

## Research Article

Effects of Cigarette Smoke on the Human Oral  
Mucosal Transcriptome

Jay O. Boyle<sup>1</sup>, Zeynep H. Gümüş<sup>2,8</sup>, Ashutosh Kacker<sup>3</sup>, Vishal L. Choksi<sup>1</sup>, Jennifer M. Bocker<sup>1</sup>, Xi Kathy Zhou<sup>4</sup>, Rhonda K. Yantiss<sup>5</sup>, Duncan B. Hughes<sup>6</sup>, Baoheng Du<sup>7</sup>, Benjamin L. Judson<sup>1</sup>, Kotha Subbaramaiah<sup>7</sup>, and Andrew J. Dannenberg<sup>7</sup>

## Abstract

Use of tobacco is responsible for ~30% of all cancer-related deaths in the United States, including cancers of the upper aerodigestive tract. In the current study, 40 current and 40 age- and gender-matched never smokers underwent buccal biopsies to evaluate the effects of smoking on the transcriptome. Microarray analyses were carried out using Affymetrix HGU133 Plus 2 arrays. Smoking altered the expression of numerous genes: 32 genes showed increased expression and 9 genes showed reduced expression in the oral mucosa of smokers versus never smokers. Increases were found in genes involved in xenobiotic metabolism, oxidant stress, eicosanoid synthesis, nicotine signaling, and cell adhesion. Increased numbers of Langerhans cells were found in the oral mucosa of smokers. Interestingly, smoking caused greater induction of aldo-keto reductases, enzymes linked to polycyclic aromatic hydrocarbon-induced genotoxicity, in the oral mucosa of women than men. Striking similarities in expression changes were found in oral compared with the bronchial mucosa. The observed changes in gene expression were compared with known chemical signatures using the Connectivity Map database and suggested that geldanamycin, a heat shock protein 90 inhibitor, might be an antimimetic of tobacco smoke. Consistent with this prediction, geldanamycin caused dose-dependent suppression of tobacco smoke extract-mediated induction of CYP1A1 and CYP1B1 *in vitro*. Collectively, these results provide new insights into the carcinogenic effects of tobacco smoke, support the potential use of oral epithelium as a surrogate tissue in future lung cancer chemoprevention trials, and illustrate the potential of computational biology to identify chemopreventive agents. *Cancer Prev Res*; 3(3); 266–78. ©2010 AACR.

## Introduction

More than a billion people smoke cigarettes daily worldwide. Tobacco use is responsible for ~30% of all cancer-related deaths in the United States (1). Exposure to tobacco causes multiple human malignancies, including cancers of the lung, oral cavity, pharynx, esophagus, stomach, liver, pancreas, kidney, bladder, and cervix (2). More than 60 carcinogens are found in mainstream cigarette smoke and most of these are also found in sidestream

smoke (3). In addition to being a major cause of cancer, smoking alters the activity of chemopreventive agents (4, 5), stimulates the clearance of selected targeted anticancer therapies (6), reduces the efficacy of cancer treatment (7–10), and increases the risk of second primary tumors (11). Women have been suggested to be at increased risk of lung, oral, and oropharyngeal cancer compared with men who had similar cigarette smoking exposure levels (12–14). The mechanisms underlying this apparent gender-dependent difference in risk are poorly understood.

Numerous studies have been carried out to elucidate the carcinogenic effects of tobacco smoke on the bronchial epithelium. In histologically normal airway epithelial cells, smoking causes a range of abnormalities, including *P53* mutations (15), changes in promoter methylation (16, 17), and allelic loss (18). Transcriptome profiling showed that smoking induced the expression of genes involved in xenobiotic metabolism and redox stress in large airway epithelial cells (19). Importantly, a profile of bronchial airway gene expression in cytologically normal large airway epithelial cells was found to be potentially useful as a biomarker of lung cancer (20). In theory, the successful development of a transcriptome-based biomarker to identify high-risk smokers could provide the basis for risk reduction strategies, including chemoprevention. Although sampling

**Authors' Affiliations:** <sup>1</sup>Department of Surgery (Head and Neck Service), Memorial Sloan-Kettering Cancer Center; Departments of <sup>2</sup>Physiology and Biophysics, <sup>3</sup>Otorhinolaryngology, <sup>4</sup>Public Health, <sup>5</sup>Pathology and Laboratory Medicine, <sup>6</sup>Surgery, and <sup>7</sup>Medicine and <sup>8</sup>HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, New York

**Note:** Supplementary data for this article are available at Cancer Prevention Research Online (<http://cancerprevres.aacrjournals.org/>).

J.O. Boyle and Z.H. Gümüş contributed equally to this work.

**Corresponding Author:** Andrew J. Dannenberg, Department of Medicine and Weill Cornell Cancer Center, 525 East 68th Street, Room F-206, New York, NY 10065. Phone: 212-746-4403; Fax: 212-746-4885; E-mail: [ajdann@med.cornell.edu](mailto:ajdann@med.cornell.edu).

doi: 10.1158/1940-6207.CAPR-09-0192

©2010 American Association for Cancer Research.

the bronchial epithelium to identify potential biomarkers of cancer risk has yielded significant insights, it would be very useful if similar information could be obtained using less invasive tissue collection methods. Recently, Sridhar and colleagues (21) compared the effects of smoking on the transcriptome of extrathoracic (buccal and nasal) versus intrathoracic (bronchial) epithelium. The results of gene expression profiles from buccal ( $n = 10$ ) and nasal ( $n = 15$ ) epithelial cells indicated that many of the smoking-related changes in the bronchial epithelium were also present in buccal and nasal epithelium. Possibly, sampling of extrathoracic epithelial cells will yield information that can help to define individual susceptibility to smoking-related diseases of the upper aerodigestive tract, including the lung.

In the current study, 40 current smokers and 40 age- and gender-matched never smokers underwent buccal biopsies. We had four objectives: (a) to define the effects of smoking on the transcriptome of oral epithelial cells, (b) to determine if any of the effects of tobacco smoke on the transcriptome are gender dependent, (c) to compare the effects of tobacco smoke exposure on the transcriptome in oral versus bronchial epithelium, and (d) to identify agents with the potential to suppress the effects of tobacco smoke on the transcriptome. We show that smoking altered the expression of genes involved in xenobiotic metabolism, oxidant stress, eicosanoid synthesis, nicotine signaling, and cell adhesion. Smoking-mediated induction of aldo-keto reductases (AKR), enzymes linked to polycyclic aromatic hydrocarbon (PAH)-induced genotoxicity (22), was greater in women than in men. Most smoking-related changes in gene expression in oral epithelial cells also occur in airway epithelial cells. Collectively, these data provide new insights into the carcinogenic effects of tobacco smoke and offer insights that may prove useful in developing preventive strategies.

