UNIVERSITA' DEGLI STUDI DELL'INSUBRIA

PhD IN BIOTECHNOLOGIES, BIOSCIENCES AND SURGICAL TECHNOLOGIES (XXXI CYCLE)

Curriculum: Biotechnologies, Biosciences and Surgical Technologies

Serum miRNAs as biomarkers for early diagnosis of non-small cell lung cancer (NSCLC)

miRNAs sierici come biomarcatori per la diagnosi precoce del tumore al polmone non a piccole cellule (NSCLC)

Docente guida: Prof. **Paola Campomenosi**
Co-tutor: Prof. **Lorenzo Dominioni**
Co-tutor: Prof. **Douglas Noonan**

Thesis of:
D'Antona Paola
ID: 713586

Tables of contents

SUMMARY

Lung cancer, composed predominantly by non-small cell lung cancer (NSCLC), is the leading cause of cancer-related deaths worldwide. NSCLC comprises adenocarcinomas (AD, 50% of cases) and squamous cell carcinomas (SCC, 40% of cases), both characterized by a high degree of heterogeneity due to an interplay of genetic and microenvironmental factors.

Tobacco smoking and some pathological conditions, such as Chronic Obstructive Pulmonary Disease (COPD), are associated with the onset of lung cancer.

The high mortality of lung cancer is contributed to by late diagnosis, that is in turn attributable to the lack of non-invasive screening methods.

Recently, microRNAs (miRNAs) attracted the attention of scientists as possible lung cancer biomarkers since they are stable, retrievable and dysregulated in many types of cancers.

miRNAs, which belong to the class of small RNAs, are endogenous 19-24 nucleotides long non-coding RNAs. They have an intracellular function in post-transcriptionally regulating the expression of target mRNAs, but they have also been shown to be secreted in the extracellular environment and reach bloodstream, potentially regulating gene expression in distant cells.

Like standard genes, some microRNAs can act as oncogenes or oncosuppressors and their levels can be increased or reduced in cancer

patients compared to controls. However, there is no consensus about which miRNAs are the best biomarkers for lung cancer, because their quantification depends upon clinico-pathological and methodological factors, that are different in each study. One of the factors possibly underlying differences in the identification or quantification of miRNA biomarkers is the selection of subjects that are included in the "controls" in different studies, since smoking habits or inflammatory conditions, such as Chronic Obstructive Pulmonary Disease (COPD) can influence miRNA levels and impact on their diagnostic use as biomarkers.

To identify candidate miRNAs among those published as early stage lung cancer biomarkers, we performed a critical review of the literature and selected eight miRNAs, four of which with high sensitivity/AUC and four with high specificity as stage I-II NSCLC biomarkers. These miRNAs compose a two-step screening: in the first step the high sensitivity miRNAs are measured (miR-223, miR-20a, miR-145, miR-448), whereas in the second step the high specificity miRNAs are measured (miR-210, miR-628-3p, miR-29c, miR-1244). We aim to identify the two best performing miRNAs (one with high sensitivity and one with high specificity) in distinguishing between lung cancer patients and controls, to be applied as biomarkers for large scale screenings.

In the second part of this PhD project we optimized the conditions for measurement of the miRNAs composing the two-step screening by Droplet Digital PCR (ddPCR). DdPCR is a method that, by combining a traditional PCR amplification together with the use of fluorescent detection and Poisson statistics, enables sensitive and precise absolute

quantification of target, without the requirement for standard curves. As a result of the optimization phase, we excluded two miRNAs (miR-448 and miR-628-3p) from further analyses, due to the requirement of special conditions for their ddPCR, that make them less suitable as biomarkers for large scale screening.

We then investigated if the remaining miRNAs could discriminate patients with stage I-II NSCLC from controls; the latter included three different control subgroups: non-smokers, smokers and subject with COPD, to highlight possible differences among them.

All analyzed miRNAs so far (miR-223, miR-20a, miR-29c and miR-210) were expressed at significantly higher levels in NSCLC patients compared to controls; for three out of four miRNAs analyzed (miR-210, miR-20a, miR-29c) there was no significant difference among control subgroups, whereas miR-223 was significantly higher in the non-smoker subgroup compared to the remaining control subgroups. ROC curves were built for each miRNA to determine their ability in distinguishing NSCLC patients from controls; AUC values were respectively 0.71 for miR -20a, 0.80 for miR-223, 0.72 for miR-29c and 0.63 for miR-210. Accordingly, among the high-sensitivity miRNAs, miR-223 performed best; among high- specificity miRNAs, miR-29c performed best.

Therefore, we decided to use miR-223 and miR-29c for the next part of the project, aimed at determining the best combination and cut-offs of these miRNAs for early lung cancer diagnosis. To do this, we measured the levels of miR-223 and miR-29c in a Training set of 80 subjects [40 stage I-II lung cancer patients and 40 controls (20 smokers and 20 non-

smokers)] and tested combinations and cut-off values that allowed to obtain the best separation between cancer patients and controls. We found that the formula “miR-223>500 copies/µl OR miR-29c>50 copies/µl” had the highest sensitivity (75%) and acceptable specificity (50%). The formula was then applied to a blind Validation set, again composed of 80 subjects [40 stage I-II lung cancer patients and 40 controls (20 smokers and 20 non-smokers)] and found that it had a very high sensitivity of 92.5%, despite a poor specificity of 37.5%. As high sensitivity is a fundamental prerequisite for a first line, large scale screening test, our results suggest that our test holds great potential for screening of patients at risk for stage I-II NSCLC. Our test is also amenable for improvement, indeed other high specificity miRNAs from our panel, such as miR-1244, may be included in the formula to verify if they can increase specificity of our test.

RIASSUNTO

Il tumore polmonare, di cui il tumore al polmone non a piccole cellule (NSCLC) rappresenta il tipo più frequente, è la principale causa di morte correlata al cancro in tutto il mondo. Il NSCLC comprende gli adenocarcinomi (ADC, 50%) ed i carcinomi spino-cellulari (SCC, 40%), entrambi contraddistinti da un alto grado di eterogeneità, dovuta all'interazione di fattori genetici e microambientali.

Sia il fumo di sigaretta che alcune condizioni patologiche, come la broncopneumopatia cronica ostruttiva (BPCO), sono associate all'insorgenza del tumore polmonare. All'alta mortalità che caratterizza questo tipo di tumore contribuisce la diagnosi tardiva, a sua volta attribuibile alla mancanza di metodi di screening poco invasivi.

Recentemente, i microRNA hanno attratto l'attenzione della comunità scientifica come possibili biomarcatori, poiché è stato dimostrato che i livelli di molti microRNA sono deregolati in vari tipi di tumori e anche nei fluidi corporei in conseguenza della patologia. Nei biofluidi inoltre, i microRNA hanno dimostrato grande stabilità e sono facili da ottenere.

I microRNA sono prodotti endogenamente e appartengono alla classe dei piccoli RNA, sono infatti lunghi dai 19 ai 24 nucleotidi. All'interno della cellula fungono da regolatori post-trascrizionali dell'espressione genica, modulando la traduzione o la degradazione di RNA messaggeri target, ma possono anche essere secreti nell'ambiente extracellulare e giungere nel torrente circolatorio regolando potenzialmente anche cellule distanti.

Come altri geni, i microRNA possono agire come oncogeni o come oncosoppressori ed i loro livelli possono essere aumentati o ridotti negli individui con tumore rispetto a individui sani. Tuttavia, non c'è un consenso su quali miRNA siano i migliori biomarcatori del tumore al polmone, poiché la loro quantificazione dipende sia da fattori clinico-patologici sia da fattori metodologici, che sono diversi in ogni studio condotto.

Uno dei possibili fattori che possono influire sull'identificazione di specifici microRNA come biomarcatori è la selezione dei soggetti inclusi nei "controlli" dei diversi studi, poiché il fumo di sigaretta, o condizioni infiammatorie come la BPCO, potrebbero influenzare i livelli dei microRNA e quindi influire sul loro potenziale diagnostico come biomarcatori.

Per selezionare alcuni microRNA tra quelli già pubblicati come possibili biomarcatori del NSCLC in stadio precoce (stadio I e II), abbiamo condotto una revisione critica della letteratura e abbiamo selezionato 8 possibili candidati, quattro con alta sensibilità/AUC e quattro con alta specificità. Questi miRNA vengono proposti come parte di un sistema di screening a due fasi, in cui nella prima fase si misurano i microRNA ad alta sensibilità (miR-223, miR-20a, miR-145, miR-448) e nella seconda quelli ad alta specificità (miR-210, miR-628, miR-29c, miR-1244). In particolare, le caratteristiche di tali miRNA come biomarcatori verrebbero analizzate per identificare i due miRNA (uno ad alta sensibilità ed uno ad alta specificità) più performanti nel distinguere individui con tumore da controlli da applicare per uno screening su larga scala.

Nella seconda parte di questo progetto di dottorato abbiamo ottimizzato le condizioni di lavoro per la misurazione dei microRNA selezionati; la

misurazione avverrà mediante droplet digital PCR (ddPCR) che, combinando una tradizionale reazione di amplificazione di PCR fluorescente con il partizionamento della reazione e l'applicazione dell'analisi statistica di Poisson, consente una quantificazione assoluta e precisa del target senza la necessità di costruire curve standard.

I risultati della fase di ottimizzazione ci hanno consentito di escludere due microRNA da ulteriori analisi (miR-448 e miR-628-3p), in quanto avevano bisogno di condizioni particolari per la loro amplificazione, che mal si adattano ad uno screening su larga scala.

Successivamente, l'obiettivo è stato quello di verificare se i microRNA rimanenti fossero in grado di discriminare pazienti con NSCLC in stadio precoce (stadio I e II) da individui controllo, a loro volta suddivisi in 3 sottogruppi di controllo: non fumatori, fumatori e soggetti affetti da BPCO; infatti, un ulteriore obiettivo era evidenziare possibili differenze tra i sottogruppi di controlli.

Tutti i microRNA analizzati finora (miR-223, miR-20a, miR-210 e miR-29c) sono espressi a livelli maggiori nei pazienti affetti da NSCLC rispetto ai controlli; per tre dei quattro miRNA analizzati (miR-20a, miR-210 e miR-29c) non è stata riscontrata alcuna differenza significativa tra i sottogruppi di controllo mentre il miR-223 era significativamente maggiore nel gruppo dei non-fumatori rispetto agli altri due.

Sono state costruite le curve ROC per i microRNA in esame, per evidenziare la loro capacità di discriminare pazienti con NSCLC da individui controllo e i valori di AUC ottenuti sono stati rispettivamente 0.71 per miR-20a, 0.80 per miR-223, 0.72 per miR-29c e 0.63 per miR-210. Quindi tra i miRNA ad alta sensibilità, miR-223 è il migliore candidato; miR-29c è il più

performante tra quelli ad alta specificità. Abbiamo quindi deciso di utilizzare miR-223 e miR-29c per la successiva parte del progetto, con l'obiettivo di ricercare la migliore combinazione e i migliori valori di cut-off di questi due miRNA per la diagnosi precoce del cancro del polmone. Abbiamo misurato i livelli del miR-223 e miR-29c in un "Training set" di 80 soggetti [40 pazienti con carcinoma polmonare di stadio I-II e 40 controlli (20 fumatori e 20 non fumatori)] e abbiamo valutato combinazioni e valori di cut-off che ci permettessero di ottenere la migliore separazione tra pazienti con tumore e individui controllo. La formula "miR-223> 500 copie / µl OR miR-29c> 50 copie / µl" aveva la massima sensibilità (75%) e una specificità accettabile (50%). La formula è stata quindi applicata ad un "Validation set" in cieco, composto ancora da 80 soggetti [40 pazienti con carcinoma polmonare da stadio I-II e 40 controlli (20 fumatori e 20 non fumatori)]. Abbiamo riscontrato un'elevata sensibilità, pari al 92,5% ed una più scarsa specificità, pari al 37,5%. Poiché l'alta sensibilità è un prerequisito fondamentale per un test di screening su larga scala, i nostri risultati suggeriscono che il nostro test abbia comunque un grande potenziale per lo screening dei pazienti a rischio di NSCLC in stadio I-II. Ci prefiggiamo comunque di migliorare il nostro test, includendo altri microRNA ad alta specificità dal nostro pannello (miR-1244) e non, che possono essere inseriti nella formula verificando così se sono in grado di aumentare la specificità del nostro test.

INTRODUCTION

Lung cancer

- ***Lung cancer epidemiology***

Primary cancer of the lung ranks among cancers with the highest incidence and mortality both in the US and worldwide. [Siegel et al., 2018]. Indeed, lung cancer is the leading cause of cancer death in men and the second leading cause of cancer death after breast cancer in women. The gap between mortality in men and women has been reducing recently due to increasing diffusion of smoking habit among women in the recent past. Cigarette smoking is a major risk factor for lung cancer. It was estimated that 1.8 million new lung cancer cases occurred in 2012 globally, accounting for about 20% of all cancer deaths [Ferlay et al., 2015; Siegel et al., 2018] (Figure 1). The five-year survival rate of lung cancer is low worldwide (10-15%), despite recent advances in therapy of the disease, mainly because the majority of cases are diagnosed at an advanced stage, when surgery is not applicable and therapies ineffective [Lazar et al., 2013]. Usually, diagnosis of lung cancer occurs incidentally during routine tests, such as sputum cytology or chest-X-rays, performed for other indications. However, these tests showed low sensitivity and did not yield a satisfactory mortality reduction when used for lung cancer screening purposes [Dominioni et al., 2013]. Several studies aiming to diagnose lung cancer at an early stage, conducted in high-risk patients (older than 50

years old and heavy smokers) have provided promising results by using spiral computerized tomography (CT) [Pegna et al., 2009], showing a 20% reduction of lung cancer mortality (The National Lung Screening Trial Reaseach Team, 2011). Nevertheless, lung cancer screening by CT scan presents disadvantages, including radiation exposure, high cost/benefit ratio, high false positive rate and overdiagnosis. Yet early diagnosis of lung cancer is an important point in the management of this disease.

Estimated New Cases

Males				Females		
Prostate	164,690	19%		Breast	266,120	30%
Lung & bronchus	121,680	14%		Lung & bronchus	112,350	13%
Colon & rectum	75,610	9%		Colon & rectum	64,640	7%
Urinary bladder	62,380	7%		Uterine corpus	63,230	7%
Melanoma of the skin	55,150	6%		Thyroid	40,900	5%
Kidney & renal pelvis	42,680	5%		Melanoma of the skin	36,120	4%
Non-Hodgkin lymphoma	41,730	5%		Non-Hodgkin lymphoma	32,950	4%
Oral cavity & pharynx	37,160	4%		Pancreas	26,240	3%
Leukemia	35,030	4%		Leukemia	25,270	3%
Liver & intrahepatic bile duct	30,610	4%		Kidney & renal pelvis	22,660	3%
All Sites	**856,370**	**100%**		**All Sites**	**878,980**	**100%**

Estimated Deaths

Males				Females		
Lung & bronchus	83,550	26%		Lung & bronchus	70,500	25%
Prostate	29,430	9%		Breast	40,920	14%
Colon & rectum	27,390	8%		Colon & rectum	23,240	8%
Pancreas	23,020	7%		Pancreas	21,310	7%
Liver & intrahepatic bile duct	20,540	6%		Ovary	14,070	5%
Leukemia	14,270	4%		Uterine corpus	11,350	4%
Esophagus	12,850	4%		Leukemia	10,100	4%
Urinary bladder	12,520	4%		Liver & intrahepatic bile duct	9,660	3%
Non-Hodgkin lymphoma	11,510	4%		Non-Hodgkin lymphoma	8,400	3%
Kidney & renal pelvis	10,010	3%		Brain & other nervous system	7,340	3%
All Sites	**323,630**	**100%**		**All Sites**	**286,010**	**100%**

Figure 1. *Ten leading cancer types for the Estimated New Cancer Cases and Death by Sex United States, 2017.* *[Siegel et al., 2017]*

- ***Risk factors***

Several risk factors can be accounted for in the genesis of lung cancer, the principal being tobacco consumption; other factors, such as genetic factors, poor diet, occupational exposures and air pollution may act independently or in concert with tobacco smoking to determine the epidemiology of lung cancer [Malhotra et al., 2016].

It has been estimated that heavy smokers have approximately a 20-fold increase in the risk of developing lung cancer compared to non-smokers and that the duration of smoking and number of cigarettes smoked should be considered as the strongest determinant of lung cancer risk in smokers [US Department of Health and Human Consequences of smoking: 50 years of progress, 2014]. Nevertheless, also non-smokers, exposed to second-hand smoking, show an increased relative risk to develop lung cancer during their lives [Hackshaw et al., 1997; Boffetta, 2002].

Besides cigarette smoking, there are other environmental risk factors that are involved in 10-15% of total cases of lung cancer: chromium, silica, polycyclic aromatic hydrocarbons, ionizing radiation, outdoor and indoor air pollution and hormonal factors.

Chromium [VI] compounds, crystalline silica and polycyclic aromatic hydrocarbons, groups of chemicals formed during combustion of organic material, increase the risk of lung cancer among workers employed in a number of industries and represent a major source of occupational exposure [Malhotra et al., 2016].

Ionizing radiation increases the risk of lung cancer, as demonstrated by studies on Hiroshima and Nagasaki atomic bomb survivors, but there is a

difference between this single massive exposure and the smaller doses that the population may receive during X-ray exams or computed tomography scans [Schwartz et al., 2016].

Moreover, indoor air pollution is considered the major risk factor for lung cancer arisen in never-smoking women in Asia, due to fumes from high temperature cooking using crude vegetable oils, as rapeseed oil, as well as coal burning or wood burning in poorly ventilated houses. [IARC Monographs on the Evaluation of carcinogenic Risks to Humans, 2006].

Finally, some conditions, such as chronic obstructive pulmonary disease (COPD), correlate with lung cancer. COPD is a type of obstructive lung disease characterized by long-term poor airflow, often caused by cigarette smoking. So, COPD and lung cancer share common risk factors, but other studies also suggest that COPD itself is a risk factor for lung cancer, independently from smoking habits, increasing the risk of developing lung cancer up to 3 times, even among never smokers [Schwartz et al., 2016].

About the genetic risk factors, independent genome-wide association studies identified a susceptibility region in 15q25.1 [Amos et al., 2008; Thogeirsson et al., 2008]. This genomic region contains six identified ORFs encoding nicotinic acetylcholine receptors in neuronal and other tissues [Amos et al., 2008]. 15q25 is the only susceptibility locus confirmed to be implicated in all types of lung cancer, independently of their histology [Timofeeva et al., 2012].

- ***Lung cancer classification***

The majority of primary lung cancers are lung carcinomas and can be classified into two groups, based on histology: small cell lung carcinoma (SCLC, about 15% of cases) and non-small cell lung carcinoma (NSCLC, about 85% of cases) [Chen et al., 2014]. SCLC is an aggressive neuroendocrine tumor consisting of small tumor cells deriving from epithelial and neuroendocrine cells. This type of lung cancer is strongly associated with smoking and has very poor prognosis. Patients with SCLC are rarely operated due to fast spread of the tumor; this however is more responsive to chemotherapy and radiotherapy than other types of lung cancer [Rekhtman et al., 2010; Travis et al., 2011].

NSCLC has a more favorable prognosis and is subdivided into adenocarcinomas (ADC, 50% of the cases), squamous cell carcinomas (SCC, 40% of the cases) and large cell carcinomas (LCC, 10% of the cases). However, with advances in molecular typing of lung cancer, the LCC subtype is gradually disappearing, because it does not seem to be genetically distinct from the other two [Chen et al., 2014]. Lately, the ADC subtype has been found more frequently than SCC both in men and women [Devesa et al., 2005]; the shift towards this subtype seems to be related to higher concentrations of certain carcinogens [Stellman et al., 1997].

NSCLC is staged from IA to IV, IA having the best prognosis and IV the worst, based on the degree of spreading from the primary tumor [Mountain et al., 2003].

Five-year relative survival of NSCLC patients is inversely related to stage

at the time of diagnosis: whereas five-years survival ranges from 70% to 85% in patients with early stage disease (stage I and II), survival drastically decreases to 4% in patients diagnosed with stage IV lung cancer [Siegel et al., 2017].

NSCLC subtypes differ from one another for various hallmarks (Fig. 2): ADCs commonly arise in distal airways whereas SCCs originate in more proximal airways and are more strictly correlated to smoking habit and inflammation than ADCs.

ADC is a malignant epithelial tumor with glandular differentiation, that expresses biomarkers typical of an origin in the distal lung, such as thyroid transcription factor 1 (TTF-1, also known as NKX2-1) and keratin 7 (KRT7). In contrast, SCCs have a squamous differentiation, that reminds of the pseudostratified columnar epithelium of the trachea and upper airways. SCCs can be distinguished from ADCs because of their positivity to cytokeratin 5 and 6 and/or to the transcription factors SRY-box 2 and p40, an isoform of p63, by immunostaining (Figure 2) [Chen et al., 2014].

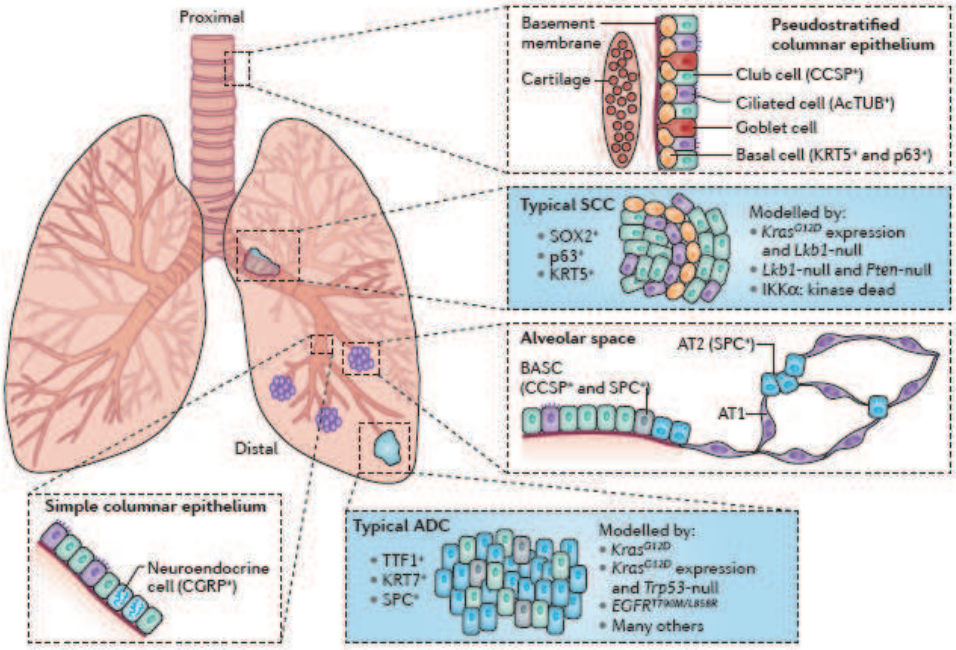


Figure 2. *A diagram of proximal and distal lung cells, indicating markers that are retained in ADC and SCC carcinomas and suggest their putative cells of origin.* *[Chen et al., 2014]*

- ***Lung cancer and genomics***

Lung cancer is a heterogeneous disease, but understanding the multiple complex combinations of morphological, genetic and molecular alterations underlying its formation can help to define tumor subtype and to guide treatment decisions [Travis WD, 2011; Hanahan et al, 2000].

For example, driver oncogenic mutations suitable as targets for therapy in ADC, are those found in Epidermal growth factor receptor (EGFR) and Kirsten rat sarcoma viral oncogene homolog (KRAS), comprising between 5-15% of the cases. Other well defined genetic mutations appearing in 5% of ADCs are the Echinoderm microtubule-associated protein-like 4 (EML4)

and anaplastic lymphoma kinase (ALK) fusion gene, and mutations involving estrogen-related receptor beta type 2 (ERRB2), NRAS, v-raf murine sarcoma viral oncogene homolog B1 (BRAF), phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA), met protooncogene (MET) and cadherin-associated protein beta 1 (CTNNB1) genes [Pao et al., 2011; Seo et al., 2012].

The genomic landscape of the SCC seems to have different specific mutations [Cancer genome Atlas Research Network, 2012].

The KRAS, EGFR mutations and ALK rearrangements are rare in comparison to ADC, whereas the *ERBBs*, Fibroblast growth factor receptor 1 (FGFR1), the tyrosine kinase DDR2 genes and the JAK/STAT pathway are frequently altered by mutation or amplification in SCC. This suggests that subtype specific alterations play a crucial role in therapy decisions in lung cancer.

A common feature for both ADC and SCC is a strong correlation between smoking status and number and type of mutations. Smokers have a 10 fold higher mutation rate compared to non-smokers. Mutations in BRAF, JAK2, JAK3, TP53 and mismatch repair genes are strongly associated with smoking, whereas EGFR, ROS1, and ALK rearrangements appear in never smokers as well [Govindad et al., 2012; Kandoth et al., 2013].

It is also conceivable that the biology of the cell of origin is what drives the presence of different mutations in the two subtypes of NSCLC. Chen et al. showed that ADCs arise from AT2 cells or club cells in the alveolar space. SCCs, instead, have long been hypothesized to arise from basal cells (Figure 2) [Chen et al., 2014]; however, these hypotheses still need to be proven.

- ***Lung cancer and biomarkers***

Biomarkers are molecules that can be used to discriminate between normal and abnormal statuses, for example healthy from cancerous conditions. Tumor biomarkers can be very different in nature: they include nucleic acids (including search for specific mutations, gene copy number alterations, gene expression profiles,), epigenetic changes (changes in DNA methylation profiles), proteins (alteration in level and profiles of protein expression), metabolic markers (changes in level and spectrum of low molecular weight metabolites), DNAs and RNAs circulating in the bloodstream, including miRNAs, circulating tumor cells (CTCs) and immune, stromal and endothelial cells. Cancer biomarkers can be used for screening, detection, diagnosis, prognosis, prediction, stratification, and monitoring of therapy response. [I and Cho, 2015; Villalobos et al., 2017]. Ideally, a biomarker should be stable, available, measurable and cost-saving; its diagnostic efficacy can be measured by parameters such as sensitivity and specificity. Sensitivity is defined as the number (percentage) of subjects testing positive for the biomarker among those affected by the condition (true positives), whereas specificity is defined as the number (percentage) of true negative, subjects correctly testing negative among those not showing the condition.

Proteins are considered suitable biomarkers for cancer due to their involvement in cellular processes leading to cancer development. However, despite the great advances made in lung cancer biomarker discovery, no protein biomarkers with high enough sensitivity and specificity have been found. This is likely attributable to multiple reasons:

genetic heterogeneity of tumors, poor performance of techniques applied to biomarker identification, poor reproducibility of laboratory tests and low concentration of biomarkers.

Examples of proteins commonly used as biomarkers include: cytokeratine 19 fragments (CIFRA 21-1), EPCAM (epithelial cell adhesion molecule), ProGRP (pro-gastrin-releasing peptide), CEACAM (carcinoembryonic antigen); however, in clinical practice these protein biomarkers fail to provide sufficient sensitivity for optimal screening. Some clinically used biomarkers, such as CEACAM, CYFRA 21-1 and ProGRP have low concentration in the serum; thus, single biomarkers cannot be reliably used for early lung cancer diagnosis and biomarker combinations are sometimes used.

More recently, microRNAs have attracted the attention of researchers as possible biomarkers for screening of early stage lung cancer, because they possess some characteristics that satisfy the requirement for good biomarkers, mainly their stability in biofluids, as will be described better in the following chapter.

MicroRNAs

- ***Introduction***

For long time the field of molecular biology has been ruled by the central dogma that DNA is transcribed into RNA and the latter translated into proteins, but recent discoveries have added layers and new information to this principle; a group of RNAs termed “non-coding RNAs” have been found to play a role in regulating both transcription and translation.

MicroRNAs, that together with siRNAs and piRNAs, belong to the class of small non-coding RNAs, are small endogenous molecules defined by their length of 19-24 nucleotides and their association with Argonaute (AGO) proteins to guide target-specific gene regulation [Garzon et al., 2010; Krol et al., 2010].

MicroRNAs were discovered in 1993 in *Caenorhabditis elegans* (*C. elegans* [Lee et al., 1993]. Since then the field has evolved rapidly: thousands of microRNAs have been discovered and their involvement in physiology and pathology has been described. Currently, microRNAs are among the most studied molecules [Gyoba et al., 2016].

In 2000, Reinhart et al. and Pasquinelli et al. reported the discovery of a second microRNA, termed let-7, highlighting that microRNAs were evolutionarily conserved and not specific only to Nematodes [Reinhart et al., 2000; Pasquinelli et al., 2000]. This finding marked the beginning of a wave of studies on microRNAs biogenesis, role and significance in various normal and pathologic conditions.

- ***miRNA genomic localization***

MiRNA genes are found across all chromosomes and can be in intergenic regions or within genes [Lagos-Quintana et al., 2001; Lau et al., 2001]. Some miRNA genes can be distant place from other genes of the same miRNA family, whereas some others may be in neighbor places or be even grouped in clusters, defined by miRbase (http://www.mirbase.org/) as a group of miRNA genes that are located within 10 kb of each other; microRNAs belonging to the same cluster can either be co-transcribed or transcribed independently [Hausser et Zavolan, 2014; Ramalingam et al., 2013].

- ***miRNA structure***

MiRNA molecules are made of a single strand of about 22 nucleotides; the sequence at the 5′ end of the strand that spans nucleotide positions 2 to 7 is essential for target recognition and is termed the “miRNA seed”. The nucleotides located downstream (particularly nucleotide 8 and at a minor extent nucleotides 13–16) concur to base pairing with the targets. MiRNA-binding sites are usually located in the 3′ untranslated region (UTR) of mRNAs [Rolle et al., 2016]

- ***miRNA biogenesis***

The biogenesis of miRNAs begins in the nucleus where miRNA genes are transcribed by either RNA polymerase II or III into long primary transcripts

(pri-miRNAs) that are polyadenylated at the 3’ end and capped at the 5’ end. These long transcripts are characterized by the presence of a hairpin like structure and are further processed in the nucleus by RNase III/DROSHA complex, which crops the primary transcript down to a precursor miRNA (pre-miRNA), a small hairpin-shaped RNA of 70-120 nucleotides long; for correct processing by the RNase III/DROSHA complex, the cofactor DGCR8, also named Pasha, is essential. The pre-miRNA has a protruding 3’ end, terminating with the hydroxyl group, and a phosphate group at the 5’ end [Ha et al., 2014].

The pre-miRNA is then transported to the cytoplasm with a mechanism involving the transporter Exportin-5, which recognizes the nucleotides protruding at the 3’ end; this transport is active and based on the Ran-GTP complex [Murchison et Hannon, 2004]. The pre-miRNA is further processed by the cytoplasmic RNase III/DICER in the mature miRNA molecule of about 22 nucleotides [Fazi et al., 2008]. This duplex is unwound by helicase and just one mature strand enters the multicomponent complex termed RNA-induced silencing complex (RISC) which includes AGO proteins, while the complementary strand is degraded. Sometimes, the strand with poorer stability at the 5’ end is kept incorporated within the RISC [Schwarz et al., 2003].

MiRNAs regulate gene expression preferentially by binding to the complementary strand in the 3´UTR of the mRNA, leading to mRNA degradation, translational inhibition or destabilization. However, it is now known that miRNAs can bind to 5´UTRs or open reading frames (ORF) and eventually also upregulate their targets [Vasudevan et al., 2007].

An alternative pathway for miRNA biogenesis, without the DROSHA

mediated cleavage, occurs for miRNAs located within introns and takes place during splicing of pre-mRNAs. These miRNAs can also undergo other maturation processes before leaving the nucleus [Piva et al., 2013].

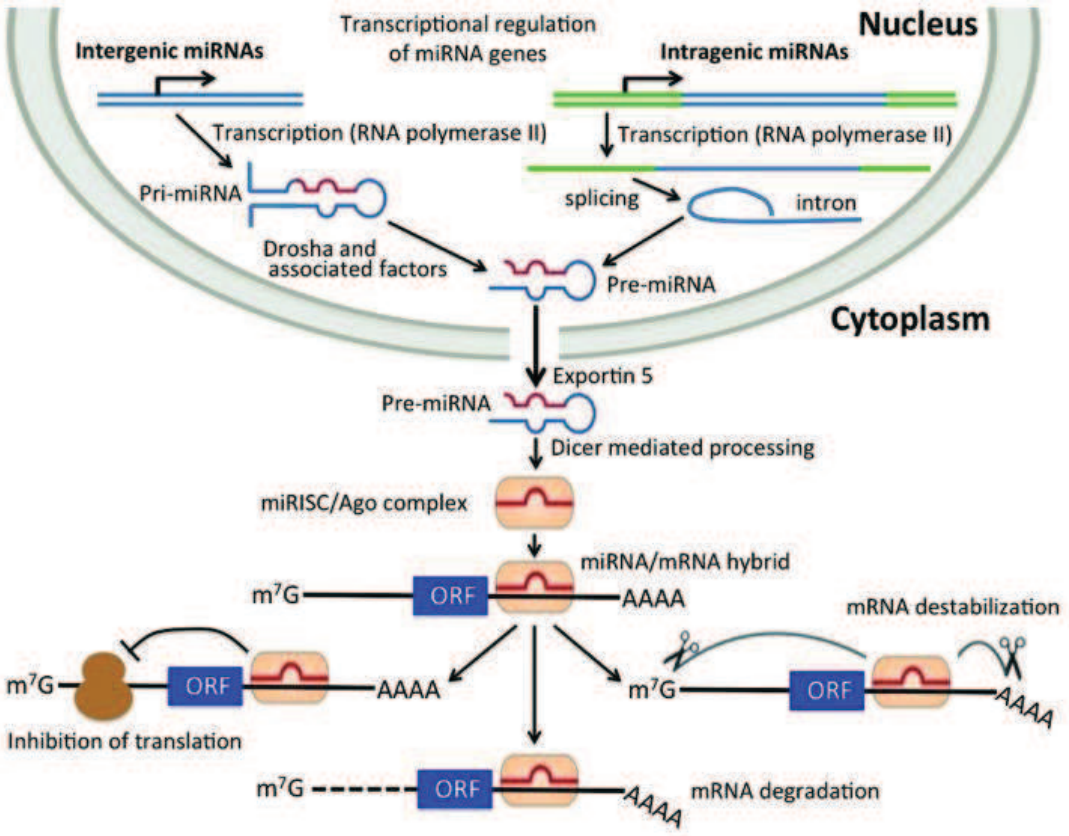


Figure 3. *microRNA biogenesis.* *[Piva et al. 2013]*

- **miRNA functions**

The intracellular function of miRNAs is the regulation of gene expression at the post-transcriptional levels by the gene silencing mechanism by the RISC complex [Rana et al., 2007]. A single miRNA guide can regulate several mRNA targets and conversely multiple miRNAs can cooperatively regulate a single mRNA target [Bartel, 2004]. MiRNAs guide miRISC to

specifically recognize messenger RNA and downregulate gene expression by one of two mechanisms: translational repression and mRNA cleavage. The degree of miRNA–mRNA complementarity is a main factor guiding the choice between the two regulatory mechanisms. A high degree of complementarity allows Ago-catalyzed degradation of target mRNA sequences through the mRNA cleavage mechanism. Conversely, the presence of a central mismatch favors the translational repression process. The mechanism for repression of target mRNA translation by miRISC is still unknown and whether repression occurs at the translational initiation or post-translational level has yet to be defined. One main difference between the downstream consequences is reversibility: mRNA degradation is an irreversible process while translation inhibition is a reversible mechanism because the mRNA can be translated following elimination of translation repression [Valencia-Sanchez et al., 2006; Maroney et al., 2006].

