemolysis is a critical issue for the analysis of circulating miRNA, since false results could be obtained due to a nonspecific release of miRNAs by blood cells. In the proposed methodological workflow, a pre-analytical quality control step (QC1 in **Table 1**) to identify hemolyzed, and thus non-analyzable, plasma samples was described. **Figure 1** reports a spectrophotometric analysis of hemolyzed (**Figure 1A**) and non-hemolyzed (**Figure 1B**) plasma samples with respective A414/A375 values. In a clinical context where the result of a molecular test is crucial for volunteer management, this QC1 allows repeating the blood sampling and, in most cases, restoring the sample.

After molecular processing, the RT-qPCR performance should be accurately evaluated to identify any technical issues. **Figure 2A** shows a RT-qPCR with a low performance due to a poor passive reference signal. In this case reloading the pre-amplification product to a new microfluidic card is recommended. When the RT-qPCR performs well, as in **Figure 2B**, the card proceeds to the first post analytical quality control step to ensure that the miRNA expression values are comparable with the historical measurement (QC2 in **Table 1**). If the QC2 step fails, a technical problem may have occurred during reverse transcription or in the pre-amplification reactions, and it is convenient to repeat the entire molecular processing starting from reverse transcription.

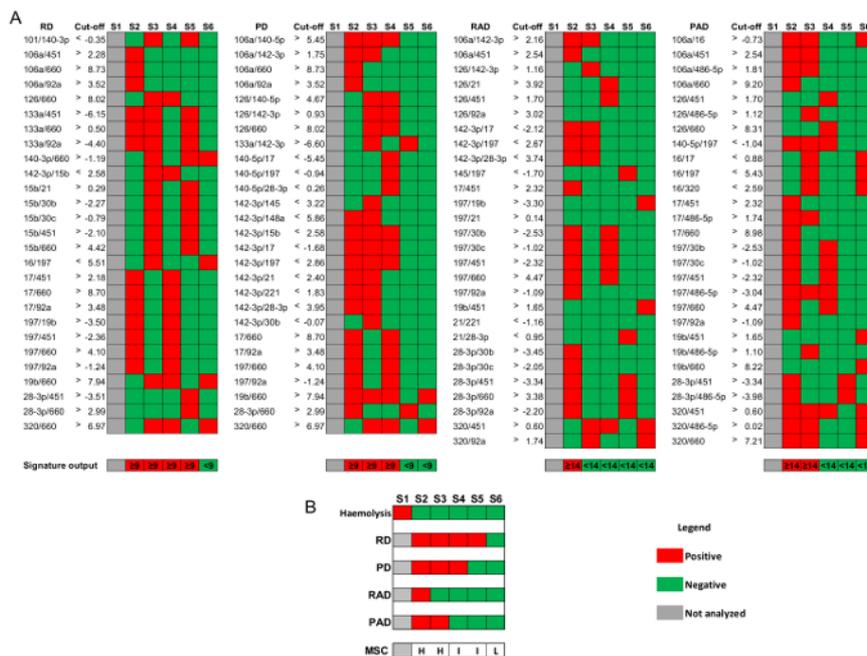
The last quality control step (QC3 in **Table 1**) evaluates the hemolysis also at the molecular level. Expression levels of 4 erythrocyte-specific miRNAs (mir-16, mir-451, mir-486-5p, and mir-92a) were compared to those of 4 non-hemolysis related miRNAs (mir-126, mir-15b, mir-221, and mir-30b), generating the hemolysis-related miRNA signature composed by the 16 (4 x 4) miRNA ratios. Positive samples are classified as hemolyzed, while negative samples proceed to the final MSC risk level classification as described in Protocol Section 9 and **Figure 3**.



**Figure 1: Spectrophotometric analysis of fresh plasma samples.** Spectrophotometric profiles of (A) 2 hemolyzed and (B) 2 non-hemolyzed plasma samples, measuring the absorbance ratio between wavelengths A414 and A375. [Please click here to view a larger version of this figure.](#)



**Figure 2: RT-qPCR performance in microfluidic cards analysis.** Amplification and multicompartment plots comparing (A) poor and (B) good quality microfluidic cards. All 24 miRNAs and a single exemplifying miRNA are reported in each panel. The multicompartment plots illustrate the passive reference signals occurring in these RT-qPCR reactions. [Please click here to view a larger version of this figure.](#)