## Materials and Methods

### Materials

Keratinocyte basal and growth media were obtained from Lonza. Antibody to  $\beta$ -actin and Lowry protein assay kits were obtained from Sigma Chemical. Antiserum to CYP1B1 was a gift of Dr. Craig B. Marcus (Oregon State University, Corvallis, OR). Antibody to CYP1A1 was obtained from Santa Cruz Biotechnology. CD1a mouse monoclonal antibody (clone MTB1) was from Novocastra Laboratories Ltd. Western blot analysis detection reagents (enhanced chemiluminescence) were from Amersham Biosciences. Nitrocellulose membranes were from Schleicher and Schuell. Geldanamycin was purchased from Calbiochem. Murine leukemia virus reverse transcriptase, oligo(dT)<sub>16</sub>, and RNase inhibitor were from Roche Applied Science, and Taq polymerase was from Applied Biosystems. HGU133 Plus 2 microarrays were from Affymetrix.

### Study design

Forty never smokers (<100 cigarettes per lifetime) and 40 active smokers ( $\geq 15$  pack-year exposure) were recruited

(see Supplementary Table S1). Subjects were age and gender matched. Eligible subjects were healthy volunteers recruited from the community and hospital. Subjects were excluded if they had gross evidence of oral inflammation, a history of heavy alcohol consumption, or recent use of nonsteroidal anti-inflammatory drugs or other anti-inflammatory medications. The study was approved by the Weill Cornell Medical College Institutional Review Board and the Clinical and Translational Science Center. All subjects provided written informed consent for participation.

### Human tissue

After topical anesthesia, 5-mm punch biopsies were obtained from grossly normal-appearing buccal mucosa. Tissue samples were immediately divided into two parts. Approximately, two thirds of each specimen was snap frozen in liquid nitrogen. Total RNA was then isolated with an RNeasy Mini kit (Qiagen, Inc.) and stored at  $-80^{\circ}\text{C}$  until analysis. The remaining one third of the biopsy was formalin fixed for immunohistochemical analysis.

### Microarray procedures

Biotinylated cRNA was prepared according to the standard Affymetrix protocol from 2.5  $\mu\text{g}$  of total RNA.<sup>9</sup> Following fragmentation, 10  $\mu\text{g}$  of cRNA were hybridized for 16 h at  $45^{\circ}\text{C}$  on GeneChip HGU133 Plus 2 arrays. GeneChips were washed and stained in the Affymetrix Fluidics Station 450 and scanned using the Affymetrix GeneChip Scanner 3000 7G.

### Microarray data analysis

The scanned image of each array was checked for significant artifacts. One sample was excluded from the study based on this quality measure, leaving 79 arrays for analysis.

**Preprocessing.** Raw image data were background corrected, normalized, and summarized into probe set expression values using the Robust Multichip Average algorithm (23, 24) within GeneSpring 7.2 software (Agilent Technologies). Data from each chip were normalized for interarray comparisons by first setting measurements of  $<0.01$  to 0.01 and then normalizing to 50% of the measurements taken from that array. Probe sets that were not reliably detected were filtered out. From the complete set of  $\sim 54,675$  probe sets on the HGU133 Plus 2 array, genes were filtered for minimum raw expression level of 50 in at least 16 of 79 conditions. Genes with low confidence were filtered out based on  $t$  test  $P$  value of  $<0.05$  in at least one of two conditions (smoker or never smoker). The cross-gene error model was active. The  $\sim 24,103$  probe sets that passed these tests were defined as expressed and were statistically analyzed.

**Statistical analysis.** To identify differentially expressed gene groups between smoker and never smoker groups, one-way ANOVA was done using parametric test, variances

<sup>9</sup> <http://www.affymetrix.com>

not assumed equal (Welch *t* test) with *P* value cutoff of 0.05, and Benjamini-Hochberg multiple testing correction to maintain false discovery rate (FDR) at 5%. Genes with normalized smoker versus never smoker expression values that changed by a factor of 1.5 were deemed significant and listed in Table 1.

**Clustering.** An unsupervised hierarchical clustering analysis across all samples of the microarray data was done for the probe sets found to be differentially expressed in the oral mucosa of smokers and never smokers (using log-transformed, normalized, gene median-centered data). Pearson correlation (uncentered) similarity metric and average linkage clustering was done with CLUSTER and TREEVIEW software obtained online<sup>10</sup> and shown in Fig. 1A.

**Functional analysis.** The effects of tobacco smoke were examined in the context of detailed molecular interaction networks using Ingenuity Pathway Analysis (IPA), a web-delivered application used to discover, visualize, and explore relevant networks.<sup>11</sup> Affymetrix probe identifiers and fold values were uploaded to IPA, and each identifier was mapped to its corresponding gene object in the IPA Knowledgebase. Interactions were then queried between these gene objects and all other gene objects stored within IPA to generate a set of direct interaction networks that were merged. Putative transcription regulator hubs that directly interact with a minimum of three differentially expressed genes were included in the network. Because UDP glucuronosyltransferases (UGT) and AKR1C probes map to multiple genes in these families, all members of these families were included to identify their individual interconnections. The regulations of ALDH3A1, UGT1A1, UGT1A3, and UGT1A4 by the aryl hydrocarbon receptor (AHR) and AKR1Cs by Nrf2 were manually added (25–28) to Fig. 1B.

**Significantly altered groups.** Significantly differentially expressed genes between smokers and never smokers were mined for statistically overrepresented gene groups using EASE software (29). Functional gene groups in Gene Ontology (GO)<sup>12</sup> database were queried, and the likelihood of overrepresentation of each gene group in the differentially expressed gene set with respect to the HGU133 Plus 2 microarray was scored (using Affymetrix identifiers). Relevant gene sets with Bonferroni *P* < 0.05 are reported in Table 2.

**Effects of gender on the transcriptome of smokers.** The approach that was used to carry out this analysis is detailed in Supplementary Materials and Methods.

**Modest and consistent alterations.** The entire 54,675 microarray probe sets from each of the 79 subjects were mined for statistically significant, concordant functional gene group differences between smokers and never smokers using Gene Set Enrichment Analysis (GSEA) version

2 (30). GSEA helps functionally interpret modest but consistent changes in the gene expression data and focuses on groups of genes that share common biological function. Normalized ratio expression values were analyzed using default parameter settings. Relevant gene sets with FDR of <0.25 were deemed significant.

**Comparison of effects of smoking on the oral and bronchial epithelium.** Smoking-related changes in the transcriptome of the oral and airway epithelium were compared using the current data as well as previously reported smoker and never smoker airway transcriptome data (analyzed as described in the Statistical analysis section; ref. 19). Overlapping genes are listed in Table 4. The relationship between the gene expression patterns in response to tobacco smoke in the oral and bronchial epithelium was identified by Enrichment Analysis as described in Supplementary Materials and Methods.

**Gene expression signature–based chemical genomic prediction.** Differentially expressed genes were separated into up-regulated and downregulated gene sets and converted to their HGU133A identifiers,<sup>13</sup> which were queried to identify drugs with antimimetic gene expression signatures within the Connectivity Map<sup>14</sup> (31).

