microRNA and non-targeted proteome analysis of liquid biopsies from the distal lung collected by

Particles in Exhaled Air (PExA) reveals presence from extracellular vesicles

M. Bonatti[1,2]*, M. Ezerskyte[1], P. Lindberg[1,2], Benedikt Zöhrer[1,2], Iryna Kolosenko[1,2], Á. Végvári[3], Å.M.

Wheelock[1,2]*.

[1]Division of Immunology and Respiratory Medicine, Dept. of Medicine Solna and Center for Molecular

Medicine, Karolinska Institutet, Stockholm, Sweden.

[2]Dept. of Allergy and Respiratory Medicine, Karolinska University Hospital Solna, Stockholm, Sweden.

[3]Division of Chemistry I, Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet,

Stockholm, Sweden.

*Corresponding authors

Martina Bonatti: martina.bonatti@ki.se

Åsa M Wheelock: asa.wheelock@ki.se

ABSTRACT

Background. Early detection is essential to ensure prompt treatment selection for many lung

disorders, which are becoming more prevalent. Unfortunately, invasive methods are often used to

examine the distal parts of the lungs for diagnosis and to establish disease biomarkers. The growing

need to explore the bio-molecular mechanisms in various lung diseases highlights the importance of

non-invasive methods, particularly in longitudinal studies. Thus, the use of particles in exhaled air

(PExA), a non-invasive technique for sampling epithelial lining fluid from distal airways, is gaining

attention. MicroRNAs (miRNAs) are crucial in modulating protein expression both intracellularly and

intercellularly, often transported via extracellular vesicles. Dysregulated miRNAs have been linked to

many pulmonary diseases, and their relative stability, especially when encapsulated in EVs, makes

them promising biomarkers. Although underreported, exhaled miRNA identified with other non-

invasive techniques shows potential as a molecular tool. Therefore, here we report the initial steps in

the miRNA detection in PExA, offering opportunities to study miRNA roles in both normal and diseased

states in a non-invasive manner.

Methods. Exhaled particles were collected from healthy subjects using the PExA 2.0 instrument

following prescribed guidelines. Different samples were processed for non-targeted proteome

analysis using mass spectrometry and RNA extraction followed by miRNAseq workflow for low input

starting material. The detection of some miRNAs was confirmed with the miRCURY LNA miRNA PCR assays.

The quality of the whole workflow was evaluated with a RNAses activity test and with synthetic

controls miRNAs (spike-in kit).

Results. Proteomic analysis identified over 50 proteins across experiments and samples, with a

notable enrichment of proteins associated with extracellular vesicles, including blood microparticles,

and secretory granules. miRNA-seq revealed 39 mature miRNAs, the majority of which have been

previously reported to be detected in the airways and respiratory disorders. Some were also reported

to be secreted by primary human airway epithelial cells via extracellular vesicles. Notably, miRNA-

125b and members of the let-7 family were among the most abundant. Fluorometric assays showed

significant RNase activity in both PExA and other lung-related samples, such as bronchoalveolar lavage

fluid, suggesting that this activity originates from the airways and is independent of the sampling

techniques used. The workflow for extraction and processing of the PExA collection membrane, tested

with abundant synthetic miRNAs and analyzed using the miRCURY LNA miRNA PCR assay, yielded

results comparable to control samples, indicating that the membrane material does not interfere with

the assay.

Conclusions. Using PExA, we identified several miRNAs that are known to be dysregulated in various

pulmonary disorders. The enrichment of extracellular secretory components in the core protein list

from our proteomic analysis, along with the elevated RNAse activity in the respiratory tract, suggests

that the detected miRNAs may be encapsulated within extracellular vesicles. These miRNAs are

particularly intriguing due to their potential role in intercellular communication during both normal

and pathological lung processes. Our findings suggest that PExA holds significant potential as a non-

invasive tool for studying extracellular vesicle-mediated miRNA signalling in the small airways.

<u>INTRODUCTION</u>

Early detection of lung diseases is important to avoid loss of lung function and holds the potential to

significantly enhance the quality of life for affected individuals. Timely intervention has the capacity

to slow down or even halt the progression of the disease, leading to superior long-term outcomes.

Today, we rely on invasive techniques such as lung biopsies or bronchoalveolar lavage (BAL) for

biomarker discovery, and in some cases for diagnosis. While invasive methods provide valuable

insights, their resource intensiveness may be prohibitive, may be inappropriate or impractical for

longitudinal studies, and not be well tolerated for some patient groups. It is thus essential to balance

the benefits with the potential risks and patient discomfort associated with these procedures [1]. As

such, the development of non-invasive sampling techniques of the distal lung is pivotal to facilitate

biomarker discovery, longitudinal monitoring or large-scale screening of respiratory disease. Particles

in exhaled air (PExA) represents a unique sampling method in that it permits non-invasive sampling of

the epithelial lining fluid and airway exudates from the small airways and alveoli [2], a.k.a. liquid

biopsies from the distal airways. The approach involves impaction of exhaled aerosols onto comprising

ultrafine droplets of respiratory tract lining fluid (RTLF) onto a membrane[3]. These droplets are

generated and exhaled after a breathing manoeuvre that facilitates the closure and reopening of the

small airways. PEx samples have been shown to contain several types of biomolecules, including

proteins [2, 4] and lipids [5] known to be present in the RTLF in the distal airways and alveolar region.

Accordingly, this type of sampling method can be especially relevant for disorders involving the small

airways, such as asthma and COPD [2, 6, 7]. Despite a number of recent advancements of the sampling

method and associated publications, our understanding of the material collected and its relevance to

disease diagnosis and monitoring remains limited. As such, further benchmarking and analysis is needs

to unlock the full potential of this non-invasive technique.