**Figure 3: MSC algorithm for risk stratification.** Representative results considering 6 plasma samples from volunteers enrolled in BioMILD screening trial. (A) miRNA-ratios' signatures of risk of disease (RD), risk of aggressive disease (RAD), presence of disease (PD), presence of aggressive disease (PAD), and corresponding cut-off values (log<sub>2</sub>) for plasma samples stored at -80 °C for at least 1 week and up to 5 weeks. (B) Combination of the four signatures to stratify analyzable samples in high (H), intermediate (I), and low (L) MSC risk level. [Please click here to view a larger version of this figure.](#)

## Discussion

### Pitfalls and Challenges of the Methodological Workflow

As BioMILD is the first prospective study analyzing circulating miRNAs as a diagnostic tool, thresholds and cut-offs utilized for data-analysis were generated from retrospective series. In order to overcome the issue of the different storage interval of plasma samples, the parameters can be

refined to possibly obtain an increase of the MSC performance. Nevertheless, before starting miRNAs analysis, several aspects of the overall process need to be carefully evaluated, such as the inclusion criteria, the sample collection and processing, and the bioinformatics algorithm adopted for data analysis. The following sections will focus on the most critical points of the workflow described in the present paper, addressing the major issues of each single step and then proposing strategies to overcome them.

### Study Population

Enrollment criteria for the BioMILD screening trial are: 50-75 years old, heavy current or former (less than 10 years) smokers with at least 30 pack-years and no prior history of cancer within 5 years. In this population, the estimated lung cancer prevalence and incidence per year were about 1% and 0.7%, respectively. In order to apply the MSC test as described in the proposed protocol, the population analyzed should maintain similar characteristics. In particular, the QC2 step described above involves comparison with the given historical data obtained from a similar population. While a correlation below 97.5 could be indicative of technical problems, on the other hand, most individuals with altered circulating miRNA levels reflecting the risk of having lung cancer are actually below that threshold. In other words, if the QC2 step with the parameters reported here is applied to a population at higher risk of developing lung cancer, it may exclude acceptable samples. New, proper reference parameters should then be defined that are specific for each population analyzed according to expected results.

### Blood Collection and Hemolysis

The collection of biological samples is a crucial pre-analytical step, since it may impair sample quality and characteristics, and thus modify the final results of analyses. Regarding blood samples, it is important to collect fresh blood and process the specimen as quickly as possible. Samples should neither be stored at 4 °C nor frozen after blood withdrawal, since freezing and thawing passages may induce cell lysis (hemolysis) and RNA degradation. To minimize hemolysis process, it is suggested to separate plasma within 2 h from blood collection<sup>31</sup>.

Hemolysis is largely due to breakdown of red blood cells (RBCs) and consequent release of their content in the surrounding environment. This could be a problem for circulating miRNAs testing because release of RBCs intracellular content dramatically alters the microRNA profile in blood, potentially affecting the accuracy of plasma-based analysis<sup>30</sup>.

Being able to identify also minimal levels of hemolysis in plasma samples is crucial for studies based on circulating miRNAs as biomarkers. Hemolysis can initially be assessed by visual inspection, since after the two centrifugation steps the color of plasma can range from yellow to red for highly hemolyzed samples. However, mere visual inspection may not be sensitive enough in the context of molecular biomarker research.

A simple, pre-analytical method to identify hemolyzed plasma samples is the spectrophotometric measurement at wavelength ( $\lambda$ ) = 414 nm of the main oxyhemoglobin peak absorbance, after appropriate adjustment for any other source of interference. For that reason, we normalized the 414 nm peak to the absorbance value at 375 nm, indicative of the lipidic content<sup>33</sup>. Alternatively, or better, in combination, a lipemia-independent hemolysis score (HS) can be used<sup>34</sup>. This spectrophotometrically-based procedure can identify at least 6.1 mg/dL of free hemoglobin. Otherwise, hemolyzed specimens can be identified through the detection of erythrocyte-specific miRNAs. Since even a very low percentage of hemolysis can elicit a considerable increase in erythrocyte-specific miRNA levels<sup>32</sup>, Fortunato *et al.* identified a specific hemolysis-related signature based on 16 miRNAs-ratios that can efficiently classify hemolysis in plasma samples<sup>33</sup>. All the mentioned assays may be used in combination since the absorbance ratio 414/375 and/or the HS are adopted in a pre-analytical phase saving further costs, whereas measuring hemolysis-related miRNAs can be used in a post-analytical phase as a quality control step.