#### Additional information

The complete results from the gender and GSEA analyses are available through an interactive Web site<sup>15</sup> established as a resource of the Institute for Computational Biomedicine. The microarray data have been deposited at the National Center for Biotechnology Information Gene Expression Omnibus<sup>16</sup> under Gene Expression Omnibus Series accession no. GSE17913.

#### Quantitative PCR validation

Samples from 10 never smokers and 10 smokers were chosen at random. Total RNA was isolated using RNeasy Mini kit. RNA (1 µg) was reverse transcribed using murine leukemia virus reverse transcriptase and oligo(dT)<sub>16</sub> primer. The resulting cDNA was then used for amplification. Each PCR was 20 µL and contained 5 µL cDNA, 2× SYBR Green PCR master mix, and forward and reverse primers (see Supplementary Table S2 for list of primers). Experiments were done using a 7500 real-time PCR system (Applied Biosystems). β-Actin served as an endogenous normalization control. Relative fold induction was determined by ΔΔC<sub>T</sub> (relative quantification) analysis.

#### Immunohistochemistry

Formalin-fixed, paraffin-embedded oral mucosal tissue sections from 54 subjects (27 smokers and 27 never smokers) were evaluated for the presence and distribution of Langerhans cells using antiserum directed against CD1a, a Langerhans cell marker. Four-micrometer-thick

<sup>10</sup> <http://rana.lbl.gov/EisenSoftware.htm>

<sup>11</sup> <http://www.ingenuity.com>

<sup>12</sup> <http://www.geneontology.org>

<sup>13</sup> <http://www.affymetrix.com/analysis/netaffx/index.affx>

<sup>14</sup> <http://www.broad.mit.edu/cmapp>

<sup>15</sup> <http://physiology.med.cornell.edu/go/smoke>

<sup>16</sup> <http://www.ncbi.nlm.nih.gov/geo/>

tissue sections were immunohistochemically stained with the CD1a mouse monoclonal antibody as described below. Unstained tissue sections were baked, deparaffinized, and rehydrated on the Vision Biosystems/Leica BondMax autostainer. Tissue sections were pretreated using the heat-induced epitope retrieval solution-1 (Vision Biosystems/Leica) and incubated with the primary antibody (1:20 dilution) for 25 min. The Refine Detection kit supplied by the manufacturer was used to block endogenous peroxidase activity and enhance the staining reaction. Positive (skin) and negative (replacement of the primary antibody with immunoglobulin) controls were included in the experiment. Cells that displayed moderate to strong cytoplasmic staining for CD1a in dendritic-type cellular processes were separately evaluated in three regions of the mucosa: the peripapillary, interpapillary, and superficial epithelium. The total and mean number of CD1a-positive cells present in the peripapillary mucosa of four well-oriented papillae, and four high-magnification (400× objective) fields of the interpapillary and superficial mucosa, were recorded for each of the 54 cases. Comparisons between smokers and never smokers were made by Student's *t* test. A difference between groups of  $P < 0.05$  was considered significant.

#### Tissue culture

The MSK-Leuk1 cell line was established from a dysplastic leukoplakia lesion adjacent to a squamous cell carcinoma of the tongue (32). Cells were routinely maintained in keratinocyte growth medium supplemented with bovine pituitary extract. Cells were grown in basal medium for 24 h before treatment.

#### Preparation of tobacco smoke extract

Cigarettes (2R4F, Kentucky Tobacco Research Institute) were smoked in a Borgwaldt piston-controlled apparatus (model RG-1) using a Federal Trade Commission standard protocol. Cigarettes were smoked one at a time in the apparatus and the smoke was drawn under sterile conditions into premeasured amounts of sterile PBS (pH 7.4). This smoke in PBS represents whole trapped mainstream smoke (TS). Quantitation of smoke content is expressed in puffs/mL of PBS, with one cigarette yielding about 8 puffs drawn into a 5 mL volume. The final concentration of TS in the cell culture medium is expressed as puffs/mL medium. All treatments were carried out with 0.03 puffs/mL TS because this concentration was previously found to induce CYP1A1 and CYP1B1 (33).

#### Western blot analysis

Cell lysates were prepared by treating cells with lysis buffer [150 mmol/L NaCl, 100 mmol/L Tris (pH 8.0), 1% Tween 20, 50 mmol/L diethyldithiocarbamate, 1 mmol/L phenylmethylsulfonyl fluoride, 10 μg/mL aprotinin, 10 μg/mL trypsin inhibitor, and 10 μg/mL leupeptin]. Lysates were sonicated for  $3 \times 10$  s on ice and centrifuged at  $14,000 \times g$  for 10 min at 4°C to sediment the particulate

material. The protein concentration of the supernatant was measured by the method of Lowry et al. (34). SDS-PAGE was done under reducing conditions on 10% polyacrylamide gels. The resolved proteins were transferred onto nitrocellulose sheets and then incubated with antisera to CYP1A1, CYP1B1, and β-actin. Secondary antibody to immunoglobulin G conjugated to horseradish peroxidase was used. The blots were then reacted with the enhanced chemiluminescence Western blot detection system, according to the manufacturer's instructions.

## Results

### Smoking status is a determinant of the transcriptome in the oral mucosa

A total of 80 subjects (40 smokers and 40 never smokers) underwent biopsies of the buccal mucosa. One female smoker was excluded from the study because of problems processing the biopsy sample. Hence, samples from 79 subjects were available for analysis. Demographic data for these 79 subjects are presented in Supplementary Table S1. The never smoker group included 20 males (median age, 45 years) and 20 females (median age, 45 years). The smoker group included 20 males (median age, 45.5 years; median pack-years, 32.5) and 19 females (median age, 43 years; median pack-years, 25). mRNA from 79 subjects (40 never smokers and 39 smokers) was suitable in quantity and quality for microarray analysis. The gene probes that were differentially expressed at least 1.5-fold between smokers and never smokers are listed in Table 1. Smoking altered the expression of numerous genes. Forty probes representing 32 genes showed increased expression and 14 probes representing 9 genes showed reduced expression in the oral mucosa of smokers versus never smokers. Increases were found for genes involved in xenobiotic metabolism (*CYP1A1*, *CYP1B1*, *AKR1C1/AKR1C2*, *UGT1A*, *NQO1*, and *AHRR*), oxidant stress (*ALDH3A1* and *GPX2*), eicosanoid synthesis (*PTGES*, *ALOX12B*, and *ALOX15B*), nicotine signaling (*CHRNA3*), and cell adhesion (*CEACAM7*). Decreased expression was detected for genes including *CCL18*, *SOX9*, *IGF2BP3*, and *LEPR*. Subsequently, quantitative PCR was used to validate the microarray findings for a subset of 11 differentially expressed genes. Importantly, the observed changes in expression were quantitatively consistent with the microarray results for all 11 genes evaluated (Supplementary Table S3). Figure 1A shows the unsupervised hierarchical clustering analysis of smokers versus never smokers based on genes that were differentially expressed in the two groups. The majority of subjects clustered accurately into the two groups.

