Application of Systems Pharmacology to Identify Exposure Response Markers in Peripheral Blood after Switching to a Candidate Modified-risk Tobacco Product: The Tobacco Heating System 2.2 (THS 2.2) – REXC 04 and REXA 07 Japan Clinical Studies

TABLE OF CONTENTS

1 ABSTRACT 3
2 ABBREVIATIONS 4
3 INTRODUCTION 4
4 MATERIALS AND METHODS 6
4.1 PMI reduced exposure (REX) clinical studies conducted in Japan 6
4.2 REXC-04 study 6
4.3 REXA-07 study 7
4.4 RNA isolation 7
4.5 RNA preparation and Affymetrix hybridization 7
4.6 Affymetrix data processing 8
5 RESULTS 8
5.1 Performance of the signature under cessation and THS 2.2/THS 2.2M switching scenarios 11
6 DISCUSSION 15
7 CONCLUSION 15
8 REFERENCES 16

TABLE OF FIGURES

Figure 1. EXC-04 study design. An open-label, 3-arm parallel group with discharge at Day 6 for each arm (CC, SA, and THS 2.2) 8
Figure 2. REXA-07 study design. A randomized, controlled, open-label, 3-arm parallel group with 5-day confinement followed by 85 days in an ambulatory setting 9
Figure 3. Volcano plots for the differentially expressed genes in the REXC-04 study 10
Figure 4. Volcano plots for the differentially expressed genes in the REXA-07 study 11
Figure 5. Schematic overview of the approach 12
Figure 6. LDA scores for the training set (BLD-SMK-01, CS and NS), test samples (BLD-SMK-01 FS, QASMC, and the REX studies) (Martin et al, 2015) 13
Figure 7. Predicted relative decrease in the REXC-04 REXA-07 studies 14

1 ABSTRACT

Gene expression signatures help to provide a global and robust view of biological processes in normal and pathological situations. Additionally, such signatures can be used in toxicology to assess both the level and impact of toxicant exposure, which is in line with the 21st century vision of toxicology. In this study, we present a novel 11-gene signature derived from whole blood that can distinguish current smokers from either non-smokers or former smokers. We have developed this signature to assist us in the evaluation of the reduction in exposure response in subjects enrolled in PMI clinical trials for THS 2.2 and THS 2.2 menthol assessment. The signature was applied to gene expression profiles from blood after 5-day cessation confinement studies, which showed a consistent decrease of the signature scores. Similarly, upon switching to THS2.2 for 5 days, a significant decrease in the signature score was consistently observed. Finally, in a study where a 5-day confinement period was followed by an 85-day ambulatory period, a similar decrease was observed for both cessation and switching to mentholated THS2.2.

Keywords: gene expression, signature, blood, surrogate, smoking

2 ABBREVIATIONS

CC	Conventional Cigarettes
CS	Current Smoker
DEG	Differentially Expressed Gene
FS	Former Smoker
LDA	Linear Discriminant Analysis
mCC	mentholated Conventional Cigarettes
NCS	Non-current Smoker
NS	Never Smoker
QASMC	Queen Ann Street Medical Center
REX	Reduced Exposure
SA	Smoking abstinence
Se	Sensitivity
Sp	Specificity
THS2.2	Tobacco Heating System 2.2
THS2.2M	Tobacco Heating System 2.2 Menthol

3 INTRODUCTION

Cigarette smoke exposure has a negative impact on human health and is linked to the development of several fatal diseases (US Department of Health, 2004). The response to cigarette smoke exposure has been monitored by the use of biomarkers, such as the levels of "nicotine equivalents" (Hecht et al, 2010) in urine or the metabolites of the tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (Carmella et al, 1995). However, the use of these biomarkers is associated with several drawbacks, including their short half-life and inter-individual differences in metabolism (Acosta et al, 2004; Carmella et al, 1995; Hecht et al, 2010; Jacobson & Ferguson, 2014). Hair nicotine and cotinine have been used to monitor smoke exposure among infants and adults, and have been found to be a more precise measurement of exposure than urine cotinine levels (Al-Delaimy, 2002; Al-Delaimy et al, 2002; Tzatzarakis et al, 2012). Salivary cotinine concentrations have also been proposed as a non-invasive biomarker for environmental tobacco exposure in children (Phillips et al, 1999). While nicotine and cotinine measurements from hair and saliva may provide accurate verification of nonsmoking status or secondhand smoke exposure, they are restricted to a single constituent present in cigarette smoke. Moreover, such biomarkers do not offer insights into the biological mechanisms that are affected in response to cigarette smoke exposure, a feature advocated by 21st century toxicity testing principles (Krewski et al, 2010).

Consequently, new technologies, such as whole-genome microarrays, have been incorporated into toxicity testing to increase efficiency and to provide a more data-driven approach to exposure response assessment (Thomas et al, 2013). For example, several

studies have described chemical-specific gene expression profiles associated with the adverse effects of active substances in various tissues (Hamadeh et al, 2002; Morgan, 2002; Waring et al, 2001). The large airway transcriptome of smokers has also been characterized, and as stated in the field of injury theory (Dakubo et al, 2007), molecular changes in response to smoke exposure can be detected even when no histological abnormalities are visible (Brody & Steiling, 2011; Gower et al, 2011; Gustafson et al, 2010; Spira et al, 2007; Steiling et al, 2008). Moreover, sample acquisition from the primary site of exposure, i.e., the airways, is usually invasive and is therefore not convenient for exposure assessment and monitoring. As a minimally invasive alternative, peripheral blood sampling can be employed in the general population to establish systemic biomarkers. Several blood-based biomarkers of potential harm have been proposed, including those related to inflammation, oxidative stress, and platelet activation (Liu et al, 2011). A more global picture of the affected biology can be achieved through gene expression studies, and various exposures have been shown to alter gene expression profiles in blood (Bushel et al, 2007b; Chauhan et al, 2014a; Forrest et al, 2005). Indeed, several studies have shown that gene expression in blood can distinguish among subject populations, such as those with early stage non-small cell lung cancer from non-malignant lung disease (Rotunno et al, 2011b; Showe et al, 2009), subjects with chronic obstructive pulmonary disease (COPD) from healthy smokers (Poliska et al, 2011), and even smokers with no detectable disease from non-smokers (Büttner et al, 2007; Charlesworth et al, 2010; Dumeaux et al, 2010; Lampe et al, 2004; Lodovici et al, 2007; Van Leeuwen et al, 2007; Votavova et al, 2011). Furthermore, blood-based gene signatures have been used to identify diseases (Poliska et al, 2011; Rotunno et al, 2011a; Zander et al, 2011; Zeller & Blankenberg, 2013) or subjects exposed to harmful environmental toxicants (Bushel et al, 2007a; Chauhan et al, 2014b; Dressman et al, 2007; Forrest et al, 2005).

