http://informahealthcare.com/bmk
ISSN: 1354-750X (print), 1366-5804 (electronic)

Biomarkers

Biomarkers, 2015; 20(2): 123–131
© 2015 British American Tobacco. Published by Taylor & Francis. DOI: 10.3109/1354750X.2014.1000970

informa healthcare

RESEARCH ARTICLE

Quantification of plasma microRNAs in a group of healthy smokers, ex-smokers and non-smokers and correlation to biomarkers of tobacco exposure

Anisha Banerjee, David Waters, Oscar M. Camacho#, and Emmanuel Minet

BAT, Group Research and Development, Southampton, UK

Abstract

The stability of circulating miRNAs, their non-invasive sampling techniques and deregulation in diseases make them potential candidate biomarkers of biological effect. Here, we profiled the level of 84 plasma miRNAs in 30 smokers, 20 non-smokers and 20 ex-smokers. A robust statistical strategy was applied with replicate samples to account for reproducibility of the results. We identified differential expression of miR-124 and let-7a between the smoking and control groups. We further explored the dose–response relationship of miR-124 and let-7a with two biomarkers of tobacco exposure and found that this relationship was affected by adjustments based on age, pack-year and gender.

Keywords

Biomarkers, microRNA, plasma, smoking, TNEQ

History

Received 26 November 2014
Revised 17 December 2014
Accepted 17 December 2014
Published online 19 January 2015

Introduction

MicroRNAs (miRNAs) are a family of small, endogenous, non-coding functional RNA molecules of 18–25 nucleotides in length. These regulatory molecules function to modulate the activity of specific mRNA targets either by translational repression or by mRNA degradation (Takahashi et al., 2013). The sequences of miRNAs are evolutionarily conserved across species which suggests an important biological function (He & Hannon, 2004). miRNAs are key regulators of various biological processes including development, differentiation, proliferation, cell death and metabolism (Shen et al., 2013). More than 940 mature miRNAs have been characterized to date in humans (Duttagupta et al., 2011) and it has been hypothesized that nearly 30% of protein-coding genes could be regulated by miRNAs through post-transcriptional mechanisms (Leidinger et al., 2011). The expression of many miRNAs is tissue specific and altered miRNA profiles may reflect abnormalities in developmental regulation or tissue functions (Mi et al., 2013; Sharma et al., 2010).

The presence of miRNAs in circulating body fluids such as plasma or serum was first reported in 2008 (Mitchell et al., 2008). Plasma miRNAs were found to be highly stable even after exposure to severe conditions such as high temperatures, low or high pH, prolonged storage at room temperature and multiple freeze thaw cycles (Creemers et al., 2012). This stability has been attributed to packaging of miRNA into lipoprotein complexes such as exosomes, microvesicles and apoptotic bodies which prevents their degradation (Valadi et al., 2007; Zernecke et al., 2009). Since the discovery of extracellular miRNA in body fluids, a number of studies have reported the association of levels of specific circulating miRNA in blood with various pathophysiological conditions including multiple sclerosis (Fenoglio et al., 2013), coronary artery disease (Fichtlscherer et al., 2010), colorectal cancer (Huang et al., 2010) and liver disease (Wang et al., 2009). Evidence also indicate that unique patterns of altered miRNA expression provide valuable information such as tumor origin, tumor stage and other pathological factors (Iorio & Croce, 2009). Additionally, the predictive value of circulating miRNAs has been reported in breast cancer (Krell et al., 2012), rheumatoid arthritis (Duroux-Richard et al., 2014) and early stage non-small cell lung cancer (Foss et al., 2011). These findings demonstrate the potential for using the levels of specific miRNA expression in body fluids as biomarkers of biological effects that could potentially be predictive of disease.

Cigarette smoking is a high-risk factor for diseases such as lung cancer (Shields, 1999), chronic obstructive pulmonary disease (COPD) and cardiovascular disease (CVD) (Takahashi et al., 2013). MicroRNAs have been reported to be deregulated in smoking-related diseases (Banerjee & Luettich, 2012) and their expression profiles differ between healthy and diseased tissues. Examples include association of plasma miR-21 (Wei et al., 2011) and miR-155, miR-197 and miR-182 as potential biomarkers for early diagnosis of lung

#Oscar M. Camacho is responsible for statistical analysis. E-mail: Oscar_M_Camacho@bat.com
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/Licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
Address for correspondence: Anisha Banerjee, BAT, Group Research and Development, Regents Park Road, Southampton SO15 8TL, UK. Tel: +44 2380 588 481. E-mail: Anisha_Banerjee@BAT.com

cancer (Zheng et al., 2011). The role of circulating miRNAs has also been suggested in tobacco-related inflammatory diseases such as COPD and CVD (Akbas et al., 2012; Dickinson et al., 2013). The increasing diversity of nicotine delivery products and regulatory scrutiny for safer alternatives to combustible cigarettes is driving the discovery of biomarkers that could be used to evaluate adverse biological effects in the absence of epidemiological evidence. Therefore, our objective was to conduct a miRNA screen to identify differentially expressed plasma miRNA among healthy smokers, ex-smokers and non-smokers and potential novel biomarkers of biological effect. In order to identify the most robust miRNA candidates, we implemented a design where duplicate samples were processed and normalized to mitigate the effect of operator and day of collection. Finally, we investigated the correlation of differentially expressed miRNAs with tobacco exposure biomarkers of internal dose (total nicotine equivalent (TNEQ) in urine) and effective dose Cyanoethyl valine (CEVal), an acrylonitrile haemoglobin adduct.