### Interpreting the global transcriptome changes in terms of biological pathways and functions

Several databases and tools were used to classify the differentially expressed genes into relevant molecular and physiologic categories. Interactions within IPA Knowledgebase<sup>11</sup> and other known literature (25–28) were used to define potential smoking-induced effects

**Table 1.** Differentially expressed genes in the oral mucosa of smokers versus never smokers with corresponding fold changes and *P* values

Gene name	Affymetrix ID	Fold	<i>P</i>	Gene title
<i>S100A7</i>	205916_at	4.4	3.1E-02	S100 calcium binding protein A7
<i>RPTN</i>	1553454_at	4.3	1.7E-02	Repetin
<i>CYP1B1</i>	202437_s_at	4.2	1.4E-11	Cytochrome <i>P450</i> , family 1, subfamily B, polypeptide 1
<i>CYP1B1</i>	202436_s_at	3.2	6.9E-10	Cytochrome <i>P450</i> , family 1, subfamily B, polypeptide 1
<i>LOR</i>	207720_at	3.2	9.4E-03	Loricrin
<i>CEACAM7</i>	206198_s_at	3.0	3.1E-02	Carcinoembryonic antigen-related cell adhesion molecule 7
<i>CYP1A1</i>	205749_at	2.5	1.1E-07	Cytochrome <i>P450</i> , family 1, subfamily A, polypeptide 1
<i>CYP1B1</i>	202435_s_at	2.5	2.7E-08	Cytochrome <i>P450</i> , family 1, subfamily B, polypeptide 1
<i>HTR3A</i>	216615_s_at	2.2	7.7E-03	5-Hydroxytryptamine (serotonin) receptor 3A
<i>GPX2</i>	202831_at	2.1	3.1E-04	Glutathione peroxidase 2 (gastrointestinal)
<i>FCGBP</i>	203240_at	2.0	2.9E-05	Fc fragment of IgG binding protein
—	227452_at	2.0	7.8E-10	Full-length cDNA clone CS0DD005YM12 of neuroblastoma Cot 50-normalized of <i>Homo sapiens</i> (human)
<i>CCL26</i>	223710_at	1.9	2.8E-02	Chemokine (C-C motif) ligand 26
<i>PNLIPRP3</i>	1558846_at	1.9	6.2E-03	Pancreatic lipase-related protein 3
<i>ALOX12B</i>	207381_at	1.9	1.2E-02	Arachidonate 12-lipoxygenase, 12R type
<i>LOC388610</i>	227862_at	1.9	2.9E-02	Hypothetical LOC388610
<i>CD207</i>	220428_at	1.8	1.3E-05	CD207 molecule, langerin
<i>CHRNA3</i>	210221_at	1.7	2.0E-04	Cholinergic receptor, nicotinic, $\alpha$ 3
<i>CYTL1</i>	219837_s_at	1.7	1.6E-07	Cytokine-like 1
<i>NQO1</i>	201468_s_at	1.7	6.7E-04	NAD(P)H dehydrogenase, quinone 1
<i>NQO1</i>	210519_s_at	1.6	2.1E-03	NAD(P)H dehydrogenase, quinone 1
<i>NQO1</i>	201467_s_at	1.5	4.9E-03	NAD(P)H dehydrogenase, quinone 1
<i>CLEC7A</i>	1555756_a_at	1.6	1.6E-02	C-type lectin domain family 7, member A
<i>CLEC7A</i>	221698_s_at	1.6	9.4E-04	C-type lectin domain family 7, member A
<i>LOC344887</i>	241418_at	1.6	6.2E-03	Similar to hCG2041270
<i>PTGES</i>	210367_s_at	1.6	4.9E-02	Prostaglandin E synthase
<i>KRT10</i>	207023_x_at	1.6	3.5E-02	Keratin 10 (epidermolytic hyperkeratosis; keratosis palmaris et plantaris)
<i>C10orf99</i>	227736_at	1.6	9.4E-03	Chromosome 10 open reading frame 99
<i>C10orf99</i>	227735_s_at	1.6	1.2E-02	Chromosome 10 open reading frame 99
<i>ALDH3A1</i>	205623_at	1.6	2.9E-05	Aldehyde dehydrogenase 3 family, memberA1
<i>ALOX15B</i>	206714_at	1.6	3.5E-02	Arachidonate 15-lipoxygenase, type B
<i>UGT1A6</i> /// <i>UGT1A8</i> /// <i>UGT1A9</i>	221305_s_at	1.5	1.1E-02	UDP glucuronosyltransferase 1 family, polypeptide A6 /// UDP glucuronosyltransferase 1 family, polypeptide A8 /// UDP glucuronosyltransferase 1 family, polypeptide A9
<i>MUC1</i>	213693_s_at	1.5	4.7E-02	Mucin 1, cell surface associated
<i>AKR1C1</i> /// <i>AKR1C2</i>	1555854_at	1.5	1.9E-02	Aldo-keto reductase family 1, member C1 /// aldo-keto reductase family 1, member C2

(Continued on the following page)

**Table 1.** Differentially expressed genes in the oral mucosa of smokers versus never smokers with corresponding fold changes and *P* values (Cont'd)

Gene name	Affymetrix ID	Fold	<i>P</i>	Gene title
AHRR /// PDCD6	229354_at	1.5	3.8E-08	Aryl-hydrocarbon receptor repressor /// programmed cell death 6
LYPD5	236039_at	1.5	4.7E-02	LY6/PLAUR domain containing 5
UGT1A1 /// UGT1A3 to UGT1A10	215125_s_at	1.5	2.3E-03	UDP glucuronosyltransferase 1 family, polypeptide /// UDP glucuronosyltransferase 1 family, polypeptide A3 to A10
CCL5	1405_i_at	1.5	1.7E-02	Chemokine (C-C motif) ligand 5
UGT1A1 /// UGT1A4 /// UGT1A6 /// UGT1A8 to UGT1A10	207126_x_at	1.5	3.1E-04	UDP glucuronosyltransferase 1 family, polypeptide A1 /// UDP glucuronosyltransferase 1 family, polypeptide A4 /// UDP glucuronosyltransferase 1 family, polypeptide A6 /// UDP glucuronosyltransferase 1 family, polypeptide A8 to A10
CD1a	210325_at	1.5	9.6E-03	CD1a molecule
LYVE1	220037_s_at	-1.5	1.0E-02	Lymphatic vessel endothelial hyaluronan receptor 1
YOD1	215150_at	-1.5	3.0E-02	YOD1 OTU deubiquinating enzyme 1 homologue ( <i>S. cerevisiae</i> )
CCL18	209924_at	-1.5	3.2E-02	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)
ANKRD37	227337_at	-1.5	3.7E-02	Ankyrin repeat domain 37
SOX9	202936_s_at	-1.5	2.4E-03	SRY (sex determining region Y)-box 9
SOX9	202935_s_at	-1.5	1.1E-02	SRY (sex determining region Y)-box 9
LEPR	211355_x_at	-1.5	4.4E-03	Leptin receptor
LEPR	211354_s_at	-1.6	4.3E-03	Leptin receptor
LEPR	211356_x_at	-1.6	2.4E-03	Leptin receptor
IGF2BP3	203819_s_at	-1.7	5.4E-04	Insulin-like growth factor 2 mRNA binding protein 3
IGF2BP3	203820_s_at	-1.6	1.4E-02	Insulin-like growth factor 2 mRNA binding protein 3
CCL18	32128_at	-1.6	3.4E-02	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)
HIG2	1554452_a_at	-1.9	3.8E-02	Hypoxia-inducible protein 2
PEG3	209242_at	-2.1	1.1E-02	Paternally expressed 3

