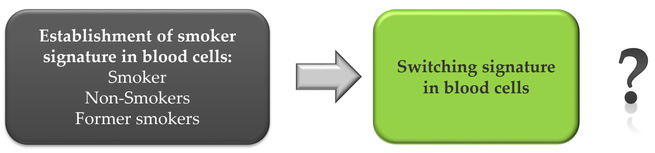
Application of Systems Pharmacology to Identify Exposure Response Markers in Peripheral Blood After Switching to a Candidate Modified Risk Tobacco Product: the Tobacco Heating System 2.1 (THS 2.1)

F. Martin, M. Talikka, C. Haziza, J. Ancerewicz, J. Hoeng, M.C. Peitsch
Biological Systems Research, PMI Research & Development, Neuchatel, Switzerland

Introduction and Objective

Establishing exposure response markers is necessary for the assessment of candidate modified risk tobacco products (MRTPs) against conventional cigarettes (CC). Biomarkers derived from the primary site, such as the airway, require invasive sampling, whereas blood offers a minimal-invasive alternative for the general population.

Various diseases and exposures, including cigarette smoke, have been shown to alter the molecular profile of the blood. To identify exposure response, we have conducted a whole genome Affymetrix microarray analyses from blood samples from current Smokers and never Smokers.



The aim of this study was to derive a blood-based current smoker (CS) gene signature with the potential for distinguishing subjects who smoked from those that have quit or have never smoked. Using human biobanking samples, we have developed and applied a 'blood signature' which can differentiate between smokers and non-smokers.

Taking advantage of the lessons from the IMPROVER Diagnostic Signature Challenge [1], we have developed a new methodology to derive a blood-based gene signature that is robust. The key feature of our methodology was to build a prediction model by jointly using high fold-change genes extracted from several publicly available gene expression datasets that profiled blood samples from CS and non-smokers (NS) or former smokers (FS). Pre-selecting genes based on high fold-change genes from various independent studies has the potential to enforce the robustness of the signature across studies.

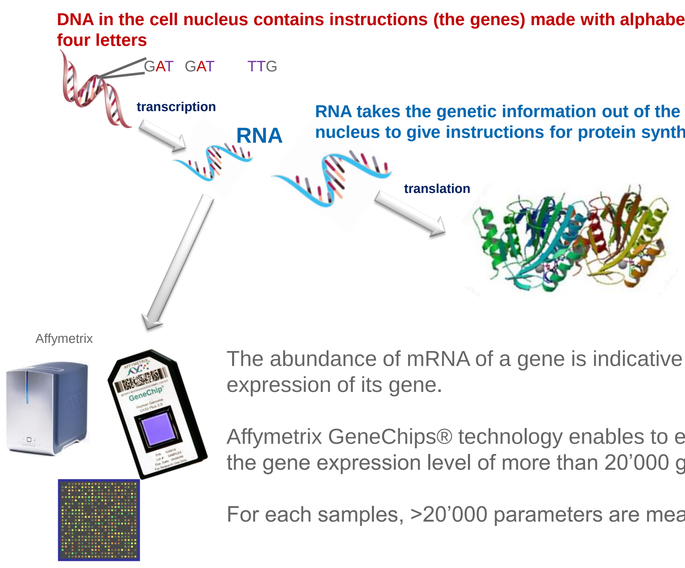
The validation was performed with an independent dataset derived from a clinical study initiated by Philip Morris International that aimed to discover novel biomarkers for COPD. Finally, from an additional clinical study sponsored by Philip Morris International, we evaluated the blood transcriptome of smokers who switched from CC to Tobacco Heating System (THS) 2.1 for 5 consecutive days and compared with those who continued to smoke CC.

Deriving a Predictive Smoke-Exposure Signature from Blood Transcriptomics

Blood samples for BLD-SMK-01 were obtained from a banked repository (BioServe Biotechnologies Ltd, Beltsville, MD 20705 USA). At the time of sampling, the subjects were between 23 and 65 year of age. Subjects with no disease history and anyone taking prescription medications were excluded. Current smokers had smoked at least 10 cigarettes per day for at least 3 years. Former smokers had ceased smoking at least 2 years prior to sampling and before quitting had smoked at least 10 cigarettes per day for at least 3 years. Current smokers and non-smokers were matched by age and gender. A total of:

1. 31 blood samples were obtained from **current smokers (CS),**
2. 30 from **never smokers (NS)**,
3. 30 from **former smokers (FS)**.

A Very Short Primer on Transcriptomics



The abundance of mRNA of a gene is indicative of the level of expression of its gene.

Affymetrix GeneChips® technology enables to estimate the gene expression level of more than 20'000 genes in a single run.

For each samples, >20'000 parameters are measured

PROBLEM STATEMENT

Based on the expression of >20'000 genes measured in the samples of the study, can we find a small set of genes and a rule combining their expression levels able to predict the smoking status of the patient (calles a "signature")?

