The American Journal of Pathology, Vol. 185, No. 11, November 2015

The American Journal of
PATHOLOGY
ajp.amjpathol.org

BIOMARKERS, GENOMICS, PROTEOMICS, AND GENE REGULATION

Smoking-Induced Expression of the *GPR15* Gene Indicates Its Potential Role in Chronic Inflammatory Pathologies

Gea Kõks,* Mari-Liis Uudelepp,*† Maia Limbach,* Pärt Peterson,* Ene Reimann,*‡ and Sulev Kõks*‡

From the Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu; the Department of Genetics,† Tartu University Hospital, Tartu; and the Department of Reproductive Biology,‡ Estonian University of Life Sciences, Tartu, Estonia*

Accepted for publication
July 21, 2015.

Address correspondence to Sulev Kõks, M.D., Ph.D., Institute of Biomedicine and Translational Medicine, University of Tartu, 19 Ravila St., Tartu 50411, Estonia. E-mail: sulev.koks@ut.ee.

Despite the described clear epigenetic effects of smoking, the effect of smoking on genome-wide gene expression in the blood is obscure. We therefore studied the smoking-induced changes in the gene-expression profile of the peripheral blood. RNA was extracted from the whole blood of 48 individuals with a detailed smoking history (24 never-smokers, 16 smokers, and 8 ex-smokers). Gene-expression profiles were evaluated with RNA sequencing, and results were analyzed separately in 24 men and 24 women. In the male smokers, 13 genes were statistically significantly (false-discovery rate <0.1) differentially expressed; in female smokers, 5 genes. Although most of the differentially expressed genes were different between the male and female smokers, the G-protein–coupled receptor 15 gene (*GPR15*) was differentially expressed in both male and female smokers compared with never-smokers. Analysis of *GPR15* methylation identified significantly greater hypomethylation in smokers compared with that in never-smokers. GPR15 is the chemoattractant receptor that regulates T-cell migration and immunity. Up-regulation of *GPR15* could explain to some extent the health hazards of smoking with regard to chronic inflammatory diseases. *(Am J Pathol 2015, 185: 2898–2906; http://dx.doi.org/10.1016/j.ajpath.2015.07.006)*

Tobacco smoking is considered the leading preventable cause of morbidity and mortality. Smoking affects >1 billion individuals worldwide and accounts for an estimated 3 million deaths per year.[1] The prevalence of smoking remains high despite the wide knowledge of its negative effects on health.[1,2] In developed countries, the prevalence of smoking is decreasing, but in developing countries, the number of smokers is still high.[2] Smoking cessation has been shown to be effective in reducing all-cause mortality in ex-smokers.[3–5] Survival curves are similar between never-smokers and individuals who quit smoking before the age of 35 years.[5]

Smoking-induced molecular alterations and their effects on diseases have been analyzed in several previous studies.[6,7] Peripheral gene-expression profiling allows for the detection of early molecular signatures relevant to the toxic effects of smoking.[8,9] Smoking also induces prominent up-regulation of a subtype of regulatory T cells in bronchoalveolar lavage fluid.[10] The association of tobacco use with various chronic diseases is smoking-duration dependent, and the effects of tobacco depend on the amount (pack-years) of smoking.[11,12] These findings suggest roles for epigenetic reprogramming in the modulation of the biological effects of smoking and in the development of smoking-induced signature. Indeed, previous studies have identified an association between global DNA methylation and tobacco smoking in cancer-related tissues.[13,14]

In other studies, an association between global DNA methylation and tobacco smoking was not identified in normal samples.[15,16] Focus on the methylation of particular gene loci has identified several regions as differentially methylated between smokers and never-smokers. The CpG site in the *COMT* gene at position -193 is methylated in

Supported by institutional research funding (IUT20-46) from the Estonian Ministry of Education and Research, the Centre of Translational Genomics of the University of Tartu (SP1GVARENG), and the European Regional Development Fund, Center of Translational Medicine, University of Tartu.

Disclosures: None declared.

http://dx.doi.org/10.1016/j.ajpath.2015.07.006

22.2% of smokers and 18.3% of never-smokers.[17] Similarly, in smokers, significant hypomethylation of the nontranslated region in the *MAOA* gene has been reported.[18] In addition, a clear correlation between methylation and smoking status (smokers, ex-smokers, and never-smokers) has been observed; smokers had a significantly lesser amount of methylation in the *MAOA* locus compared with that in never-smokers, and the pattern in ex-smokers was in between.[18]

Sex differences in the methylation response to smoking have also been addressed.[18] Smoking induced hypomethylation in the *MAOB* promoter region, with a high monoamine oxidase protein concentration.[19]

Studies of genome-wide methylation have identified distinct regions affected by smoking and smoking cessation.[20,21] In one study, 27,000 sites in the peripheral blood DNA were analyzed for methylation. Factor II receptor–like 3 gene (*F2RL3*) expression was robustly associated with smoking status, and this finding was replicated in two independent samples of European ancestry.[22] In another study with the same methylation array (27K BeadChip; Illumina, San Diego, CA), the *F2RL3* finding was replicated, and additionally a novel association at the G-protein–coupled receptor 15 gene (*GPR15*) locus was identified.[21]

Recently, higher-density arrays have been used. An epigenome-wide association study in 374 Europeans replicated the smoking-related hypomethylation of *F2RL3* and identified three additional loci, including the aryl hydrocarbon receptor repressor gene (*AHRR*).[23] Another study in a larger cohort that used the same 450K methylation chip identified several smoking-dependent loci.[24] In addition to the confirmation of the *AHRR* methylation, at least eight additional loci were found. Methylation was found to depend on the cessation time and pack-years of smoking.[24] All of these studies indicate the broad effects of smoking on the genome and cell physiology.

Although the clear effects of smoking on epigenetics have been described, the effects of smoking on genome-wide gene expression in blood is not often studied. Our goals were to describe the smoking-induced changes in the gene-expression profile of the blood and to identify methylated regions that match the altered RNA levels. We analyzed RNA and DNA extracted from the whole blood of donors at the Estonian Genome Center, University of Tartu (Tartu, Estonia).