### Circulating miRNAs Isolation and RNA Purity

The extraction RNA procedure is a very influential step for the success of the subsequent analyses. Starting from a low input sample, such as plasma, extraction needs to be as efficient as possible to ensure an accurate quantification of miRNAs. Plasma is characterized by high concentrations of many proteins, including nucleases, and other components which can interfere with downstream enzymatic reactions of RT-qPCR and could affect the circulating-miRNAs detection. Several extraction methods have been used in these types of studies, but in general, organic extraction followed by silica-based RNA enrichment allows the removal of plasma inhibitors better than TRIzol alone<sup>35</sup>. Nevertheless, in the last few years, manual extraction has been replaced by automated extraction. The minor risk of contamination, the higher reproducibility, and the economic time use are the main advantages of automated systems compared to manual isolation protocols. Moreover, automated extraction methods yielded the highest miRNA amounts from blood samples and the highest efficiency compared to manual protocols<sup>36</sup>.

### Normalization Issue

Normalization of plasma miRNA levels is still a controversial issue. Indeed, differences in sample quality or quantity can interfere with the identification of true changes in miRNAs expression levels caused by the presence or risk of diseases, and thus leading to ambiguous data interpretations and misleading conclusions. So far, the lack of consensus has resulted in various normalization strategies<sup>37</sup>.

Total RNA extracted from plasma samples is usually below the threshold of standard quantification methods such as UV spectrophotometry or fluorescence-based spectrophotometry, thus precluding RNA mass standardization<sup>38</sup>. Consequently, a fixed volume rather than a fixed amount of eluted RNA should be chosen as input for the RT reaction<sup>39</sup>.

Housekeeping transcripts used for miRNA analysis in tissue samples, like RNU48 and RNU6, are not suitable for normalizing extracellular miRNA measurements, as they both are not always detectable in circulation, due to RNase-mediated degradation<sup>40</sup>, and are also deregulated in a disease-specific manner<sup>37</sup>. It is unclear if any circulating miRNAs could effectively be used as reference controls. For example, miR-16 is frequently proposed for housekeeping, but different studies have shown that it is deregulated in plasma samples of cancer patients<sup>28,37</sup>. Moreover, miR-16 levels are highly affected by hemolysis<sup>30</sup>.

While for high-throughput assays measuring several miRNAs and normalizing on the mean expression value is generally accepted<sup>41</sup>, when a limited number of miRNAs needs to be analyzed, this normalization approach cannot be applied. In order to bypass the normalization issue, Boeri *et al.* previously established an algorithm based on reciprocal ratios among 24 miRNAs<sup>28</sup>. Such an approach, despite being very useful for the development of the diagnostic tool, does not easily allow distinguishing effectively altered miRNAs that possibly having a role in tumor development, and thus identifying miRNAs as functioning calibrators. Another effective strategy, adopted by the group of Bianchi *et al.* that

developed a serum miR-Test for lung cancer early detection, was based on normalizing on the geometric mean of a group of 6 "housekeeping serum-miRNAs" varying less among samples and classes of patients<sup>42</sup>.

#### Features of miRNA-based Liquid Biopsy and Potential Application

Plasma-based MSC is a molecular non-invasive test studied for early diagnosis of lung cancer in heavy smokers. The MSC test can identify intermediate or high-risk subjects before lung cancer detection by LDCT, with high sensibility (SE, 87%) and negative predictive value (NPV, 99%). Specifically, MSC could implement LDCT screening by reducing false-positive results to 3.7%, from 19.7% for LDCT alone<sup>29</sup>. The MSC test has also shown to have a good prognostic performance when considering only lung cancer patients, and could be used as a monitoring tool alone or in combination with other clinical pathological parameters<sup>43,44</sup>. Potentially, the MSC test could enable greater consistency of lung cancer diagnostic algorithms, thus decreasing screening cost-effectiveness.

Besides MSC, other non-invasive tests for early lung cancer detection have been reported. Bianchi *et al.* proposed a circulating miR-Test based on a 34-miRNA signature<sup>45</sup>, then reduced it to a 13 miRNA signature<sup>42</sup>. Serum samples for discovery and validation were obtained from the Continuous Observation of Smoking Subjects (COSMOS) LDCT screening trial, performed at the European Institution of Oncology (IEO, Milan, Italy). The miR-Test showed 78% accuracy for lung cancer detection in high-risk volunteers. Between the 13 serum-miRNAs and the 24 composing the MSC, only 5 miRNAs were overlapping (miR-92a, miR-30b, miR-30c, miR-148a, and miR-140-5p). This could be due to the different type of starting samples (serum vs. plasma) and to data elaboration. In fact, for the miR-Test, the authors selected 6 miRNAs behaving as housekeeping in serum, 3 of which were included also among the MSC miRNA ratios (miR-15b, miR-19b, and miR-197). Nevertheless, the MSC in plasma and the miR-Test in serum samples more deeply validate the use of miRNA-based liquid biopsy for detection of lung cancer.