Finding such signature was the focus of the first challenge of the SBV IMPROVER initiative [1]. Lessons from SBV IMPROVER Diagnostic Signature Challenge [2] indicated that internal cross-validation is sometimes miss-leading and that robustness across datasets (population, laboratory, sample analysis) is the key toward a useful signature.

IDEA Leverage public datasets to enforce robustness:

1) NOWAC study [3] (GSE15289, 211 Never Smokers, 74 Smokers samples from blood, post-menopausal women)
2) Bahr study [4] (GSE42057, PBMC's 42 control subjects and 94 subjects with varying severity of COPD had PBMC (! Gene expression profiles generated. All subjects were non-hispanic white, current or former smokers.

By analyzing the whole blood transcriptome of the study subjects and leveraging previously identified genes that exhibited a response to smoking in blood, we derived a classification tool using statistical learning methods.

We obtained a prediction model based on the following core genes: ***LRRN3***, ***SASH1***, ***PALLD***, ***RGL1***, ***TNFRSF17*** and ***CDKN1C***. The 5-fold cross-validation (100 times) of the model led to a Sensitivity (Se) of 0.91 and a Specificity (Sp) of 0.85 when classifying current smoker samples vs. never smoker samples.

When studying predictive models obtained by leveraging each list of high fold-change genes individually, ***IGJ***, ***RRM2***, ***ID3***, ***SERPING1*** and ***FUCA1*** were repeatedly identified as potential candidates in signatures having a high specificity and sensitivity. These five genes were also among the high fold-change genes in the blood transcriptomes of both the NOWAC (current smokers vs. never smokers) and Bahr et al. (current smokers vs. former smokers) studies and were used to extend the core gene signature to an extended signature. The cross-validation performance of the model based on the extended signature (***LRRN3***, ***SASH1***, ***PALLD***, ***RGL1***, ***TNFRSF17***, ***CDKN1C***, ***IGJ***, ***RRM2***, ***ID3***, ***SERPING1*** and ***FUCA1***) Se=0.88 and Sp=0.84 when classifying current smokers vs. never smokers (Figure 1).

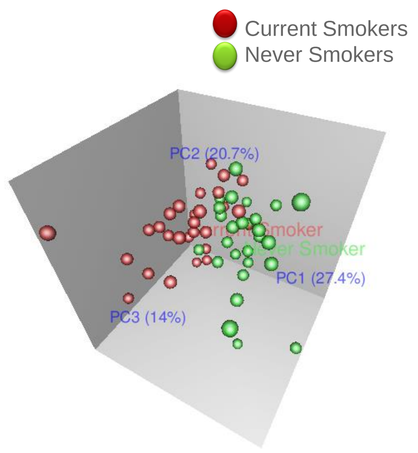


Figure 1: PCA on the extended signature showing the separation between CS and NS

Validation on Independent Clinical Samples

The Queen Ann Street Medical Center clinical study (referred to as the QASMC study) was a case-control study conducted at The Heart and Lung Centre in London, UK, according to Good Clinical Practice (GCP) and was registered on ClinicalTrials.gov with the identifier, NCT01780298. It aimed to identify biomarker or a panel of biomarkers that would enable the differentiation between subjects with

1. 60 COPD (current smokers with a ≥ 10 pack year smoking history at GOLD Stage 1 or 2)

and three control groups of matched non-smoking subjects:

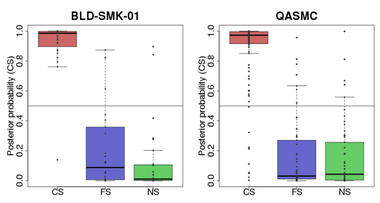
1. 60 never smokers (NS),
2. 60 ex-smokers (FS)
3. 60 current smokers (CS).

(All groups were matched on age and gender)

Additional goals of this study were to assess standard biomarkers of inflammation and to compare inflammatory cell responses and selected markers of inflammation in blood, induced sputum and nasal samples. In this study, male and female subjects aged between 40 and 70 years were included. All subjects were matched by ethnicity, gender and age (within 5 years) with the COPD subjects recruited in the study. Blood samples were sent to AROS Applied Biotechnology AS (Aarhus, Denmark) where they were further processes and hybridized to Affymetrix Human Genome U133 Plus 2.0 GeneChips..

All the COPD patients are Currents Smokers. Therefore it enable not only the validation of eth signature only, but also its robustness with respects to the COPD disease status of the patients

	Truth / Predicted	Core CS	Core NCS	Core True Rate	Extended CS	Extended NCS	Extended True Rate
BLD SMK-01	Former Smoker	3	23	0.88	4	22	0.85
QASMC	Current Smokers	99	11	0.90	100	10	0.91
	Non-Current smoker	12	105	0.90	12	105	0.90



Application to the THS 2.1 Switching Study

The REX-EX-01 study was an open-label, randomized, controlled, two-arm parallel group study (Figure 1) that recruited 42 healthy smokers of both genders, aged between 23 and 65 years. It was conducted to compare smokers switching CC to smokers switching to a candidate Modified Risk Tobacco Product: the Tobacco Heating System 2.1 (THS 2.1) over 5 consecutive days (Figure 2).