Materials and Methods

Study Cohort

The study cohort was derived from the Estonian Biobank of the Estonian Genome Center.[25] The Ethics Review Committee on Human Research of the University of Tartu approved the protocols and informed-consent forms used in this study. All of the participants signed a written informed-consent form.

The Estonian Biobank cohort is a volunteer-based sample of the Estonian resident adult population (aged over 18 years). The current number of participants (52,000) represents 5% of the adult population of Estonia, making it ideally suited to population-based studies. The Biobank stores DNA and peripheral blood mononuclear cells from donors, along with data on lifestyle (eg, smoking and alcohol-intake habits, physical activity). Forty-eight samples (from 24 men and 24 women) were used for gene-expression profiling in the present study.

RNA Extraction

For collecting whole-blood samples, Tempus blood RNA tubes (Thermo Fisher Scientific Inc., Waltham, MA) were used. The samples were stored at −20°C until RNA extraction. For total RNA extraction, the combination of TRIzol reagent (Thermo Fisher Scientific) and the RNeasy Mini Kit (Qiagen, Hilden, Germany) was used. The protocol was as follows—after samples were thawed and mixed the in Tempus tube, blood was transferred to an empty 50-mL tube. The Tempus tube was additionally washed with 3 mL of phosphate-buffered saline. The sample in the 50-mL tube was mixed and centrifuged at 4°C for 60 minutes at 3000 × *g*. The supernatant was discarded (the invisible precipitation was at the bottom of the tube), and the tube without the cap was placed onto the clean tissue, upside down, for 2 minutes. TRIzol reagent (1 mL) was added into the tube, mixed, and incubated for 5 minutes at room temperature. The sample was lifted to a new 1.5-mL tube, and 200 μL of chloroform was added. The sample was mixed with vortexing for 15 seconds, incubated at room temperature for 2 to 3 minutes, and centrifuged at 4°C for 15 minutes at 12,000 × *g*. Five hundred microliters of the upper (clear) phase was lifted to a new 1.5-mL tube, and an equal amount of isopropanol was added to the sample. The sample was mixed with vortexing for 15 seconds, incubated at room temperature for 10 minutes, and centrifuged at 4°C for 10 minutes at 12,000 × *g*. The invisible RNA sediment was at the bottom. The supernatant was discarded, and 1 mL of freshly prepared 75% ethanol was added and centrifuged at 4°C for 5 minutes at 7500 × *g*. The supernatant was discarded and the ethanol wash step was repeated once more.

After the second wash, the tube was dried with open cap at room temperature for 5 minutes. The RNA was eluted in 50 μL of nuclease-free water and incubated at 55°C for 5 minutes. The DNase treatment was conducted with an Ambion Turbo DNA-free kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The final volume of DNase-treated total RNA was 50 μL. The total RNA was then cleaned with an RNeasy Mini Kit. Three hundred fifty microliters of RLT buffer was added and the sample was mixed, followed by the addition of 1225 μL of 100% ethanol and mixing. The sample was centrifuged through RNeasy Mini Kit columns for 15 seconds at 8000 × *g* and washed twice with 500 μL of buffer RPE. The RNA was eluted in 50 μL of nuclease-free water. The quality of total RNA was evaluated with an Agilent 2100 Bioanalyzer and the RNA 6000 Nano kit (Agilent Technologies Inc., Santa Clara, CA); the RNA Integrity Number of all of the samples was >5.

Table 1 Characteristics of the Subjects Enrolled in the Present Study

Characteristic	Men	Women
Number of participants	24	24
Age, years	52.3 ± 17.9	53.0 ± 17.9
Smoking status, n		
Smoker	8	8
Ex-smoker	8	0
Never-smoker	8	16
Years of smoking	33 ± 17.8	20 ± 13.9
Height, cm	177.6 ± 7.2	163.2 ± 6.4
Weight, kg	88.0 ± 21.6	69.5 ± 14.1
Body mass index, kg/m^2	27.8 ± 5.8	26.1 ± 5.4

Quantitative data are expressed as means ± SEM.

Total RNA extraction was conducted at the Estonian Genome Center. From there, 2 μg of total RNA per individual was transported to the Core Facility of Clinical Genomics, University of Tartu, where all of the following procedures were conducted.

Globin Clear Treatment

Total RNA from whole blood consists of up to 70% of Ig mRNA; a Globin Clear Human kit (Thermo Fisher Scientific) was applied to purify the samples from globin mRNA. After Globin Clear treatment, nearly 1.5 μg of RNA was left. The RNA Integrity Number remained >5. The RNA quality was assessed using the Agilent 2100 Bioanalyzer and the RNA 6000 Nano kit (Agilent Technologies).

SOLiD WT RNAseq Library Preparation and Sequencing

Whole transcriptome RNAseq libraries for 48 RNA samples were prepared. Fifty nanograms of each Globin Clear kit–treated total RNA sample was taken as library input. The rest of the RNA material was stored at −80°C. For library preparation, the Ovation RNAseq V2 Kit (NuGen, Emeryville, CA) together with the 5500 Series Fragment Library Core Kit (Thermo Fisher Scientific) were used according to the manufacturers' protocols.

An automated SOLiD EZ Bead System and EZ Bead E80 System Consumables (Thermo Fisher Scientific) were applied for emulsion PCR. With each template preparation, the pool of 12 libraries was used (marked with barcoding sequences to distinguish the samples on data analysis). All together, four template preparations were achieved.

The samples were sequenced with SOLiD 5500 xl platform (Thermo Fisher Scientific) on two flowchips. For each sample, at least 40 million mappable reads were received, which is enough for gene-expression analysis and also fusion and exon junction analysis. Paired-end chemistry for barcoded libraries was used, which provides up to 110 Bp (75 Bp forward and 35 Bp reverse) per one paired-end read.