A transcriptome-based exposure response signature could be as simple as the presence or absence of the expression of a single gene. More commonly, it could be characterized by the expression levels of a collection of genes, each contributing to a specific diagnosis. Therefore, such a signature, which may be based on differentially expressed genes (DEGs) between case and control populations, would be distinct and more robust than a molecular biomarker. Furthermore, because of the large inter-individual variations that can be expected in human populations, signatures should be rigorous and well designed, maintaining high specificity (Sp) and sensitivity (Se) across independent subject cohorts, laboratories, and nucleic acid extraction methods. However, many published signatures lack proper validation, making them overly optimistic (Barbash & Soreq, 2013; Zeller & Blankenberg, 2013).

Previously, we have identified a whole-blood-based gene signature with the potential to distinguish subjects who smoked (current smokers, CS) from those who stopped smoking (former smokers, FS), or never smoked (non-smokers, NS). The signature consists of 11 genes, *LRRN3*, *SASH1*, *PALLD*, *RGL1*, *TNFRSF17*, *CDKN1C*, *IGJ*, *RRM2*, *ID3*, *SERPING1* and *FUCA1*, some of which have been previously described in the context of smoking. Validated using independent datasets, the signature of smoking exposure response was able to distinguish CS from non-current smokers (NCS) with remarkable accuracy, and the signature translated well across species (Martin et al, 2015).

The goal of this study was to apply the previously developed signature to determine whether the impact of switching to a heat-non-burn tobacco product could be readily detected in the whole-blood transcriptome after 5 days of switching, thus providing a sensitive and non-invasive tool for assessing exposure response in clinical trials. Rather than characterizing the gene expression profile of a smoker who has switched to THS 2.2 and generating a signature that is specific only to acute changes (i.e., by extracting the signature from the 5-day confinement study data), we set out to find a transcriptomics-based exposure response signature that could also serve as an indicator of a longer-term switching pattern. This was achieved by establishing the signatures that distinguished CS samples from NCS samples (Martin et al, 2015).

In this study, we have applied the signature on whole-blood gene expression from smokers who switched to THS 2.2 in a randomized controlled study that was conducted in confinement (ZRHR-REXC-04-JP), as well as another study in which the confinement was followed by an 85-day ambulatory period (ZRHM-REXA-07-JP). Herein, we present the results from blood transcriptomics analysis of subject samples at baseline and after 5 days, as well as at 90 days for the ambulatory study of different study arms, including smoking abstinence (SA), continuous smoking, and switching to THS2.2. The prediction scores derived from the signature Linear Discriminant Analysis (LDA) model were further compared to a similar confinement study previously reported (REXC03EU_1762_SysTox_Pub). Finally, we show that the exposure response signature indicates that the subjects who switch to THS 2.2 are more similar in their exposure response to the abstinence groups than the continuous smoking groups, in both cohorts.

4 MATERIALS AND METHODS

4.1 PMI reduced exposure (REX) clinical studies conducted in Japan

In addition to the Informed Consent Form (ICF) for participation in the study, subjects were provided with information and asked for their consent to collect blood samples for bio-banking for transcriptomics profiling. The blood sampling for transcriptomics and the data related to these samples were anonymized. Anonymized data and samples were initially single- or double-coded where the link between the subjects' identifiers and the unique code(s) was subsequently deleted.

4.2 REXC-04 study

The ZRHR-REXC-04-JP study was a randomized, controlled, open-label, 3-arm parallel group, single-center study that monitored exposure to selected smoke constituents in conventional cigarettes (CC). The study included healthy smokers switching to the THS 2.2 or SA, compared with continuing the use of CC, for 5 days in confinement. The CC and SA arms involved 40 participants each, whereas the THS2.2 arm involved 80 (Figure 1). This study was registered at ClinicalTrials.gov, Bethesda (MD): National Library of Medicine (US). 2013-2014 [cited 2015 Jun 16]. Available from: http://clinicaltrials.gov/show/NCT01970982 NLM Identifier: NCT01970982. This study follows an equivalent design to REXC03EU_1762_SysTox_Pub, while involving a separate study population.

4.3 REXA-07 study

The ZRHM-REXA-07-JP study was a randomized, controlled, open-label, 3-arm parallel group, multi-center study that monitored exposure to selected smoke constituents in smokers switching to THS2.2 Menthol (THS 2.2M) or SA, compared with smoking menthol conventional cigarettes (mCC), for 90 days. Specifically, the study included healthy smokers switching to the THS2.2M or those observing SA, compared with continuing to use mCC, for 5 days in confinement and then for 85 days in an ambulatory setting (Figure 2). This was an ad libitum smoking study. During the confinement period, compliance to product/regimen allocation (exclusive use of THS2.2M and mCC in THS2.2M and mCC arms, respectively, and full abstinence from smoking in the SA arm) was ensured by strict distribution of each mCC when requested by the subject. During the ambulatory period, the subjects randomized to the THS2.2M arm were instructed to exclusively use THS2.2M, and subjects randomized to the SA arm were instructed to abstain from smoking. This study was registered in: ClinicalTrials.gov, Bethesda (MD): National Library of Medicine (US). 2013-2014 [cited 2015 Jun 16]. Available from: http://clinicaltrials.gov/show/NCT01970995 NLM Identifier: NCT01970995.

4.4 RNA isolation

Total RNAs (including microRNAs) were isolated using the PAXgene Blood miRNA Kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. The concentration and purity of the RNA samples were determined using an ultraviolet (UV) spectrophotometer (NanoDrop ND1000; Thermo Fisher Scientific, Waltham, MA, USA) by measuring the absorbance at 230, 260, and 280 nm. The RNA integrity was further checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Only RNAs with an RNA integrity number >6 were processed for further analysis.

4.5 RNA preparation and Affymetrix hybridization

Affymetrix probe sets targeting the 3′ ends of transcripts were prepared from 50 ng of RNA using the Ovation® Whole Blood Reagent and Ovation® RNA Amplification System V2 (NuGEN, San Carlos, CA). The quantity of cDNA was measured with a NanoDrop® 1000 or 8000 spectrophotometer (Thermo Fisher Scientific) or a SpectraMax® 384Plus microplate reader (Molecular Devices, Sunnyvale, CA, USA). The cDNA quality was determined by assessing the size of un-fragmented cDNA using an Agilent 2100 Bioanalyzer. The size distribution of the final fragmented and biotinylated product was also monitored using electropherograms. After labeling the cDNA, the fragments were hybridized to the GeneChip® Human Genome U133 Plus 2.0 Array (Affymetrix) according to the manufacturer's guidelines. Samples for target preparation were fully randomized for the Affymetrix gene expression microarray. After the investigation of artifacts on the chip scan, the data were processed through a standard quality control pipeline. Briefly, raw data files were read using the ReadAffy function of the *affy* package (Gautier et al, 2004) from the Bioconductor suite of microarray analysis tools (Gentleman et al, 2004) available for the R statistical environment (R Development Core Team, 2007).