MicroRNAs can also exert their functions in the extracellular environment, and their presence has been detected in all biological fluids [Lawrie et al., 2008; Mitchell et al., 2008].

It has been shown that biofluids from patients with specific pathological states show distinct miRNA expression profiles. This indicates that circulating miRNAs are not passively liberated from necrotic or injured cells; rather miRNAs are selectively released by cells [Mar-Aguilar et al., 2013; Noferesti et al., 2015].

Circulating miRNAs are remarkably stable and can survive under unfavorable conditions for a long time, in contrast to intracellular miRNAs, which are degraded in extracellular environment within few seconds. The

mechanisms underlying the remarkable stability of circulating miRNAs in the RNase-rich environment of biofluids are not well clarified. It has been hypothesized that miRNAs are conjugated with molecules that would protect them from RNase activity, such as lipids or high-density lipoprotein complexes [El-Hefnawy et al., 2004]; alternatively, miRNAs could be packaged into membrane-bound vesicles like exosomes, microvesicles or apoptotic bodies (Figure 4) [Valadi et al., 2007]. Exosomes are vesicles of endosomal origin of 50-100 nm in diameter, that contain lipids, proteins and RNAs; microvesicles are larger than exosomes (100-1000 nm in diameter) and are released into the extracellular space by shedding of the plasma membrane.

Studies demonstrated that microvesicles, like exosomes, can be involved in cell-cell communication, transferring transmembrane proteins and cytosol components, such as microRNAs, from one cell to another [Kinet et al., 2013].

Finally, also apoptotic bodies have been shown to contain microRNAs; they are a heterogeneous population of membrane vesicles, containing organelles, nuclear fragments and parts of the cytoplasm, that are subsequently taken up by macrophages and dendritic cells.

[Bayraktar et al., 2017].

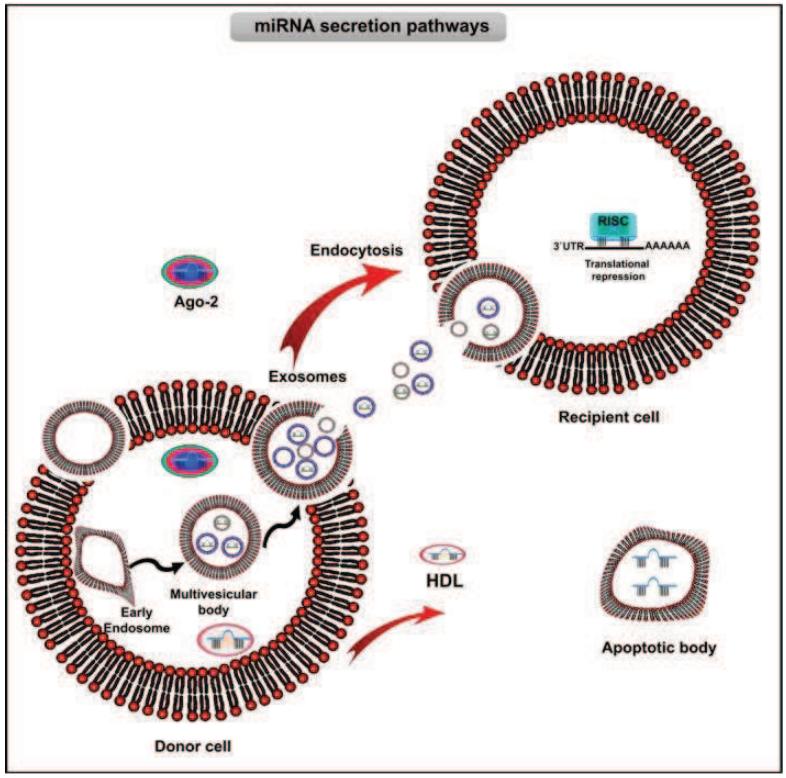


Figure 4. *Mechanisms of microRNA release and uptake between donor and recipient cells.* *[Bayraktar et al., 2017]*

- **miRNAs and lung cancer**

The main function of cellular miRNAs is to regulate cellular processes such as development, differentiation, proliferation and apoptosis [Bartel, 2004]. While little is known about the specific targets and biological functions of miRNA molecules, it is evident that miRNAs play a crucial role in the regulation of gene expression [Ambros, 2004]. Takamizawa and colleagues published the first evidence for the role of miRNAs in lung cancer in 2004 [Takamizawa et al., 2004]. This study found that let-7 expression correlated with post-surgery survival in non small cell lung cancer (NSCLC). Moreover, overexpression of let-7 in A549 cells inhibited

their proliferation, indicating a tumor suppressor function of this miRNA. Later studies on the role of let-7 in lung cancer confirmed this hypothesis, by demonstrating that let-7 targets the *RAS* genes, an oncogenic gene family that is frequently mutated and upregulated in lung adenocarcinomas [Seo et al., 2012; Johnson et al., 2005].

Other examples of oncogenes regulated by let-7 are *cdc25a*, *cdk16* and *cyclin D,* involved in the G1/S transition, and BCL-2, an anti-apoptotic gene [Xiong et al., 2011].

Proliferation is an important hallmark of cancer, including lung cancer. Feng and colleagues demonstrated that miR-192 overexpression in the NSCLC cell lines A549 and H460 inhibits cell proliferation and tumorigenesis *in vivo* [Feng et al., 2011].

The first oncogenic miRNA identified in lung cancer was the miR cluster mir-17-92, frequently amplified in small cell lung cancer (SCLC) [Hayashita et al., 2005]. miR-21 is another well studied oncogenic microRNA in several types of cancer, including lung cancer: in lung adenocarcinoma, miR-21 was shown to be upregulated by EGFR signaling and to target tumor-suppressor PTEN [Li et al., 2012].

- **miRNAs as lung cancer biomarkers**

Lung cancer is a relevant public health problem, causing the majority of cancer deaths in developed countries; about 85% of lung cancer cases is represented by non-small cell lung cancer (NSCLC). Despite advances in standard treatments, the high mortality of this tumor is related to the fact

that most patients are diagnosed in the late stage of the disease, when they develop clinical symptoms such as coughing, haemoptysis, chest pain and shortness of breath. The lung cancer five-year survival rate is low worldwide (10-15%) [Koike et al., 2013].

However, if non-small cell lung cancer were discovered at an early stage, the probability of survival would dramatically increase (Figure 5): this is why it is of fundamental importance to find lung cancer screening tests and diagnostic biomarkers research for this disease.

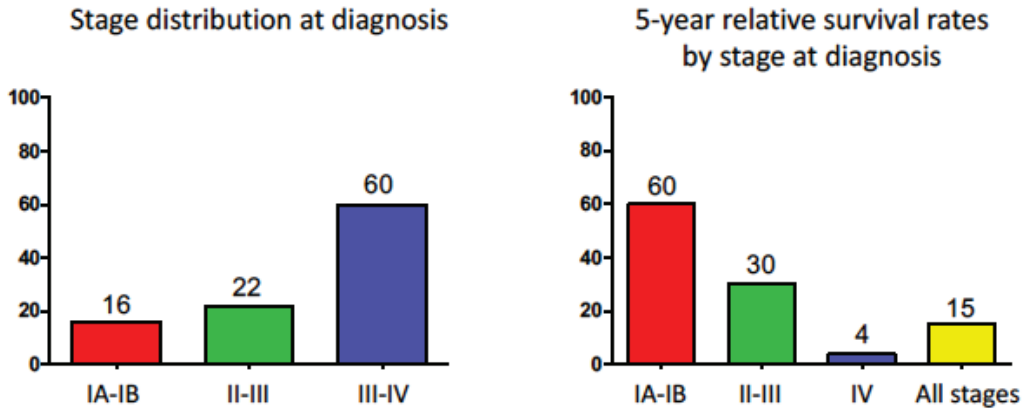


***Figure 5**. Left: the average percentage of cases diagnosed at the specified stage is reported. Right: relative 5-year survival by stage. [Modified from Siegel et al., 2017]*

Recently, microRNAs (miRNAs) have been suggested as a novel class of tumor biomarkers, as the levels of some of them are altered in various human cancers, including NSCLC [Esquela-Kerscher et al. 2006, Iorio et al. 2012, Negrini et al. 2014]. miRNAs presence and stability in biofluids, together with the evidence that some of them correlate with clinical-pathological parameters, are strong points suggesting their application as circulating biomarkers [Blondal et al. 2013].

The first evidence that microRNAs could be used as biomarkers of solid tumors was suggested by Mitchell and colleagues in 2008; they showed that miRNAs are present in the bloodstream in a stable form and that extracellular miRNAs originating from human prostate cancer xenografts enter in the bloodstream. Furthermore, the Authors showed that the levels of miR-141 were higher in patients with prostate cancer than in healthy individuals [Mitchell et al., 2008].

The first analysis of miRNA panels in the sera of patients with NSCLC was conducted by Chen and colleagues in 2008; 28 microRNAs, present in healthy individuals, were not found in patients and 63 microRNAs out of 132 were detected in patients and not in healthy individuals suggesting that specific miRNA panels detectable in serum or plasma samples could discriminate between tumor and healthy subjects [Chen et al., 2008].

Although many studies have described specific panels of microRNAs as biomarkers in lung cancer, there is no overlap between the miRNAs identified in independent studies. This can be due to several reasons, attributable both to clinicopathological and to methodological factors (Table 1) [Moretti et al., 2017].

Sample type (blood, serum, plasma or other biofluids), sample preparation procedures, technique used for miRNA quantification, normalization method, could play a role. Indeed, one of the technical issues is the lack of reliable endogenous miRNAs to be used as reference for normalization in RT-qPCR, the most common method used: for example, some of the microRNAs used as reference molecules in some studies, are described as disease biomarkers in other reports [Moretti et al., 2017].

Another very important issue relates to the choice of subjects to be used as “healthy controls”: only few of the published articles on biomarker identification are prospective studies, whereas most are case control studies, comparing miRNA levels between lung cancer patients and “controls”. The composition of the control group in lung cancer is a critical issue because chronic disease, such as Chronic Obstructive Pulmonary Disease (COPD) or other factors, such as smoking history, cardiovascular disorders, diabetes, pregnancy etc, may influence circulating miRNA levels and may impact on the evaluation of miRNAs as biomarkers. Unfortunately, in the majority of studies the tumor group and the control group are typically matched only by age, gender and rarely by smoking status, with scarce or no information about their medical history [Moretti et al., 2017; Zandberga et al. 2013].

Table 1. *Factors potentially affecting circulating miRNA quantification in NSCLC patients. [Moretti et al., 2017]*

Clinicopathological factors	Ethnicity
	Gender/age
	Smoking status
	Stage of disease (early/advanced)
Methodological factors	Type of sample (plasma/serum/whole blood)
	Hemolysis
	RNA extraction method
	Reverse transcription method
	miRNA quantification method
	Normalization

AIM OF THE STUDY

Lung cancer is the main cause of cancer death in men and women worldwide. The available diagnostic methods and screening tools are not sensitive enough and the mortality of this tumor is very high because symptoms appear at an advanced stage. To increase the survival of lung cancer patients the main goal is to find specific and sensitive biomarkers to diagnose the lung cancer at an early stage.

MicroRNAs have been considered as possible cancer biomarkers since they are stable, retrievable, measurable and dysregulated in many types of tumors, including NSCLC.

Therefore, we performed a critical review of the literature to identify circulating miRNAs suitable for non-invasive screening of stage I-II NSCLC and we proposed a two-step screening based on miRNA panels with high sensitivity and high specificity, respectively.

Then, we measured the serum levels of the miRNAs composing the two-step screening to ascertain whether they could be used to discriminate patients with NSCLC stage I-II from controls. Within this context, we also aimed at verifying if some factors, such as smoke or chronic disease, may influence miRNA levels and affect their application as biomarkers. To do so, we evaluated serum levels of miRNAs of interest in three different control subgroups (non-smokers, smokers and subjects with Chronic Obstructive Pulmonary Disease), as well as in patients with stage I-II NSCLC, making all comparisons.

We believe that this rigorous strategy will allow us to identify a panel of

miRNA biomarkers with high repeatability and reproducibility among laboratories, still lacking in this type of research.

RESULTS

The results of my PhD work are included in the following section as manuscripts that we already published or are about to submit for publication:

1) Moretti F, D'Antona P, Finardi E, Barbetta M, Dominioni L, Poli A, Gini E, Noonan DM, Imperatori A, Rotolo N, Cattoni M, Campomenosi P. Systematic review and critique of circulating miRNAs as biomarkers of stage I-II non-small cell lung cancer. Oncotarget. 2017; 8(55):94980-94996.

2) Screening with serum miR-223 and miR-29c for early diagnosis of Non Small Cell Lung Cancer, under submission.

Moreover, at the beginning of my PhD I contributed to a previous paper published on BMC Biotechnology:

Campomenosi P, Gini E, Noonan DM, Poli A, D'Antona P, Rotolo N, Dominioni L, Imperatori A. A comparison between quantitative PCR and droplet digital PCR technologies for circulating microRNA quantification in human lung cancer. BMC Biotechnol. 2016; 16(1):60.

www.impactjournals.com/oncotarget/ Oncotarget, 2017, Vol. 8, (No. 55), pp: 94980-94996

Review

Systematic review and critique of circulating miRNAs as biomarkers of stage I-II non-small cell lung cancer

Francesca Moretti[1], Paola D'Antona[2], Emanuele Finardi[1], Marco Barbetta[1], Lorenzo Dominioni[3], Albino Poli[1], Elisabetta Gini[2], Douglas M. Noonan[2,4], Andrea Imperatori[3], Nicola Rotolo[3], Maria Cattoni[3] and Paola Campomenosi[2,5]

[1]Department of Diagnostic and Public Health, University of Verona, Verona, Italy

[2]Department of Biotechnology and Life Sciences, DBSV, University of Insubria, Varese, Italy

[3]Department of Medicine and Surgery, DMS, Center for Thoracic Surgery, University of Insubria, Varese, Italy

[4]Scientific and Technological Pole, IRCCS MultiMedica, Milan, Italy

[5]The Protein Factory, Centro Interuniversitario di Ricerca in Biotecnologie Proteiche, Politecnico di Milano, ICRM-CNR Milano and University of Insubria, Varese, Italy

Correspondence to: Paola Campomenosi, **email:** paola.campomenosi@uninsubria.it

Keywords: *microRNA, biomarkers, circulating, non-small cell lung cancer, stage I-II NSCLC*

Received: *August 09, 2017* ***Accepted:*** *September 22, 2017* ***Published:*** *October 11, 2017*

ABSTRACT

Selected circulating microRNAs (miRNAs) have been suggested for non-invasive screening of non-small cell lung cancer (NSCLC), however the numerous proposed miRNA signatures are inconsistent.

Aiming to identify miRNAs suitable specifically for stage I-II NSCLC screening in serum/plasma samples, we searched the databases "Pubmed", "Medline", "Scopus", "Embase" and "WOS" and systematically reviewed the publications reporting quantitative data on the efficacy [sensitivity, specificity and/or area under the curve (AUC)] of circulating miRNAs as biomarkers of NSCLC stage I and/or II. The 20 studies fulfilling the search criteria included 1110 NSCLC patients and 1009 controls, and were of medium quality according to Quality Assessment of Diagnostic Accuracy Studies checklist. In these studies, the patient cohorts as well as the control groups were heterogeneous for demographics and clinicopathological characteristics; moreover, numerous pre-analytical and analytical variables likely influenced miRNA determinations, and potential bias of hemolysis was often underestimated. We identified four circulating miRNAs scarcely influenced by hemolysis, each featuring high sensitivity (> 80%) and AUC (> 0.80) as biomarkers of stage I-II NSCLC: miR-223, miR-20a, miR-448 and miR-145; four other miRNAs showed high specificity (> 90%): miR-628-3p, miR-29c, miR-210 and miR-1244. In a model of two-step screening for stage I-II NSCLC using first the above panel of serum miRNAs with high sensitivity and high AUC, and subsequently the panel with high specificity, the estimated overall sensitivity is 91.6% and overall specificity is 93.4%. These and other circulating miRNAs suggested for stage I-II NSCLC screening require validation in multiple independent studies before they can be proposed for clinical application.

INTRODUCTION

Lung cancer is the most common cause of cancer death worldwide, globally accounting for an estimated 1.5 million deaths in 2012 [1, 2]. In Europe, every year lung cancer causes about 353000 deaths, which represent nearly 20% of total cancer deaths [3]. Approximately 15% of lung cancers are histologically classified as small cell lung cancer, a very aggressive and generally incurable tumor; the remaining 85% are cumulatively classified as Non-

Small Cell Lung Cancer (NSCLC). The latter contains two main histological subtypes, adenocarcinoma (AC) and squamous cell carcinoma (SCC), and can often be cured if diagnosed at an early stage [4, 5].

The five-year survival rate of lung cancer is low worldwide (10–15%), mainly because the majority of cases are diagnosed with advanced stage, when treatment is rarely curative. In the NSCLC cases that are diagnosed at an early stage (stage I and II), the five-year survival rate dramatically improves, ranging from 70% to 85% for surgically resected stage I disease [6], lobectomy being the established and most effective therapeutic approach [7]. However, less than one third of NSCLC cases are diagnosed in early stage [8–10] and the methodologies currently available for early diagnosis present several limitations. Chest X-rays have low sensitivity for lung cancer detection, whereas low-dose chest computed tomography (CT) scan has high sensitivity but low specificity [11–14]. The latter is a relevant limitation of CT scan for screening, considering that among individuals at risk for lung cancer (heavy smokers and former smokers) 20–60% of chest CT exams show pulmonary nodules, the vast majority of which are eventually diagnosed as benign after completion of work up [15, 16]. Moreover, in many areas of the world chest CT is a rather expensive and not widely available screening tool [17].

Newer, minimally invasive and effective methods of screening for lung cancer are needed. MicroRNAs (miRNAs) are small non-protein-coding RNA molecules 18–25 nucleotide long that play an important role in eukaryotic gene expression regulation. They have been shown to be dysregulated in human diseases, including cancer [18, 19]. The aberrant expression of specific miRNAs in body fluids from individuals with cancer has suggested their possible application as cancer biomarkers [20–25]. The quantification of selected miRNAs in plasma or serum of high risk individuals has been proposed as a simple and potentially effective screening tool for early detection of NSCLC. Unfortunately, the miRNA signatures identified by numerous published studies of lung cancer patients are largely inconsistent, the reported miRNA profiles being incoherent [23, 26–33]. These studies have been the subject of several reviews and meta-analyses [34–40]. However, these reviews were not focused on circulating miRNAs in cancer stage I and II, potentially amenable to radical cure. Moreover, the accuracy of miRNA quantification in plasma/serum is known to be affected by several methodological variables, including modality of sample preparation, hemolysis, RNA isolation procedures, method of cDNA preparation and method used for miRNA measurement. These factors, that likely contribute to the puzzling inconsistency of the published miRNA profiles of NSCLC, were only partially addressed in the aforementioned reviews.

Here we aimed to review the literature in order to identify circulating miRNAs proven to be valuable and highly accurate for diagnosis of early NSCLC (stage I and II). Further, based on our analysis, we propose two panels of miRNAs for diagnosis of stage I-II NSCLC, with a two-tier screening method.

RESULTS

Included studies

Our literature search identified a total of 1712 articles, from which duplicates were removed, yielding 1239 papers. After reviewing titles, abstracts and full texts, 17 papers fulfilling our search criteria were finally included. Manual search of the bibliography of these papers led to include 3 additional records, yielding a total of 20 articles (Figure 1). Among these, 8 papers studied single miRNAs only, 6 explored both single miRNAs and panels, and 6 focused on miRNA panels only. For the 20 studies included in the review, Supplementary Table 1A indicates the main characteristics of patients and controls, and the investigated individual miRNAs or panels; Supplementary Table 1B provides information on methods used for miRNA quantification.

The selected studies, all published in the years 2011–2017, included 2119 individuals in total (1110 NSCLC patients and 1009 controls). The sample size ranged between 11 and 126 for NSCLC cohorts and between 11 and 110 for controls. The median sample size was 56 patients (interquartile range 30–79) with 1.1 case/control ratio. In all selected papers the sample mean age ranged 60–65 years, except in the study by Shi and colleagues that was carried out in a younger patient group (patient mean age, 50) [41].

Of the 20 studies (10 from China, 4 from USA, 2 from Italy and 1 each from Poland, Norway, Russia and France), 8 were conducted on Caucasian patients (2 studies included African American subjects), 7 on Asian patients and 5 did not provide information on ethnicity (Supplementary Table 1A).

The NSCLC patient groups differed by clinicopathological status across the studies and some relevant data were missing. Regarding the patients' smoking status and comorbidities, four studies did not report any data on smoking [41–44]; two studies included only smokers (with > 20 mean pack-years) [26, 30]; five studies included ≥ 85% of smokers among NSCLC patients [27, 29, 45–47]. Fourteen of the 20 studies provided no information on comorbidity of the patient cohort; the other 6 studies indicated that patients had no history of other cancers (Supplementary Table 1A) [26, 48–52]. A mix of the two main subtypes of NSCLC, AC and SCC, was present in all the selected papers, however only in 11 studies the accuracy of the miRNA profile of NSCLC was separately evaluated for AC and SCC [26, 29, 30, 41, 44, 46, 48–50, 53, 54]. The composition of control groups was also varied (Supplementary Table 1A). Four

studies [30, 44, 53, 55] provided no medical information on the control group. In 4 studies, history of no tumor and negative chest imaging (X-rays or CT scan) were used to identify healthy controls [26, 46, 47, 51]. In the other 12 studies, individuals broadly defined "healthy subjects" or "non-neoplastic subjects" based on medical history, served as controls; 3 of these studies included patients with chronic obstructive pulmonary disease (COPD) [29, 45, 47] and 4 studies included controls with benign pulmonary nodules or non-cancerous lung disease [45, 48, 50, 52].

miRNA extraction

miRNAs were extracted from serum samples in 10 studies, from plasma samples in 9 studies, from whole blood in 1 study (Supplementary Table 1B). A training and a validation set were both described in 8 studies; of these, 4 reported two different procedures to quantify miRNAs in the training and validation sets. For miRNA extraction (Supplementary Table 1B), the mirVana PARIS RNA kit (Ambion, ThermoFisher) was used in 7 studies, the miRNeasy mini kit (Qiagen) in 3 studies, and in one study each the miRCURY RNA isolation kit (Exiqon), the RNA extraction kit (Applied Biosystems, AB), the NucleoSpin miRNA Plasma kit (Macherey-Nagel) were used. In the study by Yuxia et al. [42], RNA extraction was not performed, whereas 2 studies used phenol and guanidine isothiocyanate reagents only [27, 49]. Addition of spike-ins as a quality control step was reported in 5 papers (Supplementary Table 1B).

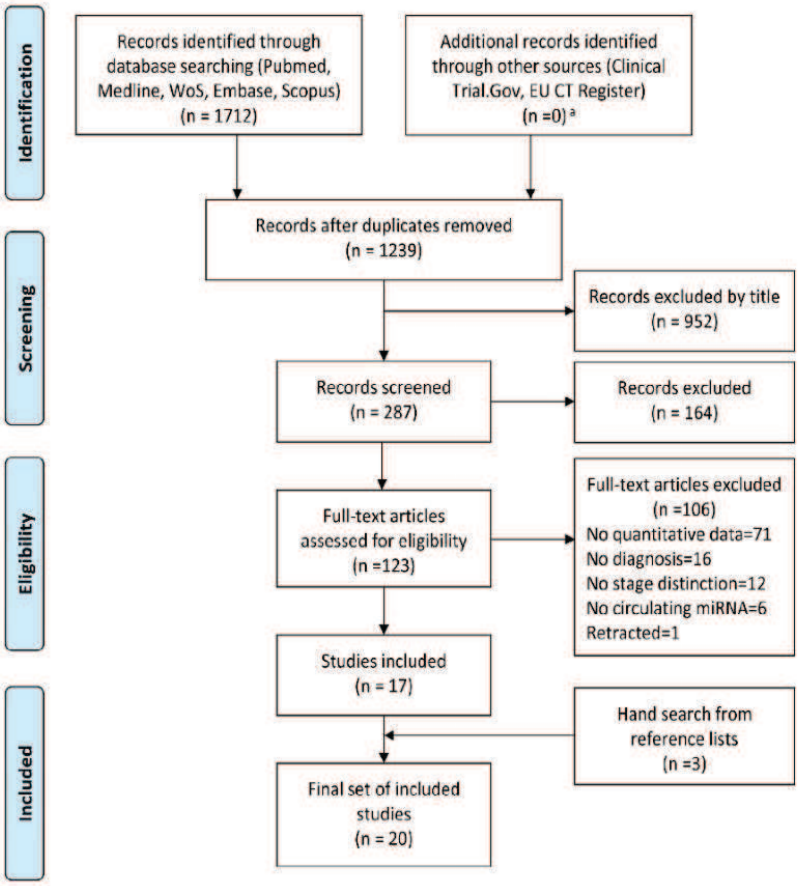


Figure 1: PRISMA flow diagram illustrating the study selection process. From the 1712 initially identified studies, duplicates were removed and records were screened by title, abstract, full text, leading to inclusion of 17 studies. Manual search of these papers led to inclusion of 3 further studies, for a total of 20 studies finally included in our review. [a]Among completed studies, no protocol satisfying the inclusion criteria was retrieved. Protocols in the recruiting stage were excluded.

miRNA retrotranscription and quantification

In 13 studies "Taqman" stem&loop primers and kits (AB) were used for retrotranscription, combined with two quantification methods [Taqman Low Density Arrays microRNA signature panel (TLDA, AB)] or another array -6 papers- and/or probe based relative quantitative PCR (qPCR) -12 papers-, as detailed in Supplementary Table 1B. Absolute miRNA quantification by Droplet Digital PCR (ddPCR) was performed in only 1 study, after retrotranscription with stem&loop primers [28]. In 3 studies, qPCR based on intercalating dyes was used for miRNA quantification. In 3 studies insufficient details of the procedures were provided [29, 41, 48]; (Supplementary Table 1B).

Normalization

As shown in Supplementary Table 1B, in the 6 studies using TLDA for quantification, the data were normalized with a geometric mean of different miRNAs, or with global normalization, or with quantile normalization. In the 14 studies using qPCR quantification, a single endogenous reference molecule (miR-16 or U6) was used in 7 studies; 4 studies used means of at least two endogenous reference genes; 2 studies used a single exogenous spike-in; 1 study did not provide details about normalization [42] (Supplementary Table 1B).

Individual miRNAs

In the 20 selected studies, altogether 27 miRNAs were individually reported (Table 1). Overall the studies were of medium quality as assessed by Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist, where "patient selection" and "index test" resulted the most critical domains. The diagnostic performance of individual miRNAs varied widely: sensitivity ranged between 30.4% and 96.1%, specificity between 38.2% and 100%; moreover, for some miRNAs (miR-223, miR-21, miR-145, miR-125b) the AUC differed substantially between independent studies (Table 1), suggesting that variability of patients, controls or methods may affect miRNA levels.

Identification of highly sensitive and highly specific miRNAs

Aiming to identify within Table 1 the individual miRNAs suitable for possible clinical application as non-invasive screening tool, we needed first to eliminate miRNAs influenced by hemolysis, a major source of bias. Therefore, we decided to exclude those miRNAs described as influenced by hemolysis in 3 or more of the relevant independent studies reporting hemolysis-induced miRNA dysregulation [56–63]. Accordingly, five miRNAs listed in Table 1 were excluded from possible clinical use: miR-486, miR-21, miR-21-5p, miR-126, miR-15b (see Supplementary Table 2). Among all miRNAs listed in Table 1 and considered unaffected by hemolysis, only miR-223, miR-20a, miR-448 and miR-145 displayed AUC value > 0.80 and sensitivity > 80% as stage I-II NSCLC biomarkers in at least one study. Moreover, miR-628-3p, miR-29c, miR-210 and miR-1244, despite low sensitivity and modest AUC value, showed specificity > 90% in at least one study. For these two sets of miRNAs, highly sensitive and highly specific respectively, the diagnostic accuracy data are summarized in Table 2. The seven studies analyzing these miRNAs [44, 46, 48, 50–53] are of medium overall quality according to the QUADAS-2 checklist; they included 649 NSCLC patients (median, 87 patients per study), predominantly smokers (62%) and represent 58% of the 1110 patients evaluated overall in the 20 selected studies. The miRNAs shown in Table 2 have important biological functions related to tumorigenesis; although the analysis of these functions is beyond the purpose of this review, they are briefly described in Supplementary File 1.

miRNA panels

In the 20 selected studies, 12 miRNA panels featuring high sensitivity (> 80%) and/or high AUC (> 0.80) as stage I-II NSCLC biomarkers were reported (Table 3). Five of these panels showed AUC> 0.90 and seven had AUC between 0.80 and 0.90; however some panels included miRNAs documented to be influenced by hemolysis (Table 3).

Table 4 illustrates miRNAs that were described (either individually or within miRNA panels) as biomarkers of stage I-II NSCLC in more than one of the selected studies.

miRNAs and NSCLC subtypes

None of the 20 studies included in this review separately evaluated miRNA signatures in SCC or AC. However, 11 of the 20 selected studies evaluated the performance of circulating miRNAs in distinguishing NSCLC subtypes, and the investigated individual miRNAs and miRNA panels completely differed across the studies. As summarized in Table 5, the proposed miRNA signatures revealed: greater accuracy in identifying SCC than AC in 7 studies [26, 29, 41, 46, 48, 49, 53]; similar accuracy in 2 studies [44, 54]; higher sensitivity for diagnosing AC in one study [30]. Wang et al. [50] reported that circulating levels of miR-425-3p and miR-628-3p were significantly higher in AC than SCC, while miR-532 was significantly lower in AC than SCC (Table 5). The AUCs for the miRNAs proposed as subtype-specific biomarkers were reported only in 4 studies [26, 48, 53, 54]. Altogether these findings show no consistent alterations of circulating miRNAs that may more accurately identify AC or SCC.

Table 1: Sensitivity, specificity and AUC of the 27 individual miRNAs described in the selected studies

miRNA	Reference	N	Sensitivity (%)	Specificity (%)	AUC
223	Geng et al., 2014 [48]	186	**87.0**	86.0	**0.94**
	Zhang et al, 2017 [51]	172	69.8	84.3	0.81
486*	Li et al., 2015 [43]	22	91.0	82.0	0.93
20a	Geng et al., 2014 [48]	186	**83.0**	81.0	**0.89**
	Zhang et al, 2017 [51]	172	79.8	88.0	089
448	Powrozek et al., 2016 [46]	114	**85.0**	77.0	**0.89**
	Sun et al., 2016 [55]	82	--	--	0.88
21*	Zhang et al, 2017 [51]	172	77.5	85.5	0.84
	Geng et al., 2014 [48]	186	67.0	68.0	0.77
21-5p*	Ma et al., 2013 [28]	74	--	--	0.79
145	Zhang et al, 2017 [51]	172	**80.6**	89.2	**0.89**
	Geng et al., 2014 [48]	186	70.0	68.0	0.77
141	Nadal et al., 2015 [45]	135	--	--	0.88
193b	Nadal et al., 2015 [45]	135	--	--	0.86
200b	Nadal et al., 2015 [45]	135	--	--	0.85
126*	Zhu et al., 2016 [52]	127	62.1	97.5	0.85
301	Nadal et al., 2015 [45]	135	--	--	0.84
328	Ulivi et al., 2013 [49]	78	--	--	0.82
4478	Powrozek et al., 2016 [46]	114	75.0	68.4	0.82
125b	Shi et al., 2017 [41]	210	30.4	83.9[§]	0.81[§]
	Yuxia et al., 2012 [42]	186	96.1	38.2	0.66
1244	Wang W et al., 2016 [53]	69	53.8	**100.0**	0.80
182	Zhu et al., 2016 [52]	127	67.8	85.0	0.78
425-3p	Wang Y et al., 2016 [50]	173	67.1	68.1	0.73
628-3p	Wang Y et al., 2016 [50]	173	42.7	**91.2**	0.73
29c	Zhu et al., 2014 [44]	84	50.0	**95.8**	0.73
429	Zhu et al., 2014 [44]	84	94.4	41.7	0.72
22	Shi et al., 2017 [41]	210	43.5	86.3[§]	0.72[§]
335-3p	Ma et al., 2013 [28]	74	--	--	0.71
532	Wang Y et al., 2016 [50]	173	53.7	80.2	0.66
210	Zhu et al., 2016 [52]	127	35.6	**100.0**	0.65
183	Zhu et al., 2016 [52]	127	41.4	82.5	0.64
15b*	Shi et al., 2017 [41]	210	41.3	82.4[§]	0.62[§]

N: sample size.

*asterisk denotes miRNA influenced by hemolysis; miRNA was considered hemolysis-influenced when documented in three or more relevant independent studies reporting hemolysis-induced miRNA dysregulation (see Supplementary Table 2).

miRNAs shown in **bold** are those uninfluenced by hemolysis (see Additional Table 2) and with high sensitivity (> 80%) and high AUC (> 0.80) reported in at least one study shown in Reference column.

miRNAs shown in **bold underlined** are those uninfluenced by hemolysis (see Additional Table 2) and with high specificity (> 90%) reported in at least one study shown in Reference column.

[§]Data are calculated among the total sample, that includes advanced stages (III and IV).