DNA Methylation Analysis

We also compared methylation patterns between the smokers and never-smokers using MassArray EpiTyper DNA methylation technology (Agena Bioscience, San Diego, CA). Samples were prepared using an EpiTyper T Complete Reagent Set according to the manufacturer's instructions (Agena Bioscience). The bisulfite-treated DNA (25 ng) was amplified with Hot FirePol DNA Polymerase (Solis BioDyne, Tartu, Estonia), and CpG methylation was determined by the MassArray Analyzer 4 (Agena Bioscience). The specific primers were designed with the EpiDesigner software beta version (Agena Bioscience), and the primer sequences for *GPR15* were 5′-AGGAAGAGAGTATTGTTTTTTTGGGTGGATAAAGA-3′ and 5′-CAGTAATACGACTCACTATAGGGAGAAGG-CTCAATAACAAATCACAATACTCAACAAAA-3′.

Bioinformatics and Statistical Analysis

Raw reads were color-space mapped to the human genome hg19 reference using a Maxmapper algorithm implemented in Lifescope software version 2.5.1 (Life Technologies Corporation, Carlsbad, California). Mapping to multiple locations was permitted. The quality threshold was set at 10, providing a mapping confidence of >90. Reads with a Phred score <10 were filtered out. The mean mapping quality was 30. RNA content and gene-based annotation were analyzed with whole-transcriptome workflow. Raw sequencing data with appropriate experimental information are available from the Gene Expression Omnibus repository (*www.ncbi.nlm.nih.gov/geo*; accession number GSE68549).

We analyzed RNA samples of blood from 24 smokers and 24 never-smokers with SOLiD 5500xl RNA sequencing technology. There were three main groups: smokers (n = 16), ex-smokers (n = 8) and individuals who had never

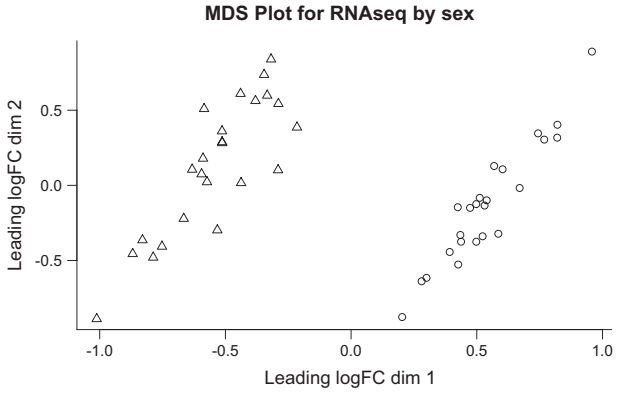

Figure 1 Multidimensional scaling (MDS) plot of expression data illustrates a significant effect of sex on variations in global gene expression. Normalized RNAseq counts were used in the MDS analysis to find the similarities of individual cases of a data set. Based on the results from MDS, we decided to analyze data from men (triangles) and women (circles) separately. logFC, log fold-change.

Table 2 Gene Expression of Smokers and Never-Smokers among Women

Gene name	Symbol	logFC	logCPM	*P*	FDR
G-protein–coupled receptor 128	*GPR128*	4.36	1.27	1.31×10^{-7}	0.003
G-protein–coupled receptor 15	*GPR15*	1.75	2.35	3.36×10^{-7}	0.004
LINE-1 type transposase domain-containing protein 1	*L1TD1*	1.63	1.36	8.88×10^{-6}	0.063
SPARC-related modular calcium binding protein 1	*SMOC1*	−1.97	−0.37	1.08×10^{-5}	0.063
FK506 binding protein 10, 65 kDa	*FKBP10*	−2.04	−0.43	2.11×10^{-5}	0.097
Long intergenic nonprotein–coding RNA 518	*C6orf218*	−3.15	−1.09	4.29×10^{-5}	0.142
Prostaglandin D_2 synthase, 21 kDa (brain)	*PTGDS*	−1.46	1.87	3.81×10^{-5}	0.142
Uncharacterized locus MGC21881	*MGC21881*	−2.16	−0.25	6.56×10^{-5}	0.189

RNAseq data were analyzed with edgeR software package version 3.2.0.

FDR, false-discovery rate (*P* value corrected for multiple testing); LINE, long interspersed nuclear element; logFC, log fold-change (smokers − never-smokers); logCPM, gene expression level in log of counts per million; SPARC, secreted protein, acidic, cysteine-rich protein.

smoked ($n = 14$) (Table 1). We analyzed men and women separately and left the data from the ex-smokers out of the initial statistical comparison. The data from ex-smokers were used only for illustrative purposes.

Non-normalized raw counts were used for the edgeR package version 3.2.0 (Bioconductor, *http://bioconductor.org*) to perform differential gene-expression analysis after quality control of samples. edgeR performs model-based scale normalization, estimates dispersions, and applies negative binomial modeling. edgeR is a flexible tool for RNAseq data analysis to identify differentially expressed genes.[26,27] It implements negative binomial model fitting, followed by testing procedures for determining differential expression.

To detect differentially expressed genes, we used negative binomial fitting followed by Fisher exact testing. False-discovery rate adjustment was used for multiple-testing correction.[28] A false-discovery rate threshold of 0.1 for statistical significance was applied. Genes with greater differential expression were defined with a threshold of log fold-change 0.5 (ie, 50% change between experimental conditions). We analyzed men and women as separate data sets (24 individuals each) (Table 1).