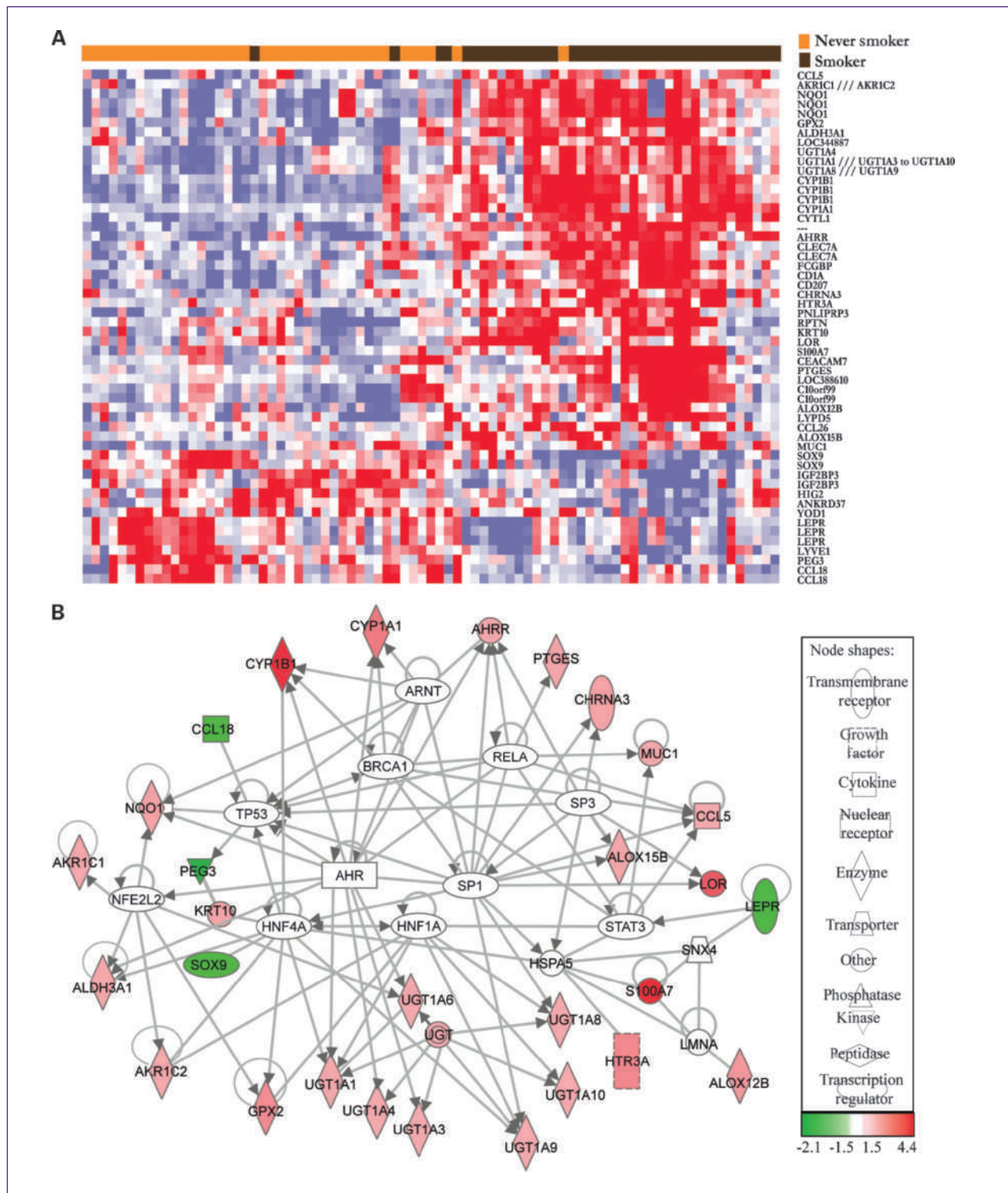
NOTE: Detailed annotations are provided at <http://physiology.med.cornell.edu/go/smoke>.

on molecular interaction networks (Fig. 1B). The likely role of the AHR, a PAH-activated transcription factor, was evident because increased levels of CYP1A1, CYP1B1, and AHR repressor (AHRR) mRNAs were found in the oral mucosa of smokers. PAH-activated AHR stimulates the transcription of each of these genes (35). NFE2L2 (Nrf2), a transcription factor activated by oxidative stress, can induce AKR1C1/2, NQO1, GPX2, and ALD3A1 (36–38). Each of these genes was overexpressed in the oral mucosa of smokers, strongly suggesting the involvement of Nrf2 (Table 1; Fig. 1B). IPA

network analysis also suggested the involvement of other regulators of transcription, including ARNT, RELA, and SP1. The genes were further classified in terms of relevant functional categories to identify additional effects of tobacco smoke. Pathways within the KEGG<sup>17</sup> and GenMAPP<sup>18</sup> databases were queried using GSEA version 2. The following pathways were enriched in smokers:

<sup>17</sup> <http://www.genome.jp/kegg>

<sup>18</sup> <http://www.genmapp.org>



**Fig. 1.** A, unsupervised hierarchical clustering of the expression of probe sets differentially expressed in the oral mucosa of smokers versus never smokers. Smokers and never smokers cluster primarily into two distinct groups. Each column corresponds to the expression profile of an oral mucosal biopsy, and each row corresponds to an mRNA. The color in each cell reflects the level of expression of the corresponding mRNA relative to its mean level of expression in the entire set of biopsy samples. In this heat map, the increasing intensities of red signify that a specific mRNA has a higher expression in the given sample, whereas the increasing intensities of blue mean that this mRNA has lower expression. White indicates mean level of expression. B, direct interaction network of differentially expressed genes generated using IPA and other known interactions. The white nodes represent genes with no significant expression change that potentially contribute to the effects of smoking.

metabolism of xenobiotics by cytochrome *P450*, androgen and estrogen metabolism, eicosanoid synthesis, prostaglandin and leukotriene metabolism, and glutathione metabolism (Table 2). Query of GO<sup>12</sup> functional databases using EASE suggested that electron transporter activity and oxidoreductase activity were increased in the oral mucosa of smokers (Table 2).