MiRNAs profiling for the identification of lung cancer biomarkers has been performed also on sputum samples<sup>46</sup>. Sputum can be easily collected and the miRNA expression levels appear to be extremely stable in it. However, lung cancer screening volunteers are older than 50-55 years and are heavy smokers with co-morbidities such as COPD, thus making sputum difficult to obtain. Indeed, among the different body fluids, plasma is one of the best alternatives to investigate the potential role of miRNAs as (but not limited to) lung cancer biomarkers.

Overall, the procedure described in the present paper can be relevant in the context of various diseases, with no restriction in the type of the pathology investigated (from cancer, to nervous system and cardiovascular disorders, or diabetes) and in the purpose of clinical use (a biomarker for diagnosis, prognosis, or even response to treatment).

## Disclosures

Gabriella Sozzi, Mattia Boeri, and Ugo Pastorino are co-inventors of three patent applications licensed to Gensignia Life Sciences, regarding the miRNA signature disclosed in this article. All the remaining authors declare no conflicts of interest in relation to the work submitted.

## Acknowledgements

The authors thank Paola Suatoni, Elena Bertocchi, Carolina Ninni, Annamaria Calanca, and Chiara Banfi for handling volunteers in the trials and for administrative assistance; and Claudio Jacomelli and Claudio Citterio for data management. The work was supported by the Italian Association for Cancer Research [Investigator Grants No. 15928 to UP, 14318 to GS, and 12162 (Special Program "Innovative Tools for Cancer Risk Assessment and early Diagnosis", 5x1000)]; the Italian Ministry of Health [Grant No. RF-2010]; Grant UO1 CA166905 from the National Cancer Institute (USA). MB was supported by a Fondazione Pezcoller Fellowship.

## References

1. Siegel, R.L., Miller, K.D., & Jemal, A. Cancer Statistics, 2017. *CA Cancer J Clin.* **67** (1), 7-30 (2017).
2. Malvezzi, M. *et al.* European cancer mortality predictions for the year 2017, with focus on lung cancer. *Ann Oncol.* **10** (2017).
3. Miller, K.D. *et al.* Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin.* **66** (4), 271-289 (2016).
4. Kanne, J.P. Screening for lung cancer: what have we learned? *AJR Am. J Roentgenol.* **202** (3), 530-535 (2014).
5. Aberle, D.R. *et al.* Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* **365** (5), 395-409 (2011).
6. Wender, R. *et al.* American Cancer Society lung cancer screening guidelines. *CA Cancer J Clin.* **63** (2), 107-117 (2013).
7. Molina-Vila, M.A. *et al.* Liquid Biopsy in Non-Small Cell Lung Cancer. *Front Med (Lausanne)*. **69** (2016).
8. Sozzi, G., & Boeri, M. Potential biomarkers for lung cancer screening. *Transl Lung Cancer Res.* **3** (3), 139-148 (2014).
9. Shepherd, F.A., Douillard, J.Y., & Blumenschein, G.R., Jr. Immunotherapy for non-small cell lung cancer: novel approaches to improve patient outcome. *J Thorac Oncol.* **6** (10), 1763-1773 (2011).
10. Ajona, D. *et al.* Investigation of complement activation product c4d as a diagnostic and prognostic biomarker for lung cancer. *J Natl Cancer Inst.* **105** (18), 1385-1393 (2013).
11. Boyle, P. *et al.* Clinical validation of an autoantibody test for lung cancer. *Ann Oncol.* **22** (2), 383-389 (2011).
12. Hayes, J., Peruzzi, P.P., & Lawler, S. MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol Med.* (2014).
13. Croce, C.M. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet.* **10** (10), 704-714 (2009).
14. Iorio, M.V., & Croce, C.M. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med.* **4** (3), 143-159 (2012).
15. Schwarzenbach, H., Nishida, N., Calin, G.A., & Pantel, K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol.* **11** (3), 145-156 (2014).
16. Mitchell, P.S. *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A.* **105** (30), 10513-10518 (2008).
17. Hanke, M. *et al.* A robust methodology to study urine microRNA as tumor marker: microRNA-126 and microRNA-182 are related to urinary bladder cancer. *Urol Oncol.* **28** (6), 655-661 (2010).