MicroRNAs (miRNAs) are small, conserved non-coding RNAs that regulate gene expression post-

transcriptionally [8]. They play a crucial role in various biological processes, and their dysregulation is

linked to many pulmonary diseases [9, 10]. miRNAs are found in numerous body fluids and

extracellular spaces, including RTLF, either associated with other biomolecules or encapsulated in

extracellular vesicles (EVs) [11-13]. Acting as functional vehicles of intercellular communication, EVs

transport miRNAs during pathological processes, protecting them from degradation and thereby

enhancing their stability [14]. Given their importance in diseases processes [9, 10], as well as relative

stability, miRNAs hold great potential for various clinical applications, including longitudinal

monitoring. Non-invasive molecular tools for studying lung-derived miRNAs are emerging, though not

yet rigorously evaluated or in clinical use [15]. Existing publications on non-invasive approaches focus

on exhaled miRNAs, primarily using exhaled breath condensate (EBC) for sampling [15-17]. However,

it is still not fully understood where EBC originates from, and there are uncertainties on its formation.

Radioactive markers studies have revealed that EBC predominantly captures central airways, rather

than providing a clear picture of the RLTF in the smaller airways [18].

Given the pivotal roles of extracellular vesicles (EVs) and miRNAs in intercellular communication and

regulatory processes in lung diseases, in this publication we describe the development of a miRNA

analysis approach for non-invasive interrogation of the distal lung using the PExA technology in

healthy subjects. To our knowledge we were the first to report detection of miRNA in PExA samples

[19]. By exploring the feasibility of detecting miRNAs in PExA samples, our study contributes to the

advancement of respiratory research and disease diagnostics.

 MATERIALS AND METHODS

Sample collection

PEx were collected using the PExA 2.0 instrument (PExA AB, Gothenburg, Sweden), according to

previously described methods [20, 21]. The collection was performed by impaction on dried blood

spot (DBS) membranes (Whatman 903 Protein saver, specially tailored for PExA, PExA AB, Gothenburg,

Sweden) or on Flinders Technology Associates (FTA) membranes (QIAcard FTA Elute Indicating Micro,

Qiagen, specially tailored for PExA, PExA AB), a chemically coated cellulose membrane type that lyse

cells on contact, denature proteins and protect nucleic acids from degradation. All subjects performed

a standardised breathing manoeuvre [22, 23] with a mouthpiece and a two-way, non-rebreathing

valve (Series 1420; Hans Rudolph, Shawnee, KS, USA), wearing a nasal clip during all process. Prior to

the sampling, the subjects took a minimum of three tidal breaths of HEPA-filtered air via the PExA

instrument to eliminate particle contamination from the surrounding ambient air. From each subject,

240 ng of exhaled particles were collected. Following the collection process, the sample holder

containing the impact membrane was directly placed in a dust- and RNAse-free container, and the

area of the membrane impacted by the particles (PEx sPOT) were excised using a puncher specifically

designed for this purpose. Subsequently, the PEx sPOTs were transferred to 1.5 ml Eppendorf® DNA

LoBind tubes for miRNA analysis, or Protein LoBind tubes for proteomic analyses, and stored at -80°C

until use. All the procedure to prepare the sample holder and to collect the PEx sPOTs were performed

under a laminar flow hood under sterile conditions to prevent sample contamination.

Protein workflow

Two distinct PEx proteome analyses were design to evaluate both the technical reproducibility, the

detection limit, and the reproducibility of repeated sampling the same individual on different days. In

Experiment 1, we collected 240 ng PEx sample from one healthy subject and divided it into three

replicates composed of three PEx dots (3 replicates of 72 ng PEx). In Experiment 2, we collected

another sample from the same subject on a different day, maintaining the same time of the day,

dividing it into five different replicates, each composed of a single PEx sPOT (5 replicates of 24 ng PEx).

Subsequently, we proceeded with the protein analysis workflow.

Non-targeted proteome analysis

Excised PEx membranes spots were wetted with a total volume of 20 µL lysis buffer composed of 0.2%

n-dodecyl-β-D-maltoside (DDM; Sigma-Aldrich) in 100 mM triethylammonium bicarbonate (TEAB;

Fluka) and sonicated in water bath for 10 min. Suspensions were vortexed and centrifuged at low

speed for 5 min at room temperature to sediment PEx sPOTs and lysed proteins. Subsequently,

samples were digested by adding 10 μL of 20 ng/uL sequencing grade trypsin (Promega) into the lysis

buffer, with a final trypsin concentration of 6.7 ng/μL. Following a short vortexing samples were

incubated for 4 h at 37˚C and centrifuged at 10,000x *g* for 5 min. The supernatants were transferred

to sample vials and dried in a vacuum concentrator (Eppendorf) until dry.

Peptides were reconstituted in solvent A (0.1% formic acid, FA in 2% acetonitrile, ACN) and 2 μL of

each sample was injected using an Ultimate 3000 nano-flow UHPLC system (Thermo Fisher Scientific).

Peptides were captured on a 2-cm Acclaim PepMap trap column (Thermo Fisher Scientific) and

separated on a heated 25-cm long EASY-Spray C18 column (Thermo Fisher Scientific) applying a 90

min long gradient: 4-26% of solvent B (0.1% FA in 98% ACN) in 90 min, 26-95% in 5 min, and 95% of

solvent B for 5 min at a flow rate of 300 nL/min.

Mass spectra were acquired on an Orbitrap Eclipse tribrid mass spectrometer (Thermo Fisher

Scientific) equipped with high-field asymmetric ion mobility spectrometry (FAIMS) in *m/z* 300 to 1500

at resolution of R=120,000 (at *m/z* 200) for full mass, targeting $8x10^5$ ions in maximum 100 ms,

followed by data-dependent higher-energy collisional dissociation (HCD) fragmentations of precursor

ions with a charge state 2+ to 6+, using 20 s dynamic exclusion. The tandem mass spectra of the top

precursor ions were acquired in 1 s cycle time by the linear ion trap in rapid mode, targeting $3x10^4$ ions

for maximum injection time of 20 ms, setting quadrupole isolation width to 1 Th and normalized

collision energy (HCD) to 30%. FAIMS was operated at three different compensation voltages in each

cycle using either -40, -60 or -80 V.