Materials and methods

Clinical study design

A detailed clinical protocol has been described previously in Shepperd et al. (2013) and the study was registered in the Current Controlled Trials database under the reference ISRCTN81286286. In brief, never-smokers ($n = 20$), ex-smokers ($n = 20$) and current smokers ($n = 30$) were recruited from Hamburg (Germany). Inclusion criteria for the smoking groups were that their age was between 23 and 55 years, current smokers of between 10 and 30 (6–8 mg) ISO tar cigarettes per day (CPD) and smokers for at least 5 years before screening. Criteria for the ex-smoking group were that their age was between 28 and 55 years, not having smoked for at least 5 years but having been a regular smoker of between 10 and 30 CPD for at least 5 years. Inclusion criteria for the never smoking group were that their age was between 28 and 55 years, never having smoked more than 100 cigarettes during his/her lifetime and none in the previous 5 years.

Subjects who participated in the study were monitored over a period of 6 months. Participants recorded in electronic diaries daily cigarette consumption, exercise, diet and medications. Cigarette butts were also collected and compared with cigarette per day reported in the diaries. Five in clinic evaluations for smokers and three for non- and ex-smokers were also conducted over the 6 months period where total nicotine equivalent (nicotine + five metabolites) was measured in 24 h urine. All the smokers were provided with a 7 mg ISO Tar Lucky Strike cigarette for the duration of the study to have consistency with the ISO Tar band of the product used. Two plasma samples were collected, the first at day 182 and the second at day 183 for the smokers and the first at day 163 and the second at day 164 for the non-smokers. One plasma sample was collected for ex-smokers at day 163. Corresponding 24 h urine samples and whole blood samples were collected for TNEQ and CEVal biomarkers analysis. Different days were used for these groups for practical reasons including clinical capacity and sample collection. Operator one processed the samples from days 163 and 182 and operator two processed the samples from days 164 and 183. The samples from each smoking status group were divided into random batches and were processed independently by the two operators at different times.

This study was conducted in compliance with the ethical principles of the Declaration of Helsinki, Good Clinical Practice and German Law, including informed consent and was approved by the Ethics Committee of Arztekammer Hamburg, Germany, 29 November 2011 Ref.: PV3824 (Scherer et al., 2014).

miRNA analysis

Blood was collected in an EDTA-containing tube. The samples were immediately centrifuged at 1000*g* for 10 min at 4 °C. The plasma was collected and stored in K2 EDTA tubes at −80 °C until further use.

Total RNA containing small miRNA was extracted from 200 μl of plasma using the miRNeasy® Serum/Plasma kit (Qiagen, Hilden, Germany) which combines phenol/guanidine-based lysis of samples and silica membrane-based purification. At the beginning of this extraction phase, a carrier MS2 RNA (used to increase the extraction yield) and an exogenous spike in control (*Caenorhabditis elegans* miR-39 which enables normalization of non-specific losses incurred during miRNA purification) was added to the samples. The subsequent phase separation, phenol extraction and filter cartridge elution steps were carried out according to the instructions of the manufacturer (Qiagen, Hilden, Germany). Total RNA including miRNA was purified from the aqueous phase using ethanol and eluted from the column with 14 μl RNAse free water.

Reverse transcription was performed using the miScript® reverse transcription kit (Qiagen, Hilden, Germany) in a 10 μl reaction. All procedures were performed following the instructions of the manufacturer.

Pre-amplification was carried out using the miScript® PreAmp PCR kit (Qiagen, Hilden, Germany) in a 25-μl reaction by means of a 12-cycle PCR reaction. Since the concentration of input cDNA was unknown, pre-amplification control experiments were performed with representative samples to determine the optimal dilution factor for pre-amplified cDNA, according to the instructions of the manufacturer (Qiagen, Hilden, Germany) and based on what was described by Kroh et al. (2010).

The expression of miRNAs was determined by quantitative real-time PCR using the SYBR® green-based detection technology (Applied Biosystems, Waltham, MA). A 20-fold dilution of pre-amplified cDNA was used as the input for a 40-cycle qRT-PCR reaction with miRNA-specific primers. A custom pathway focused panel of miRNA containing 84 pre-dispensed miRNA targets commonly found in plasma was used for this purpose (miRNA name and gene card references reported as Supplementary information in Table 1). All 25 μl reactions were performed on the Applied Biosystems® 7500 Fast Real-Time PCR system (Applied Biosystems, Waltham, MA) using the comparative CT ($\Delta\Delta$CT) method of quantitation.

Gene expression data analysis

After completion of the qRT-PCR, the CT values were visually inspected using the fast PCR 7500 software v.2.0.5

(Applied Biosystems, Waltham, MA). An automatic baseline setting of 3–15 and a threshold of 0.03 were applied by default and were manually adjusted when required to ensure maintenance of optimal sensitivity.

Normalization of the plasma miRNA expression level

There are currently no standard endogenous controls for serum miRNA studies (Mo et al., 2012). Therefore, as a first step, the CT values obtained using the spike-in control (*C. elegans* miR-39) was used to calibrate the data sets. A mathematical correction factor was calculated for each plasma sample using the difference between the groups' mean CT values of the spike-in control (*C. elegans* miR-39). This was followed by normalization using the mean of commonly expressed miRNA targets across all the samples. For further data analysis, only those miRNAs with a CT value equal to or below 30, a cut-off recommended by the manufacturer (Qiagen, Hilden, Germany), were taken into account. All CT values above 30 were truncated to 30 to ensure that fold changes are not over-represented.

Biomarkers of exposure

TNEQ and CEVal quantification in urine and blood, respectively, was described in Scherer et al. (2014).