### Increased numbers of langerhans cells are found in the oral mucosa of smokers

Generally, changes in the transcriptome reflect altered gene expression. We note, however, that changes in the cellular composition of a biopsy can also affect the transcriptome. Increased levels of both CD207 (langerin) and CD1a mRNAs, transcripts that are abundant in Langerhans cells, were found in the oral mucosa of smokers (Table 1; Supplementary Table S3). This suggested the possibility that the number of Langerhans cells might be increased in the oral mucosa of smokers. Because CD1a is a marker for Langerhans cells, immunohistochemistry was carried out to quantify the number of CD1a-positive cells in the oral mucosa of smokers versus never smokers. A significant increase in the number of CD1a-positive cells was found in the oral mucosa of smokers versus never smokers (Fig. 2A-E).

### Gender-dependent differences in smoking-mediated changes in gene expression

Previously, a somewhat higher risk of cancers of the lung, oral cavity, and oropharynx was found in women than men at comparable pack-years of smoking (12–14). It was of interest, therefore, to determine if levels of gene

expression differed in the oral mucosa of males versus females. In never smokers, the genes that were differentially expressed in males versus females primarily reflected gender-dependent differences in genes of X and Y chromosomes (Supplementary Table S4). The effects of smoking were also evaluated. Interestingly, smoking had a greater effect on both the induction (AKR1C2/3, UGT family members) and suppression (IGFL1) of several genes in women than in men (Table 3).

### Comparison of oral mucosa and airway epithelial transcriptome of smokers versus never smokers

Smoking modulates gene expression in the airway epithelium. Hence, we compared our findings for oral mucosa with previously reported data for airway epithelium (19). Striking similarities in expression changes were found in the oral mucosa and bronchial epithelial cells of smokers (Table 4). For example, smoking was associated with increased expression of a variety of genes (*CYP1A1*, *CYP1B1*, *NQO1*, *ALDH3A1*, and *UGTs*) involved in xenobiotic metabolism. Increased levels of *GPX2* and *CEACAM* family members were found in both the oral and bronchial epithelium of smokers. Interestingly, smoking was associated with increased levels of *FCGBP* in oral mucosa but reduced expression in bronchial mucosa. GSEA also suggested that the inductive effects of smoking are similar in both the oral and bronchial epithelium (Supplementary Table S5).

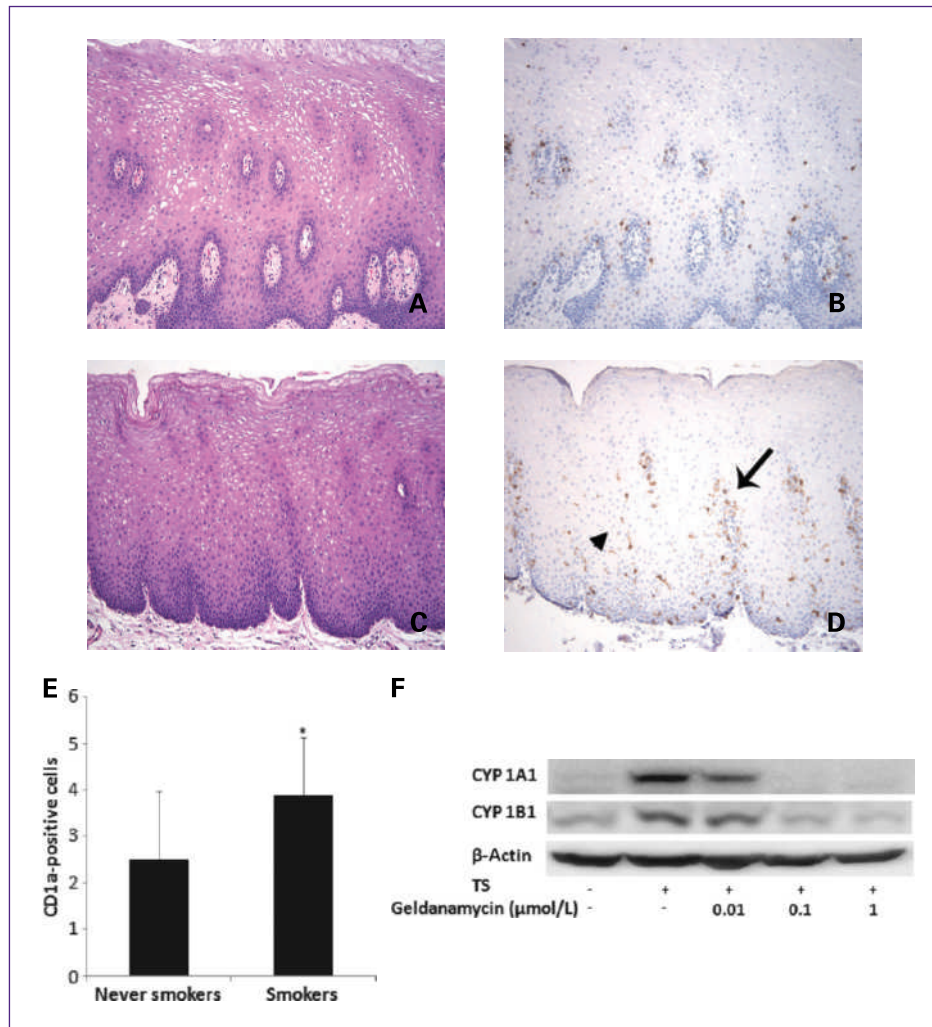
### Targeting heat shock protein 90 can attenuate the activation of AHR-Dependent gene expression

Agents that suppress tobacco smoke-mediated effects on the transcriptome are likely to possess chemopreventive properties. Accordingly, we next attempted to identify a small molecule with the potential to attenuate some of the changes in the transcriptome found in the oral mucosa of smokers. To achieve this goal, a computational approach was used in combination with an *in vitro* model that has been used in previous tobacco studies (33). The mRNA profile that was observed in the oral mucosa of smokers versus never smokers was compared with known signatures of pharmaceutical and small-molecule treatments using the Connectivity Map database (31). This computational analysis suggested that geldanamycin, an inhibitor of heat shock protein 90 (Hsp90), might be an antimimetic of tobacco smoke ( $P = 0.0003$ ). As detailed above, the AHR, a client protein of Hsp90, mediates the induction of *CYP1A1* and *CYP1B1* transcription in response to PAHs (39). *CYP1A1* and *CYP1B1* were among the genes most overexpressed in the oral mucosa of smokers (Table 1). Given this background, we determined whether geldanamycin suppressed the induction of *CYP1A1* and *CYP1B1* by TS in MSK-Leuk1 cells, a cellular model of oral leukoplakia (32). Consistent with the findings in the computational analysis, geldanamycin caused dose-dependent suppression of TS-mediated induction of both *CYP1A1* and *CYP1B1* (Fig. 2F).