Table 2: miRNAs with high sensitivity and high AUC (a), and miRNAs with high specificity (b).

a) miRNAs with sensitivity > 80% and AUC > 0.80

miRNA	Sensitivity (%)	AUC	Specificity (%)	Reference
miR-223	87.0	0.94	86.0	Geng et al., 2014 [48]
miR-20a	83.0	0.89	81.0	Geng et al., 2014 [48]
miR-448	85.0	0.89	77.0	Powrozek et al., 2016 [46]
miR-145	80.6	0.89	89.2	Zhang et al., 2017 [51]

b) miRNAs with specificity > 90%

miRNA	Sensitivity (%)	AUC	Specificity (%)	Reference
miR-628-3p	42.7	0.73	91.2	Wang Y et al., 2016 [50]
miR-29c	50.0	0.73	95.8	Zhu et al., 2014 [44]
miR-210	35.6	0.65	100.0	Zhu et al., 2016 [52]
miR-1244	53.8	0.80	100.0	Wang W et al., 2016 [53]

Proposal of a two-step screening with miRNAs

Based on the critique of the reviewed studies, we propose a model for screening of stage I-II NSCLC, using the two above indicated sets of individual miRNAs with the highest sensitivity/specificity (Table 2), that were selected as detailed in the Methods section. Panels of miRNAs were arbitrarily excluded from this model as the AUC or specificity data were based on the panels, and to simplify its possible clinical application. Accordingly, screening with miRNAs should be carried out in two steps. The four miRNAs with high sensitivity (miR-223, miR-20a, miR-448 and miR-145) should be used for the first screening step (Test 1), and the four miRNAs with high specificity (miR-628-3p, miR-29c, miR-210 and miR-1224) for the second step (Test 2). In this model the two panels of miRNAs are combined in series, and Test 2 is run only if Test 1 is positive, as described in Supplementary File 2.

The final estimated performance of these miRNAs for the two-step screening of serum samples is overall sensitivity of 91.6% and overall specificity of 93.4%. The selected two sets of miRNAs with highest sensitivity/specificity are intended for preliminary screening of the general population at high risk of lung cancer, dominated by smokers. Subjects positive to miRNA screening should be offered low-dose CT-screening, thus possibly reducing the logistic/economic burden and harms of upfront CT-screening [12,14].

DISCUSSION

Diagnosing lung cancer at an early stage is a major clinical concern that in recent years has stimulated extensive research on non-invasive screening methods, including miRNAs as lung cancer biomarkers in circulating body fluids. Dysregulated miRNA profiles in cell-free blood were shown to indicate the presence of lung cancer many months ahead of the occurrence of symptoms [64], and even before the disease was detected by CT screening [26, 65]. Therefore, miRNAs are potentially interesting biomarkers for screening of lung cancer [66]. According to "Medline Trend", the number of publications on the topic "miRNA and NSCLC" has dramatically increased in the last 10 years, however the plethora of circulating miRNA profiles proposed as lung cancer signatures are inconsistent. Six systematic reviews [34–37, 39, 40] have summarized the main findings of these studies, but have failed to clearly identify circulating miRNAs possessing high proficiency specifically for the diagnosis of stage I-II NSCLCs, which are the cancers potentially amenable to radical cure. Considering that miRNA signatures of early and late lung cancer stages frequently differ [50, 64, 67–69], we exclusively reviewed papers reporting miRNAs biomarkers of stage I-II NSCLC. We focused on miRNA molecules of high diagnostic accuracy and whose measurement is scarcely influenced by hemolysis. Among the initially retrieved 1712 papers fulfilling the search criteria, we only found 20 studies clearly reporting quantitative data on miRNA diagnostic proficiency specifically for stage I-II NSCLC. Our review confirms the variability of miRNAs proposed by many authors as lung cancer signatures. Notably, there were only 18 miRNAs identified as biomarkers of stage I-II NSCLC in more than one published paper (Table 4).

For the 20 selected studies we highlighted demographics, clinicopathological characteristics and smoking habit of patients and controls; moreover, we evaluated the main pre-analytical and analytical variables known to influence circulating miRNA levels. Notably, the training set in the selected studies consisted of a median of only 56 NSCLC patients, meaning that the training sample frequently was of smaller size than that suggested by guidelines for studies of biomarkers for early detection of cancer [70]. Moreover, in some of the selected studies the training sample appears definitely undersized if one considers the rather low precision of miRNA assays [63] and the expected diversity of miRNA signatures due to

Table 3: Sensitivity (Se), specificity (Sp) and AUC of miRNA panels described in the selected studies

	miRNA Panel	Reference	*N*	Se(%)	Sp(%)	AUC
1	miR-141, miR-200b, miR-193b, miR-301	Nadal et al., 2015 [45]	135	N.R.	N.R.	0,99
2[1]	24 miRNAs*	Wozniak et al., 2015 [54]	121	N.R.	N.R.	0,98
2 bis[2]	24 miRNAs*	Wozniak et al., 2015 [54]	149	N.R.	N.R.	0,96
3	miR-182, miR-183, **miR-210**, *miR-126*, CEA	Zhu et al., 2016 [52]	127	88,5	92,5	0,98
4	miR-532, **miR-628-3p**, miR-425-3p	Wang Y. et al. 2016 [50]	173	91,5	97,8	0,97
5	**miR-448**, miR-4478	Powrozek et al., 2016 [46]	114	90	76.3	0,90
6	**miR-145**, **miR-20a**, *miR-21*, miR-223	Zhang et al., 2017 [51]	172	81,8	90,1	0,90
7	34 miRNAs**	Bianchi et al., 2011 [26]	52	59	90	0,89
8	miR-125b, miR-200b, miR-34b, miR-203, miR-205, miR-429	Halvorsen et al., 2016 [47]	158	85	74	0,88
9	*miR-21-5p*, miR-335-3p	Ma et al., 2013 [28]	74	N.R.	N.R.	0,86
10[1]	12 miRNAs***	Sanfiorenzo et al., 2013 [29]	33	N.R.	N.R.	0,85
10 bis[2]	12 miRNAs***	Sanfiorenzo et al., 2013 [29]	42	N.R.	N.R.	0,81
11	miR-1254, miR-574-5p	Foss et al., 2011 [27]	53	73	71	0,75
12[2]	*mir-21*, *miR-126*, **miR-210**, *miR-486-5p*	Shen et al., 2011 [30]	44	73,3	96,5	N.R.
12 bis[1]	*mir-21*, *miR-126*, **miR-210**, *miR-486-5p*	Shen et al., 2011 [30]	44	86,7	96,5	N.R.

N: sample size.
N.R.: Not Reported.
Panels are listed in decreasing order of AUC value. The studies by Wozniak et al., Sanfiorenzo et al., Shen et al. [54, 29, 30], separately described findings in stage I[2] and stage II[1] non-small cell lung cancer.
*let-7c, miR-122, miR-182, miR193a-5p, miR-200c, miR-203, miR-218, miR-155, let-7b, miR-411, miR-450b-5p, miR-485-3p, miR-519a, miR-642, miR-517b, miR-520f, miR-206, miR-566, miR-661, miR-340, miR-1243, miR-720, miR-543, miR-1267.
**_miR-92a_, miR-484, _miR-486-5p_, miR-328, miR-191, miR-376a, miR-342-3p, miR-331-3p, miR-30c, miR-28-5p, miR-98, _miR-17_, miR-26b, miR-374a, miR-30b, miR-26a, miR-142-3p, _miR-103_, _miR-126_, let-7a, let-7d, let-7b, miR-32, miR-133b, miR-566, mir-432, miR-223, _miR-29a_, miR-148a, miR-142-5p, miR-22, miR-148b, miR-140-5p, miR-139-5p.
***miR-155-5p, **miR-20a-5p**, miR-25-3p, miR-296-5p, miR-191-5p, _miR-126-3p_, miR-223-3p, miR-152-3p, **miR-145-5p**, miR-199a-5p, _miR-24-3p_, and let-7f-5p.
miRNAs influenced by hemolysis are indicated in *italics*; miRNA was considered hemolysis-influenced if influence was documented in three or more relevant independent studies reporting hemolysis-induced miRNA dysregulation (see Supplementary Table 2).
miRNAs indicated in **bold** are also accurate individual predictors, included in Table 2.

molecular heterogeneity of NSCLC subtypes [36, 71–74]. Population ethnicity has been suggested as a potential source of miRNA level variability and of inconsistent miRNA signatures of lung cancer [77]. However, Shen et al. found no association between changes in circulating miRNA levels and patient ethnic group (African-American or Caucasian) [43]. As regards gender and age, several papers have documented that these variables do not significantly impact on lung cancer miRNA signatures [30, 44, 50, 75, 76].

The proportion of smokers varied among the studies and sometimes markedly differed between cases and controls within the same study. In 5 of the 20 papers, cigarette smoking data were not reported, an important lack of information because some circulating miRNAs are significantly dysregulated by smoking [34, 77, 78].

Across the 20 studies, the control groups were also very different. The majority of studies defined the control group as "healthy subjects" not otherwise specified, or "non-neoplastic subjects" based on medical history. Notably, the composition of control groups is a critical issue, because diseases of liver, heart, prostate and various other comorbidities in the control group may influence the diagnostic sensitivity and specificity of miRNA candidate

Table 4: miRNAs indicated as stage I-II NSCLC biomarkers in more than one of the selected studies

miRNA	Number of studies[a]	References
miR-21	4	Sun et al., 2016 [55]; Zhang et al., 2017 [51]; Geng et al., 2014 [48]; Shen et al., 2011 [30]*
miR-223	4	Zhang et al., 2017 [51]; Geng et al., 2014 [48]; Sanfiorenzo et al., 2013 [29]*; Bianchi et al., 2011 [26]*
miR-126	4	Zhu et al., 2016 [52]; Sanfiorenzo et al., 2013 [29]*; Shen et al., 2011 [30]*; Bianchi et al., 2011 [26]*
miR-20a	3	Zhang et al., 2017 [51]; Geng et al., 2014 [48]; Sanfiorenzo et al., 2013 [29]*
miR-145	3	Zhang et al., 2017 [51]; Geng et al., 2014 [48]; Sanfiorenzo et al., 2013 [29]*
miR-125b	3	Shi et al., 2017 [41]; Halvorsen et al., 2016 [47]*; Yuxia et al., 2012 [52]
miR-486	3	Li et al., 2015 [77]; Shen et al., 2011 [30]*; Bianchi et al., 2011 [26]*
miR-155	3	Geng et al., 2014 [48]; Wozniak et al., 2015 [54]*; Sanfiorenzo et al., 2013 [29]*
miR-200b	2	Halvorsen et al., 2016 [47]*; Nadal et al., 2015 [45]*
miR-328	2	Ulivi et al., 2013 [49]; Bianchi et al., 2011 [26]*
miR-182	2	Zhu et al., 2016 [52]; Wozniak et al., 2015 [54]*
miR-429	2	Halvorsen et al., 2016 [47]*; Zhu et al., 2014 [44]
miR-210	2	Zhu et al., 2016 [52]; Shen et al., 2011 [30]*
miR-22	2	Shi et al., 2017 [41]; Bianchi et al., 2011 [26]*
miR-203	2	Halvorsen et al., 2016 [47]*; Wozniak et al., 2015 [54]*
let-7b	2	Wozniak et al., 2015 [54]*; Bianchi et al., 2011 [26]*
miR-566	2	Wozniak et al., 2015 [54]*; Bianchi et al., 2011 [26]*
miR-191	2	Sanfiorenzo et al., 2013 [29]*; Bianchi et al., 2011 [26]*

[a]number of reviewed studies indicating the specified miRNA as stage I-II NSCLC biomarker.
*Asterisk denotes panel including the specified miRNA.
miRNAs indicated in **bold** are included in our two-step screening (Table 2).

biomarkers [79–81]. Only 4 of 20 studies subdivided control patients by comorbidity: benign lung nodules, COPD, noncancerous disease, smoker [29, 48, 50, 52]. It is debated if "controls" for lung cancer patients should be age-matched "healthy subjects" or subjects with a history of smoking, and whether COPD patients should be included as controls. Because levels of miRNA relevant for lung cancer may be altered in smokers [34, 77] and in COPD patients [82–84], the control group composition in terms of smoking pack/years and COPD prevalence may bias the accuracy of miRNAs selected as lung cancer biomarkers. In order to avoid a COPD-based miRNA signature, in the studies by Sanfiorenzo et al. [29] and Halvorsen et al. [47], non-neoplastic COPD patients were used as controls.

Elegant experimental studies have shown that miRNAs derived from cancer tissue can enter the circulation [24]. Moreover, in lung cancer patients several overexpressed circulating miRNAs (miR-21, miR-24, miR-145, miR-20a, miR-223, miR-486, miR-574-5p, miR-1825, miR-205, miR-19a, miR-19-b, miR-30b) were generally reduced a few days after tumor resection, strongly suggesting that these molecules are of tumor origin or tumor-induced [43, 51, 75, 85, 86]. It is therefore reasonable to assume that at least some of the aberrantly expressed miRNAs in the blood of lung cancer patients are genuine biomarkers of the tumor. The measurement of circulating miRNAs faces numerous technical challenges and may be biased by multiple factors, partly explaining the inconsistency of published miRNA profiles of lung cancer [51]. Because several circulating miRNAs are blood-cell derived [56, 59–62, 87], spurious miRNA level dysregulations that may result from platelet contamination and red blood cell lysis in plasma/serum samples are a major concern. Our review suggests that potential bias of hemolysis on miRNA levels has often been underestimated, as only in 2 of the 20 reviewed studies was hemolysis of specimens ruled out [29, 54]. In order to avoid spurious effects of undetected hemolysis of samples, in agreement with Pritchard and collaborators [88], we suggest that miRNAs influenced by hemolysis should preferably not be used as NSCLC biomarkers.

For analysis of circulating miRNAs, both serum and plasma are acceptable sample types, and a good correlation between serum and plasma miRNA determinations has been documented [89]. However, serum and plasma

Table 5: Characteristics of the 11 included studies evaluating the performance of circulating miRNAs in distinguishing NSCLC subtypes

Reference	Year	Sample ethnicity	Sample Sizes		NSCLC[d] Stage	miRNAs examined	AUC[e] in discriminating NSCLC subtype from controls for the examined miRNAs		Comments on miRNA performance
			Pt[b]	C[c]			AC[f] (n)	SCC[g] (n)	
Bianchi et al. [26]	2011	Caucasian	22[a]	30	I	Panel of 34 miRNAs*	(n = 22) 0.85	(n = 12) 0.94[h]	The panel distinguished better SCCs than ACs from controls; however, the SCCs were stage II-IV cases. Sample size was small.
Geng et al. [48]	2014	Asian	126	60	I-II	5 miRNAs: miR-20a miR-223 miR-21 miR-155 miR-145	(n = 45) 0.90 0.91 0.63 0.93 0.77	(n = 64) 0.98 0.98 0.97 0.96 0.97	All 5 miRNAs differentiated NSCLC from controls with greater accuracy in SCCs. 17 cases had histology other than AC or SCC.
Powrozek et al. [46]	2016	Caucasian	29[a]	85	I-II	2 individual miRNAs: miR-448, miR-4478; combination of both miRNAs	(n = 30) NA[i]	(n = 35) NA	Both miRNAs overexpressed in NSCLC plasma samples relative to control. miR-4478 expression was higher in SCC than in AC patients ($p < 0.043$). Analysis of miRNA performance included 36 cases with stage >II.
Sanfiorenzo et al. [29]	2013	NA	35[a]	20	I-II	Panel of 12 miRNAs**	(n = 27) NA	(n = 25) NA	Panel distinguished NSCLC patients from controls (AUC=0.81). In SCC compared to AC, higher plasma levels of miR-20a-5p ($p = 0.034$) and miR-25-3p ($p = 0.013$), along with lower levels of miR-191-5p ($p = 0.008$) were found. Analysis of miRNA performance included 17 cases with stage >II.
Shen et al. [30]	2011	African American, Caucasian	30[a]	29	I-II	miR-21, miR-126, miR-210, miR-486-5p	(n = 24) NA	(n = 34) NA	Diagnostic sensitivity of the composite panel in distinguishing stage I NSCLC from controls was 73.3%. Analysis of miRNA performance in diagnosing subtypes included 28 cases with stage >II and showed higher sensitivity for diagnosing ACs (91.7%) than SCCs (82.3%) ($p < 0.05$).
Shi et al. [41]	2017	NA	46[a]	45	I-II	miR-22, miR-125b, miR-15b	(n = 69) NA	(n = 51) NA	Serum levels of the three miRNAs significantly altered in NSCLC cases compared to controls. Diagnostic sensitivity of miR-125b was significantly higher for ACs than SCCs ($p = 0.021$). Analysis of miRNA performance included 74 cases with stage >II.
Ulivi et al. [49]	2013	Caucasian	54[a]	24	I-II	miR-328	(n = 63) NA	(n = 22) NA	miR-328 discriminated well between stage I-II NSCLC and controls (AUC = 0.82). Analysis of miRNA performance for subtypes, conducted in 86 NSCLCs (63 ACs; 22 SCC; 1 sarcomatoid), 32 of which were in stage > II, indicated significantly higher expression of miR-423 in SCCs than in ACs. The miRNA analyses were performed in whole blood specimens.
Wang Y. et al. [50]	2016	Asian	82	91	I-II	miR-532, miR-628, miR-425-3p	(n = 40) NA	(n = 39) NA	Combination of the three miRNAs discriminated well NSCLC from control plasma samples (AUC = 0.97). Evaluation of miRNA performance was conducted in 40 ACs and 39 SCCs. Plasma levels of miR-425-3p ($p = 0.04$) and of miR-628-3p ($p = 0.015$) were significantly higher in AC than SCC. miR-532 was significantly lower in AC than SCC ($p < 0.001$).
Wang W. et al. [53]	2016	NA	54[a]	15	I-II	miR-1244	(n = 26) 0.79	(n = 17) 0.85	For miR-1244 the AUC was higher in SCC than AC. AUC was assessed on serum samples of 43 NSCLCs (26 ACs; 17 SCCs), 17 of which were in stage > II.
Wozniak et al. [54]	2015	Caucasian	70[a]	100	I-II	Panel of 24 miRNAs***	(n, NA) 0.94	(n, NA) 0.96	Panel showed similar accuracy for distinguishing AC and SCC from controls. AUC for the panel was assessed in 70 NSCLCs [a sub-cohort of 100 NSCLCs (35 ACs; 65 SCCs), 30 of which were in stage >II].
Zhu et al. [44]	2014	Asian	36[a]	48	I	miR-29c, miR-93, miR-429	(n =34) NA	(n = 36) NA	The evaluation of 70 NSCLCs (34 ACs; 36 SCC), 34 of which were in stage II–IV, showed non-significant difference of serum miR-29c ($p = 0.232$) and miR-429 ($p = 0.811$) between AC and SCC.

[a]The sample refers to stage I and II NSCLC.
[b]Pt=Patients; [c]C=Controls.
[d]NSCLC: non-small cell lung cancer.
[e]AUC=area under the curve.
[f]AC: adenocarcinoma.
[g]SCC=squamous cell carcinoma.
[h]SCCs were 12 additional cases with stage II-IV disease.
[i]Analysis of miRNA performance included also cases with stage > II.
[l]NA: Not available.
*miR-92a, miR-484, miR-486-5p, miR-328, miR-191, miR-376a, miR-342-3p, miR-331-3p, miR-30c, miR-28-5p, miR-98, miR-17, miR-26b, miR-374a, miR-30b, miR-26a, miR-142-3p, miR-103, miR-126, let-7a, let-7d, let-7b, miR-32, miR-133b, miR-566.
**miR-155-5p, miR-20a-5p, miR-25-3p, miR-296-5p, miR-191-5p, miR-126-3p, miR-223-3p, miR-152-3p, miR-145-5p, miR-199a-5p, miR-24-3p, and let-7f-5p.
***let-7c, miR-122, miR-182, miR-193a-5p, miR-200c, miR-203, miR218, miR-155, let-7b, miR-411, miR-450b-5p, miR-485-3p, miR-519a, miR-642, miR-517b, miR-520f, miR-206, miR-566, miR-661, miR-340, miR-1241, miR-720, miR-543, miR-1267.

determinations cannot be automatically interchanged, because differences in specimen preparation and/or measurement platform are known to influence the results. As an example, in normal subjects miR-15b and miR-16 showed higher concentrations in plasma relative to serum in one study [63], while the concentration of the same two miRNAs was higher in serum relative to plasma in another independent study using a different platform [90].

It is currently debated whether serum or plasma should be used for circulating miRNAs determination; mirroring this uncertainty, 10 of the 20 selected studies were performed with serum samples and 9 with plasma. Serum has not been generically recommended over plasma as a sample type [87]; however, serum has less platelet contamination than plasma, and this may decrease bias in miRNA determination [63]. Regarding the method for miRNA quantification, all the reviewed studies except that of Ma et al. [28] used qPCR platform and performed normalization of results predominantly with endogenous miR-16 and U6, or with spike-ins. The normalization step is likely to contribute to the scarce reproducibility of miRNA determinations, as reported by others [33]. Normalization with miR-16 can be criticized because this miRNA has been described as a lung cancer biomarker itself [91–93]. U6, a small nuclear RNA, was shown to fluctuate markedly across samples [94], and such variability may contribute to inconsistency of miRNA findings. For miRNA measurement Ma et al. [28] used ddPCR, a recently introduced technique reported to be advantageous over qPCR (greater precision; no need to normalize results; higher sensitivity to low-level miRNA expression) [28, 95]. Altogether, these considerations underscore the importance of knowing the miRNA quantification procedure details, to allow reproducibility of methods and external validation of studies.

Distinguishing between the AC and SCC lung cancer subtypes on the basis of specific circulating miRNAs' aberrant expression may provide important information, relevant both for understanding the subtypes' pathogenesis and for tailored selection of cytotoxic chemotherapy in NSCLC without a driver mutation [96]. Moreover, although histology and immunohistochemistry (IHC) currently are the gold standards for NSCLC diagnosis, the subtype classification of difficult cases (scarce biopsy sample; hazardous/difficult biopsy; uncertain IHC) could be facilitated if subtype-specific circulating miRNA signatures were available. Few studies have been conducted in stage I-II NSCLC patients to identify circulating miRNA profiles than may be more accurate for either AC or SCC. In our systematic review we only found 11 studies that separately analyzed AC and SCC cases [26, 29, 30, 41, 44, 46, 48–50, 53, 54], and none of these provided convincing evidence that a specific miRNA signature exists for each of the two subtypes. A notable methodological weakness in 9 of these 11 studies [26, 29, 30, 41, 44, 46, 48–50, 53, 54] is the inclusion of many lung cancers in advanced stage (stage > II) in the subtype analysis, likely to compensate for small sample size of the AC and SCC sub-cohorts. Of note, Bianchi et al. [26], Geng et al. [48], Powrozek et al. [46], Sanfiorenzo et al. [29], Shi et al. [41], Ulivi et al. [49] and Wang et al. [53] proposed very different miRNA signatures of early stage NSCLC, yet all these signatures better differentiated SCC than AC from controls. Altogether, the available data are insufficient to define serum/plasma miRNA profiles that may reliably discriminate between AC and SCC in stage I-II lung cancer.

Strengths and limitations

A strength of this review is the critique focused on circulating miRNA biomarkers of stage I-II NSCLC, the disease stages often amenable to radical cure and for which non-invasive screening by miRNAs may be proposed. Another strength is the assessment of factors potentially influencing miRNA levels and the evaluation of miRNAs' accuracy as stage I-II NSCLC biomarkers by quantitative data (sensitivity, specificity, AUC).

This review has important limitations. First, miRNA signatures of NSCLC may be biased by pre-analytical and analytical factors, and by clinicopathological features of patients and controls. Second, in many of the selected papers the validation sample was relatively small (median, 56 patients), with limited power to correctly identify the miRNA signature of stage I-II NSCLC. Third, lack of methodological details in some studies prevented thorough evaluation of the quality of methodology used. Data on comorbidities, some of which may affect miRNA expression [79–81], were not provided in some papers. Accordingly, at QUADAS-2, "patient selection" and "index test" resulted the most critical domains, and overall the studies were only of medium quality. Another limitation is that the miRNA panels for our two-step model of screening were obtained from studies where only the majority of lung cancer patients (62%) were smokers, while screening for lung cancer is currently recommended exclusively in smokers (11). In this review, we aimed to identify circulating individual miRNAs with sensitivity > 80% and AUC > 0.80 as biomarkers of stage I-II NSCLC, for possible clinical application as non-invasive screening tool. Based on the reviewed studies, we found four individual miRNAs that fulfilled these criteria: miR-223, miR-20a, miR-448 and miR-145; four other miRNAs showed very high specificity (> 90%): miR-628-3p, miR-29c, miR-210 and miR-1244. Among factors potentially affecting circulating miRNAs, the only two that were considered for miRNAs selection were the stage of NSCLC (all studies were stages I-II) and the impact of hemolysis (miRNAs potentially affected by hemolysis were excluded). Other factors, such as smoking habits, age, ethnicity, methodological issues of RNA extraction, could not be controlled because they varied widely among the selected studies.

Screening for lung cancer with circulating miRNAs, preliminary to CT-screening, is a minimally invasive and safe blood test that may offer several advantages over upfront CT-screening: reduction of number of CT-screens (to be performed only in miRNA screening-positive individuals) and of radiation risk; decrease of false-positive CT-screening rate and consequent reduction of complications and costs from futile lung biopsies (12,14). We have proposed a two-step model of miRNA screening for stage I-II NSCLC, based on the measurement of the serum level of the above indicated selected miRNAs (Table 2): the panel of four miRNAs with high sensitivity should be used for the first screening step, and the panel with high specificity for the second step. Based on our model, for the two miRNA panels combined in series for screening of serum samples the estimated performance is overall sensitivity of 91.6% and overall specificity of 93.4%. The estimated diagnostic accuracy of the proposed model is similar to that of the 12 miRNA panels found in the selected papers (Table 3), most of which featured high sensitivity (> 80%) and/or high AUC (> 0.80) as stage I-II NSCLC biomarkers. However, several of these panels contain miRNAs that are not ideal biomarkers; as Table 3 shows, 6 of the 12 panels included miRNAs influenced by hemolysis. Moreover, Nadal et al. and Wozniak et al. provided no sensitivity nor specificity data for their panels [45, 54] and the panels tested by Powrozek et al., Halvorsen et al., and Foss et al. showed modest specificity (76.3%, 74% and 71%, respectively) [27, 46, 47]. The panel proposed by Wang et al. featured high proficiency in diagnosing AC, without containing hemolysis-influenced miRNAs [50].

MATERIALS AND METHODS

Search strategy

A systematic review of the scientific literature was conducted using the following key words: [(NSCLC OR Non Small Cell Lung Cancer) AND (lung cancer) AND (miRNA OR MicroRNA) AND (diagnosis)]" on the search engines of the databases "Pubmed", "Medline", "Scopus", "Embase" and "WOS". The research was first performed on July 21st 2016 and results were regularly updated until April 12th 2017. Including criteria were: i) circulating miRNAs; ii) histologically/cytologically defined NSCLC stage I and/or II (studies of patients with NSCLC at any stage were included only if a sub-analysis for stage I-II was provided); iii) studies reporting quantitative data on the efficacy of specific miRNAs as tools for stage I-II NSCLC screening (sensitivity, specificity and/or AUC); iv) English language. Studies analyzing single miRNAs and/or panels of miRNAs were included. Duplicate publications were eliminated through the Mendeley software [97]. All articles of interest were then evaluated and screened for eligibility by two researchers, independently, and controversies were resolved by consensus. Bibliography of the selected papers was manually examined to retrieve further articles with eligibility criteria.

The protocol was registered at the international prospective register of systematic reviews (PROSPERO, ID: CRD42017056943). The PRISMA statement and the Cochrane Handbook for Diagnostic Test Accuracy Reviews were followed as reference protocol standards.

Data extraction

From the eligible studies the following information was collected: a) author name, year and country where the study was performed; b) sociodemographic and clinical information on population under study (ethnicity, sample size, age, smoking status, comorbidity, NSCLC stage); c) individual miRNAs and/or miRNA panels under study; d) methodological issues regarding miRNAs extraction [type of specimen (plasma/serum/whole blood), hemolysis assessment, RNA isolation and measurements procedures]; e) quantitative data of diagnostic accuracy (sensitivity, specificity, AUC) for stage I-II NSCLC.

The papers then underwent rigorous critical evaluation, taking into account: i) quality of the study, assessed by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist [98]; ii) factors identified as potentially affecting miRNA quantification (Table 6). Two investigators independently assessed the seven domains of the QUADAS-2. Any discrepancies were resolved through discussion.

Selection of circulating miRNAs for a two-step screening preliminary to CT-screening

Within studies with overall satisfactory quality by QUADAS-2, we identified individual miRNAs showing at least in one study high diagnostic proficiency as stage I-II NSCLC biomarkers (arbitrarily stated as sensitivity > 80% and AUC > 0.80, or specificity > 90%) and scarcely influenced by hemolysis according to the pertinent literature [56–63]. Altogether eight individual miRNAs revealed the aforementioned high diagnostic proficiency as stage I-II NSCLC biomarkers (miR-223, miR-20a, miR-448, miR-145, miR-628-3p, miR-29c, miR-210 and miR-1244; Table 2). These miRNAs with the highest sensitivity/specificity can be applied in a mathematical model, that we are here proposing, to estimate their overall sensitivity and specificity for stage I-II NSCLC screening. The model consists of a two-step screening test, first using the panel of selected circulating miRNAs with high sensitivity and high AUC, then the panel of selected miRNAs with high specificity, as illustrated in Supplementary File 2. We arbitrarily excluded miRNA panels from the model since the AUC or specificity data were based on the panels and not individual miRNAs, aiming to simplify possible clinical application of the

Table 6: Factors potentially affecting circulating miRNA quantification in NSCLC patients

Clinicopathological factors	Ethnicity
	Gender/age
	Smoking status
	Stage of disease (early/advanced)
Methodological factors	Type of sample (plasma/serum/whole blood)
	Hemolysis
	RNA extraction method
	Reverse transcription method
	miRNA quantification method
	Normalization

NSCLC: non-small cell lung cancer.

test. However, for comparison of our model's diagnostic accuracy, the other miRNA panels included in the review are discussed.

Statistical analysis

The proposed two-step model for estimating overall sensitivity and specificity of circulating miRNAs to be used for stage I-II NSCLC screening was developed using the formulas described in Supplementary File 2.

CONCLUSIONS

Several pre-analytical and analytical variables of circulating miRNA measurements, especially hemolysis of samples, may bias the accuracy of miRNAs as biomarkers of stage I-II NSCLC. Evidence-based data are insufficient to reach a robust conclusion as to which circulating miRNAs are the best biomarkers of early lung cancer, and also insufficient to define serum/plasma miRNA profiles that may reliably discriminate between AC and SCC.

Nevertheless, based on critical review of the literature, selected circulating miRNAs that are scarcely influenced by hemolysis could be tested for screening early lung cancer in smokers and former smokers. For our theoretical model of two-step screening for stage I-II NSCLC, first using a panel of miRNAs with high sensitivity and then a panel with high specificity, we estimated overall sensitivity of 91.6% and overall specificity of 93.4%. The circulating miRNAs we selected as potentially valuable biomarkers of early lung cancer based on this review, as well as those described by other authors, require validation in multiple independent studies before they can be proposed for clinical application.

Abbreviations

AB, Applied Biosystems; AC, adenocarcinoma; AUC, area under the curve; COPD, chronic obstructive pulmonary disease; CT, computed tomography; ddPCR, droplet digital PCR; IHC, immunohistochemistry; miRNA, microRNA; NSCLC, Non-small cell lung cancer; qPCR, quantitative (realtime) PCR; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; SCC, squamous cell carcinoma; TLDA, Taqman Low Density Arrays.

Author contributions

LD, AP, FM and PC were responsible for study conception; FM, PDA, LD and PC contributed to data collection, data extraction, preparation of tables and drafting the manuscript; PDA, EG, EF, MB, AI, MC and NR contributed to reviewing the literature and data extraction; EF, FM and AP contributed to data analysis; LD, DMN and AP critically reviewed and expanded the manuscript.

All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

PDA and MC are PhD students of the "Biotechnology, Biosciences and Surgical Technology" course at Università degli Studi dell'Insubria.

We are extremely grateful to Mr. Piero Francesco Macchi and Mrs. Carlotta Biasini for their generous donation to support this study.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This work was supported by grants from Fondazione Comunitaria del Varesotto (LD); Associazione PREDICA Onlus (LD); PRIN 2010NECHBX_003; AIRC (Associazione Italiana per la Ricerca sul Cancro IG15895)

(DMN). Co-funded by donation from Mr. Piero Francesco Macchi and Mrs. Carlotta Biasini (LD).

All these funding bodies had no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136:E359–86. https://doi.org/10.1002/ijc.29210.
2. van Meerbeeck JP, Fennell DA, De Ruysscher DK. Small-cell lung cancer. Lancet. 2011; 378:1741–55. https://doi.org/10.1016/S0140-6736(11)60165-7.
3. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013; 49:1374–403. https://doi.org/10.1016/j.ejca.2012.12.027.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7–30. https://doi.org/10.3322/caac.21332.
5. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016; 893:1–19. https://doi.org/10.1007/978-3-319-24223-1_1.
6. Koike T, Koike T, Yoshiya K, Tsuchida M, Toyabe S. Risk factor analysis of locoregional recurrence after sublobar resection in patients with clinical stage IA non-small cell lung cancer. J Thorac Cardiovasc Surg. 2013; 146:372–8. https://doi.org/10.1016/j.jtcvs.2013.02.057.
7. Goldstraw P. Surgical oncologic principles. Chest Surg Clin N Am. 2001; 11:1–16, vii.
8. Imperatori A, Harrison RN, Dominioni L, Leitch N, Nardecchia E, Jeebun V, Brown J, Altieri E, Castiglioni M, Cattoni M, Rotolo N. Resection rate of lung cancer in Teesside (UK) and Varese (Italy): a comparison after implementation of the National Cancer Plan. Thorax. 2016; 71:230–7. https://doi.org/10.1136/thoraxjnl-2015-207572.
9. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008; 83:584–94. https://doi.org/10.4065/83.5.584.
10. Ellis PM, Vandermeer R. Delays in the diagnosis of lung cancer. J Thorac Dis. 2011; 3:183–8. https://doi.org/10.3978/j.issn.2072-1439.2011.01.01.
11. Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, Fagerstrom RM, Gareen IF, Gierada DS, Jones GC, Mahon I, Marcus PM, Sicks JD, et al, and National Lung Screening Trial Research Team. Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Med. 2013; 368:1980–91. https://doi.org/10.1056/NEJMoa1209120.
12. Croswell JM, Baker SG, Marcus PM, Clapp JD, Kramer BS. Cumulative incidence of false-positive test results in lung cancer screening: a randomized trial. Ann Intern Med. 2010; 152:505–12, W176–80. https://doi.org/10.7326/0003-4819-152-8-201004200-00007.
13. Toyoda Y, Nakayama T, Kusunoki Y, Iso H, Suzuki T. Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography. Br J Cancer. 2008; 98:1602–7. https://doi.org/10.1038/sj.bjc.6604351.
14. Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, Byers T, Colditz GA, Gould MK, Jett JR, Sabichi AL, Smith-Bindman R, Wood DE, et al. Benefits and harms of CT screening for lung cancer: a systematic review. JAMA. 2012; 307:2418–29. https://doi.org/10.1001/jama.2012.5521.
15. Swensen SJ, Jett JR, Hartman TE, Midthun DE, Mandrekar SJ, Hillman SL, Sykes AM, Aughenbaugh GL, Bungum AO, Allen KL. CT screening for lung cancer: five-year prospective experience. Radiology. 2005; 235:259–65. https://doi.org/10.1148/radiol.2351041662.
16. Aberle DR, Brown K. Lung cancer screening with CT. Clin Chest Med. 2008; 29:1–14, v. https://doi.org/10.1016/j.ccm.2007.12.001.
17. Dominioni L, Poli A, Mantovani W, Pisani S, Rotolo N, Paolucci M, Sessa F, Conti V, D'Ambrosio V, Paddeu A, Imperatori A. Assessment of lung cancer mortality reduction after chest X-ray screening in smokers: a population-based cohort study in Varese, Italy. Lung Cancer. 2013; 80:50–4. https://doi.org/10.1016/j.lungcan.2012.12.014.
18. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer. 2006; 6:259–69. https://doi.org/10.1038/nrc1840.
19. Iorio MV, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. EMBO Mol Med. 2012; 4:143–59. https://doi.org/10.1002/emmm.201100209.
20. Sestini S, Boeri M, Marchiano A, Pelosi G, Galeone C, Verri C, Suatoni P, Sverzellati N, La Vecchia C, Sozzi G, Pastorino U. Circulating microRNA signature as liquid-biopsy to monitor lung cancer in low-dose computed tomography screening. Oncotarget. 2015; 6:32868–77. https://doi.org/10.18632/oncotarget.5210.
21. Boeri M, Sestini S, Fortunato O, Verri C, Suatoni P, Pastorino U, Sozzi G. Recent advances of microRNA-based molecular diagnostics to reduce false-positive lung cancer imaging. Expert Rev Mol Diagn. 2015; 15:801–13. https://doi.org/10.1586/14737159.2015.1041377.
22. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, et al. MicroRNA expression profiles classify human cancers. Nature. 2005; 435:834–8. https://doi.org/10.1038/nature03702.
23. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, et al. Characterization

of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. 2008; 18:997–1006. https://doi.org/10.1038/cr.2008.282.

24. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008; 105:10513–8. https://doi.org/10.1073/pnas.0804549105.

25. De Guire V, Robitaille R, Tetreault N, Guerin R, Menard C, Bambace N, Sapieha P. Circulating miRNAs as sensitive and specific biomarkers for the diagnosis and monitoring of human diseases: promises and challenges. Clin Biochem. 2013; 46:846–60. https://doi.org/10.1016/j.clinbiochem.2013.03.015.

26. Bianchi F, Nicassio F, Marzi M, Belloni E, Dall'olio V, Bernard L, Pelosi G, Maisonneuve P, Veronesi G, Di Fiore PP. A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. EMBO Mol Med. 2011; 3:495–503. https://doi.org/10.1002/emmm.201100154.

27. Foss KM, Sima C, Ugolini D, Neri M, Allen KE, Weiss GJ. miR-1254 and miR-574-5p: serum-based microRNA biomarkers for early-stage non-small cell lung cancer. J Thorac Oncol. 2011; 6:482–8. https://doi.org/10.1097/JTO.0b013e318208c785.

28. Ma J, Li N, Guarnera M, Jiang F. Quantification of Plasma miRNAs by Digital PCR for Cancer Diagnosis. Biomark Insights. 2013; 8:127–36. https://doi.org/10.4137/BMI.S13154.

29. Sanfiorenzo C, Ilie MI, Belaid A, Barlesi F, Mouroux J, Marquette CH, Brest P, Hofman P. Two panels of plasma microRNAs as non-invasive biomarkers for prediction of recurrence in resectable NSCLC. PLoS One. 2013; 8:e54596. https://doi.org/10.1371/journal.pone.0054596.

30. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, Guarnera M, Liao J, Chou A, Lu CL, Jiang Z, Fang H, Katz RL, et al. Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. Lab Invest. 2011; 91:579–87. https://doi.org/10.1038/labinvest.2010.194.

31. Wei J, Gao W, Zhu CJ, Liu YQ, Mei Z, Cheng T, Shu YQ. Identification of plasma microRNA-21 as a biomarker for early detection and chemosensitivity of non-small cell lung cancer. Chin J Cancer. 2011; 30:407–14.

32. Zheng D, Haddadin S, Wang Y, Gu LQ, Perry MC, Freter CE, Wang MX. Plasma microRNAs as novel biomarkers for early detection of lung cancer. Int J Clin Exp Pathol. 2011; 4:575–86.

33. Zandberga E, Kozirovskis V, Abols A, Andrejeva D, Purkalne G, Line A. Cell-free microRNAs as diagnostic, prognostic, and predictive biomarkers for lung cancer. Genes Chromosomes Cancer. 2013; 52:356–69. https://doi.org/10.1002/gcc.22032.

34. Huang Y, Hu Q, Deng Z, Hang Y, Wang J, Wang K. MicroRNAs in body fluids as biomarkers for non-small cell lung cancer: a systematic review. Technol Cancer Res Treat. 2014; 13:277–87. https://doi.org/10.7785/tcrt.2012.500377.

35. Wang H, Wu S, Zhao L, Zhao J, Liu J, Wang Z. Clinical use of microRNAs as potential non-invasive biomarkers for detecting non-small cell lung cancer: a meta-analysis. Respirology. 2015; 20:56–65. https://doi.org/10.1111/resp.12444.

36. Chen L, Jin H. MicroRNAs as novel biomarkers in the diagnosis of non-small cell lung cancer: a meta-analysis based on 20 studies. Tumour Biol. 2014; 35:9119–29. https://doi.org/10.1007/s13277-014-2188-2.

37. He WJ, Li WH, Jiang B, Wang YF, Xia YX, Wang L. MicroRNAs level as an initial screening method for early-stage lung cancer: a bivariate diagnostic random-effects meta-analysis. Int J Clin Exp Med. 2015; 8:12317–26.

38. Ulivi P, Zoli W. miRNAs as non-invasive biomarkers for lung cancer diagnosis. Molecules. 2014; 19:8220–37. https://doi.org/10.3390/molecules19068220.

39. Alipoor SD, Adcock IM, Garssen J, Mortaz E, Varahram M, Mirsaeidi M, Velayati A. The roles of miRNAs as potential biomarkers in lung diseases. Eur J Pharmacol. 2016; 791:395–404. https://doi.org/10.1016/j.ejphar.2016.09.015.

40. Nitu R, Rogobete AF, Gundogdu F, Tanasescu S, Boruga O, Sas A, Popovici SE, Hutanu D, Pilut C, Sarau CA, Candea AC, Stan AT, Moise LM. microRNAs Expression as Novel Genetic Biomarker for Early Prediction and Continuous Monitoring in Pulmonary Cancer. Biochem Genet. 2017; 55:281–90. https://doi.org/10.1007/s10528-016-9789-y.

41. Shi GL, Chen Y, Sun Y, Yin YJ, Song CX. Significance of Serum MicroRNAs in the Auxiliary Diagnosis of Non-Small Cell Lung Cancer. Clin Lab. 2017; 63:133–40. https://doi.org/10.7754/Clin.Lab.2016.160710.

42. Yuxia M, Zhennan T, Wei Z. Circulating miR-125b is a novel biomarker for screening non-small-cell lung cancer and predicts poor prognosis. J Cancer Res Clin Oncol. 2012; 138:2045–50. https://doi.org/10.1007/s00432-012-1285-0.

43. Li W, Wang Y, Zhang Q, Tang L, Liu X, Dai Y, Xiao L, Huang S, Chen L, Guo Z, Lu J, Yuan K. MicroRNA-486 as a Biomarker for Early Diagnosis and Recurrence of Non-Small Cell Lung Cancer. PLoS One. 2015; 10:e0134220. https://doi.org/10.1371/journal.pone.0134220.

44. Zhu W, He J, Chen D, Zhang B, Xu L, Ma H, Liu X, Zhang Y, Le H. Expression of miR-29c, miR-93, and miR-429 as potential biomarkers for detection of early stage non-small lung cancer. PLoS One. 2014; 9:e87780. https://doi.org/10.1371/journal.pone.0087780.

45. Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC, Ramnath N, Gotoh N, Beer DG, Chen G. A Novel Serum 4-microRNA Signature for Lung Cancer Detection. Sci Rep. 2015; 5:12464. https://doi.org/10.1038/srep12464.

46. Powrozek T, Krawczyk P, Kowalski DM, Kuznar-Kaminska B, Winiarczyk K, Olszyna-Serementa M, Batura-Gabryel H, Milanowski J. Application of plasma circulating microRNA-448, 506, 4316, and 4478 analysis for non-invasive diagnosis of lung cancer. Tumour Biol. 2016; 37:2049–55. https://doi.org/10.1007/s13277-015-3971-4.

47. Halvorsen AR, Bjaanæs M, LeBlanc M, Holm AM, Bolstad N, Rubio L, Peñalver JC, Cervera J, Mojarrieta JC, López-Guerrero JA, Brustugun OT, Helland Å. A unique set of 6 circulating microRNAs for early detection of non-small cell lung cancer. Oncotarget. 2016; 7:37250–59. https://doi.org/10.18632/oncotarget.9363.

48. Geng Q, Fan T, Zhang B, Wang W, Xu Y, Hu H. Five microRNAs in plasma as novel biomarkers for screening of early-stage non-small cell lung cancer. Respir Res. 2014; 15:149. https://doi.org/10.1186/s12931-014-0149-3.

49. Ulivi P, Foschi G, Mengozzi M, Scarpi E, Silvestrini R, Amadori D, Zoli W. Peripheral blood miR-328 expression as a potential biomarker for the early diagnosis of NSCLC. Int J Mol Sci. 2013; 14:10332–42. https://doi.org/10.3390/ijms140510332.

50. Wang Y, Zhao H, Gao X, Wei F, Zhang X, Su Y, Wang C, Li H, Ren X. Identification of a three-miRNA signature as a blood-borne diagnostic marker for early diagnosis of lung adenocarcinoma. Oncotarget. 2016; 7:26070–86. https://doi.org/10.18632/oncotarget.8429.

51. Zhang H, Mao F, Shen T, Luo Q, Ding Z, Qian L, Huang J. Plasma miR-145, miR-20a, miR-21 and miR-223 as novel biomarkers for screening early-stage non-small cell lung cancer. Oncol Lett. 2017; 13:669–76. https://doi.org/10.3892/ol.2016.5462.

52. Zhu W, Zhou K, Zha Y, Chen D, He J, Ma H, Liu X, Le H, Zhang Y. 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. https://doi.org/10.1371/journal.pone.0153046.

53. Wang W, Li W, Ding M, Yuan H, Yang J, Meng W, Jin E, Wang X, Ma S. Identification of miRNAs as non-invasive biomarkers for early diagnosis of lung cancers. Tumour Biol. 2016; 37:16287–93. https://doi.org/10.1007/s13277-016-5442-y.

54. Wozniak MB, Scelo G, Muller DC, Mukeria A, Zaridze D, Brennan P. Circulating MicroRNAs as Non-Invasive Biomarkers for Early Detection of Non-Small-Cell Lung Cancer. PLoS One. 2015; 10:e0125026. https://doi.org/10.1371/journal.pone.0125026.

55. Sun M, Song J, Zhou Z, Zhu R, Jin H, Ji Y, Lu Q, Ju H. Comparison of Serum MicroRNA21 and Tumor Markers in Diagnosis of Early Non-Small Cell Lung Cancer. Dis Markers. 2016; 2016:3823121. https://doi.org/10.1155/2016/3823121.

56. Shkurnikov MY, Knyazev EN, Fomicheva KA, Mikhailenko DS, Nyushko KM, Saribekyan EK, Samatov TR, Alekseev BY. Analysis of Plasma microRNA Associated with Hemolysis. Bull Exp Biol Med. 2016; 160:748–50. https://doi.org/10.1007/s10517-016-3300-y.

57. Shah JS, Soon PS, Marsh DJ. Comparison of Methodologies to Detect Low Levels of Hemolysis in Serum for Accurate Assessment of Serum microRNAs. PLoS One. 2016; 11:e0153200. https://doi.org/10.1371/journal.pone.0153200.

58. Landoni E, Miceli R, Callari M, Tiberio P, Appierto V, Angeloni V, Mariani L, Daidone MG. Proposal of supervised data analysis strategy of plasma miRNAs from hybridisation array data with an application to assess hemolysis-related deregulation. BMC Bioinformatics. 2015; 16:388. https://doi.org/10.1186/s12859-015-0820-9.

59. MacLellan SA, MacAulay C, Lam S, Garnis C. Pre-profiling factors influencing serum microRNA levels. BMC Clin Pathol. 2014; 14:27. https://doi.org/10.1186/1472-6890-14-27.

60. Yamada A, Cox MA, Gaffney KA, Moreland A, Boland CR, Goel A. Technical factors involved in the measurement of circulating microRNA biomarkers for the detection of colorectal neoplasia. PLoS One. 2014; 9:e112481. https://doi.org/10.1371/journal.pone.0112481.

61. Kirschner MB, Edelman JJ, Kao SC, Vallely MP, van Zandwijk N, Reid G. The Impact of Hemolysis on Cell-Free microRNA Biomarkers. Front Genet. 2013; 4:94. https://doi.org/10.3389/fgene.2013.00094.

62. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, Tait JF, Tewari M. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer Prev Res (Phila). 2012; 5:492–7. https://doi.org/10.1158/1940-6207.CAPR-11-0370.

63. McDonald JS, Milosevic D, Reddi HV, Grebe SK, Algeciras-Schimnich A. Analysis of circulating microRNA: preanalytical and analytical challenges. Clin Chem. 2011; 57:833–40. https://doi.org/10.1373/clinchem.2010.157198.

64. Chen X, Hu Z, Wang W, Ba Y, Ma L, Zhang C, Wang C, Ren Z, Zhao Y, Wu S, Zhuang R, Zhang Y, Hu H, et al. Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel noninvasive biomarkers for nonsmall cell lung cancer diagnosis. Int J Cancer. 2012; 130:1620–8. https://doi.org/10.1002/ijc.26177.

65. Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, Calabro E, Croce CM, Pastorino U, Sozzi G. MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. Proc Natl Acad Sci U S A. 2011; 108:3713–8. https://doi.org/10.1073/pnas.1100048108.

66. Keller A, Leidinger P, Gislefoss R, Haugen A, Langseth H, Staehler P, Lenhof HP, Meese E. Stable serum miRNA profiles as potential tool for non-invasive lung cancer diagnosis. RNA Biol. 2011; 8:506–16.

67. Lin Q, Mao W, Shu Y, Lin F, Liu S, Shen H, Gao W, Li S, Shen D. A cluster of specified microRNAs in peripheral

blood as biomarkers for metastatic non-small-cell lung cancer by stem-loop RT-PCR. J Cancer Res Clin Oncol. 2012; 138:85–93. https://doi.org/10.1007/s00432-011-1068-z.

68. Li M, Zhang Q, Wu L, Jia C, Shi F, Li S, Peng A, Zhang G, Song X, Wang C. Serum miR-499 as a novel diagnostic and prognostic biomarker in non-small cell lung cancer. Oncol Rep. 2014; 31:1961–7. https://doi.org/10.3892/or.2014.3029.

69. Powrozek T, Krawczyk P, Kowalski DM, Winiarczyk K, Olszyna-Serementa M, Milanowski J. Plasma circulating microRNA-944 and microRNA-3662 as potential histologic type-specific early lung cancer biomarkers. Transl Res. 2015; 166:315–23. https://doi.org/10.1016/j.trsl.2015.05.009.

70. Baker SG, Kramer BS, Srivastava S. Markers for early detection of cancer: statistical guidelines for nested case-control studies. BMC Med Res Methodol. 2002; 2:4.

71. Lebanony D, Benjamin H, Gilad S, Ezagouri M, Dov A, Ashkenazi K, Gefen N, Izraeli S, Rechavi G, Pass H, Nonaka D, Li J, Spector Y, et al. Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma. J Clin Oncol. 2009; 27:2030–7. https://doi.org/10.1200/JCO.2008.19.4134.

72. Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, Beer DG, Cope L, Creighton CJ, Danilova L, Ding L, Getz G, Hammerman PS, et al, and Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511:543–50. https://doi.org/10.1038/nature13385.

73. McDoniels-Silvers AL, Nimri CF, Stoner GD, Lubet RA, You M. Differential gene expression in human lung adenocarcinomas and squamous cell carcinomas. Clin Cancer Res. 2002; 8:1127–38.

74. Bjaanaes MM, Halvorsen AR, Solberg S, Jorgensen L, Dragani TA, Galvan A, Colombo F, Anderlini M, Pastorino U, Kure E, Borresen-Dale AL, Brustugun OT, Helland A. Unique microRNA-profiles in EGFR-mutated lung adenocarcinomas. Int J Cancer. 2014; 135:1812–21. https://doi.org/10.1002/ijc.28828.

75. Le HB, Zhu WY, Chen DD, He JY, Huang YY, Liu XG, Zhang YK. Evaluation of dynamic change of serum miR-21 and miR-24 in pre- and post-operative lung carcinoma patients. Med Oncol. 2012; 29:3190–7. https://doi.org/10.1007/s12032-012-0303-z.

76. Gao X, Wang Y, Zhao H, Wei F, Zhang X, Su Y, Wang C, Li H, Ren X. Plasma miR-324-3p and miR-1285 as diagnostic and prognostic biomarkers for early stage lung squamous cell carcinoma. Oncotarget. 2016; 7:59664–75. https://doi.org/10.18632/oncotarget.11198.

77. Takahashi K, Yokota S, Tatsumi N, Fukami T, Yokoi T, Nakajima M. Cigarette smoking substantially alters plasma microRNA profiles in healthy subjects. Toxicol Appl Pharmacol. 2013; 272:154–60. https://doi.org/10.1016/j.taap.2013.05.018.

78. Russ R, Slack FJ. Cigarette-Smoke-Induced Dysregulation of MicroRNA Expression and Its Role in Lung Carcinogenesis. Pulm Med. 2012; 2012:791234. https://doi.org/10.1155/2012/791234.

79. Dorn GW 2nd. MicroRNAs in cardiac disease. Transl Res. 2011; 157:226–35. https://doi.org/10.1016/j.trsl.2010.12.013.

80. Ji J, Wang XW. New kids on the block: diagnostic and prognostic microRNAs in hepatocellular carcinoma. Cancer Biol Ther. 2009; 8:1686–93.

81. Haider BA, Baras AS, McCall MN, Hertel JA, Cornish TC, Halushka MK. A critical evaluation of microRNA biomarkers in non-neoplastic disease. PLoS One. 2014; 9:e89565. https://doi.org/10.1371/journal.pone.0089565.

82. Akbas F, Coskunpinar E, Aynaci E, Oltulu YM, Yildiz P. Analysis of serum micro-RNAs as potential biomarker in chronic obstructive pulmonary disease. Exp Lung Res. 2012; 38:286–94. https://doi.org/10.3109/01902148.2012.689088.

83. Panganiban RP, Pinkerton MH, Maru SY, Jefferson SJ, Roff AN, Ishmael FT. Differential microRNA epression in asthma and the role of miR-1248 in regulation of IL-5. Am J Clin Exp Immunol. 2012; 1:154–65.

84. Xie L, Wu M, Lin H, Liu C, Yang H, Zhan J, Sun S. An increased ratio of serum miR-21 to miR-181a levels is associated with the early pathogenic process of chronic obstructive pulmonary disease in asymptomatic heavy smokers. Mol Biosyst. 2014; 10:1072–81. https://doi.org/10.1039/c3mb70564a.

85. Leidinger P, Keller A, Backes C, Huwer H, Meese E. MicroRNA expression changes after lung cancer resection: a follow-up study. RNA Biol. 2012; 9:900–10. https://doi.org/10.4161/rna.20107.

86. Aushev VN, Zborovskaya IB, Laktionov KK, Girard N, Cros MP, Herceg Z, Krutovskikh V. Comparisons of microRNA patterns in plasma before and after tumor removal reveal new biomarkers of lung squamous cell carcinoma. PLoS One. 2013; 8:e78649. https://doi.org/10.1371/journal.pone.0078649.

87. Cheng HH, Yi HS, Kim Y, Kroh EM, Chien JW, Eaton KD, Goodman MT, Tait JF, Tewari M, Pritchard CC. Plasma processing conditions substantially influence circulating microRNA biomarker levels. PLoS One. 2013; 8:e64795. https://doi.org/10.1371/journal.pone.0064795.

88. Pritchard CC, Cheng HH, Tewari M. MicroRNA profiling: approaches and considerations. Nat Rev Genet. 2012; 13:358–69. https://doi.org/10.1038/nrg3198.

89. Kroh EM, Parkin RK, Mitchell PS, Tewari M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods. 2010; 50:298–301. https://doi.org/10.1016/j.ymeth.2010.01.032.

90. Wang K, Yuan Y, Cho JH, McClarty S, Baxter D, Galas DJ. Comparing the MicroRNA spectrum between serum and

plasma. PLoS One. 2012; 7:e41561. https://doi.org/10.1371/journal.pone.0041561.

91. Fan L, Qi H, Teng J, Su B, Chen H, Wang C, Xia Q. Identification of serum miRNAs by nano-quantum dots microarray as diagnostic biomarkers for early detection of non-small cell lung cancer. Tumour Biol. 2016; 37:7777–84. https://doi.org/10.1007/s13277-015-4608-3.

92. Wang Y, Gu J, Roth JA, Hildebrandt MA, Lippman SM, Ye Y, Minna JD, Wu X. Pathway-based serum microRNA profiling and survival in patients with advanced stage non-small cell lung cancer. Cancer Res. 2013; 73:4801–9. https://doi.org/10.1158/0008-5472.CAN-12-3273.

93. Diaz-Garcia CV, Agudo-Lopez A, Perez C, Lopez-Martin JA, Rodriguez-Peralto JL, de Castro J, Cortijo A, Martinez-Villanueva M, Iglesias L, Garcia-Carbonero R, Fresno Vara JA, Gamez-Pozo A, Palacios J, et al. DICER1, DROSHA and miRNAs in patients with non-small cell lung cancer: implications for outcomes and histologic classification. Carcinogenesis. 2013; 34:1031–8. https://doi.org/10.1093/carcin/bgt022.

94. Xiang M, Zeng Y, Yang R, Xu H, Chen Z, Zhong J, Xie H, Xu Y, Zeng X. U6 is not a suitable endogenous control for the quantification of circulating microRNAs. Biochem Biophys Res Commun. 2014; 454:210–4. https://doi.org/10.1016/j.bbrc.2014.10.064.

95. Campomenosi P, Gini E, Noonan DM, Poli A, D'Antona P, Rotolo N, Dominioni L, Imperatori A. A comparison between quantitative PCR and droplet digital PCR technologies for circulating microRNA quantification in human lung cancer. BMC Biotechnol. 2016; 16:60. https://doi.org/10.1186/s12896-016-0292-7.

96. Chan BA, Hughes BG. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. Transl Lung Cancer Res. 2015; 4:36–54. https://doi.org/10.3978/j.issn.2218-6751.2014.05.01.

97. Singh J. Mendeley: A free research management tool for desktop and web. J Pharmacol Pharmacother. 2010; 1:62–3. https://doi.org/10.4103/0976-500X.64539.

98. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, and QUADAS-2 Group. a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155:529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009.

Supplementary Table 2: Effect of hemolysis on measurement of circulating miRNAs listed in Tables 1 and 3

miRNA	Affected by hemolysis[a] (yes/no)	References[b]
223	no	MacLellan et al. [59]
486	yes	MacLellan et al. [59], Landoni et al [58], Shkurnikov et al. [56], Kirschner et al. [61], Pritchard et al. [62]
20a	no	MacLellan et al. [59], Shkurnikov et al. [56]
448	no	
21	yes	MacLellan et al. [59], Kirschner et al. [61], Pritchard et al. [62], Yamada et al. [60]
21-5p	yes	MacLellan et al. [59], Kirschner et al. [61], Pritchard et al. [62], Yamada et al. [60]
145	no	
141	no	
193b	no	
200b	no	
126	yes	MacLellan et al. [59], Kirschner et al. [61], Pritchard et al. [62]
301	no	MacLellan et al. [59]
328	no	MacLellan et al. [59], Pritchard et al. [62]
4478	no	
125b	no	Yamada et al. [60]
1244	no	
182	no	MacLellan et al. [59]
425-3p	no	MacLellan et al. [59]
628-3p	no	MacLellan et al. [59]
29c	no	MacLellan et al. [59]
429	no	
22	no	MacLellan et al. [59], Landoni et al. [58]
335-3p	no	
532	no	MacLellan et al. [59], Kirschner et al. [61]
210	no	MacLellan et al. [59], Kirschner et al. [61]
183	no	MacLellan et al. [59]
15b	yes	MacLellan et al. [59], Kirschner et al. [61], Pritchard et al. [62], Shah et al. [57], McDonald et al. [63]
92a	yes	MacLellan et al. [59], Landoni et al. [58], Kirschner et al. [61], Pritchard et al. [62]
17	yes	MacLellan et al. [59], Shkurnikov et al. [56], Kirschner et al. [61], Pritchard et al. [62]
103	yes	MacLellan et al. [59], Kirschner et al. [61], Pritchard et al. [62]
29a	yes	MacLellan et al. [59], Pritchard et al. [62], Yamada et al. [60]
24-3p	no	MacLellan et al. [59]

[a]miRNA affected by hemolysis as documented by three or more relevant studies.
[b]Studies documenting that given miRNA is affected by hemolysis (see text).
miRNAs in bold are those described in Table 2.

Systematic review and critique of circulating miRNAs as biomarkers of stage I-II non-small cell lung cance

SUPPLEMENTARY MATERIALS

Supplementary File 1: Biological functions of microRNAs proposed as biomarkers for two-step screening of stage I-II NSCLC in the serum. See Supplementary_File_1

Supplementary File 2: Model proposed to estimate the overall sensitivity and specificity of 8 miRNAs used for a two-step screening of stage I-II NSCLC. See Supplementary_File_2

Supplementary Table 1A. Main characteristics of the 20 included studies. See Supplementary_Table_1A

Supplementary Table 1B. Materials and procedures used for miRNA quantification in the 20 included studies. See Supplementary_Table_1B

Supplementary File 1: Biological functions of microRNAs proposed as biomarkers for two-step screening of stage I-II NSCLC in the serum.

miR-223-3p

Different effects of miR-223 have been shown in different types of cancer, and information about the role in lung cancer is limited. miR-223 behaves as a tumor suppressor in non-small cell lung cancer (NSCLC) and NSCLC cell line [1-3], a cervical cancer cell line and tumor tissues [4] and in acute myeloid leukemia [5, 6]. Conversely, miR-223 is overexpressed in gastric cancer compared to normal gastric mucosa and plays a role as an oncomiR in T-cell acute lymphoblastic leukemia [7] as well as in a gastric cancer cell line [8] and in prostate cancer cell lines [9]. The different behavior of miR-223 as oncogene or oncosuppressor could depend on its targets in the different tissues and cell types [10].

miR-223 inhibits cell growth, proliferation, colony formation in Hela cells in vitro and tumor formation in vivo by targeting insulin-like growth factor 1 (IGF-1R) and its downstream PI3K/Akt/mTOR/p70S6K pathway; this suppressive mechanism has also been confirmed in leukemia and hepatoma cells [11]; targeting of this pathway by re-expression of miR-223 also sensitizes erlotinib resistant lung cancer cells to drug-induced apoptosis [1].

miR-223 inhibits metastasis of cervical cancer by controlling epithelial-mesenchymal transition (EMT), upregulating the epithelial markers E-cadherin and α-cadherin and downregulating the mesenchymal marker vimentin [4]; conversely, inhibition of miR-223 in gemcitabine resistant pancreatic cancer cells reversed EMT [12].

In a human osteosarcoma cell line, miR-223 repressed the expression of a member of HSP90 family of stress induced proteins, inducing apoptosis as well as G0/G1 arrest [13].

References

1. Zhao FY, Han J, Chen XW, Wang J, Wang XD, Sun JG, Chen ZT. miR-223 enhances the sensitivity of non-small cell lung cancer cells to erlotinib by targeting the insulin-like growth factor-1 receptor. Int J Mol Med. 2016; 38: 183-91. doi: 10.3892/ijmm.2016.2588.

2. Geng Q, Fan T, Zhang B, Wang W, Xu Y, Hu H. Five microRNAs in plasma as novel biomarkers for screening of early-stage non-small cell lung cancer. Respir Res. 2014; 15: 149. doi: 10.1186/s12931-014-0149-3.

3. Lv S, Xue J, Wu C, Wang L, Wu J, Xu S, Liang X, Lou J. Identification of A Panel of Serum microRNAs as Biomarkers for Early Detection of Lung Adenocarcinoma. J Cancer. 2017; 8: 48-56. doi: 10.7150/jca.16644.

4. Tang Y, Wang Y, Chen Q, Qiu N, Zhao Y, You X. MiR-223 inhibited cell metastasis of human cervical cancer by modulating epithelial-mesenchymal transition. Int J Clin Exp Pathol. 2015; 8: 11224-9. doi:

5. Fazi F, Racanicchi S, Zardo G, Starnes LM, Mancini M, Travaglini L, Diverio D, Ammatuna E, Cimino G, Lo-Coco F, Grignani F, Nervi C. Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO oncoprotein. Cancer Cell. 2007; 12: 457-66. doi: 10.1016/j.ccr.2007.09.020.

6. Pulikkan JA, Dengler V, Peramangalam PS, Peer Zada AA, Muller-Tidow C, Bohlander SK, Tenen DG, Behre G. Cell-cycle regulator E2F1 and microRNA-223 comprise an autoregulatory negative feedback loop in acute myeloid leukemia. Blood. 2010; 115: 1768-78. doi: 10.1182/blood-2009-08-240101.

7. Mavrakis KJ, Van Der Meulen J, Wolfe AL, Liu X, Mets E, Taghon T, Khan AA, Setty M, Rondou P, Vandenberghe P, Delabesse E, Benoit Y, Socci NB, et al. A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). Nat Genet. 2011; 43: 673-8. doi: 10.1038/ng.858.

8. Li J, Guo Y, Liang X, Sun M, Wang G, De W, Wu W. MicroRNA-223 functions as an oncogene in human gastric cancer by targeting FBXW7/hCdc4. J Cancer Res Clin Oncol. 2012; 138: 763-74. doi: 10.1007/s00432-012-1154-x.

9. Wei Y, Yang J, Yi L, Wang Y, Dong Z, Liu Z, Ou-yang S, Wu H, Zhong Z, Yin Z, Zhou K, Gao Y, Yan B, et al. MiR-223-3p targeting SEPT6 promotes the biological behavior of prostate cancer. Sci Rep. 2014; 4: 7546. doi: 10.1038/srep07546.

10. Zhang B, Pan X, Cobb GP, Anderson TA. microRNAs as oncogenes and tumor suppressors. Dev Biol. 2007; 302: 1-12. doi: 10.1016/j.ydbio.2006.08.028.

11. Jia CY, Li HH, Zhu XC, Dong YW, Fu D, Zhao QL, Wu W, Wu XZ. MiR-223 suppresses cell proliferation by targeting IGF-1R. PLoS One. 2011; 6: e27008. doi: 10.1371/journal.pone.0027008.

12. Ma J, Fang B, Zeng F, Ma C, Pang H, Cheng L, Shi Y, Wang H, Yin B, Xia J, Wang Z. Down-regulation of miR-223 reverses epithelial-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells. Oncotarget. 2015; 6: 1740-9. doi: 10.18632/oncotarget.2714.

13. Li G, Cai M, Fu D, Chen K, Sun M, Cai Z, Cheng B. Heat shock protein 90B1 plays an oncogenic role and is a target of microRNA-223 in human osteosarcoma. Cell Physiol Biochem. 2012; 30: 1481-90. doi: 10.1159/000343336.

miR-20a-5p

miR-20a plays an oncogenic role in NSCLC by targeting genes involved in several aspects of tumorigenesis. An inverse correlation was shown between expression of miR-20 and expression levels of ferroportin, thus increasing iron availability supporting cell proliferation in NSCLC cell lines [1]; another target of miR-20a, both *in vivo* and *in vitro* is transforming growth factor-β (TGF-β) type II receptor [2]. In colon cancer, miR-20a directly targets Smad4 3'UTR, promoting invasion and metastasis *in vitro* and *in vivo* by a colorectal cancer cell line [3].

In Hela cells c-Myc regulates miR-20a to target E2F1 transcription factor [4]. In osteosarcoma, a novel miRNA-target gene regulatory network was identified between miR-20a and CCND2 (G1/S-specific cyclin-D2) that may play a role during cell proliferation [5]. Finally, miR-20a expression may be affected by radiation therapy regimens during breast cancer treatment [6]. In turn, miR-20a regulates the expression of known target genes of the miR-17-92 cluster (to which this miRNA belongs), such as mitogen-activated protein kinase (MAPK), ErbB, p53, Wnt, transforming growth factor-β (TGF-β), mTOR signaling pathways and cell cycle [6].

References

1. *Babu KR, Muckenthaler MU. miR-20a regulates expression of the iron exporter ferroportin in lung cancer. J Mol Med (Berl). 2016; 94: 347-59. doi: 10.1007/s00109-015-1362-3.*

2. *Yang S, Cho YJ, Jin L, Yuan G, Datta A, Buckhaults P, Datta PK. An epigenetic auto-feedback loop regulates TGF-beta type II receptor expression and function in NSCLC. Oncotarget. 2015; 6: 33237-52. doi: 10.18632/oncotarget.4893.*

3. *Cheng D, Zhao S, Tang H, Zhang D, Sun H, Yu F, Jiang W, Yue B, Wang J, Zhang M, Yu Y, Liu X, Sun X, et al. MicroRNA-20a-5p promotes colorectal cancer invasion and metastasis by downregulating Smad4. Oncotarget. 2016; 7: 45199-213. doi: 10.18632/oncotarget.9900.*

4. *O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005; 435: 839-43. doi: 10.1038/nature03677.*

5. *He C, Gao H, Fan X, Wang M, Liu W, Huang W, Yang Y. Identification of a novel miRNA-target gene regulatory network in osteosarcoma by integrating transcriptome analysis. Int J Clin Exp Pathol. 2015; 8: 8348-57.*

6. *Leung CM, Chen TW, Li SC, Ho MR, Hu LY, Liu WS, Wu TT, Hsu PC, Chang HT, Tsai KW. MicroRNA expression profiles in human breast cancer cells after multifraction and single-dose radiation treatment. Oncol Rep. 2014; 31: 2147-56. doi: 10.3892/or.2014.3089.*

miR-448

miR-448 expression is often decreased in human cancer tissues and is considered a tumor suppressor in hepatocellular carcinoma [1], ovarian [2] and breast cancer [3]. This was recently confirmed also for lung squamocellular carcinoma, where it may regulate cell growth and metastasis, possibly by targeting DCLK1 [4]. Xu et al confirmed downregulation of this miRNA in NSCLC [5].