Statistical analysis

Distributions of samples' expression were assessed by operator for outlier detection. After outlier removal, datasets were merged and normalized by operator to account for an observed shift of samples' distributions. Hierarchical clustering based on the Euclidean distance algorithm of the normalized expression data (ΔCt) and principal component analysis (PCA) were performed to assess potential groups among samples. Differences between the levels of miRNA expression on cigarette smoke exposure were tested using a mixed model analysis with the smoker group and adjusting by age and gender as fixed effects; and day of sample processing and subject as random effects ($\Delta\text{Ct} = \text{B0} + \text{smoking status} + \text{age} + \text{gender} + \sigma^2_{\text{sample processing}} + \sigma^2_{\text{subject}} + \sigma^2_{\text{subject}} + \varepsilon$). Statistical significance was adjusted using Bonferroni correction across miRNAs. Additionally, we assessed the effects of using a less stringent statistical method (FDRq value Storey). A ΔΔCT method of relative quantification was used to calculate the difference in the expression levels between non-smokers relative to smokers and ex-smokers as a measure of biological relevance. Fold change was calculated as: $2^{-\Delta\Delta\text{Ct}}$ where ΔΔCt is the ΔCt gene of interest–ΔCt reference for each comparison of interest. An arbitrary 2-fold change cut-off was applied to identify the most significantly down- or up-regulated miRNAs.

Similarly, relationships between miRNAs and biomarkers (TNEQ and CEVal) were assessed using a linear mixed model with age, gender and pack years as fixed effects and subject as a random effect. JMP® Genomics 6.0 software (SAS Inc, Cary, NC) was used to perform the miRNA expression and statistical analysis to determine the differences between the groups.

Results

Outlier detection and normalization

Two operators processed the randomized non-smoker, ex-smoker and smoker samples collected at a two day interval. Total RNA was extracted, reverse transcribed, and 84 miRNA quantified by qRT-PCR as described in the ''Material and methods'' section. Distributions of the samples miRNA expression were plotted separately per operator. Visual inspection highlighted six outliers out of which two were from the first operator and four were from the second.

After outlier removal, samples from both operators were merged (Figure 1A). There was a clear shift of the sample distributions by operator. Samples were normalized by operator (z-score) to correct this effect. Since each operator processed samples collected on a specific day, we cannot discriminate whether the shift was due to operator or day of collection. Nevertheless, the normalization performed mitigates the effect arising from operator/day.

Post-normalization, expression data across all samples appeared to be similar (Figure 1B). There were no strong distinct patterns and the miRNA expression profile obtained by both the operators seemed to be appropriately corrected.

Profiling of the plasma miRNA profiles of smokers, non-smokers and ex- smokers

Unsupervised cluster analysis based on the Euclidean distance was used to visualize the patterns on the expression of the 84 miRNA genes. Hierarchical clustering (Figure 2) illustrated the distinct levels of miRNAs detected in human plasma with abundant miRNA in green, miRNA detected at a moderate levels in black, absent miRNA or miRNA present in very low quantities in red. No clustering based on the smoking status was observed indicating that overall all three groups shared a similar miRNA profile. These results were confirmed by the absence of clusters in a PCA plot reported in Supplementary Figure 1.

miR-124 is differentially expressed in smokers compared to non-smokers and ex-smokers

Hierarchical clustering is useful to distinguish patterns of expression but is not designed to detect discreet but significant changes in a large pool of data. Therefore, we assessed differences between smoking statuses using a mixed model ANOVA. Fold change of untransformed expression data was also used for biological significance criterion. The following pairwise comparisons for each miRNA were considered: smokers versus non-smokers; smokers versus ex-smokers and ex-smokers versus non-smokers. A single miRNA, miR-124, was significantly different ($p = 0.05$, adjusted for the Bonferroni method) between non-smokers and smokers. A greater than 2-fold increase in expression was observed in smokers relative to non-smokers (Figure 3A). A similar increase was observed in smokers compared with ex-smokers for miRNA-124 (Figure 3B). No miRNA was differentially expressed when ex-smokers and non-smokers were compared (Supplementary Figure 2).

Another less stringent adjustment method (FDRq Storey) was used to explore if there were other potential differentially

 Biomarkers, 2015; 20(2): 123–131

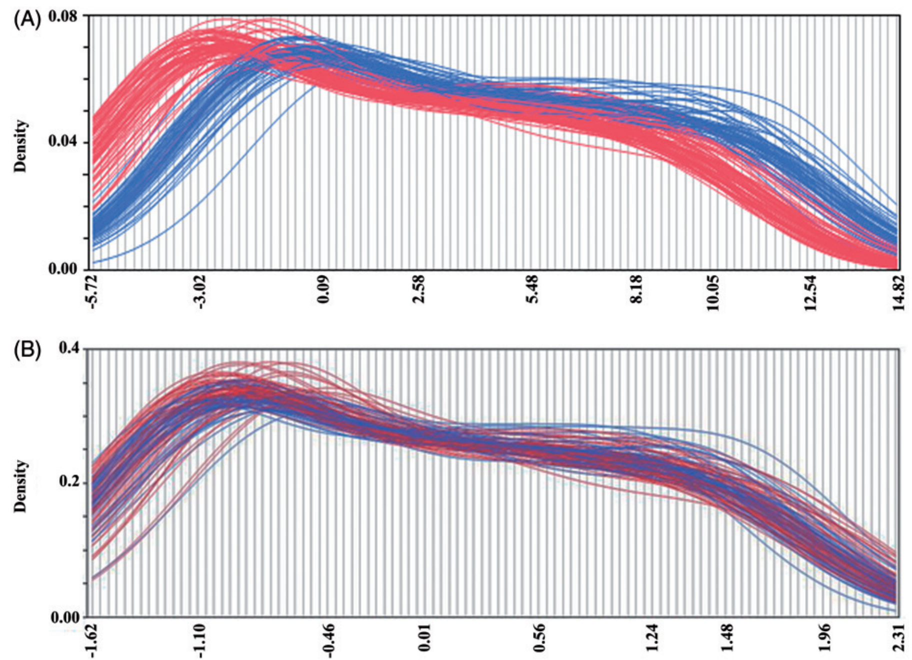

Figure 1. Sample data distribution (ΔCts) prior to normalization by operator (A) and data distribution of samples showing ΔCts post-normalization by operator (B). Colours are representative of operators. Red indicates Operator 1 and blue indicates Operator 2.

expressed miRNAs of interest. A mixed model ANOVA adjusted for days at which the samples were collected was used to assess differences between smokers' groups ($q < 0.05$). Analysis by this method showed that in addition to miR-124, let-7a was also differentially expressed between smokers and ex-smokers. However, let-7a did not appear to be statistically different between smokers and non-smokers.