4.6 Affymetrix data processing

The Affymetrix data were subsequently normalized using frozen-Robust Microarray Analysis (fRMA). Background correction and quantile normalization were used to generate microarray expression values (Irizarry et al, 2003) from all arrays passing quality control checks using the custom Chip Description File (CDF) environment HGU133Plus2_Hs_ENTREZG v16.0 (Dai et al, 2005). For the population level analysis (i.e., the study of the gene average fold-changes), an overall linear model, accounting for the longitudinal study design, was fitted for each comparison of interest to generate raw P-values for each probe set on the expression array based on moderated t-statistics (Smyth, 2004). The Benjamini–Hochberg False Discovery Rate method was used to correct for multiple testing effects.

5 RESULTS

The REXC-04 (Figure 1) and REXA-07 (Figure 2) studies have been described in detail previously (ZRHR-REXC-04-JP and ZRHM-REXA-07-JP).

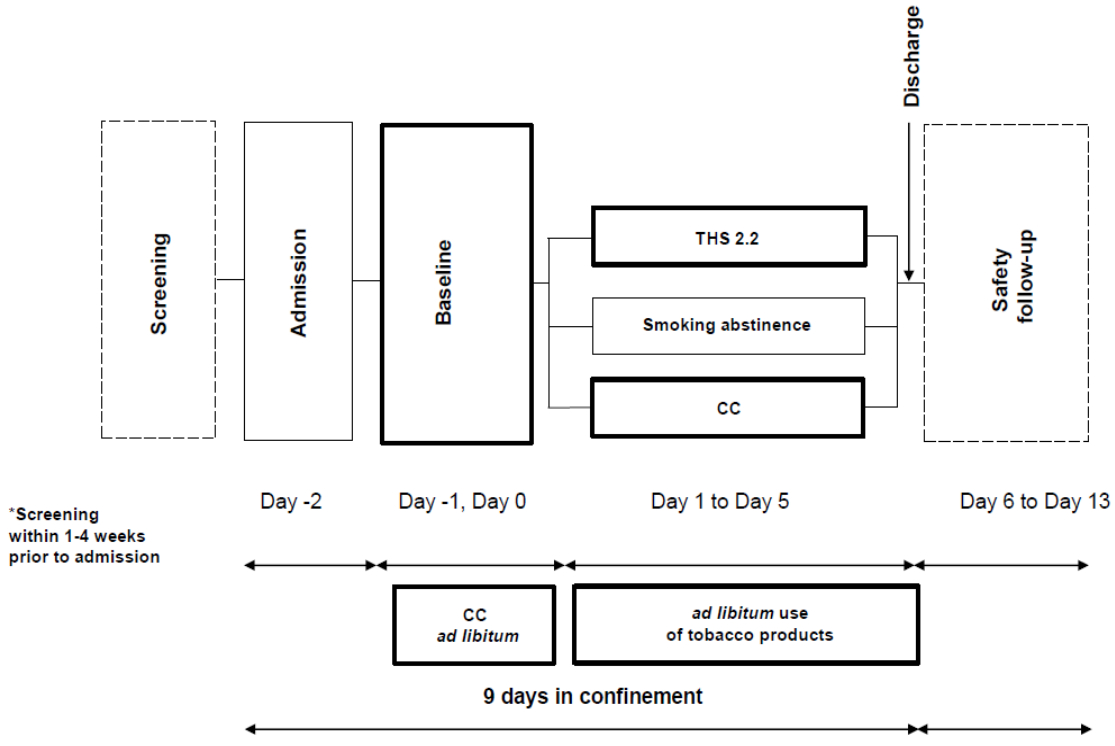


Figure 1. EXC-04 study design. An open-label, 3-arm parallel group with discharge at Day 6 for each arm (CC, SA, and THS 2.2)

Abbreviations: CC, conventional cigarette; SA, smoking abstinence; THS 2.2, Tobacco Heating System 2.2; REX, reduced exposure.

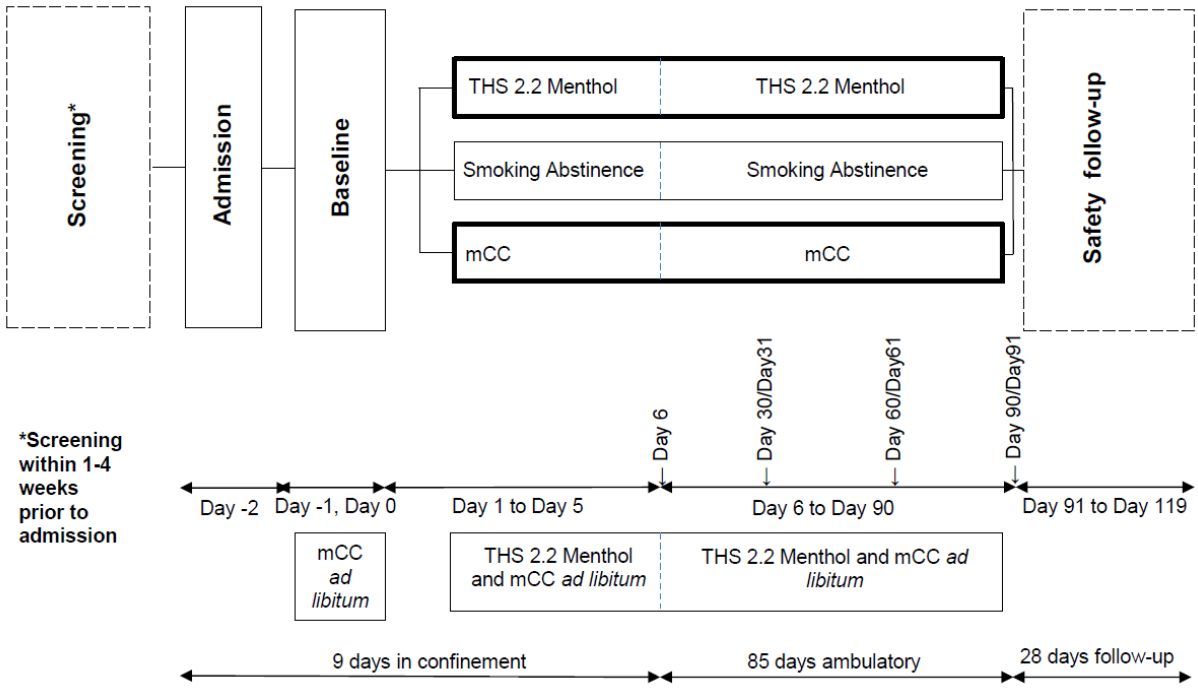


Figure 2. REXA-07 study design. A randomized, controlled, open-label, 3-arm parallel group with 5-day confinement followed by 85 days in an ambulatory setting

Abbreviations: mCC, mentholated conventional cigarette; SA, smoking abstinence; THS 2.2, Tobacco Heating System 2.2; REX, reduced exposure.