Results

RNA Sequencing

The basic characteristics of the study groups are listed in the Table 1. For each sample, at least 40 million mappable reads were received, which is enough for gene-expression analysis and also fusion and exon junction analysis. Paired-end chemistry for barcoded libraries were used, which gives up to 110 Bp (75 Bp forward and 35 Bp reverse) per one paired-end read. RNA sequencing provided high-quality reads with good similarity between different samples. Multidimensional scaling analysis of fold-change differences in gene expression indicated good separation of study groups by sex (Figure 1). Multidimensional scaling analysis showed the sample distances and illustrated the similarity of samples based on the biological coefficient of variation. As we saw strong separation of samples by sex, to reduce confounding effects, it was better to analyze samples separately by sex. Therefore, in further

Table 3 Differential Gene Expression of Smokers and Never-Smokers among Men

Gene name	Gene	logFC	logCPM	*P*	FDR
G-protein–coupled receptor 15	*GPR15*	2.61	2.96	3.21×10^{-13}	7.42×10^{-9}
Myc target 1	*MYCT1*	−1.06	4.38	1.59×10^{-6}	0.02
Coiled-coil domain-containing protein 3	*CCDC3*	−3.59	1.00	5.19×10^{-6}	0.03
Cellular repressor of E1A-stimulated genes 1	*CREG1*	−0.74	8.30	4.52×10^{-6}	0.03
Chemokine (C-C motif) receptor 8	*CCR8*	1.87	1.74	7.26×10^{-6}	0.03
Phospholipase C, η1	*PLCH1*	−1.04	2.98	1.04×10^{-5}	0.04
Complement component 4 binding protein α	*C4BPA*	−3.22	2.60	2.01×10^{-5}	0.07
Testis development–related protein	*C8orf42*	−1.24	1.41	2.48×10^{-5}	0.07
CD177 molecule	*CD177*	−2.60	2.77	4.10×10^{-5}	0.09
FERM domain-containing protein 4B	*FRMD4B*	−0.68	5.76	3.43×10^{-5}	0.09
Follistatin-like 1	*FSTL1*	−1.08	3.83	4.83×10^{-5}	0.09
Nuclear factor I/A	*NFIA*	−0.74	6.20	4.49×10^{-5}	0.09
Tropomyosin 1 (α)	*TPM1*	−0.79	6.69	3.97×10^{-5}	0.09

RNAseq data were analyzed with edgeR package.

FDR, false-discovery rate (*P* value corrected for multiple testing); FERM, 4.1, ezrin, radixin, moesin protein; logFC, log fold-change (smokers − never-smokers); logCPM, gene expression level in log of counts per million.

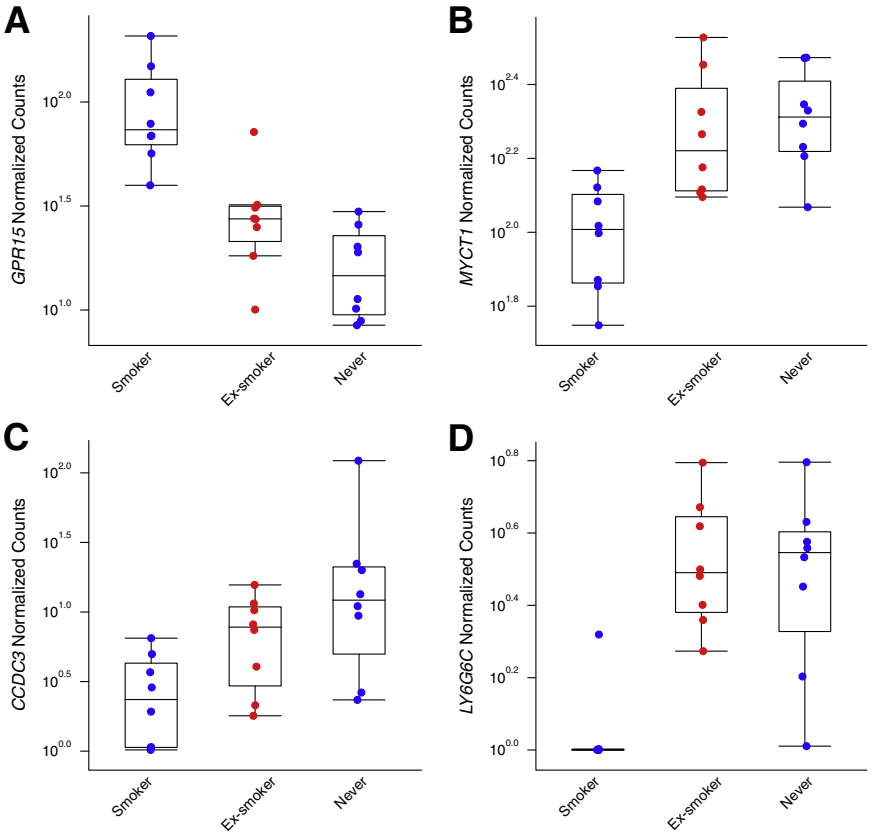

Figure 2 Boxplot illustrating the relationship between the normalized expression of the most significantly differentially expressed genes [*GPR15* (**A**), *MYCT1* (**B**), *CCDC3* (**C**), and *LY6G6C* (**D**)] and smoking status in men. The levels of expression of the genes in this figure showed the greatest statistical differences between smokers and never-smokers (Never) on comparison of the RNAseq data (Fisher exact test). For illustrative purposes, the group of ex-smokers was added.

statistical analysis we used data from men and women separately.

In women, a comparison of gene-expression profiles between smokers and never-smokers revealed differential expression (false-discovery rate <0.1) of five genes: *FKBP10*, *GPR128*, *GPR15*, *L1TD1*, and *SMOC1* (Table 2). Of these, the only gene found to be related to smoking in earlier studies was *GPR15*. Other genes were not identified in earlier studies as related to smoking.

In men, a comparison of the gene-expression profiles between smokers and never-smokers identified differential expression (false-discovery rate <0.1) of 15 genes: *C4BPA*, *C8orf42*, *CCDC3*, *CCR8*, *CD177*, *CREG1*, *FRMD4B*, *FSTL1*, *GPR15*, *MYCT1*, *NFIA*, *PLCH1*, and *TPM1* (Table 3). Again, the only gene found in earlier studies to be related to smoking was *GPR15*. Other genes were not identified in earlier studies as related to smoking. The gene-expression levels of four genes related to smoking status (smokers, ex-smokers, and never-smokers) are illustrated in Figure 2. The lymphocyte antigen 6 complex, locus G6C (*LY6G6C*) was significantly down-regulated in smokers when we analyzed the entire study sample (men and women together). *GPR15* expression correlated well with smoking status. The highest expression was in smokers, and the lowest was in never-smokers. The expression level was in between in ex-smokers (Figure 2). Similarly, expression of the *CCDC3* gene followed the smoking status with the lowest level in smokers, intermediate level in ex-smokers, and the greatest level in never-smokers (Figure 2). Interestingly, the expression levels of *MYCT1* and *LY6G6C* were reduced only in smokers and not in ex-smokers.