Association of differentially expressed miRNA with biomarkers of exposure

Next, we wanted to evaluate the potential relationship of miRNAs with biomarkers of cigarette exposure – TNEQ and CEVal. For this, we selected only the miRNAs that were differentially expressed between smokers and non-smokers or ex-smokers.

Regression models were fit between the expression of miR let-7a and miR-124 versus CEVAL and TNEQ. Only data from the smoking group was used since inclusion of the non-smokers and the ex-smokers improved the correlation based on baseline levels.

Our results indicate a significant dose–response relationship ($p = 0.0197$, $R^2 = 0.6102217$) between miR-124 expression and CEVaL with no adjustment for any confounding variables (Table 1A). However, this relationship became non-significant after adjusting for age and gender ($p = 0.0797$, $R^2 = 0.63$) and for age, gender and pack-years ($p = 0.2217$, $R^2 = 0.65$) (Table 1B and C). No significant correlation was observed between miR-124 and TNEQ irrespective of any adjustments.

There was no evidence of a relationship between let-7a and the two tested biomarkers of exposure when not adjusted for any confounding variables. After adjusting for age and gender variables, let-7a showed a significant relationship with TNEQ ($p = 0.0199$, $R^2 = 0.41$) (Table 1B). This association of let-7a remained significant when adjusted for age, gender and pack-years ($p = 0.0413$, $R^2 = 0.43$) (Table 1C). No significant correlation was observed between let-7a and CEVaL regardless of any adjustments.

Discussion

Circulating miRNAs have gained attention as potential biomarkers for detecting and monitoring different physiological and pathological conditions due to their stability, and the presence in biofluids such as sputum and plasma (Mitchell et al., 2008). In this report, we investigated the differential abundance of a panel of 84 miRNA in the plasma of smokers ($n = 30$), non-smokers ($n = 20$) and ex-smokers ($n = 20$) as potential early biomarkers of biological effect. The panel of 84 miRNA was representative of the most common miRNAs found in human plasma and serum. Samples were collected at two time points and analysed by qRT-PCR independently. To investigate the association of candidate miRNAs with two biomarkers of tobacco exposure, urinary total nicotine equivalent (TNEQ) and acrylonitrile haemoglobin adduct (CEVal) were collected for each subject at each time point.

Following qRT-PCR, outlier exclusion and data normalization (Figure 1), we determined whether subjects were grouped by smoking status based on their miRNA profile using hierarchical clustering (Figure 2). No clustering by smoking status groups was observed suggesting that ex-smokers, non-smokers and smokers share a similar miRNA profile. Since hierarchical clustering typically distinguishes patterns of expression and not necessarily isolated significant changes in a large pool of data, we sought to measure changes at individual miRNAs level by plotting mean fold change between groups with statistical significance. Interestingly, only one miRNA, miR-124, was found to be statistically different between smokers versus non-smokers ($p = 0.000074$) and ex-smokers versus smokers ($p = 0.00000154$) with a fold change >2 (Figure 3A and B). Our results indicate that the level of circulating miR-124 increases following exposure to cigarette smoke. The absence