Gene expression changes in the blood of REXC-04-JP and REXA-07-JP study subjects in the different study arms

After checking the quality of the CEL files for the REXC-04 study, 33, 25, and 56 profiles remained in the CC, SA, and THS2.2 arms, respectively. Based on the volcano plots in Figure 3, the impact of smoking abstinence or switching to THS2.2 was not apparent, as only a few genes had fold-changes with an FDR<0.05 after 5 days in confinement.

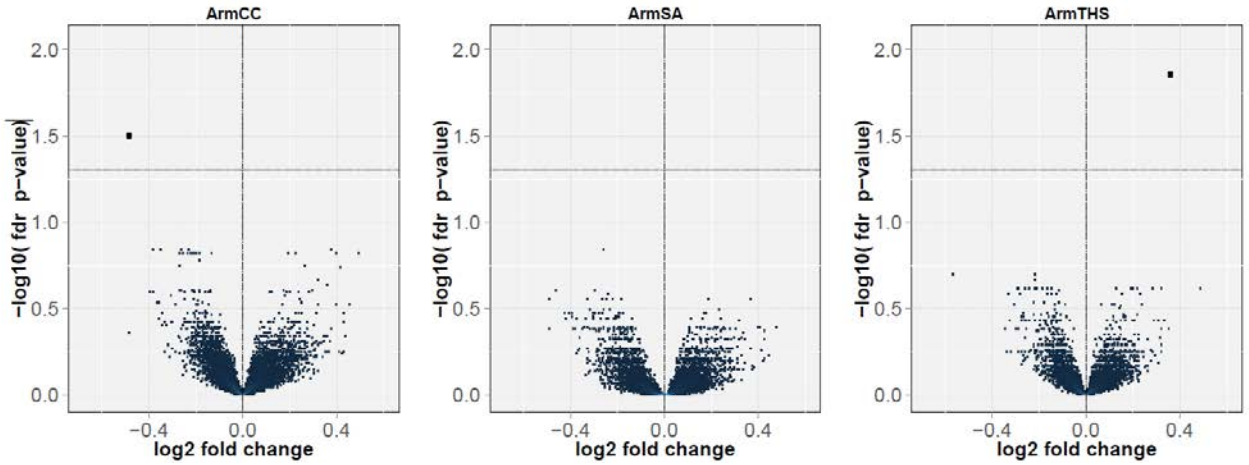


Figure 3. Volcano plots for the differentially expressed genes in the REXC-04 study

Volcano plots showing the estimated log2 (fold-change) against −log10 (adjusted p-value) for the longitudinal effect. p-values were computed based on moderated t-statistics and were adjusted by the Benjamini–Hochberg method. The line indicates the −log10 (adjusted P-value), which corresponds to the false discovery rate, FDR 0.05. Left panel: Longitudinal effect (Day 5 − Day 0) for the CC arm. Middle panel: Longitudinal effect (Day 5 − Day 0) for the SA arm. Right panel: Longitudinal effect (Day 5 − Day 0) for the THS2.2 arm. Abbreviations: CC, conventional cigarette; SA, smoking abstinence; THS2.2, Tobacco Heating System 2.2; REX, reduced exposure.

For the REXA-07 study, the mCC arm consisted of 33 profiles for mCC, 30 for SA, and 63 for THS2.2M. The volcano plots in Figure 4 show the systems response profile in the whole blood from subjects in the different study arms. The THS2.2M switching arm was the only group of subjects that showed a response in their blood gene expression in the first 5 days, which is likely owing to the different group sample size of the THS2.2M arm. Stronger changes in gene expression were observed between the end of the confinement (Day 6) and the end of the study (Day 91), with the most differentially expressed genes in the THS2.2M switching arm. This was also likely owing to the unbalanced sizes between the study arms. In the mCC arm, the differential gene expression was likely the result of release from the confinement, whereas the changes in the SA and THS2.2M arms could be due to both release from confinement and the impact of abstinence/switching.

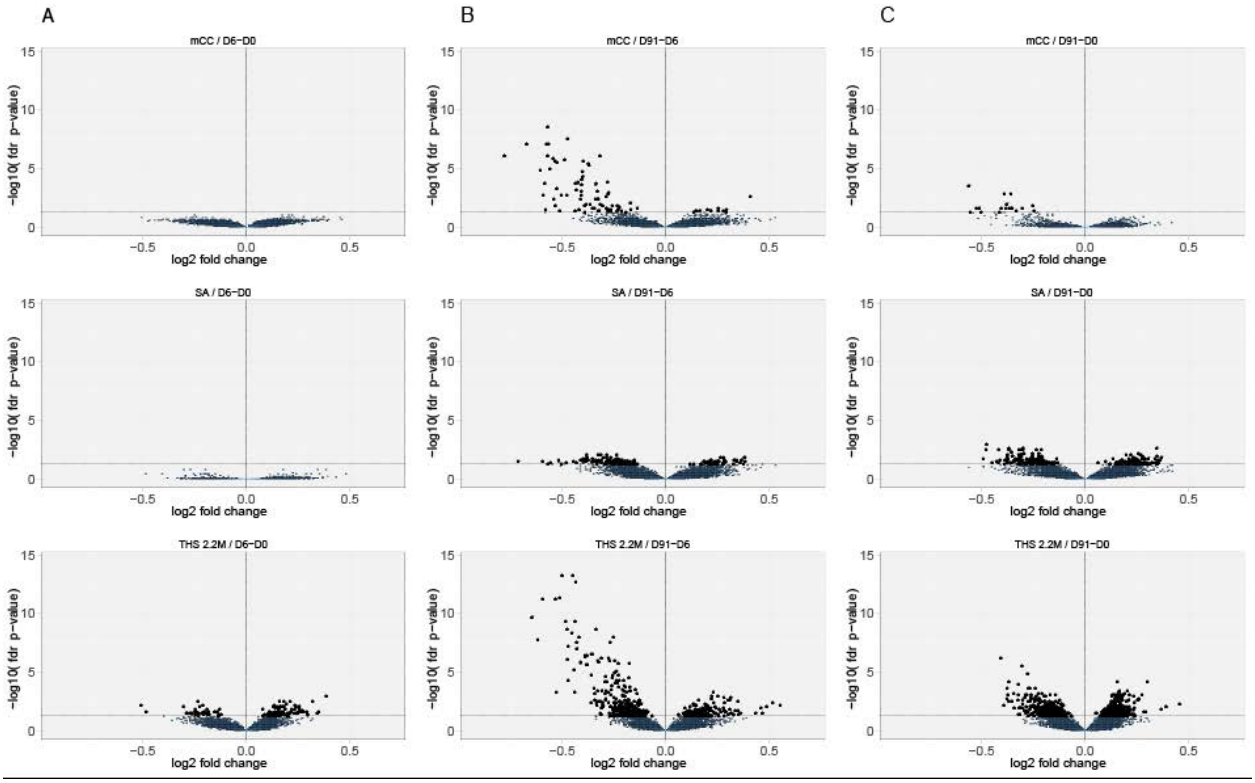


Figure 4. Volcano plots for the differentially expressed genes in the REXA-07 study