**Table 2.** Functional gene groups altered in the oral mucosa of smokers versus never smokers

Gene set	
<b>Pathways enriched in smokers (using GSEA version 2)</b>	<b>FDR</b>
Metabolism of xenobiotics by cytochrome <i>P450</i>	0.01↑ (KEGG)
Androgen and estrogen metabolism	0.110↑ (KEGG)
Eicosanoid synthesis	0.075↑ (GenMAPP)
Prostaglandin and leukotriene metabolism	0.128↑ (GenMAPP)
Glutathione metabolism	0.166↑ (KEGG), 0.093↑ (GenMAPP)
<b>GO groups enriched in smokers (using EASE)</b>	<b>Bonferroni <i>P</i></b>
GO molecular function: electron transporter activity	0.00085 (8/27)↑
GO molecular function: oxidoreductase activity	0.049 (8/27)↑





**Fig. 2.** Increased numbers of Langerhans cells were found in the oral mucosa of smokers. Nonneoplastic oral mucosae from never smokers (A) and smokers (C) were morphologically similar, but samples from never smokers showed relatively few Langerhans cells (B) compared with those from smokers (D), which contained numerous Langerhans cells in the peripapillary (arrow) and interpapillary mucosa (arrowhead). Magnification,  $\times 100$  (A-D). A and C, stained with H&E. B and D, stained with CD1a immunostain and hematoxylin. E, intraepithelial cells that displayed moderate to strong cytoplasmic staining for CD1a in dendritic-type cellular processes were quantified in the peripapillary, interpapillary, and superficial epithelium. A statistically significant increase in the number of CD1a-positive cells was found in all three regions in smokers compared with never smokers ( $P < 0.001$ ,  $P < 0.001$ , and  $P = 0.032$  for peripapillary, interpapillary, and superficial areas, respectively). Total number of CD1a-positive cells in the three regions. Columns, means ( $n = 27$  per group); bars, SE. \*,  $P < 0.001$ . F, MSK-Leuk1 cells were pretreated with vehicle or the indicated concentration of geldanamycin for 2 h. Subsequently, cells received vehicle or TS for 5 h and were then harvested for Western blot analysis. Cellular lysate protein (100  $\mu\text{g}/\text{lane}$ ) was loaded onto a 10% SDS-polyacrylamide gel, electrophoresed, and subsequently transferred onto nitrocellulose. Immunoblots were probed with antibodies specific for CYP1A1, CYP1B1, and  $\beta$ -actin.

## Discussion

This study provides new insights into the mechanisms underlying the carcinogenic effects of tobacco smoke. Multiple genes encoding enzymes (*CYP1A1*, *CYP1B1*, *AKRs*, *ALDH3A1*, *NQO1*, and *UGTs*) involved in carcinogen metabolism were overexpressed in the oral mucosa of smokers. PAHs, an important class of tobacco carcinogen, are likely to mediate some of these expression changes. The AHR, a ligand-activated transcription factor, binds with high affinity to PAHs. Following ligand binding, the AHR translocates to the nucleus where it forms a heterodi-

mer with ARNT. The AHR-ARNT heterodimer then binds to xenobiotic-responsive elements in the upstream regulatory region of target genes, resulting in the transcriptional activation of a network of genes, including *CYP1A1* and *CYP1B1* (33). The activation of AHR-mediated signaling leading to induction of xenobiotic metabolism provides a first line of defense against environmental carcinogens. However, the induction of xenobiotic-metabolizing enzymes by ligand-activated AHR may also contribute to mutagenesis. PAHs are generally biologically inert and must be metabolically activated by inducible enzymes, including *CYP1A1* and *CYP1B1*, to exert their genotoxic

actions. For example, benzo[a]pyrene (B[a]P), a potent ligand of the AHR, induces its own metabolism to noncarcinogenic B[a]P phenols (40) and a toxic metabolite anti-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene, which covalently binds to DNA, forming bulky DNA adducts that induce mutations (41). In addition to CYP1A1 and CYP1B1, PAHs induce the AHRR (35). Notably, levels of AHRR mRNA were increased in the oral mucosa of smokers. The AHR and AHRR constitute a negative feedback loop of xenobiotic signal transduction. The liganded AHR induces *AHRR* transcription, whereas expressed AHRR, in turn, inhibits the function of AHR (35).

A second pathway of PAH activation that causes mutations involves members of the AKR superfamily. PAH *trans*-dihydrodiols are oxidized by the AKRs to redox-active and electrophilic PAH *o*-quinones. The AKR-generated B[a]P-7,8-dione enters into futile redox cycles, which amplifies the formation of reactive oxygen species, resulting in oxidative DNA damage (22). Oxidative stress is caused by the presence of heavy metals and benzoquinone in tobacco smoke and AKR-mediated production of PAH *o*-quinones. Nrf2, a transcription factor that binds to antioxidant response elements in gene promoters, induces the expression of AKRs, NQO1, and ALDH3A1 (36–38). Expression of AKRs, NQO1, and ALDH3A1 was increased in the oral mucosa of smokers, suggesting a counterresponse to oxidative stress. Induction of these genes may protect against the damaging effects of harmful quinones and lipid peroxidation breakdown products. Individuals who fail to mount a normal counterresponse may be at increased risk of developing cancer. Thus, it seems that AKRs can both stimulate bioactivation of PAHs, leading to increased mutagenesis, and participate in a counterresponse to oxidative stress.

Increased levels of PTGES (prostaglandin E synthase), ALOX12B (arachidonate 12-lipoxygenase, 12R type), and ALOX15B (arachidonate 15-lipoxygenase, type B) were found in the oral mucosa of smokers. Each of these enzymes is involved in eicosanoid synthesis. These findings are potentially significant because eicosanoids, including prostaglandins, have been implicated in the development of multiple epithelial malignancies, including cancers of the upper aerodigestive tract (42). Notably, use of aspirin, a prototypic inhibitor of prostaglandin synthesis, has been associated with a reduced risk of oral cancer in smokers (43). Based on these findings, future studies are warranted to determine whether levels of eicosanoids, including prostaglandin E<sub>2</sub>, are increased in the oral mucosa of smokers.