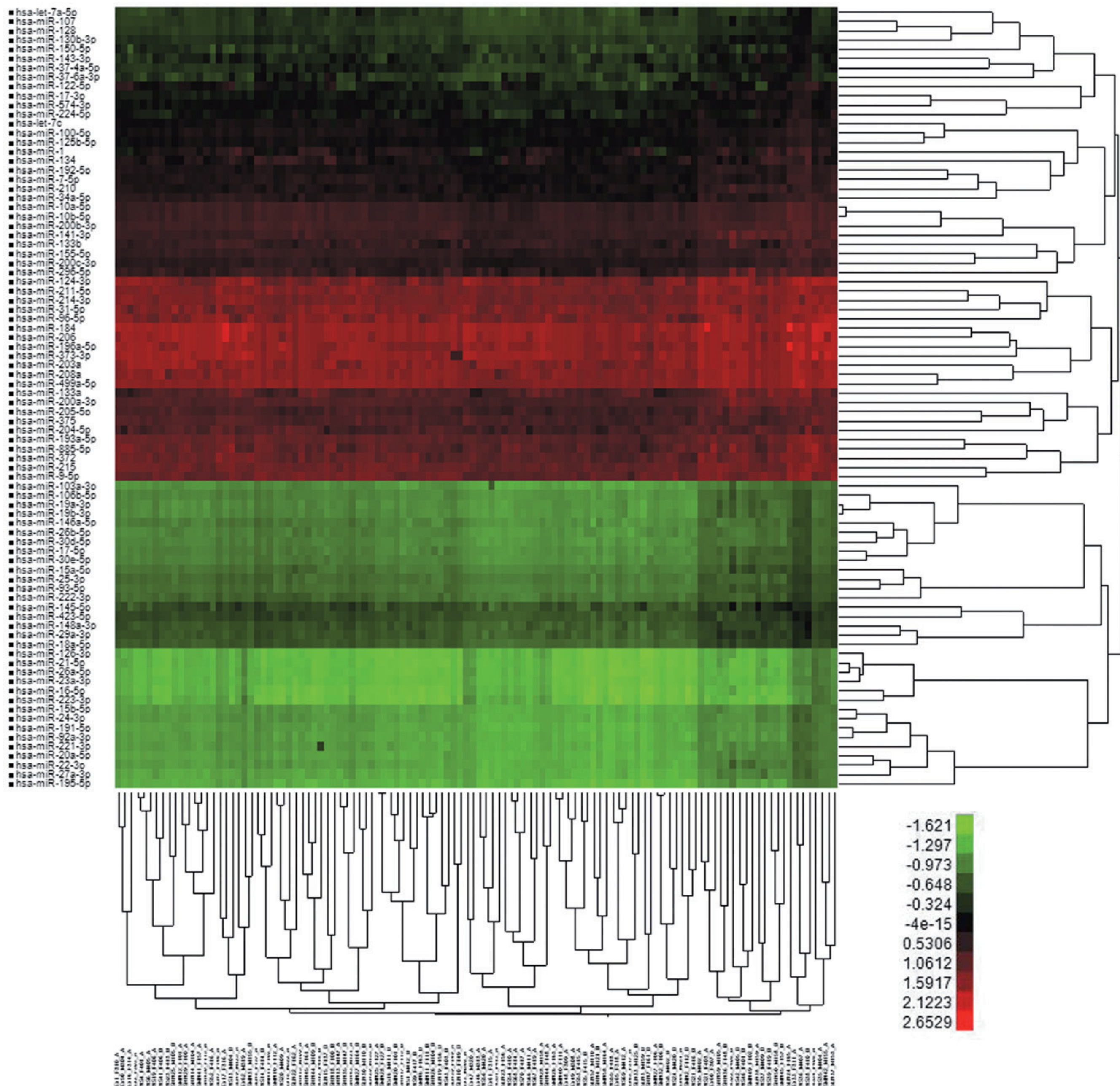

Figure 2. Hierarchical cluster representing the gene expression profiles of 84 miRNAs commonly found in plasma of 120 subjects including smokers ($n = 30$ at days 182 and 183), non-smokers ($n = 20$ at days 163 and 164) and ex-smokers ($n = 20$ at day 163). Columns represent individual subjects and rows represent miRNA. Green black and red indicate high signal intensity, moderate signal intensity and low to no signal intensity in normalized gene expression data (ΔCT).

of any differential expression of miR-124 between non-smokers and ex-smokers suggests the reversion of this miRNA to its baseline level of expression on smoking cessation (Supplementary Figure 2). Stringent statistical thresholds were established by adjusting p values for FDR (False discovery rate) and FWER (Familywise error rate) to promote confidence in the validity of the detected differences in miR-124 expression with respect to the smoker group. The Storey method revealed that miRNA, let-7a as another potential miRNA, was associated with smoking.

The limited number of differentially expressed miRNA candidates in our study was in contrast with a prior Japanese study where a significant difference in plasma miRNA profile was noticed between smokers and non-smokers (Takahashi et al., 2013). This group used a set of two microarrays covering a total of 664 miRNAs and detected 44 differentially expressed miRNA which were driving the clustering by smoking status. Seventeen of these miRNAs were included in our panel but did not show differential expression based on smoking status. The analytical methods, statistical analyses and group sizes could explain at least in part those differences. For instance, the Takahasi paper did not report adjusted p values and only recruited a small number of subjects (smokers $n = 11$; non-smokers $n = 7$; ex-smokers $n = 4$) which could lead to an increased number of false positive miRNA candidates. The microarray approach used by Takahashi and colleagues is suitable for high-throughput screening, but it is often recommended to verify