Volcano plots showing the estimated log2 (fold-change) against −log10 (adjusted P-value) for the longitudinal effect. P-values were computed based on moderated t-statistics and were adjusted by the Benjamini–Hochberg method. The vertical line indicates the −log10 (adjusted P-value), which corresponds to the false discovery rate, FDR 0.05. A) Longitudinal effect (Day 5 − Day 0) for the different study arms. B) Longitudinal effect (Day 91 − Day 5) for the different study arms. C) Longitudinal effect (Day 91 − Day 0) for the different study arms. Abbreviations: mCC, mentholated conventional cigarette; SA, smoking abstinence; THS2.2, Tobacco Heating System 2.2; REX, reduced exposure.

5.1 Performance of the signature under cessation and THS 2.2/THS 2.2M switching scenarios

The goal was to apply the previously established signature to determine whether the impact of switching to THS 2.2 could be readily detected in the whole-blood transcriptome after 5 days of switching, thus providing a sensitive and non-invasive tool for assessing the exposure response in clinical trials. Figure 5 lays out the overall approach leveraging many blood gene expression datasets to establish and validate the exposure signature. The approach is described in detail in a previous study (Martin et al, 2015).

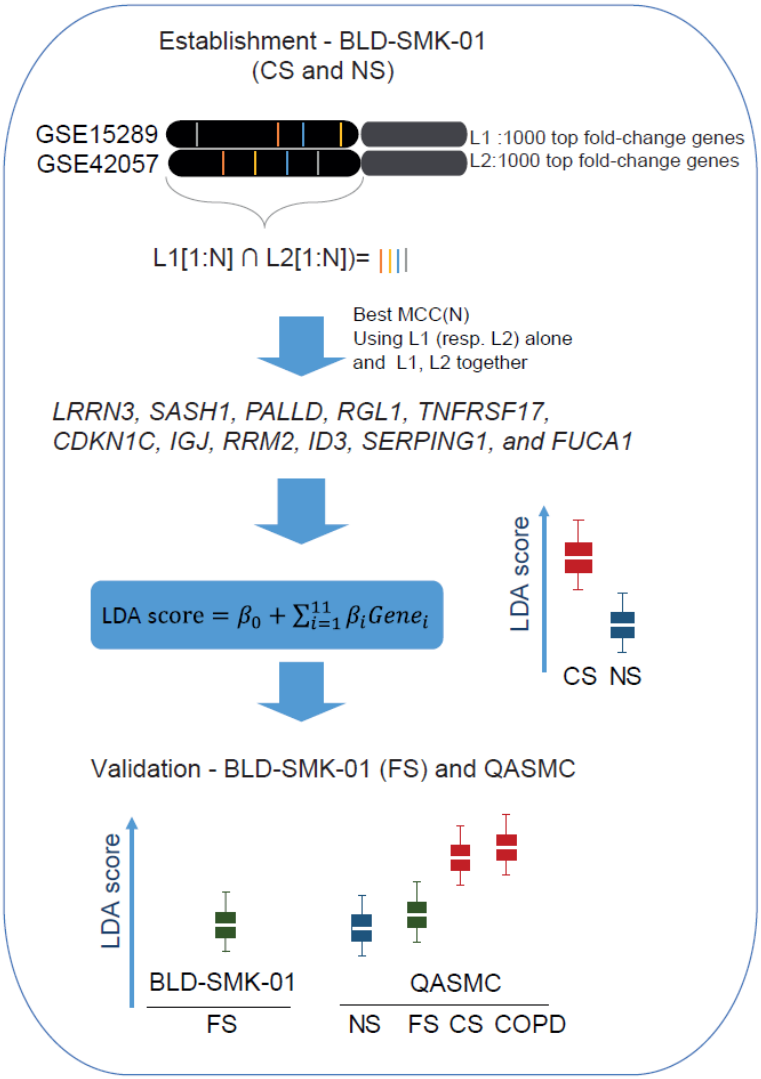


Figure 5. Schematic overview of the approach

A) The set of genes and the model were defined using two public datasets and the training set BLD-SMK-01 (CS and NS arms) (Martin et al, 2015). B) The obtained signature was verified in the FS arm of BLD-SMK-01 and in the independent study QASMC (Martin et al, 2015). Abbreviations: CS, current smoker; FS, former smoker; NS, never smoker; QASMC, Queen Ann Street Medical Center.

The underlying premise was that the blood transcriptome of smokers who switch to THS 2.2 start to resemble that of an FS. Therefore, instead of characterizing the gene expression profile of a smoker switching to THS 2.2 that is specific the acute effects of switching (i.e., by extracting the signature from the REXC study data), we set out to find a transcriptomics-based exposure response signature that could also serve as an indicator of a longer-term switching pattern. This was achieved by establishing the signatures that distinguished CS samples from NCS samples, as discussed in Martin et al. (Martin et al, 2015).

The signature scores depict a significant decrease in the score between Day 0 and Day 6 in the SA arm and in the THS arm (Figure 6). As expected, no difference was observed between Day 0 and Day 6 in the CC arm. The results between the three studies were very comparable.