More information is available on the biological role of miR-448 in other cancers. Several direct targets of miR-448 have been identified in cell lines from various tumors, including the serine/threonine kinase ROCK2 in a hepatocellular carcinoma cell line, the CXCL12 chemokine in an ovarian cancer cell line and

SATB1 (AT-rich sequence-binding protein 1), which enhances NF-κB activity and increases Twist1 expression, thereby leading to EMT in a breast cancer cell line [1, 2, 5, 6].

However, an oncogenic role of miR-448 has also been reported in gastric cancer where miR-448 was shown to suppress the expression of KDM2B that directly inhibits Myc [7].

These findings imply that miR-448 may play different and tissue-specific roles and behave differently in various tumor types.

References

1. *Zhu H, Zhou X, Ma C, Chang H, Li H, Liu F, Lu J. Low Expression of miR-448 Induces EMT and Promotes Invasion by Regulating ROCK2 in Hepatocellular Carcinoma. Cell Physiol Biochem. 2015; 36: 487-98. doi: 10.1159/000430114.*

2. *Lv Y, Lei Y, Hu Y, Ding W, Zhang C, Fang C. miR-448 negatively regulates ovarian cancer cell growth and metastasis by targeting CXCL12. Clin Transl Oncol. 2015; 17: 903-9. doi: 10.1007/s12094-015-1325-8.*

3. *Bamodu OA, Huang WC, Lee WH, Wu A, Wang LS, Hsiao M, Yeh CT, Chao TY. Aberrant KDM5B expression promotes aggressive breast cancer through MALAT1 overexpression and downregulation of hsa-miR-448. BMC Cancer. 2016; 16: 160. doi: 10.1186/s12885-016-2108-5.*

4. *Shan C, Fei F, Li F, Zhuang B, Zheng Y, Wan Y, Chen J. miR-448 is a novel prognostic factor of lung squamous cell carcinoma and regulates cells growth and metastasis by targeting DCLK1. Biomed Pharmacother. 2017; 89: 1227-34. doi: 10.1016/j.biopha.2017.02.017.*

5. *Xu C, Zheng Y, Lian D, Ye S, Yang J, Zeng Z. Analysis of microRNA expression profile identifies novel biomarkers for non-small cell lung cancer. Tumori. 2015; 101: 104-10. doi: 10.5301/tj.5000224.*

6. *Li QQ, Chen ZQ, Cao XX, Xu JD, Xu JW, Chen YY, Wang WJ, Chen Q, Tang F, Liu XP, Xu ZD. Involvement of NF-kappaB/miR-448 regulatory feedback loop in chemotherapy-induced epithelial-mesenchymal transition of breast cancer cells. Cell Death Differ. 2011; 18: 16-25. doi: 10.1038/cdd.2010.103.*

7. *Hong X, Xu Y, Qiu X, Zhu Y, Feng X, Ding Z, Zhang S, Zhong L, Zhuang Y, Su C, Hong X, Cai J. MiR-448 promotes glycolytic metabolism of gastric cancer by downregulating KDM2B. Oncotarget. 2016; 7: 22092-102. doi: 10.18632/oncotarget.8020.*

miR-145-5p

The tumor suppression function of miR-145 is recognized by the scientific community [1, 2]. The oncosuppressive functions of miR-145 have been described in lung cancer tissues [3] and in lung, breast, prostate and colorectal cancer cell lines [4-6].

The expression of miR-145-5p in NSCLC is significantly lower than that in healthy tissues, potentially leading to an increase in genes such as SMAD4, SMAD2, IRS1, FOXO1, ERBB4, NRAS, ACTB, and ACTG1 that are predicted targets in NSCLC [7]. In another study miR-145 was shown to negatively correlate with N-cadherin expression in lung adenocarcinoma tissues and, by doing so, to be able to suppress cell invasion and migration in lung adenocarcinoma cell lines [8].

miR-145 also shows antitumor activity in lung squamocellular carcinoma cell lines and downregulation of this microRNA enhances the expression of metadherin (MTDH), a downstream mediator of several signal pathways, such as PI3K/AKT, NF-κB, MAPK and Wnt/β-catenin [9].

Several independent studies suggested that miR-145 could control NSCLC cell migration and invasion. Hu et al. demonstrated that miR-145 represses TGF-β-induced EMT and invasion by targeting SMAD3 in NSCLC cells; these authors showed that SMAD3, an intracellular mediator in TGF-β signaling, affected the expression of EMT markers as E-cadherin and N-cadherin [10]. Moreover, targeting of FSCN1 could also be a mechanism by which miR-145 controls EMT transition [11].

Finally, miR-145 was shown to regulate other tumor-associated targets such as c-Myc or mucin 1 in NSCLC cells [12, 13].

References

1. *Takaoka Y, Shimizu Y, Hasegawa H, Ouchi Y, Qiao S, Nagahara M, Ichihara M, Lee JD, Adachi K, Hamaguchi M, Iwamoto T. Forced expression of miR-143 represses ERK5/c-Myc and p68/p72 signaling in*

concert with miR-145 in gut tumors of Apc(Min) mice. PLoS One. 2012; 7: e42137. doi: 10.1371/journal.pone.0042137.

2. Sachdeva M, Zhu S, Wu F, Wu H, Walia V, Kumar S, Elble R, Watabe K, Mo YY. p53 represses c-Myc through induction of the tumor suppressor miR-145. Proc Natl Acad Sci U S A. 2009; 106: 3207-12. doi: 10.1073/pnas.0808042106.

3. Lu Y, Govindan R, Wang L, Liu PY, Goodgame B, Wen W, Sezhiyan A, Pfeifer J, Li YF, Hua X, Wang Y, Yang P, You M. MicroRNA profiling and prediction of recurrence/relapse-free survival in stage I lung cancer. Carcinogenesis. 2012; 33: 1046-54. doi: 10.1093/carcin/bgs100.

4. Zhou CH, Yang SF, Li PQ. Human lung cancer cell line SPC-A1 contains cells with characteristics of cancer stem cells. Neoplasma. 2012; 59: 685-92. doi: 10.4149/neo_2012_087.

5. Wang Z, Zhang X, Yang Z, Du H, Wu Z, Gong J, Yan J, Zheng Q. MiR-145 regulates PAK4 via the MAPK pathway and exhibits an antitumor effect in human colon cells. Biochem Biophys Res Commun. 2012; 427: 444-9. doi: 10.1016/j.bbrc.2012.06.123.

6. Gotte M, Mohr C, Koo CY, Stock C, Vaske AK, Viola M, Ibrahim SA, Peddibhotla S, Teng YH, Low JY, Ebnet K, Kiesel L, Yip GW. miR-145-dependent targeting of junctional adhesion molecule A and modulation of fascin expression are associated with reduced breast cancer cell motility and invasiveness. Oncogene. 2010; 29: 6569-80. doi: 10.1038/onc.2010.386.

7. Gan TQ, Xie ZC, Tang RX, Zhang TT, Li DY, Li ZY, Chen G. Clinical value of miR-145-5p in NSCLC and potential molecular mechanism exploration: A retrospective study based on GEO, qRT-PCR, and TCGA data. Tumour Biol. 2017; 39: 1010428317691683. doi: 10.1177/1010428317691683.

8. Mo D, Yang D, Xiao X, Sun R, Huang L, Xu J. MiRNA-145 suppresses lung adenocarcinoma cell invasion and migration by targeting N-cadherin. Biotechnol Lett. 2017; 39: 701-10. doi: 10.1007/s10529-017-2290-9.

9. Mataki H, Seki N, Mizuno K, Nohata N, Kamikawaji K, Kumamoto T, Koshizuka K, Goto Y, Inoue H. Dual-strand tumor-suppressor microRNA-145 (miR-145-5p and miR-145-3p) coordinately targeted MTDH in lung squamous cell carcinoma. Oncotarget. 2016; 7: 72084-98. doi: 10.18632/oncotarget.12290.

10. Hu H, Xu Z, Li C, Xu C, Lei Z, Zhang HT, Zhao J. MiR-145 and miR-203 represses TGF-beta-induced epithelial-mesenchymal transition and invasion by inhibiting SMAD3 in non-small cell lung cancer cells. Lung Cancer. 2016; 97: 87-94. doi: 10.1016/j.lungcan.2016.04.017.

11. Zhang Y, Lin Q. MicroRNA-145 inhibits migration and invasion by down-regulating FSCN1 in lung cancer. Int J Clin Exp Med. 2015; 8: 8794-802. doi:

12. Chen Z, Zeng H, Guo Y, Liu P, Pan H, Deng A, Hu J. miRNA-145 inhibits non-small cell lung cancer cell proliferation by targeting c-Myc. J Exp Clin Cancer Res. 2010; 29: 151. doi: 10.1186/1756-9966-29-151.

13. Ye Z, Shen N, Weng Y, Li K, Hu L, Liao H, An J, Liu L, Lao S, Cai S. Low miR-145 silenced by DNA methylation promotes NSCLC cell proliferation, migration and invasion by targeting mucin 1. Cancer Biol Ther. 2015; 16: 1071-9. doi: 10.1080/15384047.2015.1046024.

miR-628-3p

Little information is available about the role of miR-628-3p in lung cancer. Bioinformatic analysis identified *ATRX* (SWI/SNF family of chromatin remodeling proteins), *SLC45A2* (transporter protein), and *TNRC6B* (Argonaute protein) *as* potential targets of miR-628-3p in NSCLC [1].

miR-628-3p dysregulation was described in some subtypes of gastric cancer and in neuroblastoma [2-4]. In the latter, this miRNA was found to negatively regulate *MYCN* gene expression [4], whereas in gastric cancer miR-628-3p appeared to regulate differentiation through as yet unknown mechanisms [2].

Another study showed that miR-628-3p is overexpressed in pancreatic cancer patient sera and tissues, compared to healthy controls [5].

References

1. Wang Y, Zhao H, Gao X, Wei F, Zhang X, Su Y, Wang C, Li H, Ren X. Identification of a three-miRNA signature as a blood-borne diagnostic marker for early diagnosis of lung adenocarcinoma. Oncotarget. 2016; 7: 26070-86. doi: 10.18632/oncotarget.8429.

2. *Gao S, Zhou F, Zhao C, Ma Z, Jia R, Liang S, Zhang M, Zhu X, Zhang P, Wang L, Su F, Zhao J, Liu G, et al. Gastric cardia adenocarcinoma microRNA profiling in Chinese patients. Tumour Biol. 2016; 37: 9411-22. doi: 10.1007/s13277-016-4824-5.*

3. *Li FQ, Xu B, Wu YJ, Yang ZL, Qian JJ. Differential microRNA expression in signet-ring cell carcinoma compared with tubular adenocarcinoma of human gastric cancer. Genet Mol Res. 2015; 14: 739-47. doi: 10.4238/2015.January.30.17.*

4. *Megiorni F, Colaiacovo M, Cialfi S, McDowell HP, Guffanti A, Camero S, Felsani A, Losty PD, Pizer B, Shukla R, Cappelli C, Ferrara E, Pizzuti A, et al. A sketch of known and novel MYCN-associated miRNA networks in neuroblastoma. Oncol Rep. 2017; 38: 3-20. doi: 10.3892/or.2017.5701.*

5. *Li A, Yu J, Kim H, Wolfgang CL, Canto MI, Hruban RH, Goggins M. MicroRNA array analysis finds elevated serum miR-1290 accurately distinguishes patients with low-stage pancreatic cancer from healthy and disease controls. Clin Cancer Res. 2013; 19: 3600-10. doi: 10.1158/1078-0432.CCR-12-3092.*

miR-29c-3p

miR-29c downregulation was associated with unfavorable prognosis in lung adenocarcinoma. miR-29c expression inhibited cell proliferation, migration and invasion in lung cancer cell lines and reduced the capability of tumor cells to promote HUVEC tube formation. VEGFA was shown to be a direct target of miR-29c [1].

Fabbri et al showed that miR-29c targets DNA methyltrasferases 3A and 3B in lung cancer cell lines and in NSCLC tissues, thus suppressing oncogenic DNA methylation [2]. Moreover, miR-29c targets Sp1 and reduces TGF-β-induced EMT, expression of epithelial markers such as TTF-1, while enhancing cell migration and invasion of two lung cancer cell lines [3].

Conversely, other studies suggested an oncogenic role for miR-29c. miR-29c level was significantly increased in sera and tissues of stage IA/B NSCLC patients compared to corresponding control sera and noncancerous tissues [4, 5].

References

1. *Liu L, Bi N, Wu L, Ding X, Men Y, Zhou W, Li L, Zhang W, Shi S, Song Y, Wang L. MicroRNA-29c functions as a tumor suppressor by targeting VEGFA in lung adenocarcinoma. Mol Cancer. 2017; 16: 50. doi: 10.1186/s12943-017-0620-0.*

2. *Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, Volinia S, Guler G, Morrison CD, et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci U S A. 2007; 104: 15805-10. doi: 10.1073/pnas.0707628104.*

3. *Zhang HW, Wang EW, Li LX, Yi SH, Li LC, Xu FL, Wang DL, Wu YZ, Nian WQ. A regulatory loop involving miR-29c and Sp1 elevates the TGF-beta1 mediated epithelial-to-mesenchymal transition in lung cancer. Oncotarget. 2016; 7: 85905-16. doi: 10.18632/oncotarget.13137.*

4. *Heegaard NH, Schetter AJ, Welsh JA, Yoneda M, Bowman ED, Harris CC. Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. Int J Cancer. 2012; 130: 1378-86. doi: 10.1002/ijc.26153.*

5. *Zhu W, He J, Chen D, Zhang B, Xu L, Ma H, Liu X, Zhang Y, Le H. Expression of miR-29c, miR-93, and miR-429 as potential biomarkers for detection of early stage non-small lung cancer. PLoS One. 2014; 9: e87780. doi: 10.1371/journal.pone.0087780.*

miR-210-3p

In lung cancer the role of this miRNA has been scarcely investigated. miR-210-3p was found overexpressed in NSCLC tissues at late stages, and targeted specific mitochondrial components, thus regulating cell death and survival and modulating HIF-1 activity [1]. The latter authors reported that miR-210 directly targets NDUFA4 (subunit of NADH dehydrogenase) and SDHD (succinate dehydrogenase complex subunit D) and induces mitochondrial dysfunction. Daugaard et al showed that increased expression of miR-210-3p was significantly associated with the presence of distant metastases [2]. However, miR-210-3p dysregulation may occur early in lung tumorigenesis, as it was found increased in plasma or serum of patients with early lung cancer [3-5].

More information is available about miR-210 functions in the context of other malignancies. miR-210 is induced by hypoxia and, in turn, it is able to increase HIF-1 transcriptional activity and expression of its target genes, VEGF and carbonic anhydrase 9, in glioblastoma specimens and cell lines [6].

In metastatic prostate cancer, overexpression of miR-210-3p positively correlates with serum PSA levels, Gleason grade and bone metastasis status in prostate cancer patients [7]. Expression of this miRNA in prostate cancer cells positively correlates with EMT transition, invasion and metastasis by targeting negative regulators of NF-κB signaling [7].

Moreover, miR-210-3p was shown to downregulate PICK1, a negative regulator of the TGF-β signaling pathway, necessary for metastasis to the bone [8].

Although the serum levels of miR-210-3p were found upregulated in clear cell renal cell carcinoma [9, 10], miR-210-3p depletion was shown to increase tumorigenesis in xenografts *in vivo* and to alter the morphology of renal cancer derived cell lines, indicating EMT, *in vitro*. TWIST1 was identified as a key target of miR-210-3p [11].

References

1. *Puissegur MP, Mazure NM, Bertero T, Pradelli L, Grosso S, Robbe-Sermesant K, Maurin T, Lebrigand K, Cardinaud B, Hofman V, Fourre S, Magnone V, Ricci JE, et al. miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity. Cell Death Differ. 2011; 18: 465-78. doi: 10.1038/cdd.2010.119.*

2. *Daugaard I, Veno MT, Yan Y, Kjeldsen TE, Lamy P, Hager H, Kjems J, Hansen LL. Small RNA sequencing reveals metastasis-related microRNAs in lung adenocarcinoma. Oncotarget. 2017; 8: 27047-61. doi: 10.18632/oncotarget.15968.*

3. *Shen J, Liu Z, Todd NW, Zhang H, Liao J, Yu L, Guarnera MA, Li R, Cai L, Zhan M, Jiang F. Diagnosis of lung cancer in individuals with solitary pulmonary nodules by plasma microRNA biomarkers. BMC Cancer. 2011; 11: 374. doi: 10.1186/1471-2407-11-374.*

4. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, Guarnera M, Liao J, Chou A, Lu CL, Jiang Z, Fang H, Katz RL, et al. Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. Lab Invest. 2011; 91: 579-87. doi: 10.1038/labinvest.2010.194.

5. Zhu W, Zhou K, Zha Y, Chen D, He J, Ma H, Liu X, Le H, Zhang Y. 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.

6. Agrawal R, Pandey P, Jha P, Dwivedi V, Sarkar C, Kulshreshtha R. Hypoxic signature of microRNAs in glioblastoma: insights from small RNA deep sequencing. BMC Genomics. 2014; 15: 686. doi: 10.1186/1471-2164-15-686.

7. Ren D, Yang Q, Dai Y, Guo W, Du H, Song L, Peng X. Oncogenic miR-210-3p promotes prostate cancer cell EMT and bone metastasis via NF-kappaB signaling pathway. Mol Cancer. 2017; 16: 117. doi: 10.1186/s12943-017-0688-6.

8. Dai Y, Ren D, Yang Q, Cui Y, Guo W, Lai Y, Du H, Lin C, Li J, Song L, Peng X. The TGF-beta signalling negative regulator PICK1 represses prostate cancer metastasis to bone. Br J Cancer. 2017; 117:685–94.

9. Petrozza V, Pastore AL, Palleschi G, Tito C, Porta N, Ricci S, Marigliano C, Costantini M, Simone G, Di Carlo A, Gallucci M, Carbone A, Fazi F. Secreted miR-210-3p as non-invasive biomarker in clear cell renal cell carcinoma. Oncotarget. 2017; 8:69551-69558. doi: 10.18632/oncotarget.18449.

10. Petrozza V, Carbone A, Bellissimo T, Porta N, Palleschi G, Pastore AL, Di Carlo A, Della Rocca C, Fazi F. Oncogenic MicroRNAs Characterization in Clear Cell Renal Cell Carcinoma. Int J Mol Sci. 2015; 16: 29219-25. doi: 10.3390/ijms161226160.

11. Yoshino H, Yonemori M, Miyamoto K, Tatarano S, Kofuji S, Nohata N, Nakagawa M, Enokida H. microRNA-210-3p depletion by CRISPR/Cas9 promoted tumorigenesis through revival of TWIST1 in renal cell carcinoma. Oncotarget. 2017; 8: 20881-94. doi: 10.18632/oncotarget.14930.

miR-1244

miR-1244 was found hypoexpressed in lung tumors and its reintroduction reduced the growth of lung cancer xenografts [1]. The same authors showed that miR-1244 targets myocyte enhancer factor 2D (MEF2D) in lung cancer cells [1].

Conversely, Wang et al. demonstrated higher levels of this miRNA in the sera of stage I-II NSCLC patients compared to patients with unidentified pulmonary nodules and healthy controls [2].

miR-1244 was found downregulated in A549 lung cancer cell line upon acquisition of resistance to cisplatin [3]. Overall survival times of cisplatin-treated NSCLC patients with high miR-1244 expression were higher than those of patients with low miR-1244 expression [4]. miR-1244 affected cisplatin-treated NSCLC via MEF2D expression [4].

References

1. *Zhang R, Zhang Y, Li H. miR-1244/Myocyte Enhancer Factor 2D Regulatory Loop Contributes to the Growth of Lung Carcinoma. DNA Cell Biol. 2015; 34: 692-700. doi: 10.1089/dna.2015.2915.*

2. *Wang W, Li W, Ding M, Yuan H, Yang J, Meng W, Jin E, Wang X, Ma S. Identification of miRNAs as non-invasive biomarkers for early diagnosis of lung cancers. Tumour Biol. 2016; 37:16287–93.*

3. *Li W, Wang W, Ding M, Zheng X, Ma S, Wang X. MiR-1244 sensitizes the resistance of non-small cell lung cancer A549 cell to cisplatin. Cancer Cell Int. 2016; 16: 30. doi: 10.1186/s12935-016-0305-6.*

4. *Li GJ, Zhao GQ, Yang JP, Zhou YC, Yang KY, Lei YJ, Huang YC. Effect of miR-1244 on cisplatin-treated non-small cell lung cancer via MEF2D expression. Oncol Rep. 2017; 37: 3475-83. doi: 10.3892/or.2017.5624.*

Supplementary File 2: Model proposed to estimate the overall sensitivity and specificity of 8 miRNAs used for a two-step screening of stage I-II NSCLC. Test 1 is performed with a panel of 4 highly sensitive miRNAs (A,B,C,D), Test 2 with a panel of 4 highly specific miRNAs (E,F,G,H) (Table 4).

Step 1) Within each panel the combined sensitivity for Test 1 (T1se) and Test 2 (T2se) and the combined specificity for Test 1 (T1sp) and Test 2 (T2sp) are calculated in parallel using the following formulas [1-3]:

-Combined sensitivity Test 1= 1 - (1-Ase) x (1-Bse) x (1-Cse) x (1-Dse)

-Combined specificity Test 1= Asp x Bsp x Csp x Dsp

-Combined sensitivity Test 2= 1 - (1-Ese) x (1-Fse) x (1-Gse) x (1-Hse)

-Combined specificity Test 2= Esp x Fsp x Gsp x Hsp

Where A-Hse, refer to the individual sensitivity value of each considered miRNA; A-Hsp refer to the individual specificity value.

Step 2) The overall sensitivity and specificity of Tests 1 and 2 are calculated in series using the following formulas:

-Overall sensitivity = T1se x T2se

-Overall specificity = T1sp + T2sp – (T1sp x T2sp)

Note: a subject is classified positive to screening if both Test 1 and Test 2 are positive; Test 2 is run only if Test 1 is positive. For either test, a single significantly dysregulated miRNA is enough to consider the test as positive.

References

1. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. AJR Am J Roentgenol. 2005; 184: 14-9. doi: 10.2214/ajr.184.1.01840014.

2. Fletcher R, Fletcher S. Clinical Epidemiology. Philadelphia [u.a.]: Lippincott Williams & Wilkins; 2009.
3. Zhou XH, Obuchowski NA, McClish DK. Statistical Methods in Diagnostic Medicine. Wiley; New York; 2002.

Supplementary **Table 1A. Main characteristics of the 20 included studies.**

Reference	Year	Country	Ethnicity	Sample Size[a]		Mean Age		Smokers		Comorbidity (case and control selection criteria)		N. of miRNAs Individual/Panel	miRNAs Examined	NSCLC[d] Stage
				Pt[b]	C[c]	Pt	C	Pt	C	Pt	C			
Bianchi et al. [26]	**2011**	Italy	Caucasian	22	30	60	matched	All smokers or former smokers (>20 pack-years)		No history of cancer disease (in the last 5 years)	Healthy controls enrolled as part of the prospective COSMOS project (Low dose-CT[e] always negative during the FU)	34, panel	Panel of 34 miRNAs*	I
Foss et al. [27]	**2011**	USA	Caucasian	31	22	62[#]	62[#]	Smokers: 96,8%	Smokers: 100%	NA[f]	Patients hospitalized for non-neoplastic and non-respiratory conditions (in ophthalmology or orthopedic department)	2, panel	miR-1254 miR-574-5p	I-II
Geng et al. [48]	**2014**	China	Asian	126	60	62.7%≤60y 37.3%>60y	58.3%≤60y 41.7%>60y	Smokers: 42.9% Former smokers: 27.8% Non-smokers: 29.3%	Smokers: 28.3% Former smokers: 18.3% Non-smokers: 53.4%	No other cancer disease	3 control groups: 1. no history of pulmonary disease; 2. Non-cancerous pulmonary disease; 3. benign pulmonary nodules	5, individual	miR-20a, miR-223, miR-21, miR-155, miR-145	I-II
Halvorsen et al. [47]	**2016**	Norway	Caucasian	79	58	62,6	57,6	Smokers: 87% Non-smokers: 2% Unknown: 11%	Smokers: 100% Non-smokers: 0%	NA	IELCA screening trial samples with no cancer at lung CT	6, panel	miR-125b, miR-200b, miR-34b, miR-203, miR-205, miR-429	I-II
Li et al. [43]	**2015**	China	Asian	11	11	59[#]	55[#]	NA		NA	No tumor history	1, individual	miR-486	I- II
Ma et al. [28]	**2013**	USA	African American, Caucasian	36	38	66,7	64,6	Smokers: 49,3% (28 pack-years)	Smokers: 19,6% (13 pack-years)	NA	No tumor history	2, individual and panel	miR-21-5p, miR-335-3p	I

Reference	Year	Country	Ethnicity	Sample Size[a]		Mean Age		Smokers		Comorbidity (case and control selection criteria)		N. of miRNAs Individual/Panel	miRNAs Examined	NSCLC[d] Stage
				Pt[b]	C[c]	Pt	C	Pt	C	Pt	C			
Nadal et al. [45]	**2015**	USA	NA	75	60	65,5[#]	60[#]	Smokers: 94% Non-smokers: 56% Unknown: 0 (0%)	Smokers: 48% Non-smokers: 26% Unknown: 26%	NA	No tumor history (mainly COPD[#], bronchiectasis, pneumonia)	4, individual and panel	miR-141, miR-200b, miR-193b, miR-301	I
Powrozek et al. [46]	**2016**	Poland	Caucasian	29	85	64[#]	57[#]	Smokers: 89.5% Non-smokers: 10.5%	matched	NA	No respiratory disease symptoms and chest X ray negative for lung disorders	2, individual and panel	miR448, miR4478, combination of both miRNA	I-II
Sanfiorenzo et al. [29]	**2013**	France	NA	35	20	65,1	67,5	Smokers: 85%	Smokers: 75%	NA (3 people died for cause other than NSCLC)	2 control groups: 1.COPD patients with no lung cancer or other malignancies; 2. healthy volunteers matched for smoking status	12, panel	Panel of 12 miRNAs**	I-II
Shen et al. [30]	**2011**	USA	African American, Caucasian	30	29	67,4[#] SCC[b] 68[#] AC[l]	66[#]	Smokers:100% (31-35 pack-years)	Smokers:100% (30 pack years)	NA		4, panel	miR-21, miR-126, miR-210, miR-486-5p	I-II
Shi et al. [41]	**2017**	China	NA	46	45	50,1	39,4	Smokers:46,7 Non-smokers: 56,4%	Smokers: 31,1% Non-smokers: 68,9%	NA	Without disease (absence of hypertension, high cholesterol, diabetes)	3, individual	miR-22, miR-125b, miR-15b	I-II
Sun et al. [55]	**2016**	China	Asian	22	60	66,9	62,4	NA		NA		1, individual	miR-21	I-II
Ulivi et al. [49]	**2013**	Italy	Caucasian	54	24	68[#]	65[#]	Smokers: 25.6% Former smokers: 53.5% Non-smokers: 12.8% Unknown: 8.1%	Smokers: 25% Former smokers: 29.2% Non-smokers: 45.8%	No history of previous diseases	No tumor history	1, individual	miR-328	I-II

Reference	Year	Country	Ethnicity	Sample Size[a]		Mean Age		Smokers		Comorbidity (case and control selection criteria)		N. of miRNAs Individual/Panel	miRNAs Examined	NSCLC[d] Stage
				Pt[b]	C[c]	Pt	C	Pt	C	Pt	C			
Wang Y. et al. [50]	**2016**	China	Asian	82	91	59,8	61,4	Smokers: 55.9%	Smokers: 47.8%	Not taking antibiotics or steroid, no HIV, HCV, HBV, no chemotherapy /radiotherapy)	4 control groups: 1. lung benign disease, 2. other subtypes of lung cancer; 3.other adenocarcinomas 4.healthy controls	3, individual and panel	miR-532, miR-628, miR-425-3p	I-II
Wang W. et al. [53]	**2016**	China	NA	54	15	60[#]	51[#]	NA		NA		1, individual	miR-1244	I-II
Wozniak et al. [54]	**2015**	Russia	Caucasian	70	100	62,6	60,1	Smokers: 54% Former Smokers:27% Non-smokers:11%:	Smokers:53% Former S.:11% Non-smokers:36%	NA	Individuals visiting two Moscow general hospitals for disorders unrelated to lung cancer and with associated risk factors	24, panel	Panel of 24 miRNAs***	I-II
Yuxia et al. [42]	**2012**	China	Asian	76	110	43.5%≤60y 56.5%>60y	51.8%≤60y 48.2%>60y	NA		NA	No tumor history	1, individual	miR-125b	I-II
Zhang et al. [51]	**2017**	China	NA	109	63	59,3	59,7	Smokers: 58,7% Non-smokers: 41,3%	Smokers: 42,9% Non-smokers: 57,1%	No previous history of cancer-related disease; no radio- or chemo-therapy prior to surgery	Without tumor-associated lesions confirmed by chest CT, blood test and other full body examinations	4, individual and panel	miR-145, miR-20a, miR-21, miR-223	I-II
Zhu et al. [44]	**2014**	China	Asian	36	48	59	matched	NA		NA		3, individual	miR-29c, miR-93, miR-429	I

Reference	Year	Country	Ethnicity	Sample Size[a]		Mean Age		Smokers		Comorbidity (case and control selection criteria)		N. of miRNAs Individual/Panel	miRNAs Examined	NSCLC[d] Stage
				Pt[b]	C[c]	Pt	C	Pt	C	Pt	C			
Zhu et al. [52]	**2016**	China	Asian	87	40	58,5	57,9	Smokers:35.7% Non-smokers: 64.3%	All non-smokers. 20 smokers individual were also collected as comparison group.	No chemotherapy / radiotherapy, no antibiotic therapy; no COPD, no others chronic conditions	4 control groups: 1. pneumonia; 2. gastric cancer; 3. healthy smokers; 4. healthy non-smokers	4, individual and panel	miR-182, miR-183, miR-210, miR-126	0-I

[a]The sample refers to stage I and II NSCLC

[b]Pt=Patients; [c]C=Controls

[d]NSCLC: non-small cell lung cancer

[e]CT: computed tomography

[f]NA: Not available

[g]COPD: chronic obstructive pulmonary disease

[h]SCC=squamous cell carcinoma; [i]AC=adenocarcinoma

[#]median age

*miR-92a, miR-484, miR-486-5p, miR-328, miR-191, miR-376a, miR-342-3p, miR-331-3p, miR-30c, miR-28-5p, miR-98, miR-17, miR-26b, miR-374a, miR-30b, miR-26a, miR-142-3p, miR-103, miR-126, let-7a, let-7d, let-7b, miR-32, miR-133b, miR-566

**miR-155-5p, miR-20a-5p, miR-25-3p, miR-296-5p, miR-191-5p, miR-126-3p, miR-223-3p, miR-152-3p, miR-145-5p, miR-199a-5p, miR-24-3p, and let-7f-5p

***let-7c, miR-122, miR182, miR193a-5p, miR200c, miR203, miR218, miR 155, let-7b, miR-411, miR450b-5p, miR-485-3p, miR519a, miR-642, miR-517b, miR-520f, miR206, miR-566, miR-661, miR-340, miR-1241, miR-720, miR-543, miR1267.