Biomarkers, 2015; 20(2): 123–131

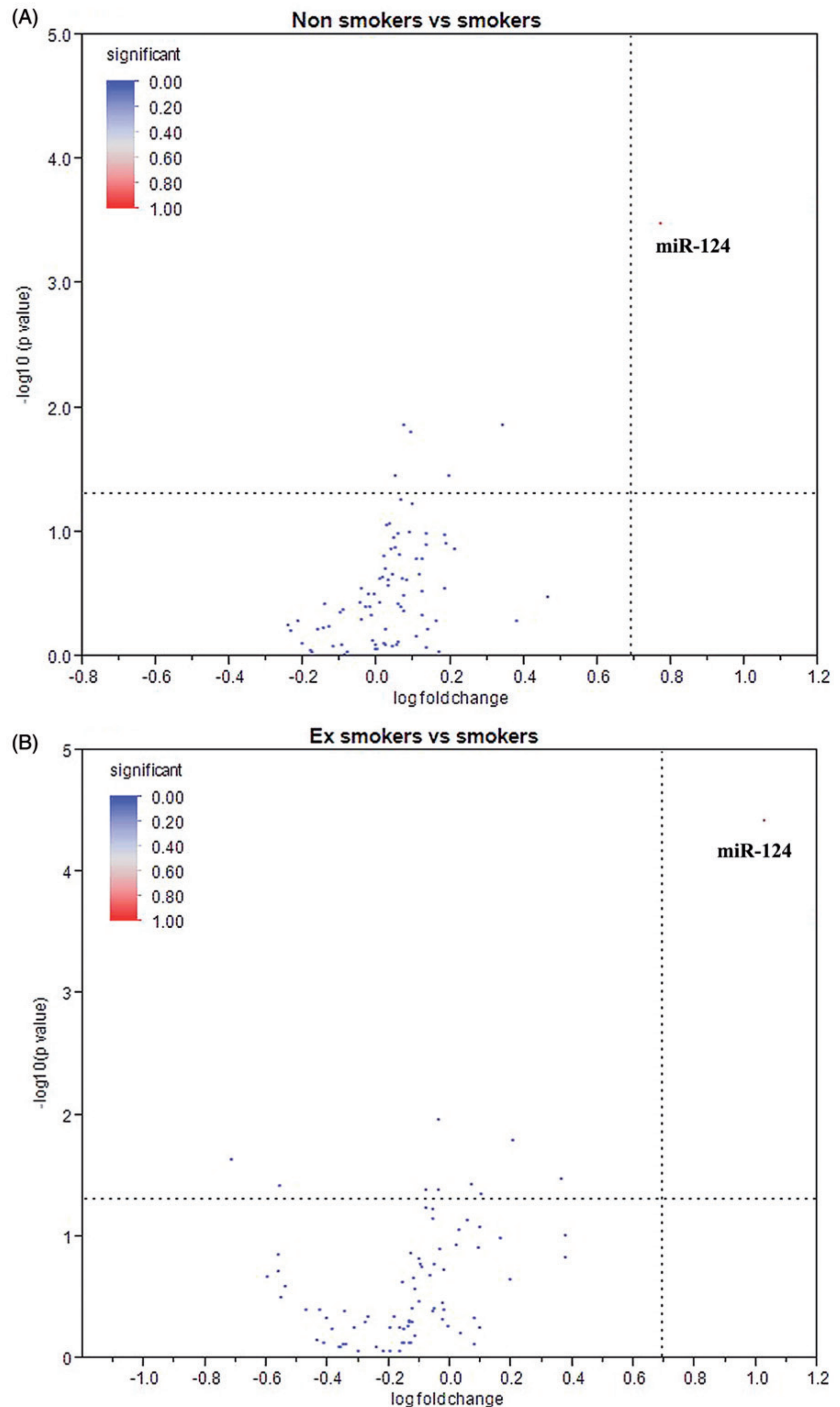

Figure 3. (A) Volcano plot showing the difference in miRNA expression between smokers and non-smokers is plotted on the *x*-axis ($\log_2$ scale) (dotted vertical line marks the 2-fold increase threshold), and false discovery rate (FDR)-adjusted significance is plotted on the *y*-axis ($-\log_{10}$ scale) (dotted horizontal line marks the *p*-value = 0.05 threshold). Up-regulated miR-124 is indicated in red. (B) Volcano plot showing the difference in miRNA expression between smokers and ex-smokers is plotted on the *x*-axis ($\log_2$ scale) (dotted vertical line marks the 2-fold increase threshold), and false discovery rate (FDR)-adjusted significance is plotted on the *y*-axis ($-\log_{10}$ scale) (dotted horizontal line marks the *p* value = 0.05 threshold). Up-regulated miR-124 is indicated in red.

the differentially expressed candidates by qRT-PCR. miRNA qRT-PCR has been proven to perform better for sensitivity and specificity when compared with microarrays (Chen et al., 2009). In our study, we opted to screen a smaller panel of miRNA candidates by qRT-PCR and including a larger number of subjects with replicate samples collected in a two day period and processed independently in order to identify the most robust miRNA candidate(s).

Our candidate miRNA biomarker miR-124 was first reported to be highly expressed in neuronal cells (Makeyev et al., 2007). Emerging evidence indicates that miR-124 is significantly down-regulated in several types of human cancer (Ando et al., 2009; Furuta et al., 2010; Lujambio et al., 2007), which can affect tumor initiation and maintenance. In the context of smoking, miR-124 was reported to be down-regulated in lung cancer tissues (Yanaihara et al., 2006). Recent studies revealed the tumor-suppressive function of miR-124 in nasopharyngeal carcinoma (Peng et al., 2014) and its prognostic value in patients with non-small cell lung cancer (Berghmans et al., 2013). miR-124 was also one of the most remarkably down-regulated miRNAs in the lungs of rats exposed to environmental cigarette smoke

Table 1. Correlation between miRNAs and biomarkers of cigarette smoke exposure.

miRNA	CEVal (*p*-values)	TNEQ (*p*-values)
(A) Not adjusted for any variables.		
miR-124	**0.0197***	0.2124
let-7a	0.5632	0.0692

Variables	let-7a (*p*-values)	miR-124 (*p*-values)	Variables	let-7a (*p*-values)	miR-124 (*p*-values)
(B) Adjusted for age and gender.					
CEVal	0.2708	0.0797	TNEQ	**0.0199***	0.2103
Age	0.1341	0.6079	Age	0.0802	0.3441
Gender	0.7346	0.4345	Gender	0.3145	0.2237
(C) Adjusted for age and gender and pack years.					
CEVal	0.5796	0.2217	TNEQ	**0.0413***	0.5481
Age	0.1020	0.8705	Age	0.3134	0.7499
Gender	0.7699	0.4633	Gender	0.2789	0.3321
Pack years	0.4547	0.6334	Pack years	0.6352	0.5677

Bold values with * sign indicate a significant dose response relationship between miRNA and biomarkers of cigarette smoke exposure.

(Izzotti et al., 2009). Those results suggest that miR-124 is associated with lung diseases and tobacco exposure; however, those reports indicated a down-regulation of miR-124 rather than an increase as observed in our study. This difference can be explained by a variety of factors. First, the cited work used biopsies of diseased or healthy lung rodent tissues for miRNA screening while we used plasma from healthy smokers. It is plausible that an increase in plasma miRNA reflects systemic stress of the respiratory tract leading to miRNA leakage in the blood stream and hence an increase of miR-124 in plasma. Second, a decrease in miR-124 could be a consequence of tumor cells escaping regulation from a variety of tumor suppressors, which might not be observed in healthy tissues or tissues at a pre-disease stage. It might be of interest to correlate lung and plasma samples from the same subjects in future. The second miRNA candidate that was identified in our screen, let-7a, is also a tumor suppressor (Long et al., 2009), which plays an important role in lung cancer development (Osada & Takahashi, 2011).