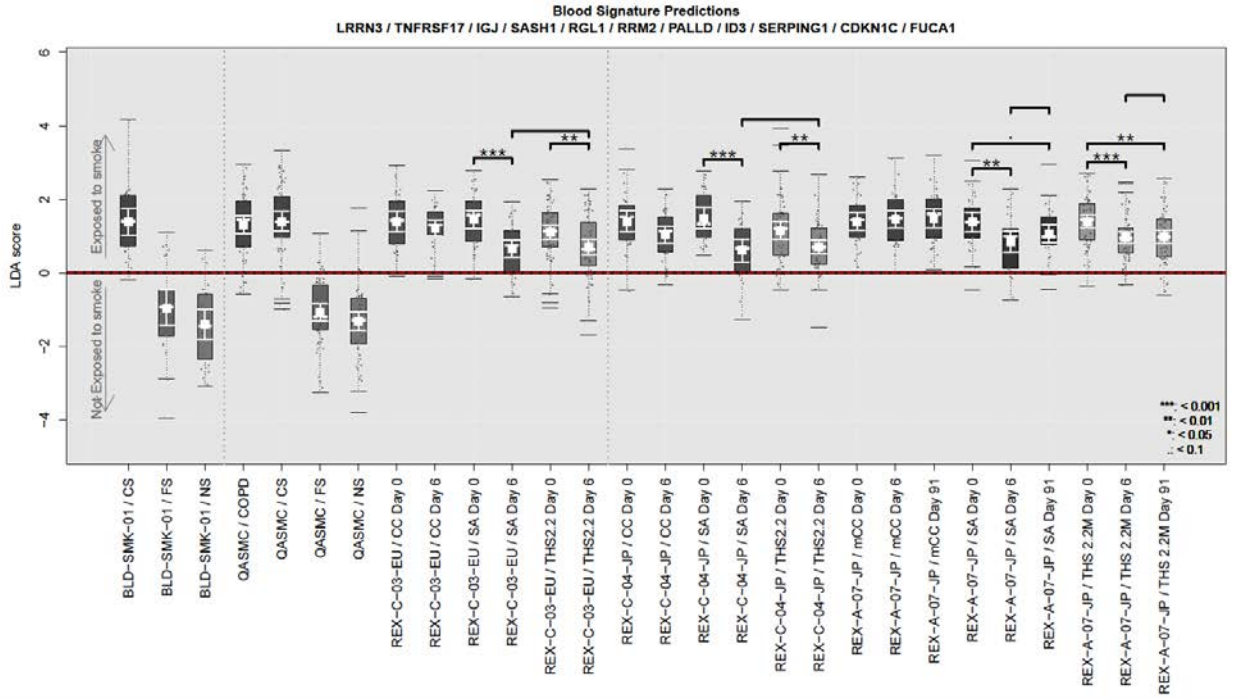


Figure 6. LDA scores for the training set (BLD-SMK-01, CS and NS), test samples (BLD-SMK-01 FS, QASMC, and the REX studies) (Martin et al, 2015)

For comparison, the REXC-03 study is also included (REXC03EU_1762_SysTox_Pub). A positive score is predictive of CS status, while a negative score indicates non-CS status. The mean prediction scores of each group are marked by a white square, together with its 95% confidence interval. The median prediction scores of each group are indicated by horizontal segments, and box notches are approximated median 95% confidence intervals. The key comparison t-test statistics are displayed on top of the boxplots. Abbreviations: CS, current smoker; FS, former smoker; LDA, linear discriminant analysis; NS, never smoker; QASMC, Queen Ann Street Medical Center; REX, reduced exposure.

The relative decrease for such a longitudinal study was computed for each subject based on the signature scores. The observed decrease in score between Day 0 and Day 6 was normalized to the observed 10th to 90th quantile range of the signature score in BLD-SMK-01 (Martin et al, 2015). Therefore, a relative decrease of 100% was achieved when a smoker was at the 90th quantile of the BLD-SMK-01 smoker group at Day 0 and ends at the 10% quantile of the BLD-SMK-01 former smoker group at Day 6. Within this conservative range, the observed SA decreases were 20% for REXC-04 samples and 12% for the REXA-07 samples. Those results were comparable with the findings described in (REXC03EU_1762_SysTox_Pub), where an 18% relative decrease was reported. For the THS arms, the observed decreases were on average between 9% and 10% for the two

studies, which were comparable with the decreases observed previously (REXC03EU_1762_SysTox_Pub). While observing a higher variability in the decrease for THS arms, all the described changes were significant at a 5% Type I error level. From the REXA-07-JP study after 90 days, the SA decrease slightly increased to 7% while the decrease from the THS2.2M arm stayed stable at 9% (Figure 7).

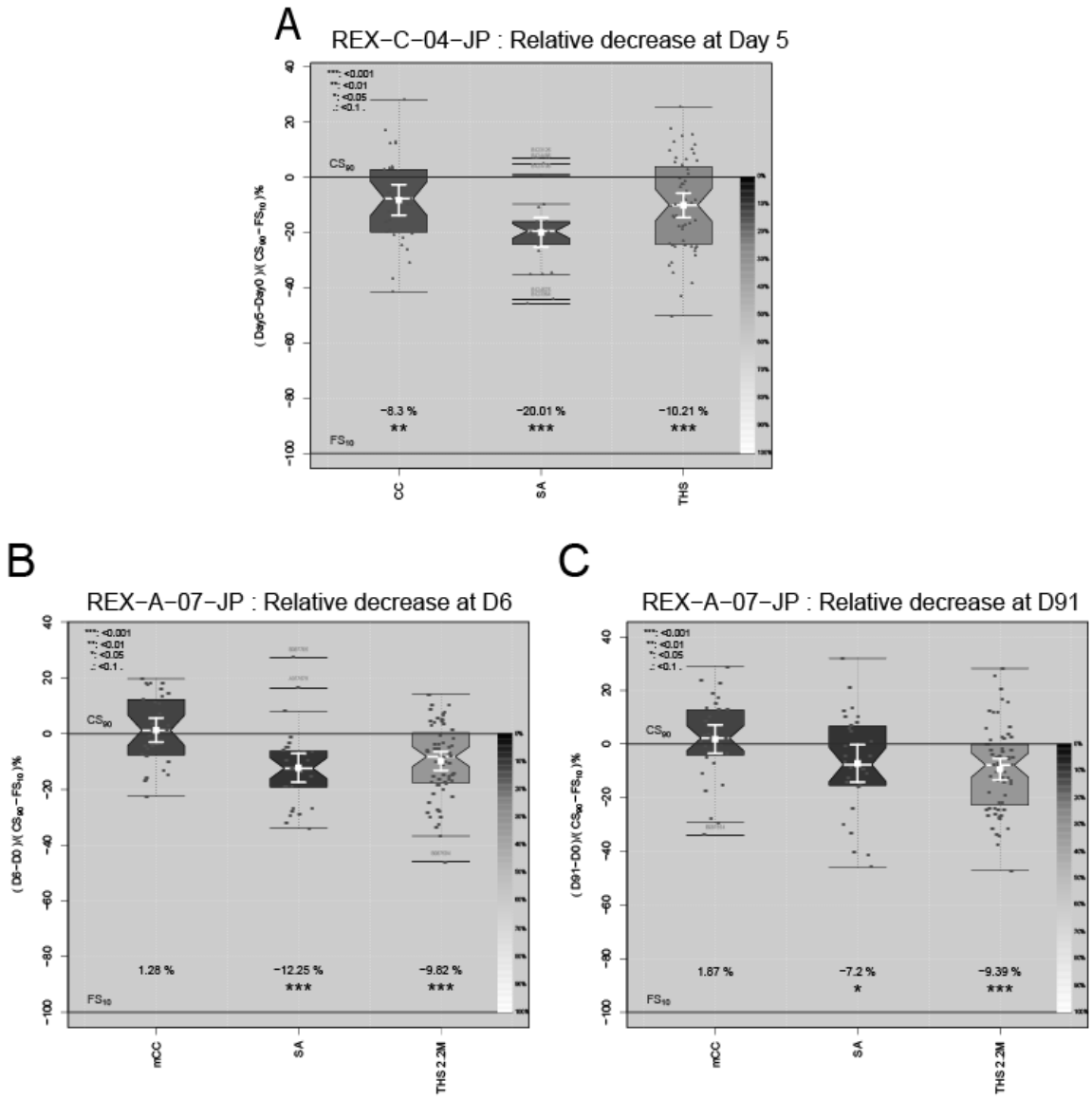


Figure 7. Predicted relative decrease in the REXC-04 REXA-07 studies

Relative change for each subject in the study based on the signature scores. The observed change is normalized to the observed 10th to 90th quantile range of the signature score in BLD-SMK-01 (Martin et al, 2015). A relative decrease of 100% is achieved when a smoker is at the 90% quantile of the BLD-SMK-01 smoker group at Day 0 and ends at the 10% quantile of the BLD-SMK-01 FS group at Day 5. The significance test assesses the departure from 0%. A) Predicted relative change in the REXC-04 at Day 5. B) Predicted relative change in the REXA-07 study at Day 6. C) Predicted relative change in the REXA-07 at Day 91. Abbreviations: FS, former smoker; REX, reduced exposure.