Supplementary Table 1B: Materials and procedures used for miRNA quantification in the 20 included studies

Reference	Year	Starting material	Study Validation (yes/no)	RNA isolation/ Spike-in	Retrotrascription/ Spike-in	Quantification method	Normalization	Analysis of results	Assessment of hemolysis
Bianchi et al. [26]	**2011**	Serum	Yes	Trizol-LS (Invitrogen) combined with mirVana miRNA Isolation Kit (Ambion)/NA[a]	TaqMan MicroRNA Reverse Transcription Kit & Taqman multiplex RT assays (AB)/NA	TLDA[b] microRNA signature panel (AB[c])	geometric mean of miR-197, miR-19a, miR-19b, miR-146, miR-15b, miR- 24	$2^{-\Delta\Delta Ct}$	NA
Foss et al. [27]	**2011**	Serum-plasma	Yes	Phenol and guanidine thiocianate (75-200 µL serum/plasma)/cel-miR-39	2) GenoExplorer miRNA First-Strand cDNA Core kit (Genosensor corp.)/NA	1) array (Genosensor corp.)(discovery set); 2) qPCR (validation set): Sybr green (Roche) & miR[d] specific forward primer+universal reverse primer)	1) PC-U6B, U6-337, 5S-rRNA, PC-HU5S; 2) RNU6 or cel-miR-39 separately	$1-(C_{tGOI}-C_{tREF})/maxC$[e] (unclear the normalization on the 2 reference miRs)	NA
Geng et al. [48]	**2014**	Plasma	Yes	NA /NA	NA & specific reverse primers /NA	Relative qPCR [MyScript SYBR green PCR kit (Qiagen)]	miR-16	$2^{-\Delta\Delta Ct}$ (where $\Delta Ct= C_{tREF}-C_{tGOI}$)	NA
Halvorsen et al. [47]	**2016**	Serum	Yes	miRCURY RNA isolation kit (Exiqon) /NA	1) TaqMan MicroRNA Reverse Transcription Kit & Megaplex RT primer pool (AB); 2) MicroRNA reverse trascription kit & stem-loop primers (AB)/NA	1) TLDA microRNA signature panel (AB); 2) Relative qPCR (TaqMan Universal PCR Master Mix, no UNG & Taqman small RNA assays (AB)	1) global normalization method; 2) geometric mean of miR-220, miR-19b, sU6, chosen from 1)	$2^{-\Delta\Delta Ct}$	NA
Li et al. [43]	**2015**	Plasma	No	miRNeasy Mini Kit (Qiagen)/cel-miR-39	Taqman MicroRNA Reverse Trascription Kit (AB) & stem-loop primers (AB)/NA	Relative qPCR with Taqman Universal Master Mix II & Taqman primers-probes (AB)	cel-miR-39	$2^{-\Delta\Delta Ct}$	NA
Ma et al. [28]	**2013**	Plasma	No	mirVana PARIS RNA kit (Ambion) /NA	Taqman MicroRNA Reverse Trascription Kit (AB) & stem-loop primers (AB)/NA	Droplet digital PCR (Taqman primers-probes)	absolute quantification	copies/µl plasma	NA
Nadal et al. [45]	**2015**	Serum	Yes	mirVana PARIS RNA kit (Ambion) /NA	TaqMan MicroRNA Reverse Transcription Kit & Megaplex RT primer pool (AB); preamplification/NA	1) Taqman Openarray Human microRNA panel (AB); 2) Relative qPCR (Taqman Master Mix & Taqman primers-probes (AB)	1) average Ct of all miRNAs as loading control; 2) U6 snRNA	$2^{-\Delta\Delta Ct}$	NA

Reference	Year	Starting material	Study Validation (yes/no)	RNA isolation/ Spike-in	Retrotrascription/ Spike-in	Quantification method	Normalization	Analysis of results	Assessment of hemolysis
Powrozek et al. [46]	**2016**	Plasma	No	miRNAeasy serum and plasma kit (Qiagen)/NA	Taqman MicroRNA Reverse Trascription Kit (AB) & stem-loop primers (AB)/NA	Relative qPCR (Taqman Universal Master Mix II with UNG & Taqman primer-probes (AB)	U6 snRNA	ΔCt, $2^{-\Delta Ct}$, $2^{-\Delta\Delta Ct}$	NA
Sanfiorenzo et al. [29]	**2013**	Plasma	No	miRNAeasy Mini Kit (Qiagen)/NA	NA /NA	Relative qPCR (NA)	Mean of miR-192-5p, and miR-16-5p	$\Delta\Delta Ct$ [(C_{tGOI}-C_{tREF})-global mean of relative expression of each miRNA]	Yes (miR-16-5p and miR-103a-3p; miR-16-5p used both as hemolysis indicator and as reference)
Shen et al. [30]	**2011**	Plasma	Yes	mirVana PARIS RNA kit (Ambion) /cel-miR-238	Taqman MicroRNA Reverse Trascription Kit (AB) & stem-loop primers (AB)/ NA	Relative qPCR (NA)	miR-16	$2^{-\Delta\Delta Ct}$	NA
Shi et al. [41]	**2017**	Serum	No	miRNA extraction kit (Tiangen Biology co.) /NA	NA	NA	miR-103	$2^{-\Delta\Delta Ct}$	NA
Sun et al. [55]	**2016**	Serum	No	RNA extraction kit (Applied Biosystem) /NA	Taqman MicroRNA Reverse Trascription Kit (AB) & Taqman microRNA assay/NA	Relative qPCR (Taqman Universal Master Mix & Taqman microRNA assay)	miR-16	$2^{-\Delta\Delta Ct}$	NA
Ulivi et al. [49]	**2013**	Blood	No	Trizol reagent (Invitrogen)/NA	Taqman MicroRNA Reverse Transcription kit (AB) & Taqman microRNA assay/NA	Relative qPCR (Taqman microRNA assay)	RNU38B, RNU58A	$2^{-\Delta\Delta Ct}$	NA
Wang Y. et al. [50]	**2016**	Plasma	Yes	mirVana PARIS kit (Ambion)/cel-miR-39	1) Taqman MicroRNA Reverse Trascription Kit (AB) & Megaplex RT primer pool (AB); preamplification (Megaplex preamp primers)(AB) 2) Taqman MicroRNA Reverse Transcription kit (AB) & Taqman microRNA assay /cel-miR-39	1) TLDA cards (AB) 2) Relative qPCR (Taqman Universal Master Mix & Taqman primer-probes (AB)	cel-miR-39	$2^{-\Delta\Delta Ct}$	NA

Reference	Year	Starting material	Study Validation (yes/no)	RNA isolation/ Spike-in	Retrotrascription/ Spike-in	Quantification method	Normalization	Analysis of results	Assessment of hemolysis
Wang W. et al. [53]	**2016**	Serum	No	miRNAeasy Serum/Plasma Kit (Qiagen)/ cel-miR-39	MyScript reverse transcription kit (Qiagen)	2) Relative qPCR	cel-miR-39	$2^{-\Delta\Delta Ct}$	NA
Wozniak et al. [54]	**2015**	Plasma	No	NucleoSpin miRNA Plasma kit (Macherey-Nagel) /Ath-miR-159a	TaqMan MicroRNA Reverse Transcription Kit & Taqman primers (AB) / NA	1) TLDA cards (AB)		Quantile normalization	Yes (none of the miRNAs described in literature as influenced by hemolysis)
Yuxia et al. [42]	**2012**	Serum	No	no extraction (quantification from serum) /NA	Taqman MicroRNA Reverse Transcription kit (AB) & Taqman microRNA assay (AB)/NA	Relative qPCR (Taqman microRNA assay)	NA	2^{50-Ct}	NA
Zhang et al. [51]	**2017**	Plasma	Yes	Trizol+mirVana PARIS kit (Ambion)/NA	Taqman MicroRNA Reverse Trascription Kit (Ambion Life Technologies) & stem&loop primers	Relative qPCR (Brilliant III Ultra-Fast SYBR-Green qPCR master mix kit (Ambion Life Technologies)	miR-16	$2^{-\Delta\Delta Ct}$	NA
Zhu et al. [44]	**2014**	Serum	No	mirVana PARIS kit (Ambion) OR miRNA isolation kit (AB) /NA	Taqman MicroRNA Reverse Transcription Kit (AB) & Taqman probes/NA	Relative qPCR (Taqman Universal Master Mix & Taqman primer-probes (AB)	U6 and U48 snRNA	$2^{-\Delta\Delta Ct}$	NA
Zhu et al. [52]	**2016**	Serum	No	mirVana PARIS kit (Ambion)/NA	Taqman MicroRNA Reverse Trascription Kit Applied Biosystems & Taqman probes/NA	Relative qPCR (Taqman Universal Master Mix & Taqman primer-probes (AB)	U6 snRNA	$2^{-\Delta\Delta Ct}$	NA

[a] NA: Not available

[b] TLDA: Taqman Low Density Array

[c] AB: Applied Biosystems

[d] miR: miRNA

[e] max C: number of maximum cycles

2) **Screening with serum miR-223 and miR-29c for early diagnosis of Non Small Cell Lung Cancer**

D'Antona Paola, Cattoni Maria, Cinquetti Raffaella, Poli Albino, Moretti Francesca, Daffrè Elisa, Douglas Noonan, Imperatori Andrea, Rotolo Nicola, Campomenosi Paola.

Abstract

Diagnosing non-small cell lung cancer (NSCLC) at an early stage is a major requirement for increasing patients' survival, but it remains challenging due to the lack of specific and robust non-invasive biomarkers. The discovery of circulating microRNAs (miRNAs) in the bloodstream has opened new perspectives for tumor diagnosis. A critical review of the literature allowed us to identify a panel of 8 miRNAs, uninfluenced by hemolysis, for a two-step screening of early lung cancer based on 4 miRNAs with high sensitivity (hsa-miR-223-3p, hsa-miR-20a-5p, hsa-miR-448, hsa-miR-145), and 4 miRNAs with high specificity (hsa-miR628-3p, hsa-miR-29c, hsa-miR-210, hsa-miR-1244). We quantified six miRNAs from our panel in stage I-II NSCLC patients and in three different control groups [non-smokers; smokers; Chronic Obstructive Pulmonary Disease (COPD) patients], since smoking habit and COPD may influence miRNA serum levels. The droplet digital PCR method was applied for quantification of miRNAs. Two of the eight miRNAs were excluded during optimization, because needed special conditions for their optimal quantification.

For 3 of the 4 remaining miRNAs (miR-20a, miR-29c and miR-210) analyzed there was no significant difference among control subgroups (non-smokers, smokers, COPD patients), whereas miR-223 was significantly higher in non-smokers. Levels of all tested miRNAs significantly differed between tumor and control groups, confirming their possible role as biomarkers. Among them, miR-223 and miR-29c had the best AUC and their measures showed to be either repeatable or stable in time. Therefore, the selected miRNAs may help to identify high risk subjects who need further investigation for the presence of early stage NSCLC; in particular, the combination of “miR-223 OR miR-29c” showed a good sensitivity (92,5%) in distinguishing tumor samples from controls.

Introduction

Lung cancer is the most common cause of cancer-related death in the world, with over 1.5 million deaths per year [Ferlay et al., 2013; Siegel et al., 2018].

Histologically, lung cancer is classified as small cell lung cancer (SCLC), a very aggressive and rarely curable tumor including approximately 15% of lung cancer cases, and non-small cell lung cancer (NSCLC), which represents approximately 85% of all lung cancers. Patients diagnosed with early stage (stage I and II) NSCLC have a better prognosis than those diagnosed in more advanced stages (60-80% versus 15% survival at 5 years, respectively) but unfortunately this only occurs in 20-25% of cases [Dominioni et al., 2000; Ganti et al., 2006]. The remaining patients are diagnosed at advances stages, sometimes during routine tests (chest X-ray or sputum cytology) that have however demonstrated limited sensitivity and several limitations [Flehinger et al., 1994, Dominioni et al., 2013]. Early detection of lung cancer is key to improve survival.

Thus, the development of novel, sensitive and non-invasive methods for screening of NSCLCs is strongly needed. Recently, microRNAs (miRNAs) have been suggested as a new class of tumor biomarkers, because the circulating levels of some miRNAs are altered in various human cancers, including NSCLC [Esquela-Kerscher et al., 2006, Iorio et al., 2012, Negrini et al., 2014].

MicroRNAs are a class of non-coding small RNAs of up to 24 nucleotides in length that are very stable in the blood [Mitchell et al., 2008; Markou

et al., 2013]; in 2008 Mitchell et al. first reported that serum miRNA-141 was upregulated in prostate cancer, suggesting that it could distinguish prostate cancer patients from healthy controls [Mitchell et al., 2008]. MiRNA molecules have important functions in different biological processes, including cell proliferation, differentiation and apoptosis and can post-transcriptionally regulate the expression of more than 30% of human protein-coding genes [Bueno et Malumbres, 2011; Ambros, 2003].

MiRNA presence and stability in biofluids, together with the demonstration that some of them correlate with clinical-pathological parameters and prognosis of lung cancer, are strong points suggesting their application as lung cancer circulating biomarkers [Blondal et al., 2013].

Although the results of many recent miRNA studies generate hope for practical application, the clinical transferability of the obtained data is uncertain because there is incoherence in miRNA signatures identifying the same disease; this is possibly due to scarce reproducibility of the methods for determining the circulating miRNA levels, the most commonly used method being real time quantitative PCR (qPCR). The accuracy of miRNA determination is affected by multiple variables, including sample storage and preparation, RNA isolation, hemolysis, retrotranscription, DNA polymerase inhibitors, quantification method and normalization [Moretti et al., 2017]. The droplet digital PCR (ddPCR) technique is currently suggested as a better method than qPCR to quantify miRNAs; ddPCR has superior precision and sensitivity compared to the qPCR method, is less affected by PCR inhibitors, and it does not need

internal/external normalization while detecting low concentrations of target nucleic acids molecules [Campomenosi et al., 2016].

To select miRNAs useful for lung cancer screening among the numerous miRNAs published as early stage NSCLC biomarkers, we previously performed a critical review of the pertinent scientific literature and identified the miRNAs showing either high sensitivity or high specificity as NSCLC biomarkers (sensitivity > 80%; AUC > 0,8; specificity > 90%) [Moretti et al., 2017]. Two miRNA panels were identified: the first panel [composed of high sensitivity miRNAs (hsa-miR-223-3p, hsa-miR-20a-5p, hsa-miR-448, hsa-miR-145)] should be used to identify subjects with high risk of having lung cancer; to eliminate false positive results, the second miRNA panel [composed of high specificity miRNAs (hsa-miR-628-3p, hsa-miR-29c, hsa-miR-210, hsa-miR-1244)], would be used to select true positive samples; patients positive for both panels should undergo further examination by Computed Tomography (CT) scan imaging [Moretti et al., 2017].

In order to currently apply the miRNAs detection to the screening process, the analysis of 8 miRNAs appears too expensive to be extended to the entire at risk population and then it needs simplification: a high sensitivity test with 2 miRNAs (1 with high sensitivity and 1 with high specificity) would be more easily applicable for very large population screenings.

The objective of this study is to apply the ddPCR technique to explore the performance of miRNAs identified in our panels in discriminating patients with early NSCLC from control subjects and select the best combination of 2 miRNAs that could be practically used for non-invasive screening of lung cancer in the clinical setting.

Another critical issue about the use of miRNAs as lung cancer biomarkers is to clarify which subjects should be included in the group of "controls". The composition of the control group is a controversial point because chronic disease, such as Chronic Obstructive Pulmonary Disease (COPD), and other important factors, such as history of smoking, cardiovascular disorders, liver disease, diabetes, pregnancy and inflammatory processes may affect circulating miRNA levels and may alter the evaluation of miRNAs candidate as biomarkers. In nearly all the relevant publications on this topic, the tumor group and the control group are typically matched by age, gender and rarely by smoking status, with scarce or no details about the subjects' clinical history [Moretti et al., 2017; Zandberga et al., 2013].

Therefore, an additional goal of this study was to compare the levels of the selected miRNAs in patients with stage I-II NSCLC and in three different control subgroups (non-smokers, smokers and subjects with COPD).

Materials and methods

Study design

Of the 8 circulating miRNAs previously identified with our critical review of the literature as potential lung cancer biomarkers [Moretti et al., 2017], 6 miRNAs were tested in this study, namely hsa-miR-223-3p, hsa-miR-20a-5p, hsa-miR-448, hsa-miR-628-3p, hsa-miR-29c, hsa-miR-210. Conditions for amplification by ddPCR were set-up for each miRNA candidate

(Supplementary file 1). **In the first part** of the work, **("miRNA selection phase",** summarized in Figure 1), we aimed to select the 2 best performing miRNAs for stage I-II NSCLC diagnosis, out of the 6 above mentioned: one highly sensitive miRNA and one highly specific miRNA. For this purpose, the levels of the 6 miRNAs were measured in serum samples from 91 stage I-II NSCLC patients and 157 tumor-free control individuals (non-smokers, smokers and COPD subjects). The miRNA measurements in NSCLC stage I-II patients were compared with those in each control subgroup. Moreover, the results in the control subgroups were compared with one another. Control subgroups that showed similar levels of a specific miRNA were pooled together for comparison with measurements obtained in the lung cancer patients.

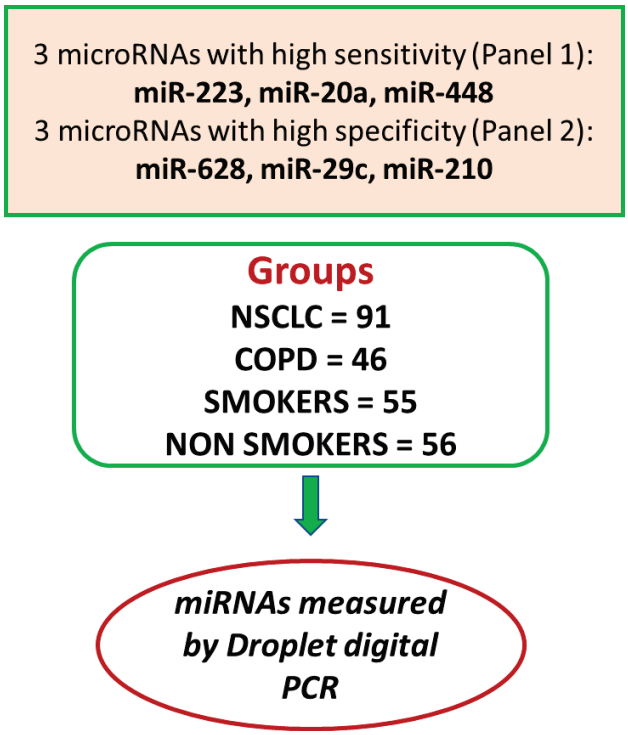


Figure 1: *Design of the first part of the study: the "miRNA selection phase".*

In the second part of the work (Testing phase), we first quantified the two best performing miRNAs (one highly sensitive and one highly specific for stage I-II NSCLC diagnosis) in a set of 80 serum samples, the "Training Set" (40 stage I-II NSCLC patients and 40 controls subjects, comprising 20 non-smokers and 20 smokers). We used the ROC curves analysis to select the miRNAs and the cut-off values for each miRNA that better discriminated NSCLC stage I-II from control sera. We then applied the selected cut-off values to classify the subjects in a blind "Validation Set", an independent new set of 80 serum samples (composed of 40 NSCLC patients and 40 controls subjects comprising 20 non-smokers and 20 smokers) [Figure 2]. This allowed us to evaluate the ability of the cut-off values identified for each miRNA to discriminate NSCLC from control sera.

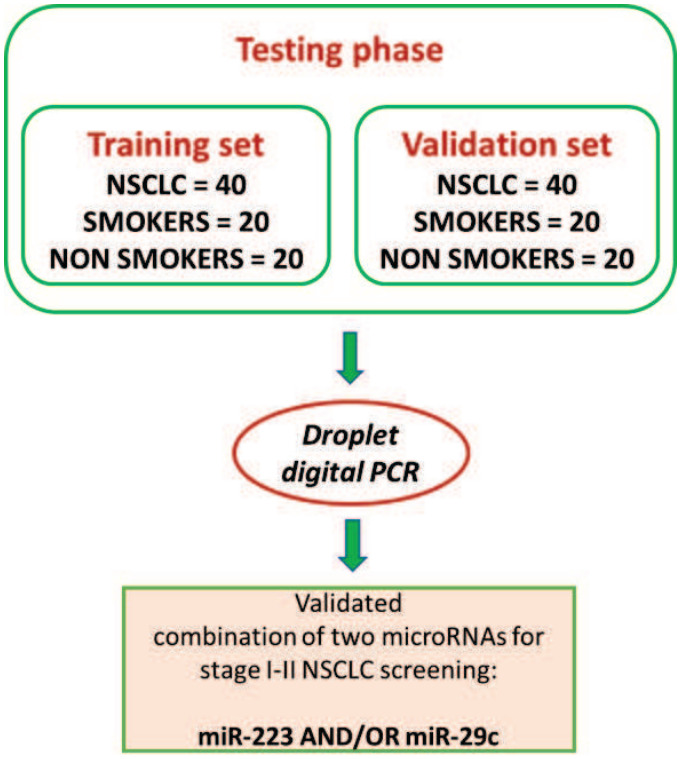


Figure 2: *Design of the second part of the study (Testing Phase).*

Study population and serum samples

This study was conducted on a total of 248 serum samples from individuals of both sexes (150 male; 98 female), aged ≥60 years. Samples were collected at the Ospedale del Circolo, Varese, Italy between January 2014 and March 2018. These samples included 91 early stage NSCLC cases (stage I-II) and 157 controls without history of cancer [56 smokers, 55 non-smokers and 46 individuals with COPD]. The patients' and controls' demographics and clinico-pathological data are shown in Table 1.

	SELECTION PHASE				TESTING PHASE		
		CONTROLS (CANCER-FREE SUBJECTS, N= 157)				CONTROLS (CANCER-FREE SUBJECTS, N=80)	
	STAGE I-II NSCLC PATIENTS	COPD	SMOKERS	NON SMOKERS	STAGE I-II NSCLC PATIENTS	SMOKERS	NON SMOKERS
	N=91	N=46	N=55	N=56	N=80	N=40	N=40
Gender							
Male	72	28	28	22	63	25	29
Female	19	18	27	34	17	15	11
Age (mean)	67	71	62	67	68	63	68
Smoking status							
Former/current smokers	49	13	22		42	29	
Pack years (mean)	46	24	44		43	29	
Ex smokers	30	31	15		27	10	
Unknown	12	2	18		11	1	
Histology							
Adenocarcinoma	70				61		
Squamous cell carcinoma	21				19		
Tumor stage							
IA/IB	71				62		
IIA/IIB	20				18		

Table 1: Characteristics of the subjects in the “miRNA selection phase” and “testing phase”.

NSCLC = Non Small Cell Lung Cancer; COPD = Chronic Obstructive Pulmonary Disease.

All samples from cancer cases were collected before initiation of anti-cancer therapy. Written consent to collect samples and to participate in the study was obtained from all subjects. This study was approved by the Varese University Hospital Ethical Committee.

Peripheral blood samples were collected in 5 mL sterile tubes (BD Vacutainer) without anticoagulant. Samples were drawn early in the morning and left at room temperature for a minimum of 40 minutes to a maximum of 90 minutes. Then the sera were separated by centrifugation at 800 x g for 8 min at room temperature, subdivided in 500 µl aliquots and stored at -80°C until use. Due to the low concentration of miRNA molecules in serum samples, we always operated with constant volumes in all subsequent procedures.

RNA extraction and Reverse transcription

Extraction of total RNA from serum samples was performed using the miRNeasy Serum/Plasma kit (Qiagen); 200 µl of serum were used and manufacturer’s instructions were followed. Spike-in mix (Exiqon) was added before the starting process, to check for loss of material during the whole procedure. 1 µg of MS2 phage carrier RNA (Sigma Aldrich) was also added to each serum sample to improve miRNA extraction. RNA was

eluted from the column with 14 µl of nuclease-free water and stored at -80°C.

For quantitative detection of miRNAs by qPCR or droplet digital PCR (ddPCR), purified serum miRNA was first converted to cDNA by reverse transcription (RT) using miRCURY LNA™ Universal RT microRNA PCR System (Exiqon). The RT reaction was set-up in a total volume of 10 µl, consisting of 4.5 µl of nuclease free water, 2 µl of 5X reaction buffer, 0.5 µl of UniSp6/cel-miR-39-3p RNA spike-in to evaluate the efficiency of the reverse transcription step, 1 µl of enzyme mix and 2 µl of RNA template.

Evaluation of hemolysis

All samples were checked for hemolysis prior to further analysis. Hemolysis can affect the levels of some miRNAs, deriving from blood cell lysis, thus causing errors in miRNA quantification [Pritchard at al., 2012]. Although the miRNAs under investigation have been selected also by their reported insensitivity to hemolysis, we excluded hemolysed sera during this development phase, to avoid this confounding factor; accordingly, all samples included in the "miRNA selection phase" and in the "Testing phase" were hemolysis-free.

To assess hemolysis we quantified by qPCR two miRNAs, respectively localized inside and outside red blood cells, and calculated the difference between the Cq of hsa-miR-451 (a miRNA highly expressed in red blood cells), and that of hsa-miR-23a-3p (a free miRNA unaffected by hemolysis) for each sample.

For each miRNA to be measured, 4 µl of cDNA template diluted 1:40 were

used in a 10 µl of reaction, adding 5 µl of ExiLENT SYBR® Green PCR Master Mix (Exiqon) and primers. Triplicate reactions were performed for each sample; we used manufacter's instruction for cycling conditions (95° for 10 min, followed by 40 cycles of 95°C for 10 sec and 60°C for 1 min (1.6 °C ramp rate).

Samples were considered at risk of hemolysis when their ΔCq was > 5.

Droplet Digital PCR and miRNA absolute quantification

The cDNAs were diluted 20 fold and the ddPCR reaction was prepared in a 20 µl volume, by adding the appropriate volume of cDNA (Supplementary file 1), 10 µl 2X Evagreen supermix (Biorad), the desired miRCURY LNA PCR primer set at the appropriate dilution (see Supplementary file 1), and nuclease-free water up to 20 µl. Preliminary tests were carried out to find the optimal conditions for ddPCR analysis, as described in Supplementary file 1. Each 20 µl ddPCR reaction was loaded onto an 8-channel droplet generation cartridge (Biorad) and placed into the QX200 Droplet Generator, that partitions samples into 20,000 nanoliter-size droplets, by creating an emulsion with the 70 µl of oil reagent (Biorad) that is added into the appropriate oil well; all droplets have the same volume and can randomly contain or not the target of interest. The resulting emulsified reactions were trasferred to a 96-well plate (Biorad) with a multichannel pipette (Rainin) in the thermal cycler and the plate sealed with Pierceable foil (Biorad). The cycling conditions were: 95°C for 5 min, followed by 40 cycles of 95°C for 30 s and 60°C for 1 min, followed by signal stabilization steps (4°C for 5 min, 90°C for 5 min).

The ramp rate was 2°C/s. All samples were run in duplicates.
After PCR, plates were placed into the QX200 Droplet Reader (Biorad) for analysis: each sample is taken up and the fluorescence of each droplet is read; the QuantaSoft™ software counts the number of positive (target containing) and negative (not target containing) droplets for each sample. Based on the Poisson distribution and keeping into account the fraction of positive *versus* negative droplets, the absolute copy number/µl of the DNA target molecules in the initial volume of reaction is determined.

Statistical analysis and ROC curves

Significance of the differences in miRNA serum levels among groups was tested by the Mann-Whitney test. For each of the tested miRNAs the Receiver Operating Characteristic (ROC) curve was constructed and the Area Under the Curve (AUC) calculated. A p value < 0.05 was considered statistically significant.

Results

First part: miRNA selection phase

In our previous critical review of the literature relevant to application of miRNAs as early lung cancer biomarkers, we proposed a two-step screening model based on eight miRNAs: four microRNAs with high

sensitivity for detecting stage I-II NSCLC (miR-223, miR-20a, miR-448 and miR-145) for the first step; four miRNAs with high specificity (miR-628, miR-29c, miR-210 and miR-1244) for the second step (Table 2) [Moretti et al., 2017]. In clinical practice, the two-step lung cancer screening should be offered to high risk individuals; we hypothesize that individuals testing positive (showing significantly aberrant miRNA level compared to controls should undergo further examinations (chest CT imaging) for suspect lung cancer.

Table 2. *MicroRNAs with high sensitivity and high AUC (a) and miRNAs with high specificity (b) selected in our critical review [Moretti et al (2017)].*

a) miRNAs with sensitivity > 80% and AUC > 0.80

miRNA	Sensitivity (%)	AUC	Specificity (%)	Reference
miR-223	87.0	0.94	86.0	Geng et al., 2014 [48]
miR-20a	83.0	0.89	81.0	Geng et al., 2014 [48]
miR-448	85.0	0.89	77.0	Powrozek et al., 2016 [46]
miR-145	80.6	0.89	89.2	Zhang et al., 2017 [51]

b) miRNAs with specificity > 90%

miRNA	Sensitivity (%)	AUC	Specificity (%)	Reference
miR-628-3p	42.7	0.73	91.2	Wang Y et al., 2016 [50]
miR-29c	50.0	0.73	95.8	Zhu et al., 2014 [44]
miR-210	35.6	0.65	100.0	Zhu et al., 2016 [52]
miR-1244	53.8	0.80	100.0	Wang W et al., 2016 [53]

Of these eight miRNAs, only 6 (3 with high sensitivity: miR-223, miR-20a, miR-448; 3 with high specificity: miR-29c, miR-210, miR-628) could be measured up to this time. In order to find the conditions yielding the best resolution in ddPCR for each of the six microRNAs, we first made a pilot experiment on a small number of tumor and control samples (10 NSCLC and 10 control subjects), working on different primer concentrations, different cDNA volumes and different annealing temperatures. The results

of this preliminary analysis are shown in Supplementary file 1, figures 1.1-1.5. Two of the six analyzed microRNAs, namely miR-448 and miR-628-3p, showed very scarce separation between positive and negative droplets (Supplementary file 1, figures 1.5) despite extensive modifications in working conditions, except when we changed the annealing temperature. However, use of specific annealing temperatures for each microRNA would be impractical for a large scale screening of individuals at risk for NSCLC. Therefore, miR-448 and miR-628-3p were excluded and the following four miRNAs were considered for subsequent analysis: miR-223, miR-20a, miR-29c and miR-210.

Comparison of miR-223, miR-20a, miR-29c and miR-210 levels in stage I-II NSCLC patients and in controls

Two miRNAs with high sensitivity (miR-20a, miR-223), and two with high specificity (miR-29c and miR-210) for early NSCLC diagnosis were measured in the serum of 157 control group subjects (55 non-smokers, 56 smokers, 46 COPD subjects) and 91 patients with stage I-II NSCLC.
These miRNAs were quantitatively measured by the ddPCR method and the significance of difference between miRNA levels of stage I-II NSCLC patients and controls was analyzed by the Mann-Whitney test. The results are summarized in the following scatter plots (Figure 3).

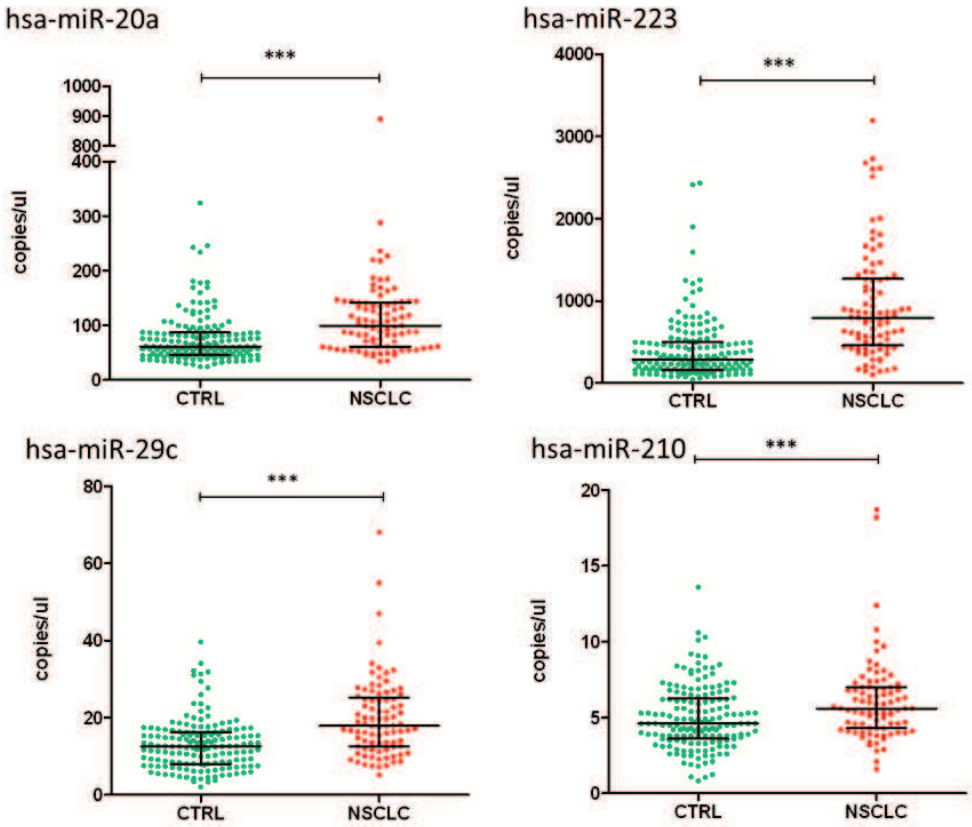


Figure 3. *Scatter plot showing the concentration of the four selected miRNAs that indicated significant differences between stage I-II NSCLC patients (NSCLC, n=91) and control group (CTRL, n=157). MiRNAs were quantified by ddPCR and data were expressed as median value with interquartile range (*** p<0.001, Mann-Whitney test).*

All four miRNAs (miR-20a, miR-223, miR-29c and miR-210) significantly discriminated between controls (green) and lung cancer patients (red); for all four miRNAs a significantly higher copy number was measured in the serum of NSCLC patients compared to cancer-free controls. The AUC values in distinguishing NSCLC patients from controls were respectively

0.71 for miR-20a, 0.80 for miR-223, 0.72 for miR-29c and 0.63 for miR-210 (Figure 4).

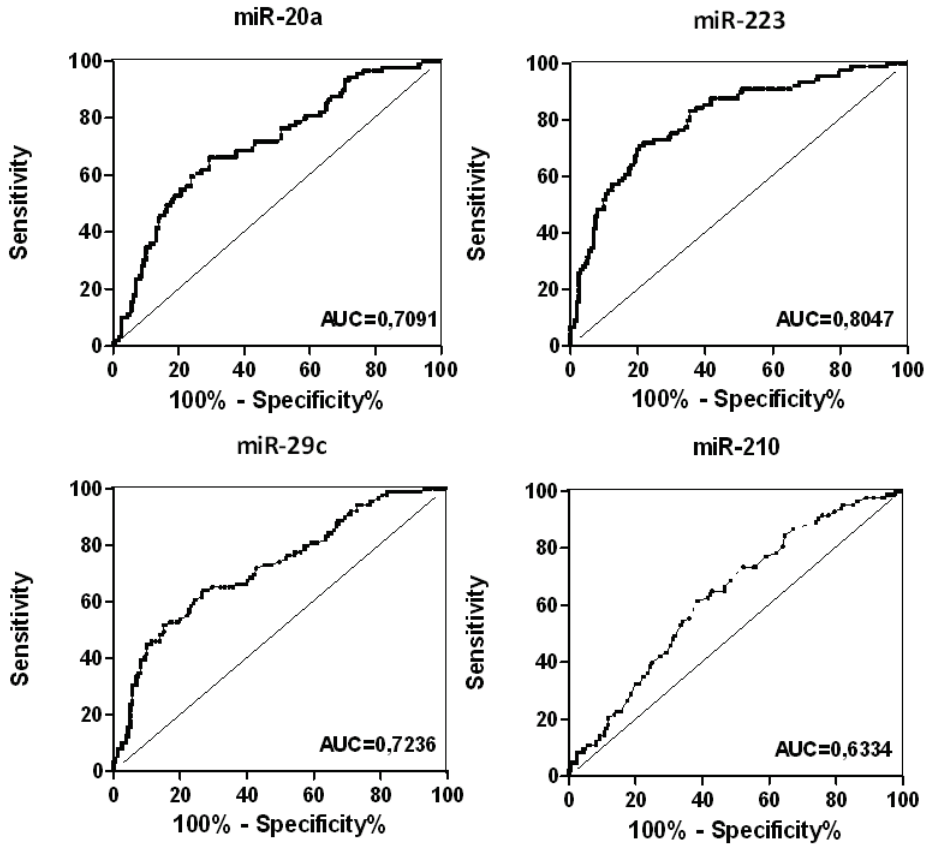


Figure 4. *Diagnostic value of miR-20a, miR-223, miR-29c and miR-210 quantified by droplet digital PCR in serum. ROC curve and AUC were used to determine sensitivity and specificity of each of these four miRNAs.*

Accordingly, among the high-sensitivity miRNAs, miR-223 performed best; among high- specificity miRNAs, miR-29c performed best. Therefore, we decided to use miR-223 and miR-29c for the second part of this work, the "Testing Phase". Moreover, we carried out a comparison across control subgroups (non-smokers, smokers, COPD) to explore the possible differences of miRNA expression in these subgroups. These comparisons

are described in Supplementary file 2. We found no significant differences across the 3 control subgroups for miR-20a, miR-210 and miR-29c, whereas miR-223 was significantly overexpressed in the non-smoker subgroup (Supplementary File 2, Figure 2.1).