In order to establish whether there was an association between miR-124, let-7a and smoking, we looked at the correlation between our miRNA candidates and two biomarkers of cigarette smoke exposure that can be used as surrogates of tobacco consumption (Table 1). The selected biomarkers of tobacco smoke exposure TNEQ in urine and CEVal (haemoglobin adduct of acrylonitrile) have very different half-life/blood residency time. While TNEQ has a short half-life of approximately 14–16 h (Hukkanen et al., 2005), CEVal has a blood residency time of 4 months (Scherer et al., 2014). Therefore, urinary TNEQ represents the smoking behaviour over a couple of days, while CEVal is representative of the smoking behaviour over multiple months and is, therefore, not subject to occasional changes in smoking habits. Furthermore, miRNAs have a plasma residency time that varies between hours and weeks based on their sequence (Ruegger & Grosshans, 2012); therefore, it is of interest to perform correlations with biomarkers of smoke exposure with different body residency times. We used a linear mixed model analysis to estimate confounder-adjusted associations of the differentially expressed miRNAs (miR-124 and let-7a with the biomarker of exposure – haemoglobin adducts of acrylonitrile (CEVal) and urinary total nicotine equivalents (TNEQ)). Regression analysis indicated a significant dose–response relationship between miR-124 expression and CEVal with no adjustment for any confounding variables (Table 1A). However, when adjusted for factors such as age, gender and pack-years, this relationship became non-significant (Table 1B and C). Regardless of any adjustment, no significant relationship was observed between miR-124 and TNEQ. let-7a, in contrast to miR-124, did not show any dose–response relationship with either of the two biomarkers when not adjusted for any variables (Table 1A). However, a significant association was seen with TNEQ when adjusted for age, gender and pack-years (Table 1B and C). These results indicate that the miRNA versus biomarker relationship is dependent on other confounding factors, those could be related to polymorphisms, diet, exercise and other lifestyle factors. Interestingly, while let-7a was found to have a correlation with TNEQ in urine, miR-124 showed a more consistent trend with CEVal indicating that there is a potential relationship with both long- and short-term smoking behaviour. Therefore, it would be of interest to investigate whether a larger cohort could reduce the influence of the confounding factors and add clarity to these conclusions. It is also worth noting that differential miRNA expression could be driven by specific smoke toxicants. For instance, differences in plasma miRNA expression have been associated with PAH exposure (Deng et al., 2014). Thus, better correlations might be obtained if a toxicant driving the miRNA response can be identified. Nevertheless, this is the first time to our knowledge that the correlation between miRNAs differentially expressed in smokers and biomarkers of exposure with very different half-life has been investigated. Based on our assessment, we would recommend using a variety of biomarkers of exposure with different body residency times to establish correlation with plasma miRNA.

Although our results appear promising and a robust methodology and statistical approach was used, there are also some limitations: (i) the panel of biomarkers were selected from a limited number of miRNA candidates, from

which other important miRNAs involved in smoking-related diseases might be excluded. Therefore, it might be worthwhile in the future to use a more comprehensive panel of candidate miRNAs. (ii) The effect of haemolysis on the quantification of miRNAs in plasma has not been investigated in great detail. The plasma samples were checked for haemolysis just by visual assessment. Based on the number of papers which have indicated that haemolysis of plasma and serum samples can be a major cause of variation in altering miRNA levels (Blondal et al., 2013; Kirschner et al., 2011), it might be desirable to assess haemolysis by other analytical techniques in future studies.

Conclusion

We have successfully identified differential serum levels of miR-124 and let-7a between the smoking groups and non-smokers. These miRNAs could be promising as potential biomarkers of biological effect after cigarette smoke exposure. It would be interesting to explore a larger sample number at different time points and also use a lung model as a tissue source to examine modified expression of miR-124 and let-7a. Although miR-124 and let-7a show a correlation with haemoglobin adduct biomarkers of tobacco exposure, we found that the relationship is dependent on other confounding factors. In future, it might be worth investigating the correlation in a larger group of subjects with a larger biomarker of exposure panel representing different chemical families.

Acknowledgements

We thank Andrew Baxter for his technical assistance in data checking and QC.

Declaration of interest

The authors report that they have no conflicts of interest. A. Banerjee, E. Minet, O. M. Camacho, and D. Waters were employees of British American Tobacco during the conduct of this study.

References

Akbas F, Coskunpinar E, Aynaci E, et al. (2012). Analysis of serum micro-RNAs as potential biomarker in chronic obstructive pulmonary disease. Exp Lung Res 38:286–94.

Ando T, Yoshida T, Enomoto S, et al. (2009). DNA methylation of microRNA genes in gastric mucosae of gastric cancer patients: its possible involvement in the formation of epigenetic field defect. Int J Cancer 124:2367–74.

Banerjee A, Luettich K. (2012). MicroRNAs as potential biomarkers of smoking-related diseases. Biomark Med 6:671–84.

Berghmans T, Ameye L, Willems L, et al. (2013). Identification of microRNA-based signatures for response and survival for non-small cell lung cancer treated with cisplatin–vinorelbine A ELCWP prospective study. Lung Cancer 82:340–5.

Blondal T, Jensby NS, Baker A, et al. (2013). Assessing sample and miRNA profile quality in serum and plasma or other biofluids. Methods 59:S1–6.

Chen Y, Gelfond JA, McManus LM, Shireman PK. (2009). Reproducibility of quantitative RT-PCR array in miRNA expression profiling and comparison with microarray analysis. BMC Genomics 10:407.

Creemers EE, Tijsen AJ, Pinto YM. (2012). Circulating microRNAs: novel biomarkers and extracellular communicators in cardiovascular disease? Circ Res 110:483–95

Deng Q, Huang S, Zhang X, et al. (2014). Plasma microRNA expression and micronuclei frequency in workers exposed to polycyclic aromatic hydrocarbons. Environ Health Perspect 122:719–25.

Dickinson BA, Semus HM, Montgomery RL, et al. (2013). Plasma microRNAs serve as biomarkers of therapeutic efficacy and disease progression in hypertension-induced heart failure. Eur J Heart Fail 15: 650–9.