The study was conducted according to Good Clinical Practices (GCP) and was registered on ClinicalTrials.gov, with the identifier NCT01780714. Blood samples were further processed and hybridized to Affymetrix Human Genome U133 Plus 2.0 GeneChips.

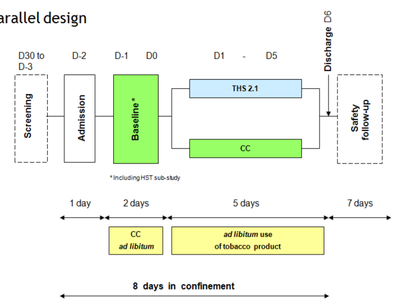
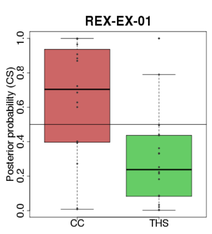


Figure2 : REX-EX-01 study design

The ultimate goal was to apply the signatures obtained to determine whether the impact of switching to THS 2.1 can be readily detected after 5 days of switching in the whole blood transcriptome, thus providing a sensitive and non-invasive tool for assessing the exposure-response in clinical trials. The underlying premise was that the blood transcriptome of smokers who switch to THS 2.1 starts to resemble that of a former smoker. Therefore, instead of characterizing the gene expression profile of a MRTP user that is specific to five days of switching (e.g., by extracting the signature from the REX-EX-01 study data), we set out to find a transcriptomic based exposure response signature that could also serve as an indicator of a longer term switching pattern. This was achieved by establishing the signatures discussed above, that distinguished current smoker samples from non-current smoker samples.
After the quality check of the CEL files of the REX-EX-01 study, 16 and 18 files remained for the CC and Switching arm at Day 5, respectively. For the extended signature, the individuals who remained on CC were mainly classified as CS (69%) while subjects in the switching arm were mostly classified as NCS (89%). For the core signature, the true rate for CC is the same and 78% of the subjects in the switching arm were classified as NCS

Truth / Predicted	Core CS	Core NCS	Core Rate	Extended CS	Extended NCS	Extended Rate
Current Smokers	11	5	0.69	11	5	0.69
Switchers to THS 2.1	4	14	0.78	2	16	0.89



Conclusion

In conclusion, our systems pharmacology approach enabled the construction of a robust whole blood based smoking gene signature. The 'signature' involves measurement of the expression of 11 different genes which respond to exposure to CC smoke and more importantly, differentiate between current smokers and never/former smokers with remarkable accuracy. Validation was performed with an independent dataset.

When applied to the blood transcriptome from the five day exploratory switching study, the signature identified the subjects who switched to a prototype MRTP as non-current smokers while the smokers were accurately classified as such. Therefore, our signature provides a sensitive tool for assessing the exposure response using minimally invasive sampling.

This signature will be used to analyze the whole blood transcriptomics in the planned clinical study ZRHM-REXA-08-US.

References

[1] Meyer, P., Hoeng, J., Rice, J.J., Norel, R., Sprengel, J., Stolle, K., Bonk, T., Corthesy, S., Royyuru, A., and Peitsch, M.C. (2012). Industrial methodology for process verification in research (IMPROVER): toward systems biology verification. Bioinformatics *28*, 1193-1201.

[2] Tarca, A.L., Lauria, M., Unger, M., Bilal, E., Boue, S., Dey, K.K., Hoeng, J., Koeppl, H., Martin, F., and Meyer, P. (2013). Strengths and limitations of microarray-based phenotype prediction: Lessons learned from the improver diagnostic signature challenge. Bioinformatic

[3] Dumeaux, V., Olsen, K.S., Nuel, G., Paulssen, R.H., Borresen-Dale, A.L., and Lund, E. (2010). Deciphering normal blood gene expression variation--The NOWAC postgenome study. PLoS Genet

[4] Bahr, T.M., Hughes, G.J., Armstrong, M., Reisdorph, R., Coldren, C.D., Edwards, M.G., Schnell, C., Kedl, R., LaFlamme, D.J., and Reisdorph, N. (2013). Peripheral Blood Mononuclear Cell Gene Expression in Chronic Obstructive Pulmonary Disease. American journal of respiratory cell and molecular biology

PMI RESEARCH & DEVELOPMENT

Philip Morris International Research & Development, Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland
T: +41 58 242 21 11, **F**: +41 58 242 28 11, **W**: www.pmi.com