Before proceeding to the Testing phase with miR-223 and miR-29c, we verified their stability and repeatability of measures. MiR-223 showed a very low intra- and inter-assay Coefficient of Variation, whereas miR-29c had a very high Coefficient of Variation. Conversely, when we analyzed stability, by measuring miR-223 and miR-29c in the same cDNA samples at 0, 2, 9 months, we observed that miR-29c was stable (the values obtained at the three timepoints were not significantly different), whereas for miR-223 the first measure (time 0) was significantly different from measure at 2 and at 9 months.

Testing phase

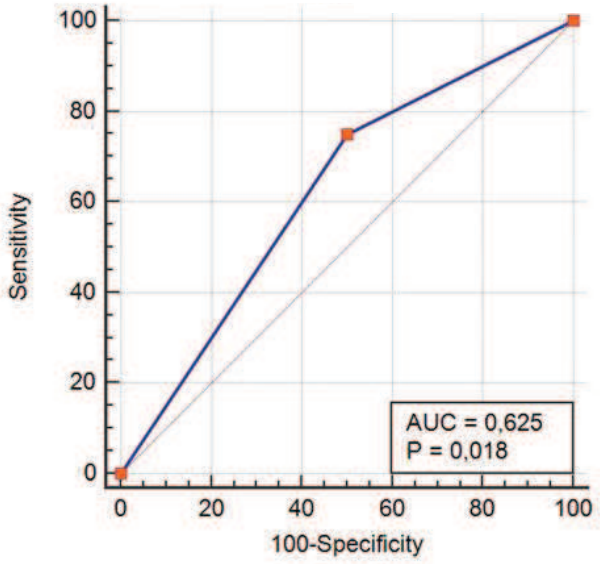
Eighty stage I-II NSCLC serum samples and 80 control samples were randomly split into a "Training set" (40 tumors and 40 controls) and a blind "Validation set" (40 tumors and 40 controls).

Training set

Analysis of ROC curves of combinations of "miR-223 AND/OR miR-29c" were investigated to identify the cut-off values best discriminating lung cancer patients from controls in the training set of 40 lung cancer patients and 40 controls (20 non-smokers and 20 smokers). The combination with

the highest sensitivity (75%) and acceptable specificity (50%) was: “miR-223 > 500 copies/µl OR miR-29c > 50 copies/µl” (Figure 5).

A



B

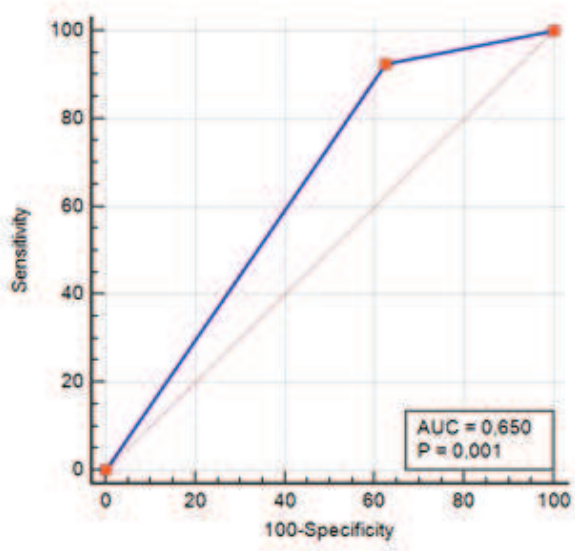
Cut-off	Sensitivity	95% CI	Specificity	95%CI
miR-223>500 copies/µl OR miR-29c>50 copies/µl	75%	0,59-0,87	50%	0,34-0,66

Figure 5: *(A) ROC curve of miR-223 in combination “OR” with miR-29c, showing discrimination between control subjects (smokers and non-smokers) and NSCLC stage I-II; (C) Sensitivity and specificity of “miR-223 OR miR-29c” combination with the selected cut-off: miR-223 > 500 copies/µl OR miR-29c >50 copies/µl.*

Validation set

The findings of the Training set were tested in a new, independent, blind, validation set including sera from 40 lung cancer patients and 40 controls (20 non-smokers and 20 smokers). The cut-off values applied for this Validation were: miR-223 > 500 copies/µl OR miR-29c >50 copies/µl.

A



B

Cut-off	Sensitivity	95% CI	Specificity	95%CI
miR-223>500 copies/µl OR miR-29c>50 copies/µl	92,5%	0,80-0,98	37.5%	0,23-0,54

Figure 6. *(A) Validation set ROC curve of miR-223 in combination "OR" with miR-29c, showing discrimination between control subjects (smokers and non-smokers) and NSCLC stage I-II; (C) sensitivity and specificity of "miR-223 OR miR-29c" combination with the selected cut-off: miR-223 > 500 copies/µl OR miR-29c >50 copies/µl.*

As shown in figure 6, in the Validation set the cut-off values miR-223 > 500 copies/ul OR miR-29c >50 copies/ul yielded an AUC of 0.65, with a high sensitivity of 92.5% and a low specificity of 37.5%.

Discussion

Although NSCLC has one of the highest mortality rates among cancers, it has long been known that patients are more likely to survive if they are diagnosed at an early stage (stage I-II NSCLC). Early lung cancer diagnosis was shown to lead to a drop in disease-specific mortality rate, from 75% to 15% [The National Lung Screening Trial Research Team, Gyoba et al., 2016].

However, new screening tools and diagnostic methods are needed to reach this goal.

Evidence has been provided of aberrant expression of miRNAs in tumors and serum samples of lung cancer patients, compared to normal subjects; such miRNA dysregulation could possibly be exploited for early diagnosis [Fan et al., 2018]. Hence, an increasing number of studies have investigated the diagnostic value of miRNAs quantification in biofluids for lung cancer screening [Del Vescovo et al., 2014].

Unfortunately, many different and inconsistent miRNA signatures of lung cancer have been published. There are several possible explanations for these discrepancies. First, the technical procedure used for miRNA quantification could influence the outcome. The starting material can be plasma, serum or whole blood; moreover, for almost all published studies, the qPCR method has been used for quantification of

circulating miRNAs and there are several pitfalls in this procedure; especially important is the normalization issue; moreover, qPCR is not sensitive enough to point out relatively small differences between samples [Campomenosi et al. 2016].

For these reasons, for measuring miRNA levels we decided to apply the droplet digital PCR technique, that overcomes the normalization issue of qPCR, and shows advantages such as greater sensitivity and precision, and absolute quantification. Whale and colleagues recently showed that the high-accuracy measurements using ddPCR will support the implementation and traceable standardization of molecular diagnostic procedures needed for advancements in precision medicine [Whale et al., 2018].

Second, ethnicity could also play a role. The majority of published works on miRNAs in lung cancer were performed in Chinese cohorts and, at present, no data are available about differences of miRNA expression due to ethnic origin.

Third, the stage of cancer may affect miRNA levels. Our cohort is exclusively composed of early stage (I and II) NSCLC, while many previously published studies included patients with advanced lung cancer. This may affect the levels of specific miRNA and therefore their identification. For example, Shell et al. showed that let-7 expression could be used as a marker of disease stage in several types of cancer [Shell et al., 2007]; we cannot rule out that the levels of other microRNAs may vary depending on lung cancer stage.

Finally, the composition of the control cohort to be used for comparisons with the tumor cohort is a relevant issue. In our critical review of the

literature, we observed that the “healthy controls” group composition was inconsistent among studies. Sometimes, within the same study heterogeneous healthy subjects, either non-smokers, smokers, or individuals affected by pulmonary diseases such as COPD, were used (Halvorsen et al., 2016; Sanfiorenzo et al., 2013; Nadal et al., 2015).

It is well known that cigarette smoking is a high-risk factor for lung cancer [Shields et al. 1999] and for Chronic Obstructive Pulmonary Disease (COPD) [Takahashi et al 2013]. Specific microRNAs have been reported to be dysregulated in smoking-related diseases [Banerjee & Luketich, 2012].

Less attention has been paid to smoking-induced molecular alterations in individuals defined as “healthy smokers” without evidence of disease; they were categorized as healthy control group in many studies. Extensive work should be done on characterizing miRNA dysregulation in those patients who have a smoking history with smoking related diseases (COPD patients) and with lung cancer, trying to find the differences in comparison with healthy non-smokers, because the relationship between circulating miRNAs and cigarette smoke-induced lung cancer is still unclear. Zhou and colleagues showed that “healthy smokers” cannot be considered completely healthy subjects [Zhou et al., 2016].

For these reasons, we explored the changes in serum miRNAs in three subgroups of possible controls, namely non-smokers, smokers and COPD patients. We found that the levels of miR-223 was different across these control subgroups. In particular, miR-223 was down-regulated in smokers and COPD, compared to non-smokers. According to Schembri et al., smoking causes the downregulation of many miRNAs in bronchial

epithelium [Schembri et al. 2009]. Downregulation of several miRNAs was also observed in the lungs of rats treated with typical components of cigarette smoke, such as tobacco-specific nitrosamines [Kalscheuer et al. 2008].

When we compared the levels of the miRNAs under investigation (miR-210, miR-20a, miR-29c and miR-223) in sera from cancer-free controls and NSCLC, we observed an increase in the levels of all four miRNAs in NSCLC patients compared to controls, in keeping with published work [Chen et al. 2012, Shen et al. 2011, Puissegur et al., 2011, Babu et al., 2016, Yang et al., 2015, Zhu et al., 2014, Zhang et al., 2016, Geng et al., 2014].

As we showed in our previous review of the literature, miR-223 and miR-20a represent highly sensitive miRNAs for early lung cancer diagnosis, whereas miR-29c and miR-210 are highly specific. In the first part of this study we found that miR-223 and miR-29c had the greatest AUC among highly sensitive and highly specific miRNAs, respectively. Therefore, we used combinations of miR-223 and miR-29c for the second part of this work, the “Testing Phase”.

In the Testing phase we explored the diagnostic performance (sensitivity, specificity and AUC) of both combinations “miR-223 AND/OR miR-29c” with different cut-offs in discriminating lung cancer patients from controls. By testing the different combinations, we found that “miR-223 OR miR-29c” had the highest sensitivity (75%), combined with good specificity (50%) in the training set. When then applied to an independent, blind validation set, the combination “miR-223 > 500 copies/µl OR miR-29c > 50 copies/µl” yielded an AUC of 0.65, sensitivity of 92.5% and

specificity of 37.5%. These results confirm the high sensitivity of the combination “miR-223 OR miR-29c”, which would effectively identify stage I-II NSCLC patients. The low specificity (37.5%) indicates that a large number of false positive findings would be expected; however, from the viewpoint of a clinical application for lung cancer screening, the importance of high sensitivity outweights the low specificity. In a screening program it is desirable to have a high sensitivity of the test which ensures the identification of the largest number of lung cancer cases, so that they can be early treated, while false positives are then highlighted by the second-level investigations (low-dose CT of the chest).

According to the literature, miR-223 plays a role as an oncomiR and is overexpressed in gastric cancer compared to normal gastric mucosa, in T-cell acute lymphoblastic leukemia [Mavrakis et al., 2011] as well as in a gastric cancer cell line [Li et al., 2012] and in prostate cancer cell lines [Wei et al., 2014]. Previous studies suggested an oncogenic role for miR-29c. miR-29c level was significantly increased in sera and tissues of stage IA/B NSCLC patients compared to corresponding control sera and noncancerous tissues [Heegard et al., 2012; Zhu et al., 2014]. The high stability of miRNAs in human biological fluids suggests that they could be ideal biomarkers for non-invasive diagnosis of NSCLC. A non-invasive test able to discriminate between tumor and control groups could have two potential uses. First, it could be a new screening method to select individuals at high risk of NSCLC (i.e. smokers) who need further clinical investigations; second, it could help to distinguish between neoplastic and non-neoplastic disease in individuals with suspicious nodules detected by

CT scan, thus avoiding further CTs or invasive biopsy. The findings of this study are preliminary and need to be confirmed with other validation sets. Moreover, in our critical review [Moretti et al., 2017] we initially identified two other miRNAs (miR-145 and miR-1244, respectively, as sensitive and specific biomarkers of lung cancer) that remain to be tested. It is necessary to explore whether miR-1244, a biomarker with reported high specificity for lung cancer diagnosis, could improve the currently low specificity of our model.

Acknowledgements

We wish to thank Mrs. Carlotta Biasini and Fondazione Comunitaria del Varesotto (FOCOVA) for contributing to funding this research project. PDA is PhD students of the "Biotechnology, Biosciences and Surgical Technology" course at Università degli Studi dell'Insubria.

Conflict of interest

The authors declare that they have no conflicts of interest.

Rererences

- Ambros V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell. 2003; 113(6):673-676.
- Babu KR, Muckenthaler MU. miR-20a regulates expression of the iron exporter ferroportin in lung cancer. J Mol Med (Berl). 2016; 94(3):347-359.
- Banerjee A, Luettich K. MicroRNAs as potential biomarkers of smoking-related diseases. Biomark Med. 2012 Oct;6(5):671-84.
- Blondal T, Jensby Nielsen S, Baker A, Andreasen D, Mouritzen P, Wrang Teilum M, Dahlsveen IK. Assessing sample and miRNA profile quality in serum and plasma or other biofluids. Methods. 2013; 59(1): S1-6.
- Bueno M and Malumbres M. MicroRNAs and the cell cycle. BBA Molecular Basis of Desease. 2011; 1812 (5): 592-601.
- Campomenosi P, Gini E, Noonan DM, Poli A, D'Antona P, Rotolo N, Dominioni L, Imperatori A. A comparison between quantitative PCR and droplet digital PCR technologies for circulating microRNA quantification in human lung cancer. BMC Biotechnol. 2016; 16(1):60.
- Chen X, Hu Z, Wang W, Ba Y, Ma L, Zhang C, Wang C, Ren Z, Zhao Y, Wu S, Zhuang R, Zhang Y, Hu H, Liu C, Xu L, Wang J, Shen H, Zhang J, Zen K, Zhang CY. Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel noninvasive biomarkers for nonsmall cell lung cancer diagnosis. Int J Cancer. 2012; 130(7):1620-1628.

- Del Vescovo V, Grasso M, Barbareschi M , Denti M. MicroRNAs as lung cancer biomarkers. Word J Clin Oncol 2014; 5(4):604-620.
- Dominioni L, Imperatori A, Rovera F, Ochetti A, Torrigiotti G, Paolucci M. Stage I non small cell lung carcinoma: analysis of survival and implications for screening. Cancer. 2000; 89(11 Suppl):2334-2344.
- Dominioni L, Poli A, Mantovani W, Pisani S, Rotolo N, Paolucci M, Sessa F, Conti V, D'Ambrosio V, Paddeu A, Imperatori A. Assessment of lung cancer mortality reduction after chest X-ray screening in smokers: A population based cohort study in Varese, Italy. Lung Cancer 2013; 80(1):50-54.
- Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 2006; 6(4):259- 269.
- Fan L, Sha J, Teng J, Li D, Wang C, Xia Q, Chen H, Su B, Qi H. Evaluation of Serum Paired MicroRNA Ratios for Differential Diagnosis of Non-Small Cell Lung Cancer and Benign Pulmonary Diseases. Mol Diagn Ther. 2018. [Epub ahead of print]
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136(5):359-386.
- Flehinger BJ, Melamed MR. Current status of screening for lung cancer. Chest Surg Clin N Am. 1994; 4(1):1- 15.
- Ganti AK, Mukshine JL. Lung cancer screening. The Oncologist. 2006; 11:481-487.

- Geng Q, Fan T, Zhang B, Wang W, Xu Y, Hu H. Five microRNAs in plasma as novel biomarkers for screening of early-stage non-small cell lung cancer. Respir Res. 2014; 15: 149.
- Gyoba J, Shan S, Roa W, Bédard EL. 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. 2016 Apr 1;17(4):494.
- Halvorsen AR, Bjaanæs M, LeBlanc M, Holm AM, Bolstad N, Rubio L, Peñalver JC, Cervera J, Mojarrieta JC, LópezGuerrero JA, Brustugun OT, Helland Å. A unique set of 6 circulating microRNAs for early detection of non-small cell lung cancer. Oncotarget. 2016; 7:37250–59.
- Heegaard NH, Schetter AJ, Welsh JA, Yoneda M, Bowman ED, Harris CC. Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. Int J Cancer. 2012; 130(6):1378-1386.
- Iorio MV, Croce CM. Causes and consequences of microRNA dysregulation. Cancer J 2012; 18(3):215-222.
- Kalscheuer S, Zhang X, Zeng Y, Upadhyaya P. Differential expression of microRNAs in early-stage neoplastic transformation in the lungs of F344 rats chronically treated with the tobacco carcinogen4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. Carcinogenesis. 2008; 29(12): 2394-2399.
- Li J, Guo Y, Liang X, Sun M, Wang G, De W, Wu W. MicroRNA-223 functions as an oncogene in human gastric cancer by targeting FBXW7/hCdc4. J Cancer Res Clin Oncol. 2012; 138: 763-774.

- Markou A, Sourvinou I, Vorkas PA, Yousef GM, Lianidou E. Clinical evaluation of microRNA expression profiling in non small cell lung cancer. Lung Cancer. 2013; 81(3):388-396.
- Mavrakis KJ, Van Der Meulen J, Wolfe AL, Liu X, Mets E, Taghon T, Khan AA, Setty M, Rondou P, Vandenberghe P, Delabesse E, Benoit Y, Socci NB, et al. A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). Nat Genet. 2011; 43: 673-678.
- Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008; 105(30):10513-10518.
- Moretti F, D'Antona P, Finardi E, Barbetta M, Dominioni L, Poli A, Gini E, Noonan DM, Imperatori A, Rotolo N, Cattoni M, Campomenosi P. Systematic review and critique of circulating miRNAs as biomarkers of stage I-II non-small cell lung cancer. Oncotarget. 2017; 8(55):94980-94996.
- Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC, Ramnath N, Gotoh N, Beer DG, Chen G. A Novel Serum 4-microRNA Signature for Lung Cancer Detection. Sci Rep. 2015; 5:12464.
- Negrini M, Cutrona G, Bassi C, Fabris S, Zagatti B, Colombo M, Ferracin M, D'Abundo L, Saccenti E, Matis S, Lionetti M, Agnelli L, Gentile M, Recchia AG, Bossio S, Reverberi D, Rigolin G, Calin GA,

Sabbioni S, Russo G, Tassone P, Morabito F, Ferrarini M, Neri A. microRNAome expression in chronic lymphocytic leukemia: comparison with normal B-cell subsets and correlations with prognostic and clinical parameters. Clin Cancer Res 2014; 20(15): 4141-4153.

- Puisségur MP, Mazure NM, Bertero T, Pradelli L, Grosso S, Robbe-Sermesant K, Maurin T, Lebrigand K, Cardinaud B, Hofman V, Fourre S, Magnone V, Ricci JE, Pouysségur J, Gounon P, Hofman P, Barbry P, Mari B. miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity. Cell Death Differ. 2011; 18(3):465-478.
- Sanfiorenzo C, Ilie MI, Belaid A, Barlesi F, Mouroux J, Marquette CH, Brest P, Hofman P. Two panels of plasma microRNAs as non-invasive biomarkers for prediction of recurrence in resectable NSCLC. PLoS One. 2013; 8: e54596.
- Schembri F, Sridhar S, Perdomo C, Gustafson AM, Zhang X, Ergun A, Lu J, Liu G,Zhang X, Bowers J, Vaziri C, Ott K, Sensinger K, Collins JJ, Brody JS, Getts R,Lenburg ME, Spira A. MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. Proc Natl Acad Sci U S A. 2009
- Shell S, Park SM, Radjabi AR, Schickel R, Kistner EO, Jewell DA, Feig C, Lengyel E, Peter ME. Let-7 expression defines two differentiation stages of cancer. Proc Natl Acad Sci U S A. 2007; 104(27):11400-11405.

- Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, Guarnera M, Liao J, Chou A, Lu CL, Jiang Z, Fang H, Katz RL, Jiang F. Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. Lab Invest. - 2011; 91(4):579-587.
- Shields PG. Molecular epidemiology of lung cancer. Ann Oncol. 1999;10 Suppl 5: S7-11.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018; 68(1):7-30.
- Takahashi K, Yokota S, Tatsumi N, Fukami T, Yokoi T, Nakajima M. Cigarette smoking substantially alters plasma microRNA profiles in healthy subjects. Toxicol Appl Pharmacol. 2013; 272(1):154-160.
- The National Lung Screening Trial Research Team. Reduced Lung-Cancer mortality with Low-Dose Computed Tomographic Screening. N Engl J Med 201; 365395-365409.
- Wei Y, Yang J, Yi L, Wang Y, Dong Z, Liu Z, Ou-yang S, Wu H, Zhong Z, Yin Z, Zhou K, Gao Y, Yan B, et al. MiR-223-3p targeting SEPT6 promotes the biological behavior of prostate cancer. Sci Rep. 2014; 4: 7546.
- Whale AS, Jones GM, Pavšič J, Dreo T, Redshaw N, Akyurek S, Akgoz M, Divieto C, PaolaSassi M, He HJ, Cole KD, Bae YK, Park SR, Deprez L, Corbisier P, Garrigou S, Taly V, Larios R, Cowen S, O'Sullivan DM, Bushell CA, Goenaga-Infante H, Foy CA, Woolford AJ, Parkes H, Huggett JF, Devonshire AS. Assessment of Digital PCR as a Primary Reference Measurement Procedure to Support Advances in Precision Medicine. Clin Chem. 2018; 64(9):1296-1307.

- Yang S, Cho YJ, Jin L, Yuan G, Datta A, Buckhaults P, Datta PK. An epigenetic auto-feedback loop regulates TGF-beta type II receptor expression and function in NSCLC. Oncotarget. 2015; 6(32):33237-33252.
- Zandberga E, Kozirovskis V, Ābols A, Andrējeva D, Purkalne G, Linē A. Cell-free microRNAs as diagnostic, prognostic, and predictive biomarkers for lung cancer. Genes Chromosomes Cancer. 2013; 52(4):356-369.
- Zhang HW, Wang EW, Li LX, Yi SH, Li LC, Xu FL, Wang DL, Wu YZ, Nian WQ. A regulatory loop involving miR-29c and Sp1 elevates the TGF-beta1 mediated epithelial-to-mesenchymal transition in lung cancer. Oncotarget. 2016; 7: 85905-85916.
- Zhou Z, Chen P, Peng H. Are healthy smokers really healthy? Tob Induc Dis. 2016; 14: 35. eCollection 2016.
- Zhu W, He J, Chen D, Zhang B, Xu L, Ma H, Liu X, Zhang Y, Le H. Expression of miR-29c, miR-93, and miR-429 as potential biomarkers for detection of early stage non-small lung cancer. PLoS One. 2014; 9(2):e87780.

Supplementary file 1

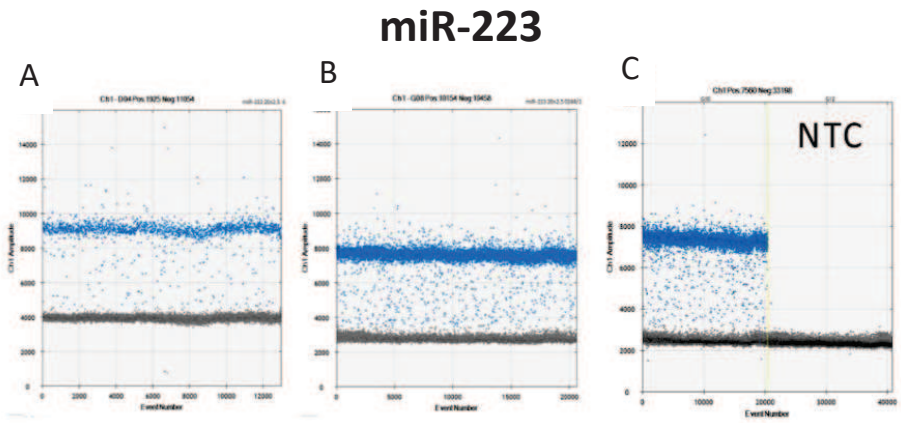
Preliminary tests were carried out to determine the appropriate volume of primers and cDNA template (and possibly amplification conditions such as annealing temperature) for each miRNA of interest, to obtain the best separation between positive and negative droplets in samples analysed by droplet digital PCR (ddPCR). This is a critical aspect, since quantification relies on calculating the ratio between them before applying Poisson's statistics to correct the estimate. For these preliminary experiments, we worked on a small number of samples (5 tumor samples and 5 controls). The cDNAs were diluted 20 fold and the ddPCR reaction was prepared in a 20 µl volume, by adding 10 µl 2X Evagreen supermix (Biorad), the desired miRCURY LNA PCR primer set at the appropriate dilution (Supplementary Table 1.1), and nuclease-free water up to 20 µl.

Supplementary Table 1.1. *Final conditions (dilution of primers and volume of cDNA) used for analysis of the indicated miRNAs by ddPCR.*

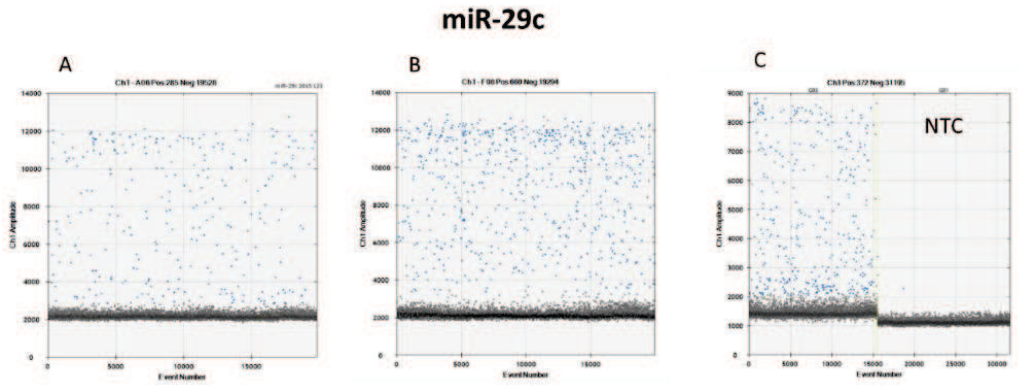
	hsa-miR-223-3p	**hsa-miR-20a-5p**	**hsa-miR-29c-3p**	**hsa-miR-210-3p**
Primer dilution	20x	20x	20x	40x
cDNA volume (1:20 dilution) (µL)	2,5	2,5	5	5

In the graphs reported in Supplementary Figures 1.1 to 1.5, typical results of ddPCR analysis for the quantitative measurement of miRNAs are shown: the x-axis reports the number of droplets in the sample

and the y-axis indicates the quantity of fluorescence emitted by each droplet. Intensity of fluorescence depends on the specific assay, on the efficiency of amplification and on the presence of inhibitors. Negative droplets do not contain the target molecule and are colored in grey; positive droplets contain the molecule and are colored in blue. The “No Template Control” sample (NTC) allows to discriminate positive from negative droplets (Supplementary Figure 1.1, C); in the example in Supplementary Figure1.1, hsa-miR-223-3p is expressed at high level in samples and positive and negative droplets separate well from negative droplets. A minority of the droplets show intermediate fluorescence (the so called “rain effect”). The “rain effect” (Supplementary figure 1.2 A, B and Supplementary figure 1.3 A, B), is observed in particular with specific assays: it is possible that suboptimal assay design due to intrinsic miRNA sequence or the presence of other members of a miRNA family is responsible for this occurrence. An alternative explanation when sporadic samples show a “rain effect” is the presence of PCR inhibitors that delay amplification in some of the droplets, thus resulting in their decreased fluorescence (hence an intermediate level of fluorescence). The rain effect is often seen with microRNAs, as the short sequence available for primer design makes the design itself difficult.



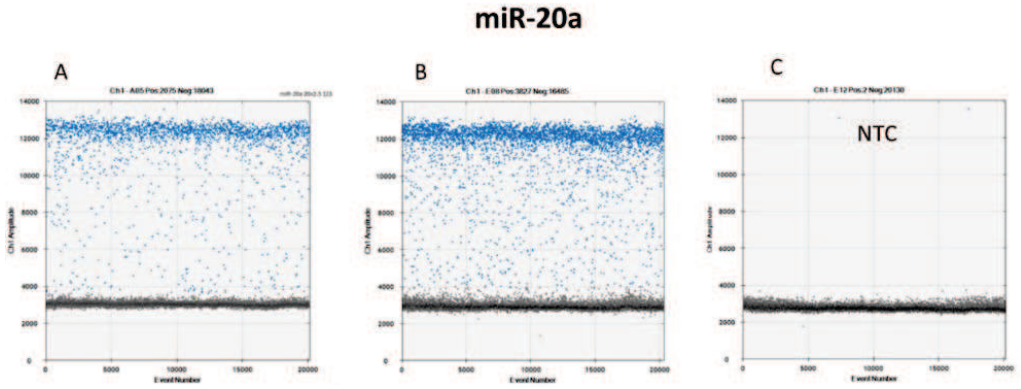
Supplementary Figure *1.1.* *Hsa-miR-223-3p levels in three different serum samples and in a NTC. The inclusion of a "No template control" (NTC, panel C) in each ddPCR run allows to better discriminate negative from positive droplets. Blue droplets contain the target, whereas grey droplets do not. The latter should be the only type of droplet in the NTC. The graph in panel A shows a sample that contains less copies of hsa-miR-223-3p than the sample in panel B.*



Supplementary *Figure 1.2.* *Hsa-miR-29c-3p levels in two samples. The inclusion of a "No template control" (NTC, panel C) in each ddPCR run allows to better discriminate negative from positive droplets. Blue droplets contain the target, whereas grey droplets do not. The latter*

should be the only type of droplet in the NTC. The graph in panel A shows a sample that contains less copies of hsa-miR-29c-3p than the sample in panel B.

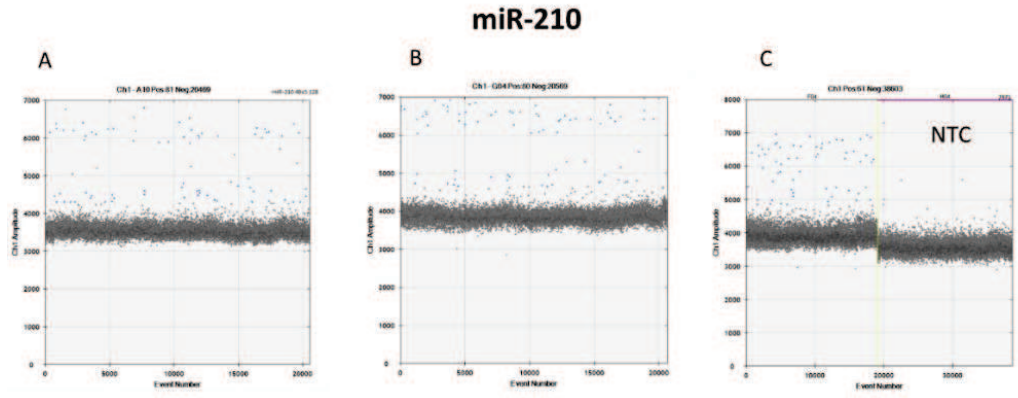
Also, hsa-miR-20a-5p showed good separation between positive and negative droplets, although the rain effect was present, as seen in Supplementary figure 1.3.



Supplementary *figure 1.3.* *Hsa-miR-20a levels in two samples: two different amounts of positive (blue) droplets are present (panel A and B, respectively) whereas blue droplets are absent in NTC (no template control, panel C). Grey droplets indicate baseline fluorescence, where target is absent. The graph in panel A shows a sample that contains less copies of hsa-miR-20a than the sample in panel B.*

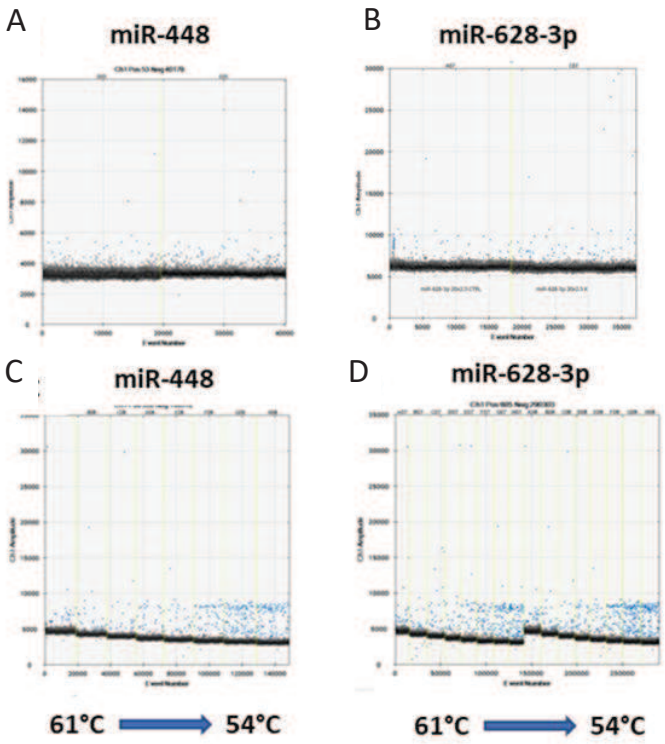
Hsa-miR-210, instead showed few positive droplets in all samples, indicating low levels of this circulating miRNA. However, it is possible to observe a demarcation line between negative and positive

droplets, in particular when samples are compared to NTC (Supplementary figure 1.4).



Supplementary figure 1.4. *Hsa-miR-210 levels in two samples: here, the number of positive droplets is low and they do not form an evident "blue band" (panel A, B, and C on the left), but it is still possible to distinguish positive from negative, especially by comparison with NTC (no template control, panel C, right). Grey droplets indicate baseline fluorescence, when target is absent.*

Two of the microRNAs initially selected for this study, namely miR-448 and miR-628-3p, did not show good separation between positive and negative droplets in none of the initial conditions tested (cDNA and primers dilutions) when the amplification reaction was performed at the usual annealing temperature (60°C). Thus, a temperature gradient during PCR was used to find the correct annealing temperature; most of the times (except for miRNA families) this led to obtain a better separation between negative and positive droplets and to decrease the rain effect, as shown in Supplementary Figure 1.5.



Supplementary Figure 1.5. *Hsa-miR-448 (panel A) and hsa-miR-628-3p (panel B) in two samples, respectively, showing scarce separation between positive (blue) and negative droplets (grey) but containing about the same number of positive droplets. Application of an annealing temperature gradient between 61°C and 54°C (panel C, D) during PCR amplification allowed us to improve the results and increase the cloud of real positive droplets.*

However, use of specific annealing temperatures for each of these microRNAs (miR-448, miR-628-3p) would be impractical for a large scale screening of individuals at risk for NSCLC. Therefore miR-448 and miR-628-3p were excluded from subsequent evaluation and in the following part of this study we report the results of the analysis of miR-20a and miR-223 (from the high sensitivity panel), miR-29c and miR-210 (from the high specificity panel).