Proteome data analysis

The raw files were searched against a human UniProt database using MS Amanda v2.0 search engine

[Ref: https://doi.org/10.1002/rcm.9088] loaded into Proteome Discoverer 2.5 software (Thermo

Fisher Scientific). MS1 precursor mass tolerance was set at 10 ppm, and MS2 tolerance was set

at 0.6 Da. Search criteria included variable modifications of oxidation (+15.9949 Da) on methionine

residues, deamidation (+0.984 Da) asparagine and glutamine residues. Search was performed with full

trypsin/P digestion and allowed a maximum of two missed cleavages on the peptides analyzed from

the sequence database. The false-discovery rates of proteins and peptides were set at 0.05. All data

were quantified using the label-free quantitation node of Precursor Ions Quantifier through the

Proteome Discoverer calculating protein abundances based on precursor abundances. The abundance

levels of all proteins were normalized using the total peptide amount normalization node in the

Proteome Discoverer.

Protein-protein interaction Network (PPI) and Pathway analysis

Pathway enrichment analysis and PPI network visualization was performed using STRING database v.

12.0 [24]. All 69 proteins identified in both examinations were analysed. Keratins were eliminated due

to possible contamination concerns, and a few protein IDs were not found in the STRING database.

The final pathway analysis and visualization network included 62 proteins. For the PPI network, the

minimum required interaction score was set to medium confidence (cutoff 0.4). The obtained network

was then transferred to Cytoscape v. 3.9.1[25] for graphical processing.

miRNA workflow

Two PExA samples were collected from two healthy subjects. We extracted all available PEx sPOTs

(10) from the first membrane and then, we separately processed two and one PEx sPOTs, from the

second membrane. Subsequently, we proceeded with the RNA extraction workflow. Details of the

protocol are reported in **Figure 2A**.

RNA extraction

The vials containing PEx sPOT were transferred from the storage state to dry ice to prevent premature

thawing. 260 µL of Buffer RLT (Qiagen) were added to each sample to disrupt it while it was in a frozen

state and allowed to incubate at room temperature for 10 min. The samples were periodically

vortexed, and pipette mixed through pipetting until PEx sPOT were disintegrated, thereby ensuring

the creation of a pseudo-uniform solution. Total RNA, including the miRNA fraction, was extracted

using the miRNeasy Tissue/Cells Advanced Micro Kit (Qiagen) according to manufacturer's instruction.

The gDNA Eliminator Spin Column step is essential as it allows to separate any paper-based particles

from the liquid, that could interfere with any downstream steps. The RNA fraction was eluted in 15 µL

nuclease-free water, and the process of elution was performed twice with an incubation time of 5 min

each to increase the elution efficiency. For samples destined for the qRT-PCR-based miRNA assays, 1µl

of RNA spike-in mix (Qiagen) was added to the lysis buffer during RNA isolation to monitor RNA

extraction efficacy.

Library Preparation and RNA sequencing

The QIAseq miRNA Library Kit (Qiagen, Hilden, Germany) was used to prepare the library, which has

integrated Unique Molecular Indices (UMIs) to facilitate measurement of pre-amplification miRNA

levels. The libraries were sequenced on the NextSeq 2000 instrument (Illumina, San Diego, US) with a

100 nt (Read1)- 8 nt (Index1) setup using 'P2' flowcell. The fastq data were uploaded to the RNA-seq

Analysis Portal (RAP) 5.0 powered by QIAGEN CLC Genomics Workbench to obtain the UMI raw counts

(reference: miRBase_v22, Homo sapiens; analysis workflow version: 1.2).

cDNA Synthesis and qRT-PCR based miRNA assays

For validation experiments using qRT-PCR based miRNA assays, a set of samples were collected in an

identical manner as for the RNAseq analysis described above. After extraction, RNA was reverse

transcribed using miRCURY Locked Nucleic Acid (LNA™) Reverse Transcription (RT) kit (Qiagen)

according to manufacturer's instruction. For all samples, the reaction mix included the maximum

allowable amount of RNA (6.5µL) in a final volume of 10 µL. The cDNA synthesis control (UniSp6) and

the PCR amplification Ctrl (cel-miR-39-3p, called UniSp3) coming from the Spike-in mix (Qiagen) were

introduced to the reverse transcription reaction to assess its efficacy. A miRCURY LNA miRNA Custom

PCR Panel (Qiagen) was used and the set of primers for the miRNAs selected for the custom plate were

all wet-lab validated for qPCR. cDNA samples were amplified using miRCURY LNA™ SYBR Green PCR

kit (Qiagen) and the real-time reaction was performed using a QuantStudio™ 7 Flex Real-Time PCR

System (Applied Biosystem Foster City, CA, USA), with the following protocol: 95°C for 2 min, followed

by 40 cycles consisting of denaturation (10 s at 95°C) and combined annealing/extension (60 s at 56°C).

Melting curve analysis from 60–95°C was performed post-amplification to evaluate product

specificity.

RNAse activity test

The presence of RNase activity in BALF and PExA samples was assayed using the RNaseAlert® Lab Test

Kit (Thermo Fisher Scientific). BALF samples were kept undiluted and thawed on ice when frozen. The

assays were conducted following the manufacturer's protocol, with fluorescence measured using the

Varioskan™ LUX multimode microplate reader at an excitation wavelength of 490 nm. As part of this

test some DBS membrane were pre-treated, before collecting the samples, with the reagents that

destroy RNases, like RNaseZap™ (ThermoFisher Scientific) and lysis buffer from the miRNeasy

Tissue/Cells Advanced Micro kit (Qiagen).

Evaluation of the potential interference of the membrane in the extraction of the RNA

To test if the membrane matrix materials interfere with the RNA extraction procedure and workflow,

5 PEx sPOTs each of DBS an FTA membrane were wetted with 0.2 µL of RNA spike-in mix (Qiagen),

making a total of 1 µL across 5 punches. RNA extraction was then carried out as described above. To

control for adsorption effects, a positive control was performed by adding 1 µL of the RNA spike-in

mix directly to lysis buffer, without including any PEx sPOTs membrane.