6 DISCUSSION

Compared with single molecule measurements, gene expression profiling provides a global and more complete view of the biological processes in normal and pathological situations. When the expression trends of multiple genes are taken together, it is possible to derive a signature, or a classifier for a given physiological state, from an exposure response to a disease state. While the tissue that is primarily affected offers a sample that more accurately represents the normal, exposed, or pathological state, it is often not realistic to classify subjects using tissue biopsies. Considering the ease of blood sampling using minimally invasive techniques, blood-based signatures hold great promise for biomarker discovery (Kittleson et al, 2004; Zeller & Blankenberg, 2013).

Prevailing rodent studies have both strengths and limitations, but the more translatable the signatures, the better they serve predictive toxicology and disease research. To comply with these demands, we showed that the whole-blood-based signature can discriminate smoke-exposed mice from sham-exposed, as well as mice that were withdrawn from smoke exposure (Martin et al, 2015).

Remarkably, in the REX studies (current and (REXC03EU_1762_SysTox_Pub)), the changes in signature scores were evident only after 5 days of abstinence and THS switching, thereby providing a very sensitive readout for the switching effects in a small number of subjects. In addition, the blood from the subjects who remained on the conventional product during the confinement showed few statistically significant changes in gene expression, confirming that the observed effects were indeed due to SA and THS switching. Upon quantifying the relative decrease of the scores, we consistently observed ~12%–20% and ~10% decreases in the SA and THS arms, respectively, across the studies after 5 days. When substantiating the study with an 85-day ambulatory period (REXA-07), the relative decrease for THS2.2 remained stable, while a slight increase for the average score for the SA arm was observed.

7 CONCLUSION

In conclusion, our systems toxicology approach enabled the construction of a robust whole-blood-based 11-gene signature of smoking exposure response that could distinguish CS from NCS with remarkable accuracy. Moreover, the signature scores for blood samples of subjects switching to THS under reduced exposure studies REXC-04 and REXA-07 predicted a significant reduction in exposure response within just 5 days, which was sustained over 90 days after switching to THS.

8 REFERENCES

Acosta MC, Buchhalter AR, Breland AB, Hamilton DC, Eissenberg T (2004) Urine cotinine as an index of smoking status in smokers during 96-hr abstinence: comparison between gas chromatography/mass spectrometry and immunoassay test strips. *Nicotine & tobacco research* **6:** 615-620

Al-Delaimy W (2002) Hair as a biomarker for exposure to tobacco smoke. *Tobacco Control* **11:** 176-182

Al-Delaimy W, Crane J, Woodward A (2002) Is the hair nicotine level a more accurate biomarker of environmental tobacco smoke exposure than urine cotinine? *Journal of Epidemiology and Community Health* **56:** 66-71

Barbash S, Soreq H (2013) Statistically invalid classification of high throughput gene expression data. *Scientific reports* **3**

Brody JS, Steiling K (2011) Interaction of cigarette exposure and airway epithelial cell gene expression. *Annu Rev Physiol* **73:** 437-456

Bushel P, Heinloth A, Li J, Huang L, Chou J, Boorman G, Malarkey D, Houle C, Ward S, Wilson R (2007a) Blood gene expression signatures predict exposure levels. *Proceedings of the National Academy of Sciences* **104:** 18211-18216

Bushel PR, Heinloth AN, Li J, Huang L, Chou JW, Boorman GA, Malarkey DE, Houle CD, Ward SM, Wilson RE, Fannin RD, Russo MW, Watkins PB, Tennant RW, Paules RS (2007b) Blood gene expression signatures predict exposure levels. *Proc Natl Acad Sci U S A* **104:** 18211-18216

Büttner P, Mosig S, Funke H (2007) Gene expression profiles of T lymphocytes are sensitive to the influence of heavy smoking: a pilot study. *Immunogenetics* **59:** 37-43

Carmella SG, Akerkar SA, Richie J, Hecht SS (1995) Intraindividual and interindividual differences in metabolites of the tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in smokers' urine. *Cancer Epidemiology Biomarkers & Prevention* **4:** 635-642

Charlesworth JC, Curran JE, Johnson MP, Goring HH, Dyer TD, Diego VP, Kent JW, Jr., Mahaney MC, Almasy L, MacCluer JW, Moses EK, Blangero J (2010) Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics* **3:** 29

Chauhan V, Howland M, Wilkins R (2014a) Identification of gene-based responses in human blood cells exposed to alpha particle radiation. *BMC medical genomics* **7:** 43

Chauhan V, Howland M, Wilkins R (2014b) Identification of gene-based responses in human blood cells exposed to alpha particle radiation. *BMC medical genomics* **7:** 1

Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic acids research* **33:** e175-e175

Dakubo GD, Jakupciak JP, Birch-Machin MA, Parr RL (2007) Clinical implications and utility of field cancerization. *Cancer cell international* **7:** 1

Dressman HK, Muramoto GG, Chao NJ, Meadows S, Marshall D, Ginsburg GS, Nevins JR, Chute JP (2007) Gene expression signatures that predict radiation exposure in mice and humans. *PLoS Med* **4:** e106

Dumeaux V, Olsen KS, Nuel G, Paulssen RH, Borresen-Dale AL, Lund E (2010) Deciphering normal blood gene expression variation--The NOWAC postgenome study. *PLoS Genet* **6:** e1000873

Forrest MS, Lan Q, Hubbard AE, Zhang L, Vermeulen R, Zhao X, Li G, Wu YY, Shen M, Yin S, Chanock SJ, Rothman N, Smith MT (2005) Discovery of novel biomarkers by microarray analysis of peripheral blood mononuclear cell gene expression in benzene-exposed workers. *Environ Health Perspect* **113:** 801-807

Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy---analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20:** 307-315