The heatmap of gene-expression data based on 50 genes with the lowest *P* values illustrated a clear smoking-related expressional pattern (Figure 3). Male smokers were clustered into two groups with characteristic gene-expression profiles. *GPR15, RTKN2*, *USP46, CCR4,* and *CCR8* formed a cluster of genes up-regulated in smokers (Figure 3). In addition, *FOXP3* and *GPR15* showed similar expression patterns, suggesting potential correlation. Indeed, analysis found a correlation coefficient of 0.59 between *GPR15* and *FOXP3*, with a *P* value of 1.15E−05.

DNA Methylation

DNA methylation analysis in the *GPR15* locus was performed, and two CpG sites were analyzed. Hypomethylation of CpG1 was significantly greater in smokers compared with that in never-smokers (Figure 4). This finding correlated with the gene-expression results, in which significant overexpression of *GPR15* was observed. We did not find differential methylation with another CpG site in the *GPR15* gene.

Discussion

Our study showed that smoking induces the overexpression of the *GPR15*, and this change was statistically significant in both men and women. The level of *GPR15* expression

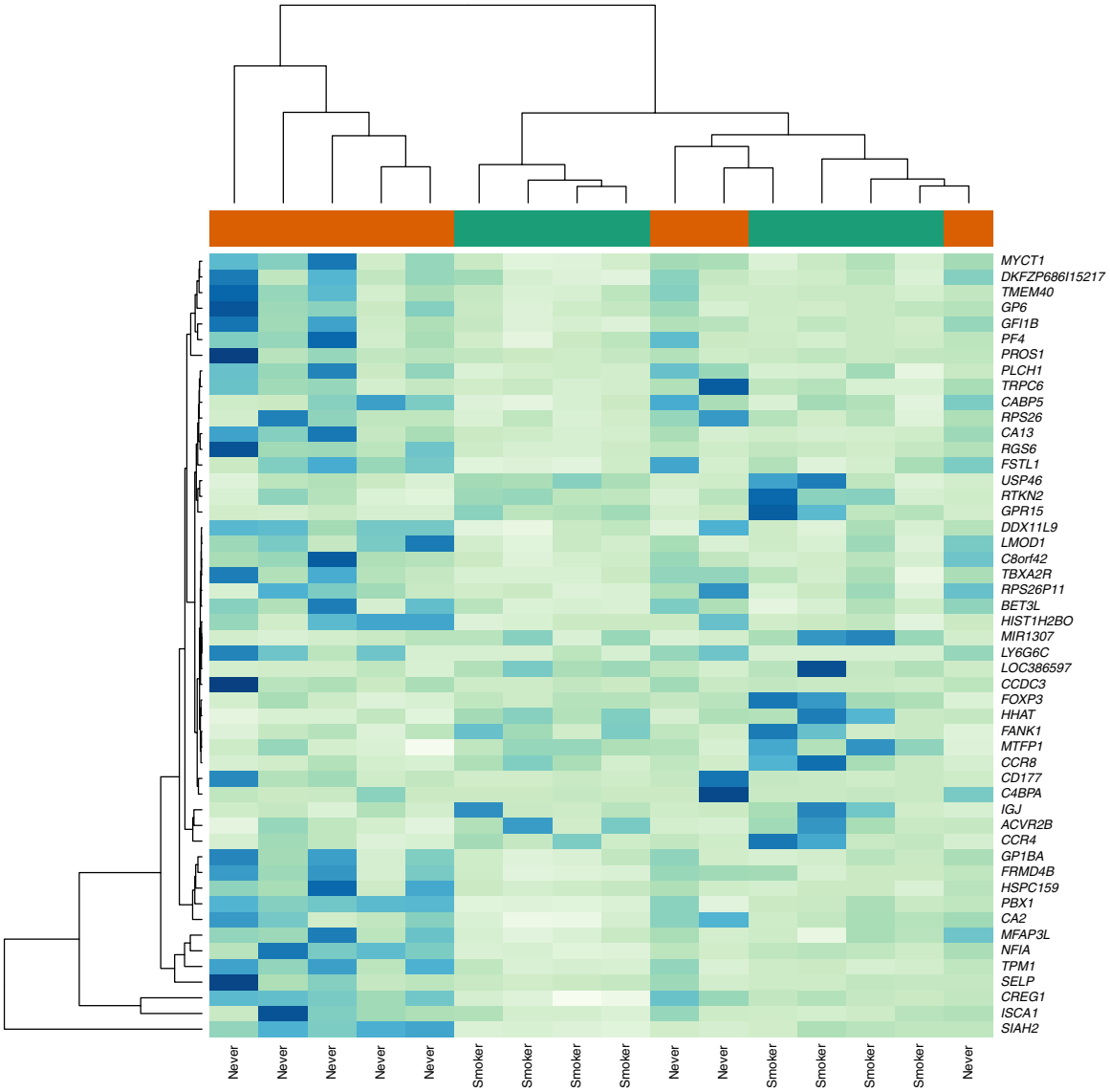

Figure 3 Heatmap illustrating RNAseq expression (normalized counts) of the 50 genes with the greatest statistical differences between smokers (green) and never-smokers (Never, red) (Fisher exact test). Deep blue indicates greater expression; light green indicates lesser expression.

followed smoking status. *GPR15* expression was greatest in smokers, in the intermediate range in ex-smokers, and least in never-smokers. Moreover, we found that increased *GPR15* gene expression was related to the hypomethylation of the CpG1 locus in the *GPR15* gene, and we conclude that hypomethylation of this locus is involved in the up-regulation of *GPR15*. Although smoking-induced hypomethylation of the *GPR15* locus was reported in earlier studies, differential expression of *GPR15* RNA caused by smoking has not been described.[20,21,29,30]

Hypomethylation of the *GPR15* locus is linked to the cumulative exposure to smoking,[21] suggesting that the longer the smoking period, the more hypomethylated the region. Smoking cessation reverses hypomethylation of the *GPR15* locus.[20] Therefore, *GPR15* seems to be an interesting target for the biological effects of smoking. The function of *GPR15* gives additional impact and significance that this finding can have in the smoking-related pathologies.