Ethical Statement

Analysis of miRNA and proteins collected by the PExA method has been approved by the Swedish Ethical Authorities (Case no. 2022-03768-01). For the purpose of methods optimizations, PEx material was stripped of all personal identifiers at the time of collection, and all downstream handling and analysis was performed in a de-identified manner.

RESULTS

PEx protein content

Non-targeted proteome analysis of all replicates of the two experiments revealed a total of 163

proteins detectable in DBS PExA membranes, with 69 proteins in common between the two

experiments (**Figure 1A**). We identified 142 proteins in the first experiment (72 with more than 1

unique peptide) and 90 in the second (44 with more than 1 unique peptide). Of the 69 common

proteins, 58 were identified in all replicates of both experiments. All the proteins in this list that are

present in the STRING database are shown in **Figure 1B**, relevant information and detailed gene names

are reported in supplementary material (**Table S1**). The provided list exhibited strong, statistically

significant enrichment (false discovery rate, FDR ≤ 0.05) for various terms associated with extracellular

secretion-related components, such as extracellular exosome, blood micro-particle, vesicle and

secretory granule (**Figure 1C**). Complete list and further details of the pathway analysis are reported

in the supplementary materials (**Table S2**). On one PEx sPOTs, 70 proteins were identified in all the 5

replicates, of which 43 were identified with more than one unique peptide, that give strong

statistically significant enrichment for terms associated with extracellular release processes (data

showed in supplementary material, **Table S3**), providing a good confidence of the analysis performed

even in one single PEx sPOT.

miRNA content evaluation on DBS membrane using miRNA NGS sequencing

A total 181 miRNAs were detected with at least 1 UMI count in at least 1 of the samples. A set of 39

miRNAs were detected with at least 10 UMI counts in at least one of the samples (**Figure 2B**). The

amount of UMI counts is higher in the samples processed with 10 PEx sPOTs and then they gradually

decrease in the samples processed with 2 and 1 PEx sPOTs for both the subjects employed for this

analysis. The number of counts in the sample with only 1 PEx sPOTs are extremely low or absent for

most of the miRNAs. Most PExA-identified miRNAs are present in respiratory tract miRNA atlases

(lungs, bronchus, and saliva[26, 27]), as reported in the right part of the **Figure 2B.** The highest number

of UMI counts was registered by miR-125b-5p in the 10 PEx sPOTs of subject 1, and then the number

gradually decreased in the 2 and 1 PEx sPOT settings. Subject 2 exhibits the same trend, but with lower

counts. We detected many members of the let-7 miRNA family (let-7a-5p, let-7b-5p, let-7c-5p, let-7e-

5p, let-7f-5p, let-7i-5p) with high UMI counts. They exhibit the same reduction trend as miR-125b-5p,

transitioning from 10 to 2 to 1 PEx sPOTs.

RNA degradation evaluation in lung related samples

During the QC analysis of miRNA-seq data (data not shown), a significant proportion of reads that
were shorter than expected mapped to RNAs other than miRNA, indicating potential RNA degradation.
We then performed a fluorimetric assay to evaluate the presence of RNAse activity in the airways and
exclude any contamination coming from the PEx tools and workflow. First, we measured RNase
activity in PExA samples collected on a DBS membrane and in fresh and frozen BALF samples to get
another biological sample coming from the airways to compare (**Figure 3A**). For the BALF we tested
the supernatant only and the complete version with cells. All samples have significantly greater
fluorescence than the negative control, indicating strong RNase contamination. We then measured
the RNAse activity on the membrane only and applying some RNAse inactivation strategies (**Figure
3B**). The fluorescence level of the DBS membrane alone was comparable to the negative control.
Samples collected on DBS without any protective reagents resulted in being contaminated by RNAses,
even if the activity was a bit less in the samples that were immediately freeze at -80°C after the
collection. In samples that were collected on DBS that had been pre-treated with lysis buffer from the
RNA extraction kit, or in samples that were collected on an FTA membrane the RNAse contamination
is minimal and comparable to the negative control.

Validation of different miRNA using real-time PCR technology
From the qRT-PCR based miRNA assays few miRNAs were identified with CT values less than 35 (**Figure
4A**). Although we currently lack controls for the regular quantification, miR-125b-5p appears to be the
most abundant, as reported in the RNAseq experiment. However, some miRNAs have not been
identified using this technique. The internal controls for RNA extraction (UniSp2 and UniSp4), cDNA
synthesis (UniSp6), and PCR amplification (UniSp3) appear to be reliable and uniform across the
different samples, ensuring a high level of reproducibility in the overall workflow (**Table S4**).

**The membranes for PExA collection do not interfere with the extraction of the RNA and the
downstream steps**
To evaluate if the membranes, DBS or FTA, interfere with the workflow we used the spike-in mix used
during the main experiment. We decided to use this mix instead of a real PEx sample considering
extremely low amount of RNA in PEx samples and the potential variability of the collection procedures
that we are trying to understand. The spike-in mix was added directly to pre-cut PEx sPOTs and
processed following the PExA RNA extraction process performed in the main experiment and miRCURY
LNA miRNA PCR assay steps. CT values for five synthetic miRNAs of the spike-in kit (UniSp5, UniSp4,
UniSp2 as RNA extraction controls, UniSp6 for cDNA synthesis, and UniSp3 for RT-PCR) were
comparable to controls where the mix was processed without membranes (**Figure 5**). Thus, the DBS

and FTA membranes do not interfere with RNA extraction or affect the enzymes used in cDNA

synthesis and RT-PCR.