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5:** R80

Gower AC, Steiling K, Brothers JF, 2nd, Lenburg ME, Spira A (2011) Transcriptomic studies of the airway field of injury associated with smoking-related lung disease. *Proc Am Thorac Soc* **8:** 173-179

Gustafson AM, Soldi R, Anderlind C, Scholand MB, Qian J, Zhang X, Cooper K, Walker D, McWilliams A, Liu G, Szabo E, Brody J, Massion PP, Lenburg ME, Lam S, Bild AH, Spira A (2010) Airway PI3K pathway activation is an early and reversible event in lung cancer development. *Sci Transl Med* **2:** 26ra25

Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Bennett L, Li L, Tennant R, Stoll R, Barrett JC, Paules RS, Blanchard K, Afshari CA (2002) Prediction of compound signature using high density gene expression profiling. *Toxicol Sci* **67:** 232-240

Hecht SS, Yuan JM, Hatsukami D (2010) Applying tobacco carcinogen and toxicant biomarkers in product regulation and cancer prevention. *Chemical research in toxicology* **23:** 1001-1008

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4:** 249-264

Jacobson GA, Ferguson SG (2014) Relationship between cotinine and trans-3'-hydroxycotinine glucuronidation and the nicotine metabolite ratio in Caucasian smokers. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals***:** 1-5

Kittleson MM, Shui QY, Irizarry RA, Minhas KM, Edness G, Conte JV, Parmigiani G, Miller LW, Chen Y, Hall JL (2004) Identification of a gene expression profile that differentiates between ischemic and nonischemic cardiomyopathy. *Circulation* **110:** 3444-3451

Krewski D, Acosta Jr D, Andersen M, Anderson H, Bailar III JC, Boekelheide K, Brent R, Charnley G, Cheung VG, Green Jr S (2010) Toxicity testing in the 21st century: a vision and a strategy. *Journal of Toxicology and Environmental Health, Part B* **13:** 51-138

Lampe JW, Stepaniants SB, Mao M, Radich JP, Dai H, Linsley PS, Friend SH, Potter JD (2004) Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke. *Cancer Epidemiol Biomarkers Prev* **13:** 445-453

Liu J, Liang Q, Frost-Pineda K, Muhammad-Kah R, Rimmer L, Roethig H, Mendes P, Sarkar M (2011) Relationship between biomarkers of cigarette smoke exposure and biomarkers of inflammation, oxidative stress, and platelet activation in adult cigarette smokers. *Cancer Epidemiology Biomarkers & Prevention* **20:** 1760-1769

Lodovici M, Luceri C, De Filippo C, Romualdi C, Bambi F, Dolara P (2007) Smokers and passive smokers gene expression profiles: correlation with the DNA oxidation damage. *Free Radical Biology and Medicine* **43:** 415-422

Martin F, Talikka M, Hoeng J, Peitsch M (2015) Identification of gene expression signature for cigarette smoke exposure response—from man to mouse. *Human & experimental toxicology* **34:** 1200-1211

Morgan KT (2002) Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* **67:** 155-156

Phillips K, Bentley MC, Abrar M, Howard DA, Cook J (1999) Low level saliva cotinine determination and its application as a biomarker for environmental tobacco smoke exposure. *Human & experimental toxicology* **18:** 291-296

Poliska S, Csanky E, Szanto A, Szatmari I, Mesko B, Szeles L, Dezso B, Scholtz B, Podani J, Kilty I, Takacs L, Nagy L (2011) Chronic obstructive pulmonary disease-specific gene expression signatures of alveolar macrophages as well as peripheral blood monocytes overlap and correlate with lung function. *Respiration* **81:** 499-510

R Development Core Team. (2007) R: A Language and Environment for Statistical Computing.

Rotunno M, Hu N, Su H, Wang C, Goldstein AM, Bergen AW, Consonni D, Pesatori AC, Bertazzi PA, Wacholder S (2011a) A gene expression signature from peripheral whole blood for stage I lung adenocarcinoma. *Cancer prevention research* **4:** 1599-1608

Rotunno M, Hu N, Su H, Wang C, Goldstein AM, Bergen AW, Consonni D, Pesatori AC, Bertazzi PA, Wacholder S, Shih J, Caporaso NE, Taylor PR, Landi MT (2011b) A gene expression signature from peripheral whole blood for stage I lung adenocarcinoma. *Cancer Prev Res (Phila)* **4:** 1599-1608

Showe MK, Vachani A, Kossenkov AV, Yousef M, Nichols C, Nikonova EV, Chang C, Kucharczuk J, Tran B, Wakeam E, Yie TA, Speicher D, Rom WN, Albelda S, Showe LC (2009) Gene expression profiles in peripheral blood mononuclear cells can distinguish patients with non-small cell lung cancer from patients with nonmalignant lung disease. *Cancer Res* **69:** 9202-9210

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* **3:** Article3

Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P, Sridhar S, Beamis J, Lamb C, Anderson T, Gerry N, Keane J, Lenburg ME, Brody JS (2007) Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* **13:** 361-366

Steiling K, Ryan J, Brody JS, Spira A (2008) The field of tissue injury in the lung and airway. *Cancer Prev Res (Phila)* **1:** 396-403

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME (2013) Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. *toxicological sciences***:** kft178

Tzatzarakis M, Vardavas C, Terzi I, Kavalakis M, Kokkinakis M, Liesivuori J, Tsatsakis A (2012) Hair nicotine/cotinine concentrations as a method of monitoring exposure to tobacco smoke among infants and adults. *Human & experimental toxicology* **31:** 258-265

US Department of Health (2004) The health consequences of smoking: a report of the Surgeon General. *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health* **62**

Van Leeuwen DM, Van Agen E, Gottschalk RW, Vlietinck R, Gielen M, Van Herwijnen MH, Maas LM, Kleinjans JC, van Delft JH (2007) Cigarette smoke-induced differential gene expression in blood cells from monozygotic twin pairs. *Carcinogenesis* **28:** 691-697

Votavova H, Dostalova Merkerova M, Fejglova K, Vasikova A, Krejcik Z, Pastorkova A, Tabashidze N, Topinka J, Veleminsky M, Sram R (2011) Transcriptome alterations in maternal and fetal cells induced by tobacco smoke. *Placenta* **32:** 763-770

Waring JF, Jolly RA, Ciurlionis R, Lum PY, Praestgaard JT, Morfitt DC, Buratto B, Roberts C, Schadt E, Ulrich RG (2001) Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol Appl Pharmacol* **175:** 28-42

Zander T, Hofmann A, Staratschek-Jox A, Classen S, Debey-Pascher S, Maisel D, Ansén S, Hahn M, Beyer M, Thomas RK (2011) Blood-Based Gene Expression Signatures in Non–Small Cell Lung Cancer. *Clinical Cancer Research* **17:** 3360-3367

Zeller T, Blankenberg S (2013) Blood-based gene expression tests promises and limitations. *Circulation: Cardiovascular Genetics* **6:** 139-140