Duroux-Richard I, Pers YM, Fabre S, et al. (2014). Circulating miRNA-125b is a potential biomarker predicting response to rituximab in rheumatoid arthritis. Mediat Inflamm 2014:342524.

Duttagupta R, Jiang R, Gollub J, et al. (2011). Impact of cellular miRNAs on circulating miRNA biomarker signatures. PLoS One 6: e20769.

Fenoglio C, Ridolfi E, Cantoni C, et al. (2013). Decreased circulating miRNA levels in patients with primary progressive multiple sclerosis. Mult Scler 19:1938–42.

Fichtlscherer S, De RS, Fox H, et al. (2010). Circulating microRNAs in patients with coronary artery disease. Circ Res 107:677–84.

Foss KM, Sima C, Ugolini D, et al. (2011). miR-1254 and miR-574-5p: serum-based microRNA biomarkers for early-stage non-small cell lung cancer. J Thorac Oncol 6:482–8.

Furuta M, Kozaki KI, Tanaka S, et al. (2010). miR-124 and miR-203 are epigenetically silenced tumor-suppressive microRNAs in hepatocellular carcinoma. Carcinogenesis 31:766–76.

He L, Hannon GJ. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5:522–31.

Huang Z, Huang D, Ni S, et al. (2010). Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. Int J Cancer 127:118–26.

Hukkanen J, Jacob III P, Benowitz NL. (2005). Metabolism and disposition kinetics of nicotine. Pharmacol Rev 57:79–115.

Iorio MV, Croce CM. (2009). MicroRNAs in cancer: small molecules with a huge impact. J Clin Oncol 27:5848–56.

Izzotti A, Calin GA, Arrigo P, et al. (2009). Downregulation of microRNA expression in the lungs of rats exposed to cigarette smoke. FASEB J 23:806–12.

Kirschner MB, Kao SC, Edelman JJ, et al. (2011). Haemolysis during sample preparation alters microRNA content of plasma. PLoS One 6: e24145.

Krell J, Frampton AE, Jacob J, et al. (2012). miRNAs in breast cancer: ready for real time? Pharmacogenomics 13:709–19

Kroh EM, Parkin RK, Mitchell PS, Tewari M. (2010). Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods 50: 298–301.

Leidinger P, Keller A, Borries A, et al. (2011). Specific peripheral miRNA profiles for distinguishing lung cancer from COPD. Lung Cancer 74:41–7.

Long XB, Sun GB, Hu S, et al. (2009). Let-7a microRNA functions as a potential tumor suppressor in human laryngeal cancer. Oncol Rep 22: 1189–95.

Lujambio A, Ropero S, Ballestar E, et al. (2007). Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res 67:1424–9.

Makeyev EV, Zhang J, Carrasco MA, Maniatis T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell 27:435–48.

Mi S, Zhang J, Zhang W, Huang RS. (2013). Circulating microRNAs as biomarkers for inflammatory diseases. Microrna 2:63–71.

Mitchell PS, Parkin RK, Kroh EM, et al. (2008). Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA 105:10513–18.

Mo MH, Chen L, Fu Y, et al. (2012). Cell-free circulating miRNA biomarkers in cancer. J Cancer 3:432–48.

Osada H, Takahashi T. (2011). let-7 and miR-17-92: small-sized major players in lung cancer development. Cancer Sci 102:9–17.

Peng XH, Huang HR, Lu J, et al. (2014). MiR-124 suppresses tumor growth and metastasis by targeting Foxq1 in nasopharyngeal carcinoma. Mol Cancer 13:186.

Ruegger S, Grosshans H. (2012). MicroRNA turnover: when, how, and why? Trends Biochem Sci 37:436–46

Scherer G, Newland K, Papadopoulou E, Minet E. (2014). A correlation study applied to biomarkers of internal and effective dose for acrylonitrile and 4-aminobiphenyl in smokers. Biomarkers 19: 291–301.

Sharma S, Kelly TK, Jones PA. (2010). Epigenetics in cancer. Carcinogenesis 31:27–36.

Shen J, Stass SA, Jiang F. (2013). MicroRNAs as potential biomarkers in human solid tumors. Cancer Lett 329:125–36.

Shepperd CJ, Newland N, Eldridge A, et al. (2013). A single-blinded, single-centre, controlled study in healthy adult smokers to identify the effects of a reduced toxicant prototype cigarette on biomarkers of exposure and of biological effect versus commercial cigarettes. BMC Public Health 13:690.

Shields PG. (1999). Molecular epidemiology of lung cancer. Ann Oncol 10:S7–11.

Takahashi K, Yokota S, Tatsumi N, et al. (2013). Cigarette smoking substantially alters plasma microRNA profiles in healthy subjects. Toxicol Appl Pharmacol 272:154–60.

Valadi H, Ekstrom K, Bossios A, et al. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol 9:654–9.

Wang K, Zhang S, Marzolf B, et al. (2009). Circulating microRNAs, potential biomarkers for drug-induced liver injury. Proc Natl Acad Sci USA 106:4402–7.

Wei J, Gao W, Zhu CJ, et al. (2011). Identification of plasma microRNA-21 as a biomarker for early detection and chemosensitivity of non-small cell lung cancer. Chin J Cancer 30:407–14.

Yanaihara N, Caplen N, Bowman E, et al. (2006). Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 9:189–98.

Zernecke A, Bidzhekov K, Noels H, et al. (2009). Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. Sci Signal 2:ra81.

Zheng D, Haddadin S, Wang Y, et al. (2011). Plasma microRNAs as novel biomarkers for early detection of lung cancer. Int J Clin Exp Pathol 4:575–86.

Supplementary material available online

Supplementary Table 1 and Figures 1 and 2.