DISCUSSION

This pilot study discusses the initial setup and optimization processes for miRNA analysis in PEx samples. It is the first exploration of miRNAs in PEx to date, and it was performed on healthy people to assess the approach's feasibility. As previously pointed out in other publications, the composition of PEx is similar to that of respiratory tract lining fluid (RTLF), as exhalation after breath-holding at residual volume results in the release of a large number of minute particles generated from the fluid coating the small airways [2, 4]. The analysis previously reported of the PEx proteome highlighted a significant enrichment of extracellular proteins associated with exosomes [2], and mass spectrometry-based identification showed stable protein content in healthy individuals, affirming its reliability [4]. Although a wealth of information regarding the protein content of PExA samples is currently available, we wanted to go further with our own analysis to establish which proteins could be used in the future as internal controls to normalize RNA or other types of data. In details, we wish to investigate both the minimum material quantity required for analysis, as well as the reliability of the assay. Our proteomic analysis performed on PEx samples correlates precisely with the findings already reported, showing extremely significant enrichments for terms related to extracellular secretory-related components, such as extracellular exosome and vesicle. Furthermore, we show that this information can be captured even when analysing a single PEx sPOT, paving the way for combinatory analysis of various biomolecules (e.g., proteins and miRNAs) using a single membrane, considering the 10 PEx sPOTs available for each of them. All these results point to the highly likely existence of diverse types of EVs in PEx samples. Numerous physiological and diseased states are known to be regulated by miRNAs contained in EVs that facilitate intercellular communication, and this has also been reported for the lungs[9, 12]. In detail, exosomes and their enclosed miRNAs have been detected in various clinical lung samples, including bronchoalveolar lavage fluid (BALF) and sputum supernatants, during the investigation of respiratory diseases, like asthma[28]. Furthermore, cells could adjust the production and composition of these EVs in response to various stimuli; therefore, examining them provides useful insights into pulmonary disorders[29]. The observation that a significant portion of the PEx proteome exhibits a statistically significant association with extracellular vesicles and granules provided us with persuasive evidence that intact miRNAs could exist in such samples, as they are protected when they enter the vesicles. Given the limitations provided by the small amount of starting material, we used a combination of kits and resources to extract and analyse miRNAs from small amounts of starting material. The methods appeared effective, as we were able to discover several mature miRNAs. The kit's incorporation of the UMI that we used enabled us to obtain a realistic panoramic view of the miRNA that could be identified in the samples, unaffected by the PCR duplicate bias, which is a frequent issue with degraded or low-

abundance RNA samples. As reported in the results, for both subjects, the number of counts

decreased with the reduction of the number of PEx sPOTs employed for the RNA extraction. In

particular, it seems that using only one PEx sPOT is not enough to get miRNA information, as it seems

to be for protein analysis. A substantial presence of RNAse activity was observed in the PEx sample

after the collection, as well as in BALF, but not in the membrane before the collection that we prepared

in a sterile and dust-free environment. Moreover, the RNAse activity is comparable to the negative

controls if the membranes contain chemicals able to protect the RNA. These findings indicate that the

membrane is RNAse-free, whereas the airways and related samples are highly contaminated with

RNAses. This supports the hypothesis that extracellular miRNAs found in the respiratory tract that are

used for intercellular communication require protection in order to reach their target, and so the

miRNAs we discovered are most likely from EVs, where they are protected from RNAse activity, since

we don't expect to collect cells in PExA samples due to the size of the particles collected (0.5–7 μm).

To confirm the presence of some of the miRNAs discovered in miRNA sequencing, and to start testing

a reproducible and fast technique to use in clinical practice, we selected the most abundant ones and

designed a custom-made miRCURY LNA™ SYBR Green PCR assay from Qiagen. This product employs

LNA-enhanced and Tm-normalised primers, which allow for accurate and reliable quantification of

individual miRNAs with exceptional sensitivity and an extremely low background. With this easy-to-

assay and commercially available tool we were able to confirm the existence of some of the selected

miRNA in PEx. Even though we can't consider this system quantitative yet, since we haven't identified

stable miRNAs in all samples that we can use in normalization, the most abundant miRNA from the

sequencing data is the one with the lowest CTs (miR-125b).

The identified miRNAs are also reported in other publications that analyse miRNA in the lung. In

particular, the let-7 miRNA family, which is well represented in PEx samples, showed relatively high

counts for both subjects, with let-7b-5p and let-7c-5p having the highest expression. The significant

level of preservation observed in let-7 family members across several animal species implies that they

likely have crucial and maybe comparable functions in the biological processes of different organisms

[30, 31]. As summarized by Lee and colleagues in their comprehensive review of the let-7 family [32],

it appears that they exert two primary biological functions: crucial regulator of terminal differentiation

and fundamental tumour suppressor. Furthermore, let-7 family members, like numerous EV-secreted

eukaryotic miRNAs, are implicated in different respiratory diseases, including chronic obstructive

pulmonary fibrosis, idiopathic pulmonary fibrosis, and asthma [33], providing a good rationale for

detecting them as biomarkers at different stages of clinical studies. For example, based on a miRNA

profile of exosomes performed on BALF, some of these family members were included in a small set

of miRNAs able to distinguish asthmatics from healthy participants at baseline [28]. Regarding the

most common let-7 member found in PEx, let-7b-5p, Koeppen and colleagues discovered that this

miRNA could control genes involved in the development of biofilms in a bacterial species such as *P.*

aeruginosa. Specifically, they demonstrated that let-7b-5p delivered by EVs from airway epithelial cells

reduces biofilm formation and increases antibiotic sensitivity in *P. aeruginosa* [33]. Moreover, PEx

samples contained four of the five most common miRNAs detected in airway epithelial cells-EVs by

Koeppen and colleagues [33], which account for more than half of all miRNA sequence reads.

Identification of miRNAs that are directly produced and encapsulated by airway epithelial cells

provides further support for the concept that PEx samples can yield insights into intra-vesicular

communication within the small airways, given that PEx particles theoretically form in the lower

respiratory tract via the mechanism of rupture of the RTLF film during airway reopening after airway

closure. Other miRNAs that were detected in PEx samples from both patients with a high number of

counts were miR-125b-5p and miR-125a-5p. Evidence from many articles suggests that the miR-125

family plays a controvert role in diseases, acting as both tumour suppressors and promoters across

different carcinomas [34]. What is clearer is that miR-125 influences the fate of immune cells by

targeting and regulating a set of target genes to enhance the effect against intra- or extracellular

pathogens or by responding to the regulation of specific signal stimulation like IL and IFN [34], key

mediators in pulmonary diseases like asthma. However, efforts to use miR-125 as a target or

biomarker are ongoing and require further research.