18. Park, N.J. *et al.* Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. *Clin Cancer Res.* **15** (17), 5473-5477 (2009).
19. Boeri, M., Pastorino, U., & Sozzi, G. Role of microRNAs in lung cancer: microRNA signatures in cancer prognosis. *Cancer J.* **18** (3), 268-274 (2012).
20. Castoldi, M. *et al.* A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). *RNA.* **12** (5), 913-920 (2006).
21. Chen, C. *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33** (20), e179 (2005).
22. Schulte, J.H. *et al.* Deep sequencing reveals differential expression of microRNAs in favorable versus unfavorable neuroblastoma. *Nucleic Acids Res.* **38** (17), 5919-5928 (2010).
23. Tiberio, P., Callari, M., Angeloni, V., Daidone, M.G., & Appierto, V. Challenges in using circulating miRNAs as cancer biomarkers. *Biomed Res Int.* **2015**, 731479 (2015).
24. Boeri, M. *et al.* Recent advances of microRNA-based molecular diagnostics to reduce false-positive lung cancer imaging. *Expert Rev Mol Diagn.* **15** (6), 801-813 (2015).
25. Benes, V., & Castoldi, M. Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods.* **50** (4), 244-249 (2010).
26. Mestdagh, P. *et al.* Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat Methods.* **11** (8), 809-815 (2014).
27. Pastorino, U. *et al.* Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet.* **362** (9384), 593-597 (2003).
28. Boeri, M. *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.* **108** (9), 3713-3718 (2011).
29. Sozzi, G. *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.* **32** (8), 768-773 (2014).
30. Pritchard, C.C. *et al.* Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila).* **5** (3), 492-497 (2012).
31. Kirschner, M.B. *et al.* Haemolysis during sample preparation alters microRNA content of plasma. *PLoS One.* **6** (9), e24145 (2011).
32. Kirschner, M.B. *et al.* The Impact of Hemolysis on Cell-Free microRNA Biomarkers. *Front Genet.* **4** (94) (2013).
33. Fortunato, O. *et al.* Assessment of circulating microRNAs in plasma of lung cancer patients. *Molecules.* **19** (3), 3038-3054 (2014).
34. Appierto, V. *et al.* A lipemia-independent NanoDrop(R)-based score to identify hemolysis in plasma and serum samples. *Bioanalysis.* **6** (9), 1215-1226 (2014).
35. Kim, D.J. *et al.* Plasma components affect accuracy of circulating cancer-related microRNA quantitation. *J Mol Diagn.* **14** (1), 71-80 (2012).
36. Kulstein, G., Marienfeld, R., Miltner, E., & Wiegand, P. Automation of DNA and miRNA co-extraction for miRNA-based identification of human body fluids and tissues. *Electrophoresis.* **37** (21), 2742-2750 (2016).
37. Schwarzenbach, H., da Silva, A.M., Calin, G., & Pantel, K. Data Normalization Strategies for MicroRNA Quantification. *Clin Chem.* **61** (11), 1333-1342 (2015).
38. Schlosser, K., McIntyre, L.A., White, R.J., & Stewart, D.J. Customized Internal Reference Controls for Improved Assessment of Circulating MicroRNAs in Disease. *PLoS ONE.* **10** (5), e0127443 (2015).
39. Kroh, E.M., Parkin, R.K., Mitchell, P.S., & Tewari, M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods.* **50** (4), 298-301 (2010).
40. Wang, K. *et al.* Comparing the MicroRNA spectrum between serum and plasma. *PLoS One.* **7** (7), e41561 (2012).
41. Mestdagh, P. *et al.* A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol.* **10** (6), R64 (2009).
42. Montani, F. *et al.* miR-Test: A Blood Test for Lung Cancer Early Detection. *J Natl Cancer Inst.* **19:107** (6), djv063 (2015).
43. Sestini, S. *et al.* Circulating microRNA signature as liquid-biopsy to monitor lung cancer in low-dose computed tomography screening. *Oncotarget.* **20** (32), 32868-32877 (2015).
44. Verri, C. *et al.* Mutational Profile from Targeted NGS Predicts Survival in LDCT Screening-Detected Lung Cancers. *J Thorac. Oncol.* **17**, 10 (2017).
45. Bianchi, F. *et al.* A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. *EMBO Mol Med.* **3** (8), 495-503 (2011).
46. Gyoba, J., Shan, S., Roa, W., & Bedard, E.L. Diagnosing Lung Cancers through Examination of Micro-RNA Biomarkers in Blood, Plasma, Serum and Sputum: A Review and Summary of Current Literature. *Int J Mol Sci.* **17** (4), 494 (2016).