GPR15 was discovered as a novel G-protein–coupled receptor in chromosome 3.[31] It is a membrane-localizing protein and acts as a chemokine receptor for human immunodeficiency viruses 1 and 2.[32,33] *GPR15* is also known as brother of Bonzo (*BOB*) and is expressed in the lymphoid tissue and colon.[32] It is involved in lymphocyte homing in the large intestine, which the inflammatory bowel disease most commonly affects. Indeed, recent studies have identified a role for *GPR15* in the development of inflammation in the colon.[34,35]

There are substantial species-specific differences in the expression of *GPR15* in T cells. In mice, *Gpr15* is expressed in regulatory T cells, and in humans, *GPR15* is expressed in Th2 cells.[35] We found a positive correlation ($r = 0.59$) between the expression levels of *GPR15* and *FOXP3*

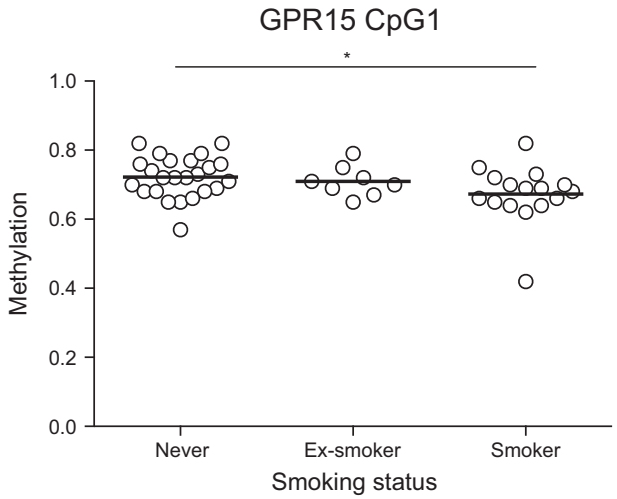

Figure 4 Methylation of the *GPR15* locus CpG1 is dependent on smoking status. Smokers have significantly greater hypomethylation compared with never-smokers (Never). Pairwise (eg, smoker versus never-smoker) approach and *t*-test were used for determining statistical significance. $*P < 0.05$ never-smokers versus smokers.

(Figure 3). This finding is in agreement with those from a recent study in which the regulatory role of GPR15 on the forkhead box P3−positive ($FOXP3^+$) regulatory cells was described.[34] In addition, *GPR15* regulates the homing of regulatory T cells ($FOXP3^+$) cells, and FOXP3 binds significantly to the *GPR15* enhancer in humans.[35] Interestingly, smoking seems to have the opposite effect on *FOXP3* activation, depending on the subjects and pathological status. In patients with ulcerative colitis, smoking increases the prevalence of $FOXP3^+$ cells, whereas in patients with Crohn disease, it increases the Th1 subsets.[36] In smokers with normal lung function, smoking prominently up-regulates the regulatory T cells in bronchoalveolar lavage fluid; this effect is absent in patients with chronic obstructive pulmonary disease.[10] Taken together, the co-regulation of *GPR15* and *FOXP3* in our study supports the functional relevance of the smoking-induced up-regulation of *GPR15*. Smoking seems to induce the appropriate transcriptional network necessary for *GPR15* induction.

As mentioned in the previous paragraph, GPR15 is involved in the inflammation of the large intestine by mediating T-cell recruitment to the colon. GPR15 is also required for the trafficking of dendritic epidermal T cells to the epidermal tissues.[37] Therefore, GPR15 is involved in the homing of T cells into the epithelial barrier tissues. In addition to skin and colon, GPR15 has been described in the synovial tissue in patients with rheumatoid arthritis (RA).[38] A further study confirmed that GPR15 is expressed in $CD68^+$ and $CD14^+$ macrophages in synovia.[39] In rheumatoid arthritis patients, GPR15 is expressed in both lining and sublining layers and is localized to the cell membrane and cytoplasm.[39] In the peripheral blood, the GPR15 was detected on monocytes and neutrophils.[39] Interestingly, GPR15 was not found to localize in T or B lymphocytes.[39] The expression level of GPR15 in the peripheral blood and synovia was dependent on the presence of inflammation. These findings taken together suggest that GPR15 is an orphan receptor with a well-recognized role in the regulation of immune response, and smoking increases its expression.

Smoking is recognized as a significant health risk factor for various chronic inflammation. In the case of skin (psoriasis) and intestinal chronic inflammation (Crohn disease and ulcerative colitis), the impact of smoking is well known, albeit controversial.[40–43] Smoking increases the risk for Crohn disease but reduces the risk for ulcerative colitis.[42] Even large-scale meta-analyses have found that current smoking is a significant risk factor for Crohn disease but is significantly protective against ulcerative colitis.[44,45] The association of smoking with psoriasis is more clear. A recent population-based, cross-sectional study found a significantly increased prevalence of psoriasis in smokers and ex-smokers.[40,46] Smokers typically have more active psoriasis than do never-smokers.[47] Therefore, smoking regulates *GPR15*, a known immunomodulator, which might explain the role of smoking in autoimmune diseases.