Looking at the miRNA atlas, the majority of the miRNAs detected in PEx are reported to be present in

the lungs and bronchi, the only two elements of the respiratory system available in the miRNA tissue

atlas. We then checked if they were also present in the respiratory glands, since some concerns had

arisen in the past given the risk of saliva contamination of the samples since PEx are exhaled through

the oral cavity [4]. MiR-4508, miR-1246, and miR-4492 are the only three miRNAs identified in PEx

that, based on the miRNA tissue atlas, exhibit expression in the bronchus and/or lung but not in the

salivary glands, but they are not the most expressed. Thus, based on these results, we cannot

completely exclude the possibility that our identified miRNAs originate from saliva; however, our

proteomic data did not detect amylase, the most prevalent protein in saliva, which has previously

been cited as evidence to deny the presence of saliva in PEx samples [4].

The investigation of miRNA detection in non-invasive samples is gaining attention since techniques

such as BALF and induced sputum have varying degrees of impact on patients and cannot be repeated

multiple times, especially in diseased patients. Shi and colleagues provided a comprehensive report on

exhaled miRNAs in lung cancer research utilizing the EBC sampling technique [15]. Specifically, they

examine the efficacy of a selected panel of 24 miRNAs in differentiating individuals with non-small cell

lung cancer from control subjects. Although further refinement of their exhaled miRNA interrogation

technique is required, the researchers have shown that the detection of miRNAs in EBC using qPCR

can lead to a modest enhancement in differentiating between cases and controls, beyond what can

be achieved with clinical factor models alone [15]. Nevertheless, the composition of biomolecules in

EBC is significantly affected by their physical characteristics as well as their relative concentrations in

the RTLF [35]. Moreover, dilution poses uncertainty in EBC, as well as in BAL, where external fluid

addition complicates concentration calculations and normalization. Conversely, PEx samples are dry

upon collection, thus enhancing the reliability of concentration measurements [4]. Therefore, this

underscores the importance and value of our findings, highlighting the efficacy of the method (PEx)

we employed for miRNA identification as a non-invasive tool of significant interest.

CONCLUSION

This proof-of-concept study aims to verify that miRNAs can be detected in PExA samples, thereby

showing its utility as a non-invasive method for collecting small airway samples for biomarker analysis

in lung disorders. Using PExA, we identified a series of miRNAs previously shown to be of relevance

for pulmonary disorders using two complementary techniques. Moreover, we strongly believe that

the identified miRNAs were detected intact due to their protection via encapsulation within EVs. This

is supported by the enrichment of extracellular secretory-related components in the core protein list

from PExA samples, along with the demonstrated high RNAse activity in the respiratory tract, which

would have otherwise rapidly degraded the miRNAs. This opens the possibility of studying EV-

mediated intracellular communication in distal small airways using PExA as a non-invasive technique.

However, this is only the first step; we are still in the process of understanding the feasibility of miRNA

detection in this type of sample. Further investigations and efforts must be made to establish internal

controls for moving forward with a quantitative evaluation, as well as to understand intra- and inter-

subject variability. Our efforts to understand the core of proteins that can always be detected in PEx

sPOTs can help in this purpose of identifying internal controls to use in the process of normalizing the

miRNAs until we will not find stable miRNAs that can be detected in the same amount in all samples.

Figure 1. Proteins observed in PEx with their biological function and location. **A.** Graphical outline of the experimental workflow. **B.** Pie chart with the number of proteins identified in both experiments. **C.** The protein-protein interaction (PPI) network of the proteins identified in both experiments; the legend describes the features of the network's elements. The edges represent the StringDB score, i.e., how likely STRING is to judge the confidence of the correspondent PPI. For the network realization, the minimum required interaction score was set at 0.4 (medium confidence). The edges are colored based on the co-expression interaction score (with no interaction of this type, the edge is grey). The yellow nodes with black border represent the proteins with the highest level of confidence for identification PExA samples, since they were identified in all the replicates with more than 1 unique peptide. **D.** Pathway analysis of the proteins identified in both experiments. In the table are reported the top five enriched terms for all three GO categories and subcellular compartments. For each

455   enriched term, the -Log10 false discovery rate (FDR) is provided. All the data was obtained from

456   stringDB[24].

Figure 2. *miRNA identified in PEx samples on DBS membrane using RNAseq.* **A.** Graphical outline of the experimental workflow. **B.** List of the miRNA identified. Data are presented as UMI counts in a colour scale from dark red to green. Each left column represents a sample, related to two subjects with varying amounts of PEx sPOTs. The right columns display gene lists sourced from or miRNATissueAtlas2 (with expression levels exceeding 10 as cut-off[26, 27]). A blue dot indicates the presence of the PExA identified miRNA in the corresponding list.

Figure 3. Evaluation of RNAse activity presence in the airways. A. The barplot shows the level of RNAse activity in different sample types coming from the airways, like PEx samples collected on DBS membrane, as well as in fresh and frozen BALF. **B.** The barplot shows the level of RNAse activity in the PEx samples using different condition and membranes (DBS Ctrl is the membrane alone, without any sample on it; LB Pretr. = the DBS membrane was pretreated with the lysis buffer from the RNA extraction kit). In both plots, according to the manufacturer, samples with fluorescence levels exceeding 2-fold of the corresponding negative control (dotted line) are considered affected by the presence of RNase activity. All the graph-bars in the two plots are reported as mean ± SD of two different replicates (except for the controls and the blank). The dotter texture of bar plots highlights samples with PExA.