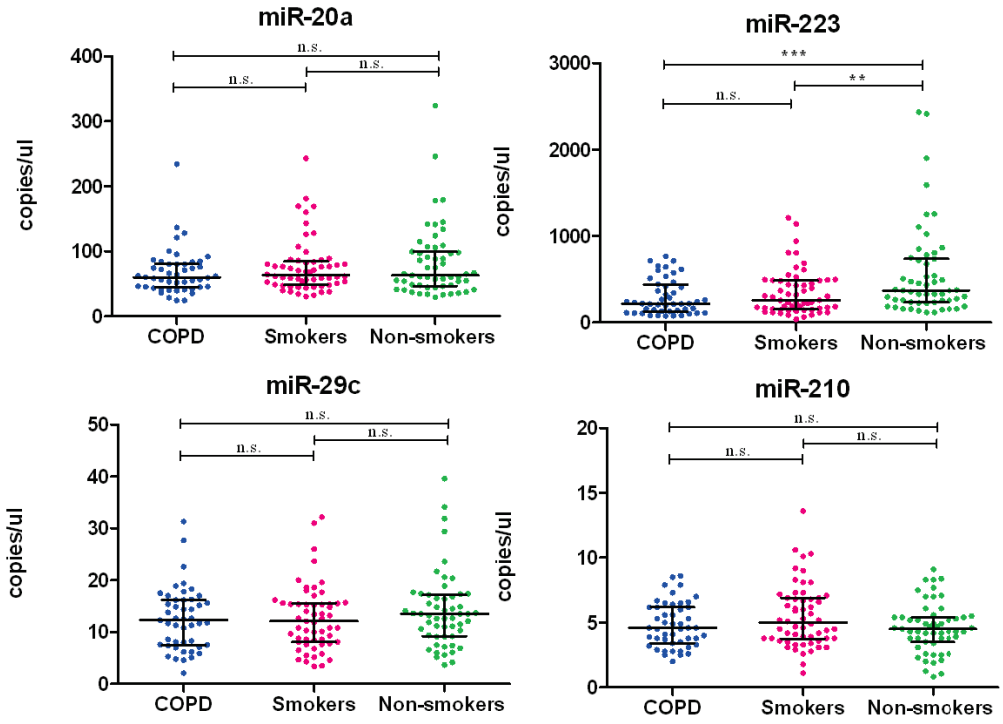
Supplementary file 2

Comparison of miR-223, miR-20a, miR-29c and miR-210 levels between control subgroups

An important issue in development of lung cancer biomarkers is the composition of the control group. Several factors, such as COPD, history of smoking, cardiovascular disorders, inflammatory processes etc. may affect circulating miRNA levels and may alter the evaluation of miRNA candidates as biomarkers. In nearly all the relevant publications on this topic, tumor and control groups are typically matched by age and gender but rarely by smoking status, with scarce or no details about the subjects' clinical history [Zhou et al., 2016].

We explored the possible differences in the levels of selected miRNAs in individuals without evidence of cancer, belonging to three subgroups relevant to lung cancer, namely smokers, non-smokers and COPD subjects. We collected serum samples from 56 smokers, 55 non-smokers and 46 COPD subjects and analysed their levels of miR-223, miR-20a, miR-29c and miR-210. The data obtained for each miRNA were compared among control subgroups, and differences were analysed with the Mann-Whitney test; results obtained with each of the four miRNAs are shown in Supplementary Figure 2.1.

For miR-20a, miR-210 and miR-29c, no significant differences were observed across the 3 control subgroups, whereas miR-223 was significantly overexpressed in the non-smoker subgroup (Supplementary figure 2.1).



Supplementary Figure 2.1. *Scatter plot showing hsa-miR-20a, hsa-miR-223, hsa-miR-210 and hsa-miR-29c levels in the three control subgroups. No significant difference of miRNA levels quantified by ddPCR was observed across COPD, smoker and non-smoker subjects for hsa-miR-20a, hsa-miR-210 and hsa-miR-29c, whereas miR-223 levels were higher in non-smokers. Data are expressed as median values with interquartile range. (** p<0.01 *** p<0,001, n.s. not significant, Mann-Whitney test).*

However, miR-223 levels were significantly higher in the non-smoker subgroup compared to the other two control subgroups (COPD and smokers; Supplementary file 1). Accordingly, the prevalence of non-smokers in a control group may partly obscure the difference of miR-223 level relative to lung cancer patients. Notably, miR-223 was described as lung cancer biomarker in two independent studies included in our recent review (Moretti et al 2017), yet with different predicted accuracy: AUC

0.94 in the work by Geng et al., and AUC 0.81 in the work by Zhang et al. [Geng et al., 2014; Zhang et al., 2017]; this could be due to different prevalence of smoking habit in the controls of the two studies.

Supplementary file 3

Repeatability of miRNAs analysis by droplet digital PCR

A critical issue in developing biomarkers for clinical use is the reliability of the biomarker quantification method. The most frequently used technique to measure microRNAs is real time quantitative PCR (qPCR), usually applying relative quantification; however, there is no agreement on the molecules and methods to be used for normalization of the qPCR results. Also, absolute quantification of qPCR findings is not devoid of pitfalls, as it relies on the use of synthetic miRNAs for construction of a calibration curve. However, these synthetic molecules turned out to be quite unstable and easily degraded when in pure form, making their measurement not completely reliable. Finally, with the qPCR method, only differences >0.5 Cq can be reliably detected as changes in the levels of the measured molecules. In addition to these specific issues related to qPCR, most studies fail to provide sufficient details on the methods used. Therefore, comparison between works from different laboratories is difficult or impossible and it is not surprising that different studies failed to identify the same microRNAs panel as lung cancer biomarker. The recent introduction of “droplet digital PCR” provides a possible way to overcome these limitations, by allowing absolute quantification of target molecules based on the principle of compartmentalization and application of Poisson statistics; the ddPCR technique can help in the development of microRNAs as cancer biomarkers.

The reliability and the robustness of the method that enable the quantification of microRNAs at the end of the procedure starting from the processing of blood samples is an important prerequisite for their application as biomarkers. One of the aims of this thesis was to test accuracy and day-to-day repeatability of microRNAs measurements in cDNA samples in our laboratory by the droplet digital PCR.

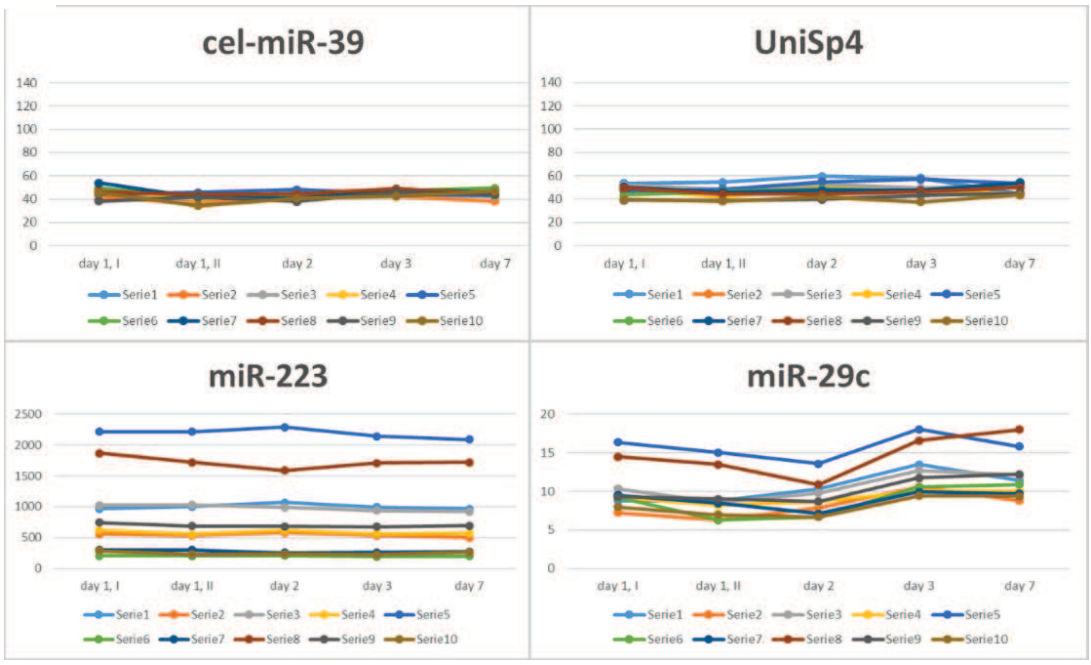
For this purpose, in preliminary experiments we tested repeatability of the measurements of the cDNA samples after 1, 2, 3 and 7 days from preparation and after applying five cycles of “freeze and thaw”. After reverse transcription of 20 cDNA samples (from 10 NSCLC patients and 10 non-smoker subjects), each reaction was split in 4 aliquots and stored at -20°C.

The first aliquot of the 20 cDNA samples was thawed after 1 day and ddPCR analysis was conducted by measuring 2 endogenous miRNAs (miR-223 and miR-29c), and 2 spike-ins (cel-miR-39 and UniSp4, the first of which was added during the initial extraction step and the second during reverse transcription) (day 1, aliquot I). All samples were analyzed in duplicate. As a single plate was not sufficient to analyse 4 targets in 20 samples in duplicate, we split the 20 samples and analysed the complete set of miRNAs in 10 samples per plate.

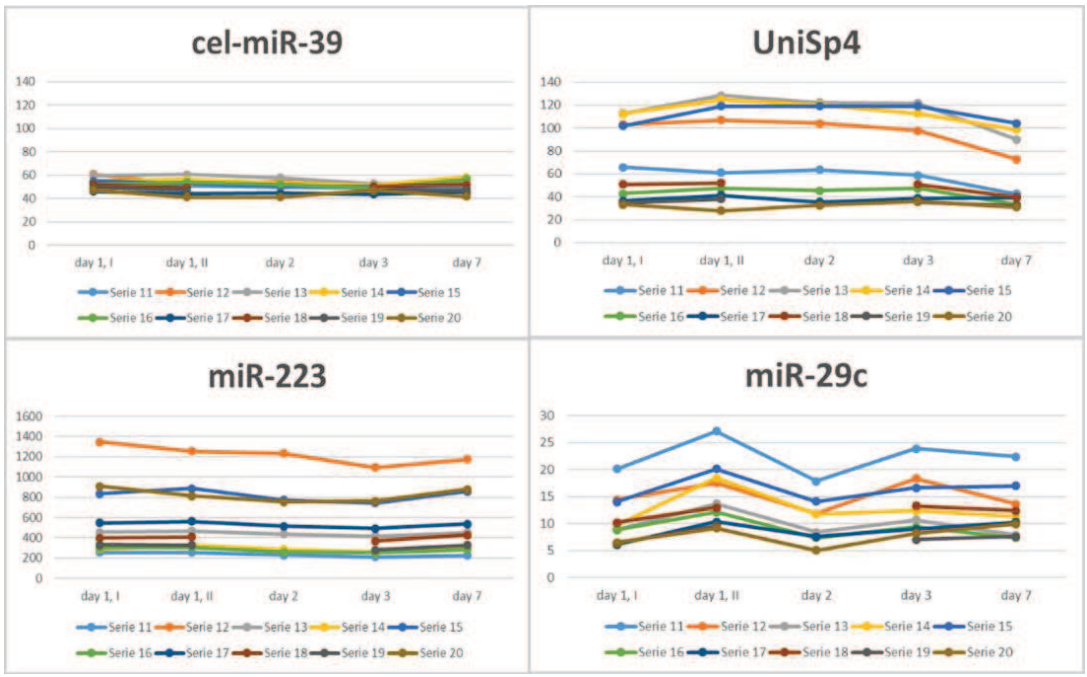
Analysis of miRNAs in the second aliquot of each sample was performed the same day in the afternoon (day 1, aliquot II), keeping constant all conditions. The third and fourth aliquots were tested on day 2 and day 3, respectively. After 7 days we carried out 5 freeze and thaw cycles of the last aliquot of each sample before measuring the same miRNAs. The data obtained show that miRNA levels of UniSp4, cel-miR-39 and miR-223 are

stable over repeated measures (days 1, 2, 3) and after repeated freeze and thaw cycles (day 7). Conversely, miR-29c shows some fluctuations in the different measurements (Supplementary Figure 3.1).

A



B



Supplementary Figure 3.1. *Repeatability of miRNA measures in cDNA samples. (A): samples 1-10; (B): samples 11-20. X-axis reports the different measurements, Y-axis reports the obtained number of copies/μl.*

The precision of miRNA quantification by droplet digital PCR was measured by the Coefficient of Variation (CV), intra-assay and inter assay (Supplementary Table 3.1).

Supplementary Table 3.1*. Coefficients of variation of cel-miR-39, UniSp4, miR-223, miR-29c measurements with the droplet digital PCR method.*

	CV intra-assay %	CV inter-assay %
cel-miR-39	3,94	6,37
UniSp4	5,08	5,40
miR-223	2,94	6,60
miR-29c	8,24	18,97

Our results demonstrate that droplet digital PCR is a robust method for miRNAs measurement, as, except for miR-29c, there was low variability both in the intra-assay and the inter-assay analyses. Accordingly, miR-223 can be considered a potentially useful lung cancer circulating biomarker, as it can be precisely measured and is significantly overexpressed in lung cancer. The lower precision of miR-29c measurements has several possible explanations: miR-29c intrinsic nucleotide sequence or sequence homology to the other two known members of this miRNA family could make the optimal design of primers difficult, resulting in the observed rain

effect (Supplementary file 1, Figure 1.2); the relatively low concentration of this miRNA compared to miR-223 could make it more subject to quantification errors.

Stability of miRNAs over time

We next sought to investigate the stability of the two miRNAs under investigation (miR-223 and miR-29c), as well as the UniSp4 spike-in, given that this is an essential prerequisite for utility as biomarker. In this experiment we tested the stability over relatively long periods of time compared to the previous experiment. In this case, the main variable that would justify a significant difference of the measurements would be the time and the relative stability/instability of the molecules we measured and not the method used.

We tested stability of the cDNA samples carrying out 3 measurements by ddPCR, at day 0 (number 0), after two months (number 2) and after 9 months (number 9). This experiment was conducted on 23 samples (14 smokers and 9 non-smokers). Paired samples t-test was used to test significance of the differences in miRNA stability. The results are summarized in Supplementary table 3.2.

Supplementary table 3.2. *Values and significance of the paired samples t-test for three measurements of UniSp4, miR-223, miR-29c by droplet digital PCR at 0, 2 and 9 months).*

		Paired Differences					t	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower	Upper		
Pair 1	UniSp4 0 - UniSp4 9	11,92	19,641	3,779	4,145	19,69	3,152	0,004
Pair 2	Mir-29c 0 - Mir-29c 9	,88	4,264	,909	-1,01	2,77	,970	0,343
Pair 3	MiR-223 0 - MiR-223 9	-280,88	235,757	78,585	-462,10	-99,67	-3,574	,007
Pair 4	UniSp4 2- UniSp4 9	,52	19,719	4,648	-9,28	10,33	,112	0,912
Pair 5	MiR-29c 2 – Mir-29c 9	5,05	6,984	3,492	-6,06	16,16	1,446	0,244
Pair 6	MiR-223 2 - MiR-223 9	-44,58	248,145	58,488	-167,98	78,82	-,762	0,456

The temporal variability of miR-223 measurements is uncertain because the first measure (MiR-223 -0) is significantly different from the third (MiR-223 -9) while the second and third measurements (MiR-223 2 – MiR-223 9) are not different. The same occurred for the UniSp4 spike-in (added during the extraction process). In these cases, the period in which analysis was performed could affect the quantification. For miR-29c the three measurements are not significantly different (MiR-29c 0 - MiR-29c 2 and MiR-29c 2 – MiR-29c 9), indicating a stability in time. Therefore, inclusion of miR-29c analysis together with miR-223 should make the combination reasonably uninfluenced by time of analysis.

CONCLUSIONS AND FUTURE PERSPECTIVES

Diagnosing lung cancer at an early stage is mandatory to increase survival rates. This requirement has motivated in recent years extensive research on non-invasive screening methods, including quantification of microRNAs as lung cancer biomarkers in circulating body fluids. However, it soon turned out that identifying circulating biomarkers is not an easy task, due to several reasons that we described in our manuscript. Reliability and reproducibility of measurements of circulating miRNAs identified as biomarkers are a major issue.

For these reasons, we performed a critical review of the literature in order to identify circulating miRNAs proven to be valuable and highly accurate for diagnosis of early NSCLC. The final result of this critical review was the proposal of a screening method based on two panels of miRNAs for diagnosis of stage I-II NSCLC: a first panel of four miRNAs, (miR-223, miR-20a, miR-448 and miR-145) reported to have high sensitivity (sensitivity > 80% and AUC > 0.80), that should identify the subjects really affected by lung cancer, but also some false positives; a second panel of four miRNAs (miR-628, miR-29c, miR-210 and miR-1244), reported to have high specificity (specificity > 90%), that would allow to eliminate most false positives. Determination of these miRNA panels should be offered as a non-invasive screening tool to subjects at risk for lung cancer (i.e. smokers or former smokers aged > 60 years).

To reduce costs however, out of these 8 miRNAs only those that perform best in terms of AUC, sensitivity and specificity, would likely be applicable in large scale screenings of at-risk individuals (i.e. smokers).

The preliminary results obtained by testing our lung cancer cohort with four miRNAs (miR-223, miR-20a, miR-29c and miR-210) of our screening model indicated that all these microRNAs were aberrantly expressed in stage I-II NSCLC patients compared to controls, making all four miRNAs good biomarkers of NSCLC. In particular, based on the results of training and validation sets, the combination of two of these miRNAs, namely "miR-223 OR miR-29c" with the selected cut-offs, appears to have very high sensitivity in detecting stage I-II NSCLC patients, albeit with low specificity. For this reason, we are now planning to improve specificity by including in the combination other miRNAs with high specificity shown in previous studies (miR-1244) [Wang et al., 2016].

Actually, the use of more than two miRNAs could offer both higher sensitivity and higher specificity for lung cancer detection. However, for the purpose of realistic clinical applicability in population screening for lung cancer, we chose to limit the analysis to two miRNAs providing altogether high sensitivity; indeed, the importance of high sensitivity of the screening tool outweighs the low specificity. The high sensitivity of the test allows to identify the vast majority (92,5%) of true positive cases; false positives can then be ruled out by a second-level investigation, such as low-dose CT of the chest. It has been shown that dysregulated miRNA profiles in cell-free blood indicate the presence of lung cancer many months before the

occurrence of symptoms [Chen et al., 2012], and even before the disease was detected by CT screening [Bianchi et al., 2011; Boeri et al., 2011]. Efficacy of CT screening for lung cancer is widely accepted; however, the possibility has been raised that CT screening may lead to over-diagnosis and over-treatment of biologically indolent cancers. Moreover, the feasibility of large-scale population screening by CT is uncertain, as this diagnostic procedure requires specialized centers, not widely available, it has high costs for healthcare systems, and there are risks related to radiation exposure [Strauss and Dominioni, 2013]. Therefore, the development of diagnostic blood tests may be relevant for effective population screenings. The test that we aim to develop is minimally invasive, as the starting material is blood, and the screening tool is less expensive than CT scans.

From the viewpoint of perspective application in clinical practice, the combination of different miRNAs with high sensitivity and high specificity may result in accurate prediction of subject affected by early stage lung cancer, with a low false positive rate; only subjects positive to the miRNA test should subsequently undergo chest CT scan, to confirm or rule out lung cancer.

The availability of a non-invasive screening test, to be offered to asymptomatic at-risk individuals would likely increase the probability of early diagnosis and would possibly increase the cure rate, decreasing the need for expensive anticancer drug therapies.

Future perspectives in the short-term include analysis of the other two miRNAs composing our screening model (miR-145 and miR-1244) aiming to improve the specificity of the test (miR-1244 is part of the second panel

with high specificity). Moreover, further experiments need to be done to ensure reproducibility of the tests, by exchanging samples and repeating analysis in other laboratories. In the long-term, we hope to move miRNA biomarkers from discovery phase to clinical application, to benefit lung cancer patients and the HealthCare System.

REFERENCES

- Ambros V. The functions of animal microRNAs. Nature. 2004; 431:350–355.
- Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X, Vijayakrishnan J, Sullivan K, Matakidou A, Wang Y, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet. 2008; 40(5):616-622.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004; 116(2):281-297.
- Bayraktar R, Van Roosbroeck K, Calin GA. Cell-to-cell communication: microRNAs as hormones. Mol Oncol. 2017; 11(12):1673-1686.
- Bianchi F, Nicassio F, Marzi M, Belloni E, Dall'olio V, Bernard L, Pelosi G, Maisonneuve P, Veronesi G, Di Fiore PP. A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. EMBO Mol Med. 2011; 3:495–503.
- Blondal T, Jensby Nielsen S, Baker A, Andreasen D, Mouritzen P, Wrang Teilum M, Dahlsveen IK. Assessing sample and miRNA profile quality in serum and plasma or other biofluids. Methods 2013; 59(1): S1-6.
- Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, Calabro E, Croce CM, Pastorino U, Sozzi G. MicroRNA signatures in tissues

and plasma predict development and prognosis of computed tomography detected lung cancer. Proc Natl Acad Sci U S A. 2011; 108:3713–3718.

- Boffetta P. Involuntary smoking and lung cancer. Scand J Work Environ Health. 2002;28 Suppl 2:30-40.
- Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012; 489(7417):519-525.
- Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, Zhang Y, Wang J, Jiang X, Xiang Y, Xu C, Zheng P, Zhang J, Li R, Zhang H, Shang X, Gong T, Ning G, Wang J, Zen K, Zhang J, Zhang CY. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. 2008; 18(10):997-1006.
- Chen X, Hu Z, Wang W, Ba Y, Ma L, Zhang C, Wang C, Ren Z, Zhao Y, Wu S, Zhuang R, Zhang Y, Hu H, et al. Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel noninvasive biomarkers for nonsmall cell lung cancer diagnosis. Int J Cancer. 2012;130:1620–1628.
- Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. Nat Rev Cancer. 2014; (8):535-46.
- Devesa SS, Bray F, Vizcaino AP*, et al.* International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. Int J Cancer 2005; 117(2):294-299.

- Dominioni L, Poli A, Mantovani W, Pisani S, Rotolo N, Paolucci M, Sessa F, Conti V, D'Ambrosio V, Paddeu A, Imperatori A. Assessment of lung cancer mortality reduction after chest X-ray screening in smokers: a population based cohort study in Varese, Italy. Lung Cancer 2013; 80(1):50-54.
- El-Hefnawy T, Raja S, Kelly L, Bigbee WL, Kirkwood JM, Luketich JD, Godfrey TE. Characterization of amplifiable, circulating RNA in plasma and its potential as a tool for cancer diagnostics. Clin Chem. 2004; 50(3):564-573.
- Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 2006; 6(4):259- 269.
- Fazi F, Nervi C. MicroRNA: basic mechanisms and transcriptional regulatory networks for cell fate determination. Cardiovasc Res. 2008; 79(4):553-561.
- Feng S, Cong S, Zhang X, Bao X, Wang W, Li H, Wang Z, Wang G, Xu J, Du B, Qu D, Xiong W, Yin M, Ren X, Wang F, He J, Zhang B. MicroRNA-192 targeting retinoblastoma 1 inhibits cell proliferation and induces cell apoptosis in lung cancer cells. Nucleic Acids Res. 2011; 39(15):6669-6678.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136(5):359-386.
- Garzon R, Marcucci G, Croce CM. Targeting microRNAs in cancer: rationale, strategies and challenges. Nat Rev Drug Discov. 2010;

9(10):775-789.

- Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, Maher CA, Fulton R, Fulton L, Wallis J, Chen K, Walker J, McDonald S, Bose R, Ornitz D, Xiong D, You M, Dooling DJ, Watson M, Mardis ER, Wilson RK. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell. 2012; 150(6):1121-1134.
- Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014; 15(8):509-524.
- Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. BMJ. 1997; 315(7114):980-988.
- Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000; 100:57-70.
- Hausser J, Zavolan M. Identification and consequences of miRNA-target interactions--beyond repression of gene expression. Nat Rev Genet. 2014; 15(9):599-612.
- Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, Yatabe Y, Kawahara K, Sekido Y, Takahashi T. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res. 2005; 65(21):9628-9632.
- http://www.mirbase.org/
- I H, Cho JY. Lung Cancer Biomarkers. Adv Clin Chem. 2015; 72:107-170.
- IARC Monographs on the Evaluation of carcinogenic Risks to

Humans. Volume 95 HouseHold use of solid fuels and high temperature fryng. Lyon, World Health Organization/IARC, 2006.

- Iorio MV, Croce CM. Causes and consequences of microRNA dysregulation. Cancer J 2012; 18(3):215-222.
- Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D, Slack FJ. RAS is regulated by the let-7 microRNA family. Cell. 2005; 120(5):635-647.
- Johnson SM, Grosshans H, Shingara J, et al. RAS is regulated by the let-7 microRNA family. Cell 2005; 120(5):635-647.
- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ, Ding L. Mutational landscape and significance across 12 major cancer types. Nature. 2013; 502(7471):333-339.
- Kinet V, Halkein J, Dirkx E, Windt LJ. Cardiovascular extracellular microRNAs: emerging diagnostic markers and mechanisms of cell-to-cell RNA communication. Front Genet. 2013; 4:214.
- Koike T, Koike T, Yoshiya K, Tsuchida M, Toyabe S. Risk factor analysis of locoregional recurrence after sublobar resection in patients with clinical stage IA non-small cell lung cancer. J Thorac Cardiovasc Surg. 2013; 146(2):372-378.
- Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. 2010; 11(9):597-610.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification

of novel genes coding for small expressed RNAs. Science. 2001; 294(5543):853-8.

- Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science. 2001; 294(5543):858-862.
- Lawrie CH, Gal S, Dunlop HM, Pushkaran B, Liggins AP, Pulford K, Banham AH, Pezzella F, Boultwood J, Wainscoat JS, Hatton CS, Harris AL. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. Br J Haematol. 2008; 141(5):672-675.
- Lazar V, Suo C, Orear C, van den Oord J, Balogh Z, Guegan J, Job B, Meurice G, Ripoche H, Calza S, Hasmats J, Lundeberg J, Lacroix L, Vielh P, Dufour F, Lehtiö J, Napieralski R, Eggermont A, Schmitt M, Cadranel J, Besse B, Girard P, Blackhall F, Validire P, Soria JC, Dessen P, Hansson J, Pawitan Y. Integrated molecular portrait of non-small cell lung cancers. BMC Med Genomics. 2013; 6: 53.
- Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993; 75(5):843-854.
- Li C, Nguyen HT, Zhuang Y, Lin Z, Flemington EK, Zhuo Y, Kantrow SP, Morris GF, Sullivan DE, Shan B. Comparative profiling of miRNA expression of lung adenocarcinoma cells in two-dimensional and three-dimensional cultures. Gene. 2012; 511(2):143-150.
- Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J. 2016; 48(3):889-

902.

- Mar-Aguilar F, Mendoza-Ramírez JA, Malagón-Santiago I, Espino-Silva PK, Santuario-Facio SK, Ruiz-Flores P, Rodríguez-Padilla C, Reséndez-Pérez D. Serum circulating microRNA profiling for identification of potential breast cancer biomarkers. Dis Markers. 2013; 34(3):163-169.
- Maroney PA, Yu Y, Nilsen TW. MicroRNAs mRNAs and Translation. Cold Spring Harb. Symp. Quant. Biol. 2006; 71:531–535.
- Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008; 105(30):10513-10518.
- Moretti F, D'Antona P, Finardi E, Barbetta M, Dominioni L, Poli A, Gini E, Noonan DM, Imperatori A, Rotolo N, Cattoni M, Campomenosi P. Systematic review and critique of circulating miRNAs as biomarkers of stage I-II non-small cell lung cancer. Oncotarget. 2017; 8(55):94980-94996.
- Mountain CF, Hermes KE. Surgical treatment of lung cancer. Past and present. Methods Mol Med 2003; 75:453-487.
- Murchison EP, Hannon GJ. miRNAs on the move: miRNA biogenesis and the RNAi machinery. Curr Opin Cell Biol. 2004; 16(3):223-229.
- National Center for Chronic Disease Prevention and Health

Promotion (US) Office on Smoking and Health. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2014.

- Negrini M, Cutrona G, Bassi C, Fabris S, Zagatti B, Colombo M, Ferracin M, D'Abundo L, Saccenti E, Matis S, Lionetti M, Agnelli L, Gentile M, Recchia AG, Bossio S, Reverberi D, Rigolin G, Calin GA, Sabbioni S, Russo G, Tassone P, Morabito F, Ferrarini M, Neri A. microRNAome expression in chronic lymphocytic leukemia: comparison with normal B-cell subsets and correlations with prognostic and clinical parameters. Clin Cancer Res 2014; 20(15):4141-4153.
- Noferesti SS, Sohel MM, Hoelker M, Salilew-Wondim D, Tholen E, Looft C, Rings F, Neuhoff C, Schellander K, Tesfaye D. Controlled ovarian hyperstimulation induced changes in the expression of circulatory miRNA in bovine follicular fluid and blood plasma. J Ovarian Res. 2015; 8:81.
- Pao W, Girard N. New driver mutations in non-small-cell lung cancer. Lancet Oncol 2011; 12(2):175-180.
- Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Müller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature. 2000; 408(6808):86-89.

- Pegna A, Picozzi G, Mascalchi M, Maria Carozzi F, Carrozzi L, Comin C, Spinelli C, Falaschi F, Grazzini M, Innocenti F, Ronchi C, Paci E. ITALUNG Study Research. Group Design, recruitment and baseline results of the iTaLUNG trial for lung cancer screening with low-dose CT. Lung Cancer 2009; 64(1):34-40.
- Piva R, Spandidos DA, Gambari R. From microRNA functions to microRNA therapeutics: novel targets and novel drugs in breast cancer research and treatment. Int J Oncol. 2013; 43(4):985-994.
- Ramalingam P, Palanichamy JK, Singh A, Das P, Bhagat M, Kassab MA, Sinha S, Chattopadhyay P. Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. RNA. 2014; 20(1):76-87.
- Rana TM. Illuminating the silence: understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol. 2007; 8(1):23-36.
- Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature. 2000; 403(6772):901-906.
- Rekhtman N. Neuroendocrine tumors of the lung: an update. Arch Pathol Lab Med 2010;134(11):1628-1638.
- Rolle K, Piwecka M, Belter A, Wawrzyniak D, Jeleniewicz J, Barciszewska MZ, Barciszewski J. The Sequence and Structure Determine the Function of Mature Human miRNAs. Plose one. 2016; 11(3): e0151246.
- Schwartz AG, Cote ML. Epidemiology of Lung Cancer. Adv Exp Med

Biol. 2016; 893:21-41.

- Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003; 115(2):199-208.
- Seo JS, Ju YS, Lee WC*, et al.* The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res 2012; 22(11):2109-2119.
- Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, Lee J, Jung YJ, Kim JO, Shin JY, Yu SB, Kim J, Lee ER, Kang CH, Park IK, Rhee H, Lee SH, Kim JI, Kang JH, Kim YT. The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res. 2012; 22(11):2109-2119.
- Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017; 67(1):7-30.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018; 68(1):7-30.
- Stellman SD, Muscat JE, Thompson S*, et al.* Risk of squamous cell carcinoma and adenocarcinoma of the lung in relation to lifetime filter cigarette smoking. Cancer 1997; 80(3):382-388.
- Strauss GM, Dominioni L. Chest X-ray screening for lung cancer: overdiagnosis, endpoints, and randomized population trials. J Surg Oncol. 2013; 108(5):294-300.
- Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T. Reduced expression of the let-7 microRNAs in human

lung cancers in association with shortened postoperative survival. Cancer Res. 2004; 64(11):3753-3756.

- The National Lung Screening Trial Research Team. Reduced Lung-Cancer mortality with Low-Dose Computed Tomographic Screening. N Engl J Med 2011; 365:395-409.
- Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsäter A, Flex A, Aben KKH, de Vegt F, Mulders PFA, Isla D, Vidal MJ, Asin L, Saez B, Murillo L, Blondal T, Kolbeinsson H, Stefansson JG, Hansdottir I, Runarsdottir V, Pola R, Lindblad B, van Rij AM, Dieplinger B, Haltmayer M, Mayordomo JI, Kiemeney LA, Matthiasson SE, Oskarsson H, Tyrfingsson T, Gudbjartsson DF, Gulcher JR, Jonsson S, Thorsteinsdottir U, Kong A, Stefansson K. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008; 452(7187):638-642.
- Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeböller H, Risch A, McKay JD, Wang Y, Dai J, Gaborieau V, McLaughlin J, Brenner D, Narod SA, Caporaso NE, Albanes D, Thun M, Eisen T, Wichmann HE, Rosenberger A, Han Y, Chen W, Zhu D, Spitz M, Wu X, Pande M, Zhao Y, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokan HE, Gabrielsen ME, Skorpen F, Vatten L,

Njølstad I, Chen C, Goodman G, Lathrop M, Benhamou S, Vooder T, Välk K, Nelis M, Metspalu A, Raji O, Chen Y, Gosney J, Liloglou T, Muley T, Dienemann H, Thorleifsson G, Shen H, Stefansson K, Brennan P, Amos CI, Houlston R, Landi MT; Transdisciplinary Research in Cancer of the Lung (TRICL) Research Team. Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. Hum Mol Genet. 2012; 21(22):4980-4995.

- Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS, Wistuba I; WHO Panel. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol. 2015; 10(9):1243-1260.
- Travis WD. Pathology of lung cancer. Clin Chest Med 2011; 32(4):669-692.
- Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol. 2007; 9(6):654-659.
- Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. 2006; 20:515–524.
- Vasudevan S, Tong Y, Steitz JA. Switching from repression to

activation: microRNAs can up-regulate translation. Science 2007;318(5858):1931-1934.

- Villalobos P, Wistuba II. Lung Cancer Biomarkers. Hematol Oncol Clin North Am. 2017; 31(1):13-29.
- Wang W, Li W, Ding M, Yuan H, Yang J, Meng W, Jin E, Wang X, Ma S. Identification of miRNAs as non-invasive biomarkers for early diagnosis of lung cancers. Tumour Biol. 2016; 37:16287–93.
- Xiong S, Zheng Y, Jiang P, Liu R, Liu X, Chu Y. MicroRNA-7 inhibits the growth of human non-small cell lung cancer A549 cells through targeting BCL-2. Int J Biol Sci. 2011; 7(6):805-814.
- Zandberga E, Kozirovskis V, Ābols A, Andrējeva D, Purkalne G, Linē A. Cell-free microRNAs as diagnostic, prognostic, and predictive biomarkers for lung cancer. Genes Chromosomes Cancer. 2013; 52(4):356-369.