In many previous studies, *AHRR* was found to be epigenetically modified by smoking.[20,24,30] We did not find changes in the gene-expression level of *AHR* or *AHRR*. However, we found significant methylation differences in the case of *AHRR*, but these differences did not have an impact on RNA expression level (data not shown). Our approach was to start from the global gene-expression data to identify the differentially expressed genes and then to find molecular reasons for these differences. In the case of *GPR15*, this approach worked nicely—we saw differences in RNA levels and in methylation. These findings are a good example in which the results of RNA analysis coincide with epigenetic results, which provides additional support that the found differences have a biological impact. At the same time, the identification of *GPR15* is functionally a relevant finding to explain, at least partially, the health effects of smoking.

Our study had some important limitations. Its cross-sectional design did not allow for determining whether the health outcomes were correlated with the expression of *GPR15*. We also did not have functional immune cells from the study groups to perform more detailed mechanistic analysis. Our results support a causative connection between smoking and *GPR15* expression via hypomethylation, but we still were not able to show a full causative pathway from smoking to GPR15 changes and to the development of pathologies. These questions will be the focus of further study.

Conclusion

We found that smoking induces overexpression of *GRP15* in the blood, and that this overexpression is caused by the hypomethylation of this locus. This change was evident in both men and women. *GPR15* is the chemoattractant receptor that directs the homing of T cells to the colon and skin and is

up-regulated in the synovial tissue in cases of rheumatoid arthritis. Therefore, *GPR15* and its up-regulation are interesting candidates for the explanation of the health hazards of smoking in chronic inflammatory diseases. The induction of *GPR15* expression might be, at least partially, the mechanism of the influence of smoking on the functions of the body.

Acknowledgments

We thank the Estonian Genome Center for providing data and samples.

G.K. analyzed data and prepared the manuscript. E.R. performed RNA sequencing. M.-L.U. and M.L. performed methylation analysis. P.P. performed methylation analysis and helped with data analysis. S.K. conceived the study, analyzed RNAseq data, and wrote the manuscript.

References

1. Peto R, Lopez AD, Boreham J, Thun M, Heath C Jr, Doll R: Mortality from smoking worldwide. Br Med Bull 1996, 52:12–21
2. WHO: WHO Report on the Global Tobacco Epidemic, 2013: Warning about the Dangers of Tobacco. Geneva, Switzerland: World Health Organization, 2013
3. Cully M: Public health: the benefits and challenges of smoking cessation. Nat Rev Cardiol 2013, 10:117
4. Huxley RR, Woodward M: Full hazards of smoking and benefits of stopping for women. Lancet 2013, 381:96–98
5. Jha P, Ramasundarahettige C, Landsman V, Rostron B, Thun M, Anderson RN, McAfee T, Peto R: 21st-century hazards of smoking and benefits of cessation in the United States. N Engl J Med 2013, 368: 341–350
6. Ryder MI, Hyun W, Loomer P, Haqq C: Alteration of gene expression profiles of peripheral mononuclear blood cells by tobacco smoke: implications for periodontal diseases. Oral Microbiol Immunol 2004, 19:39–49
7. Bahr TM, Hughes GJ, Armstrong M, Reisdorph R, Coldren CD, Edwards MG, Schnell C, Kedl R, LaFlamme DJ, Reisdorph N, Kechris KJ, Bowler RP: Peripheral blood mononuclear cell gene expression in chronic obstructive pulmonary disease. Am J Respir Cell Mol Biol 2013, 49:316–323
8. Lampe JW, Stepaniants SB, Mao M, Radich JP, Dai H, Linsley PS, Friend SH, Potter JD: Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke. Cancer Epidemiol Biomarkers Prev 2004, 13:445–453
9. van Leeuwen DM, Gottschalk RW, van Herwijnen MH, Moonen EJ, Kleinjans JC, van Delft JH: Differential gene expression in human peripheral blood mononuclear cells induced by cigarette smoke and its constituents. Toxicol Sci 2005, 86:200–210
10. Barcelo B, Pons J, Ferrer JM, Sauleda J, Fuster A, Agusti AG: Phenotypic characterisation of T-lymphocytes in COPD: abnormal CD4+CD25+ regulatory T-lymphocyte response to tobacco smoking. Eur Respir J 2008, 31:555–562
11. Pirie K, Peto R, Reeves GK, Green J, Beral V: The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. Lancet 2013, 381:133–141
12. Thun MJ, Carter BD, Feskanich D, Freedman ND, Prentice R, Lopez AD, Hartge P, Gapstur SM: 50-year trends in smoking-related mortality in the United States. N Engl J Med 2013, 368:351–364
13. Furniss CS, Marsit CJ, Houseman EA, Eddy K, Kelsey KT: Line region hypomethylation is associated with lifestyle and differs by human papillomavirus status in head and neck squamous cell carcinomas. Cancer Epidemiol Biomarkers Prev 2008, 17:966–971
14. Smith IM, Mydlarz WK, Mithani SK, Califano JA: DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. Int J Cancer 2007, 121: 1724–1728
15. Figueiredo JC, Grau MV, Wallace K, Levine AJ, Shen L, Hamdan R, Chen X, Bresalier RS, McKeown-Eyssen G, Haile RW, Baron JA, Issa JP: Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. Cancer Epidemiol Biomarkers Prev 2009, 18:1041–1049
16. Zhu ZZ, Hou L, Bollati V, Tarantini L, Marinelli B, Cantone L, Yang AS, Vokonas P, Lissowska J, Fustinoni S, Pesatori AC, Bonzini M, Apostoli P, Costa G, Bertazzi PA, Chow WH, Schwartz J, Baccarelli A: Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. Int J Epidemiol 2012, 41: 126–139
17. Xu Q, Ma JZ, Payne TJ, Li MD: Determination of Methylated CpG Sites in the Promoter Region of Catechol-O-Methyltransferase (COMT) and their Involvement in the Etiology of Tobacco Smoking. Front Psychiatry 2010, 1:16
18. Philibert RA, Beach SR, Gunter TD, Brody GH, Madan A, Gerrard M: The effect of smoking on MAOA promoter methylation in DNA prepared from lymphoblasts and whole blood. Am J Med Genet B Neuropsychiatr Genet 2010, 153B:619–628
19. Launay JM, Del Pino M, Chironi G, Callebert J, Peoc'h K, Megnien JL, Mallet J, Simon A, Rendu F: Smoking induces long-lasting effects through a monoamine-oxidase epigenetic regulation. PLoS One 2009, 4:e7959
20. Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Vinuela A, Grundberg E, Nelson CP, Meduri E, Buil A, Cambien F, Hengstenberg C, Erdmann J, Schunkert H, Goodall AH, Ouwehand WH, Dermitzakis E, Spector TD, Samani NJ, Deloukas P: Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. Epigenetics 2014, 9:1382–1396
21. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI, Agusti A, Anderson W, Lomas DA, Demeo DL: Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. Hum Mol Genet 2012, 21: 3073–3082
22. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H: Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am J Hum Genet 2011, 88:450–457
23. Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, Belvisi MG, Brown R, Vineis P, Flanagan JM: Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. Hum Mol Genet 2013, 22:843–851
24. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Lattka E, Adamski J, Peters A, Strauch K, Waldenberger M, Illig T: Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS One 2013, 8: e63812
25. Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, Perola M, Ng PC, Magi R, Milani L, Fischer K, Metspalu A: Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol 2015, 44:1137–1147
26. McCarthy DJ, Chen Y, Smyth GK: Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 2012, 40:4288–4297
27. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010, 26:139–140
28. Storey JD, Tibshirani R: Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003, 100:9440–9445