Figure 4. *miRCURY LNA™ SYBR Green PCR assay in PEXA samples.* Results of the PCR array performed on the same set of samples as in the RNAseq analysis. **A.** Visualization of miRNA presence across diverse samples. The heatmap is constructed using a color scale that represents the detected Ct values, with a threshold set at 0.07. Grey cells represent undetected miRNAs in the respective sample. **B.** Amplification plot of the internal controls of the array and the most abundant miRNA in S1 and S2 (miR125b-5p).

Figure 5. *Evaluation of potential membrane interference during RNA extraction.* PCR cycle thresholds for RNA spike-in controls (UniSp5, UniSp4, UniSp2) for RNA isolation, cDNA synthesis (UniSp6), and RT-PCR (UniSp3) show comparable results between DBS (white), FTA (purple), and no membrane (blue) conditions, with no statistically significant differences among the conditions (Kruskal-Wallis test). All graph bars are reported as mean ± SD of three different replicates.

<u>REFERENCES</u>

1. Sutedja G: **New techniques for early detection of lung cancer.** *Eur Respir J Suppl* 2003, **39:**57s-
66s.
2. Ostling J, Van Geest M, Olsson HK, Dahlen SE, Viklund E, Gustafsson PM, Mirgorodskaya E,
Olin AC: **A novel non-invasive method allowing for discovery of pathologically relevant**
proteins from small airways. *Clin Proteomics* 2022, **19:**20.
3. Larsson P, Larstad M, Bake B, Hammar O, Bredberg A, Almstrand AC, Mirgorodskaya E, Olin
AC: **Exhaled particles as markers of small airway inflammation in subjects with asthma.** *Clin*
Physiol Funct Imaging 2017, **37:**489-497.
4. Bredberg A, Gobom J, Almstrand AC, Larsson P, Blennow K, Olin AC, Mirgorodskaya E: **Exhaled**
endogenous particles contain lung proteins. *Clin Chem* 2012, **58:**431-440.
5. Larsson P, Bake B, Wallin A, Hammar O, Almstrand AC, Larstad M, Ljungstrom E,
Mirgorodskaya E, Olin AC: **The effect of exhalation flow on endogenous particle emission**
and phospholipid composition. *Respir Physiol Neurobiol* 2017, **243:**39-46.
6. Postma DS, Brightling C, Baldi S, Van den Berge M, Fabbri LM, Gagnatelli A, Papi A, Van der
Molen T, Rabe KF, Siddiqui S, et al: **Exploring the relevance and extent of small airways**
dysfunction in asthma (ATLANTIS): baseline data from a prospective cohort study. *Lancet*
Respir Med 2019, **7:**402-416.
7. Hogg JC, Pare PD, Hackett TL: **The Contribution of Small Airway Obstruction to the**
Pathogenesis of Chronic Obstructive Pulmonary Disease. *Physiol Rev* 2017, **97:**529-552.
8. Bhaskaran M, Mohan M: **MicroRNAs: history, biogenesis, and their evolving role in animal**
development and disease. *Vet Pathol* 2014, **51:**759-774.
9. Brown D, Rahman M, Nana-Sinkam SP: **MicroRNAs in respiratory disease. A clinician's**
overview. *Ann Am Thorac Soc* 2014, **11:**1277-1285.
10. Hoefel G, Tay H, Foster P: **MicroRNAs in Lung Diseases.** *Chest* 2019, **156:**991-1000.
11. Llorens-Revull M, Martinez-Gonzalez B, Quer J, Esteban JI, Nunez-Moreno G, Minguez P,
Burgui I, Ramos-Ruiz R, Soria ME, Rico A, et al: **Comparison of Extracellular Vesicle Isolation**
Methods for miRNA Sequencing. *Int J Mol Sci* 2023, **24**.
12. Gallo A, Tandon M, Alevizos I, Illei GG: **The majority of microRNAs detectable in serum and**
saliva is concentrated in exosomes. *PLoS One* 2012, **7:**e30679.
13. Levanen B, Bhakta NR, Torregrosa Paredes P, Barbeau R, Hiltbrunner S, Pollack JL, Skold CM,
Svartengren M, Grunewald J, Gabrielsson S, et al: **Altered microRNA profiles in**
bronchoalveolar lavage fluid exosomes in asthmatic patients. *J Allergy Clin Immunol* 2013,
131:894-903.
14. Shurtleff MJ, Temoche-Diaz MM, Karfilis KV, Ri S, Schekman R: **Y-box protein 1 is required to**
sort microRNAs into exosomes in cells and in a cell-free reaction. *Elife* 2016, **5**.
15. Shi M, Han W, Loudig O, Shah CD, Dobkin JB, Keller S, Sadoughi A, Zhu C, Siegel RE, Fernandez
MK, et al: **Initial development and testing of an exhaled microRNA detection strategy for**
lung cancer case-control discrimination. *Sci Rep* 2023, **13:**6620.
16. Mozzoni P, Banda I, Goldoni M, Corradi M, Tiseo M, Acampa O, Balestra V, Ampollini L, Casalini
A, Carbognani P, Mutti A: **Plasma and EBC microRNAs as early biomarkers of non-small-cell**
lung cancer. *Biomarkers* 2013, **18:**679-686.
17. Pinkerton M, Chinchilli V, Banta E, Craig T, August A, Bascom R, Cantorna M, Harvill E, Ishmael
FT: **Differential expression of microRNAs in exhaled breath condensates of patients with**
asthma, patients with chronic obstructive pulmonary disease, and healthy adults. *J Allergy*
Clin Immunol 2013, **132:**217-219.
18. Larsson P: **Exhaled particles for monitoring of airway inflammation - Sampling and analysis**
of endogenous particles from breath. *Doctoral Thesis.* University of Gothenburg, Department

of Public Health and Community Medicine - Institute of Medicine - Sahlgrenska Academy at
University of Gothenburg; 2016.

19. Bonatti M, Ezerskyte M, Lindberg P, Wheelock ÅM: **miRNA identification in Particles in**
Exhaled Air (PExA) to study lung diseases. *Eur Respir J* 2023, **62**.

20. Almstrand AC, Ljungstrom E, Lausmaa J, Bake B, Sjovall P, Olin AC: **Airway monitoring by**
collection and mass spectrometric analysis of exhaled particles. *Anal Chem* 2009, **81:**662-
668.

21. Larstad M, Almstrand AC, Larsson P, Bake B, Larsson S, Ljungstrom E, Mirgorodskaya E, Olin
AC: **Surfactant Protein A in Exhaled Endogenous Particles Is Decreased in Chronic**
Obstructive Pulmonary Disease (COPD) Patients: A Pilot Study. *PLoS One* 2015, **10:**e0144463.

22. Almstrand AC, Bake B, Ljungstrom E, Larsson P, Bredberg A, Mirgorodskaya E, Olin AC: **Effect**
of airway opening on production of exhaled particles. *J Appl Physiol (1985)* 2010, **108:**584-
588.

23. Holmgren H, Gerth E, Ljungstrom E, Larsson P, Almstrand AC, Bake B, Olin AC: **Effects of breath**
holding at low and high lung volumes on amount of exhaled particles. *Respir Physiol*
Neurobiol 2013, **185:**228-234.

24. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T,
Doncheva NT, Pyysalo S, et al: **The STRING database in 2023: protein-protein association**
networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic*
Acids Res 2023, **51:**D638-D646.

25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker
T: **Cytoscape: a software environment for integrated models of biomolecular interaction**
networks. *Genome Res* 2003, **13:**2498-2504.

26. **miRNA Tissue Atlas** [https://ccb-web.cs.uni-saarland.de/tissueatlas2/]

27. Keller A, Groger L, Tschernig T, Solomon J, Laham O, Schaum N, Wagner V, Kern F, Schmartz
GP, Li Y, et al: **miRNATissueAtlas2: an update to the human miRNA tissue atlas.** *Nucleic Acids*
Res 2022, **50:**D211-D221.

28. Sastre B, Canas JA, Rodrigo-Munoz JM, Del Pozo V: **Novel Modulators of Asthma and Allergy:**
Exosomes and MicroRNAs. *Front Immunol* 2017, **8:**826.

29. Fujimoto S, Fujita Y, Kadota T, Araya J, Kuwano K: **Intercellular Communication by Vascular**
Endothelial Cell-Derived Extracellular Vesicles and Their MicroRNAs in Respiratory Diseases.
Front Mol Biosci 2021, **7**.

30. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE,
Degnan B, Muller P, et al: **Conservation of the sequence and temporal expression of let-7**
heterochronic regulatory RNA. *Nature* 2000, **408:**86-89.

31. Hertel J, Bartschat S, Wintsche A, Otto C, Students of the Bioinformatics Computer L, Stadler
PF: **Evolution of the let-7 microRNA family.** *RNA Biol* 2012, **9:**231-241.

32. Lee H, Han S, Kwon CS, Lee D: **Biogenesis and regulation of the let-7 miRNAs and their**
functional implications. *Protein Cell* 2016, **7:**100-113.

33. Koeppen K, Nymon A, Barnaby R, Bashor L, Li Z, Hampton TH, Liefeld AE, Kolling FW, LaCroix
IS, Gerber SA, et al: **Let-7b-5p in vesicles secreted by human airway cells reduces biofilm**
formation and increases antibiotic sensitivity of P. aeruginosa. *Proc Natl Acad Sci U S A* 2021,
118.

34. Sun YM, Lin KY, Chen YQ: **Diverse functions of miR-125 family in different cell contexts.** *J*
Hematol Oncol 2013, **6:**6.

35. Bicer EM, Mudway IS, King's College L: *Compositional characterisation of human respiratory*
tract lining fluids for the design of disease specific simulants. Great Britain: King's College
London; 2015.

A.

Subject: **S1 & S2**

Two Collection procedures of 240 ng per subject

10 PEx sPOTs   **2 & 1** PEx sPOTs

miRNeasy
Tissue/Cells
Advanced **Micro** Kit

↓

QIAseq miRNA Library Kit +
miRNA Sequencing

B.

UMI counts

N° PEx sPOTs
1   2   10

Not detected

S1			S2			LungAtlas	BronchusAtlas	SaliGlandsAtlas	id
						●	●	●	hsa-miR-125b-5p
						●	●	●	hsa-let-7b-5p
						●	●	●	hsa-let-7c-5p
						●	●	●	hsa-miR-125a-5p
						●	●	●	hsa-let-7i-5p
						●	●	●	hsa-miR-16-5p
						●	●	●	hsa-let-7a-5p
						●	●	●	hsa-miR-21-5p
						●	●	●	hsa-miR-199a-3p
						●	●	●	hsa-let-7f-5p
						●	●	●	hsa-let-7e-5p
						●	●	●	hsa-miR-26a-5p
							●		hsa-miR-4508
						●	●	●	hsa-miR-143-3p
						●	●	●	hsa-miR-10a-5p
						●	●	●	hsa-miR-99b-5p
						●	●	●	hsa-miR-26b-5p
						●	●	●	hsa-miR-92a-3p
						●	●	●	hsa-miR-146b-5p
						●	●	●	hsa-miR-30d-5p
						●	●	●	hsa-miR-93-5p
						●	●	●	hsa-miR-206
						●	●	●	hsa-miR-29a-3p
						●	●	●	hsa-miR-30a-5p
						●	●	●	hsa-miR-23a-3p
						●	●	●	hsa-miR-221-3p
						●	●	●	hsa-miR-103a-3p
						●	●	●	hsa-let-7d-5p
						●			hsa-miR-4492
						●	●	●	hsa-miR-191-5p
						●	●	●	hsa-miR-424-5p
						●	●	●	hsa-miR-155-5p
						●	●	●	hsa-miR-423-5p
						●	●		hsa-miR-203a-3p
						●	●		hsa-miR-1246
						●	●	●	hsa-miR-205-5p
						●	●	●	hsa-miR-9-5p
						●	●	●	hsa-miR-24-3p
						●	●	●	hsa-miR-1-3p