29. Sun YV, Smith AK, Conneely KN, Chang Q, Li W, Lazarus A, Smith JA, Almli LM, Binder EB, Klengel T, Cross D, Turner ST, Ressler KJ, Kardia SL: Epigenomic association analysis identifies smoking-related DNA methylation sites in African Americans. Hum Genet 2013, 132:1027−1037
30. Dogan MV, Shields B, Cutrona C, Gao L, Gibbons FX, Simons R, Monick M, Brody GH, Tan K, Beach SR, Philibert RA: The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. BMC Genomics 2014, 15:151
31. Heiber M, Marchese A, Nguyen T, Heng HH, George SR, O'Dowd BF: A novel human gene encoding a G-protein-coupled receptor (GPR15) is located on chromosome 3. Genomics 1996, 32:462−465
32. Deng HK, Unutmaz D, KewalRamani VN, Littman DR: Expression cloning of new receptors used by simian and human immunodeficiency viruses. Nature 1997, 388:296−300
33. Farzan M, Choe H, Martin K, Marcon L, Hofmann W, Karlsson G, Sun Y, Barrett P, Marchand N, Sullivan N, Gerard N, Gerard C, Sodroski J: Two orphan seven-transmembrane segment receptors which are expressed in CD4-positive cells support simian immunodeficiency virus infection. J Exp Med 1997, 186:405−411
34. Kim SV, Xiang WV, Kwak C, Yang Y, Lin XW, Ota M, Sarpel U, Rifkin DB, Xu R, Littman DR: GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. Science 2013, 340:1456−1459
35. Nguyen LP, Pan J, Dinh TT, Hadeiba H, O'Hara E 3rd, Ebtikar A, Hertweck A, Gokmen MR, Lord GM, Jenner RG, Butcher EC, Habtezion A: Role and species-specific expression of colon T cell homing receptor GPR15 in colitis. Nat Immunol 2015, 16:207−213
36. Ueno A, Jijon H, Traves S, Chan R, Ford K, Beck PL, Iacucci M, Fort Gasia M, Barkema HW, Panaccione R, Kaplan GG, Proud D, Ghosh S: Opposing effects of smoking in ulcerative colitis and Crohn's disease may be explained by differential effects on dendritic cells. Inflamm Bowel Dis 2014, 20:800−810
37. Lahl K, Sweere J, Pan J, Butcher E: Orphan chemoattractant receptor GPR15 mediates dendritic epidermal T-cell recruitment to the skin. Eur J Immunol 2014, 44:2577−2581
38. Schmutz C, Hulme A, Burman A, Salmon M, Ashton B, Buckley C, Middleton J: Chemokine receptors in the rheumatoid synovium: upregulation of CXCR5. Arthritis Res Ther 2005, 7:R217−R229
39. Cartwright A, Schmutz C, Askari A, Kuiper JH, Middleton J: Orphan receptor GPR15/BOB is up-regulated in rheumatoid arthritis. Cytokine 2014, 67:53−59
40. Helmick CG, Lee-Han H, Hirsch SC, Baird TL, Bartlett CL: Prevalence of psoriasis among adults in the U.S.: 2003-2006 and 2009-2010 National Health and Nutrition Examination Surveys. Am J Prev Med 2014, 47:37−45
41. Kinahan CE, Mazloom S, Fernandez AP: Impact of smoking on response to systemic treatment in patients with psoriasis: a retrospective case-control study. Br J Dermatol 2015, 172:428−436
42. Ponder A, Long MD: A clinical review of recent findings in the epidemiology of inflammatory bowel disease. Clin Epidemiol 2013, 5:237−247
43. Ott C, Takses A, Obermeier F, Schnoy E, Muller M: Smoking increases the risk of extraintestinal manifestations in Crohn's disease. World J Gastroenterol 2014, 20:12269−12276
44. Calkins BM: A meta-analysis of the role of smoking in inflammatory bowel disease. Dig Dis Sci 1989, 34:1841−1854
45. Mahid SS, Minor KS, Soto RE, Hornung CA, Galandiuk S: Smoking and inflammatory bowel disease: a meta-analysis. Mayo Clin Proc 2006, 81:1462−1471
46. Huerta C, Rivero E, Rodriguez LA: Incidence and risk factors for psoriasis in the general population. Arch Dermatol 2007, 143:1559−1565
47. Emre S, Metin A, Demirseren DD, Kilic S, Isikoglu S, Erel O: The relationship between oxidative stress, smoking and the clinical severity of psoriasis. J Eur Acad Dermatol Venereol 2013, 27:e370−e375