Levels of CD1a and CD207 mRNAs, transcripts expressed in Langerhans cells, were increased in the oral mucosa of smokers compared with never smokers. Changes in transcript levels may occur because of either altered gene expression or a difference in cellular composition. Immunohistochemistry was carried out and revealed an increased number of Langerhans cells in the oral mucosa of smokers. This finding is consistent with previous reports (44) and may reflect a smoking-related change in mucosal immune function. PAH-mediated induction of prostaglandin E<sub>2</sub> has been suggested to stimulate the accumulation of Langerhans cells in skin (45). Lipoxygenase products (e.g., 12-HETE) have been reported to be chemotactic for Langerhans cells (46). It is reasonable to speculate, therefore, that the increased expression of enzymes involved in arachidonic acid metabolism may be causally linked to the increased number of Langerhans cells in the oral mucosa of smokers. Possibly, smoking cessation

**Table 3.** Gender-dependent differences in the effect of smoking on the expression of select genes in the oral mucosa

Gene name	Affymetrix ID	Female fold	Male fold	Interaction <i>P</i>	Gene title
<i>AKR1C3</i>	209160_at	1.6	1.1	0.0196	Aldo-keto reductase family 1, member C3 (3 $\alpha$ -hydroxysteroid dehydrogenase, type II)
<i>AKR1C2</i>	209699_x_at	1.5	1.1	0.0266	Aldo-keto reductase family 1, member C2
<i>UGT1A1 // UGT1A3</i> to <i>UGT1A10</i>	206094_x_at	1.7	1.3	0.0313	UDP glucuronosyltransferase 1 family, polypeptide A1, A3 to A10
<i>UGT1A1 // UGT1A4 // UGT1A6 // UGT1A8</i> to <i>UGT1A10</i>	204532_x_at	1.8	1.3	0.0346	UDP glucuronosyltransferase 1 family, polypeptide A1, A4, A6, A8 to A10
<i>UGT1A1 // UGT1A4 // UGT1A6 // UGT1A8</i> to <i>UGT1A10</i>	207126_x_at	1.8	1.3	0.0423	UDP glucuronosyltransferase 1 family, polypeptide A1, A4, A6, A8 to A10
<i>IGFL1</i>	239430_at	-2.1	-1.0	0.0210	Insulin-like growth factor-like family member 1

NOTE: Differentially expressed genes (fold changes) in the oral mucosa of smokers versus never smokers for females and males, respectively. Interaction *P* values indicate that the magnitude of the change in gene expression induced by smoking was greater in females than males for each of the genes shown below (annotations from November 2008 NetAffx).

**Table 4.** Genes that are overexpressed in the oral mucosa of smokers are also commonly overexpressed in the airways of smokers**Oral versus airway**

Gene name	Affymetrix ID	Oral mucosa fold (P)	Airway fold (P)	Gene title
<i>CYP1B1</i>	202437_s_at	4.2 (1.1E-11)	8.1 (1.2E-07)	Cytochrome <i>P450</i> , family 1, subfamily B, polypeptide 1
<i>CYP1B1</i>	202436_s_at	3.2 (5.5E-10)	7.1 (2.4E-07)	Cytochrome <i>P450</i> , family 1, subfamily B, polypeptide 1
<i>CYP1A1</i>	205749_at	2.5 (3.1E-05)	2.8 (5.6E-04)	Cytochrome <i>P450</i> , family 1, subfamily A, polypeptide 1
<i>GPX2</i>	202831_at	2.1 (3.1E-04)	3.3 (5.8E-14)	Glutathione peroxidase 2 (gastrointestinal)
<i>NQO1</i>	201468_s_at	1.7 (6.9E-04)	3.7 (4.1E-13)	NAD(P)H dehydrogenase, quinone 1
<i>NQO1</i>	210519_s_at	1.6 (1.9E-03)	3.6 (4.5E-14)	NAD(P)H dehydrogenase, quinone 1
<i>NQO1</i>	201467_s_at	1.5 (4.4E-03)	3.2 (6.7E-12)	NAD(P)H dehydrogenase, quinone 1
<i>ALDH3A1</i>	205623_at	1.6 (7.5E-04)	6.5 (3.4E-12)	Aldehyde dehydrogenase 3 family, member A1
<i>UGT1A1</i> /// <i>UGT1A3</i> to <i>UGT1A10</i>	215125_s_at	1.5 (2.1E-03)	2.2 (1.4E-08)	UDP glucuronosyltransferase 1 family, polypeptide A1 /// A3 to A10
<i>UGT1A1</i> /// <i>UGT1A4</i> /// <i>UGT1A6</i> /// <i>UGT1A8</i> to <i>UGT1A10</i>	207126_x_at	1.5 (3.1E-04)	1.8 (2.0E-08)	UDP glucuronosyltransferase 1 family, polypeptide A1 /// A4 /// A6 /// A8 to A10
<i>MUC1</i>	207847_s_at	(>0.05)	1.6 (0.021)	Mucin 1, cell surface associated
<i>MUC1</i>	213693_s_at	1.5 (0.043)	1.4 (0.015)	Mucin 1, cell surface associated
<i>CEACAM5</i>	201884_at	1.3 (0.011)	5.4 (3.8E-11)	Carcinoembryonic antigen-related cell adhesion molecule 5
<i>CEACAM6</i>	203757_s_at	(>0.05)	2.4 (1.3E-4)	Carcinoembryonic antigen-related cell adhesion molecule 6
<i>CEACAM6</i>	211657_at	(>0.05)	2.3 (2.1E-4)	Carcinoembryonic antigen-related cell adhesion molecule 6
<i>CEACAM7</i>	206198_s_at	3.0 (0.031)	(>0.05)	Carcinoembryonic antigen-related cell adhesion molecule 7
<i>FCGBP</i>	203240_at	2.0 (3.1E-05)	-2.3 (0.01)	Fc fragment of IgG binding protein

NOTE: Differentially expressed genes (fold changes) in the oral and airway mucosa of smokers versus never smokers with associated *P* values (ANOVA). The CEACAM family genes *CEACAM5* and *CEACAM6* are induced in the airways of smokers, whereas *CEACAM7* is induced in the oral mucosa of smokers. *FCGBP* is induced in the oral mucosa but repressed in the airway of smokers.

or treatment with inhibitors of prostaglandin synthesis will result in normalization of the number of Langerhans cells in the oral mucosa and improved immune function.

Levels of CHRNA3, the  $\alpha_3$  subunit of the nicotinic acetylcholine receptor, were increased in the oral mucosa of smokers. Nicotine binds to nicotinic acetylcholine receptors, leading to activation of Akt signaling and increased epithelial cell survival (47). The  $\alpha_3$  subunit is important for mediating these effects of nicotine in epithelial cells (47). Common variants in the nicotinic acetylcholine receptor gene cluster on chromosome 15q24-25.1 have been associated with an increased risk of lung cancer in smokers (48). This region includes the nicotinic acetylcholine receptor subunit gene *CHRNA3*. In theory, nicotine-mediated increased cell survival might lead to the accumulation of DNA adducts and increased mutagenesis and thereby stimulate carcinogenesis.

The fact that levels of CHRNA3 are increased in the oral mucosa of smokers underscores the possible role that altered nicotine signaling plays in carcinogenesis.

As mentioned above, women seem to be at increased risk of lung, oral, and oropharyngeal cancer compared with men who had similar levels of cigarette smoking exposure (12-14). Our results provide potential insights into the mechanisms underlying this gender-dependent difference in smoking-related cancer risk. *CYP1B1* was one of the genes most highly overexpressed in the oral mucosa of smokers. *CYP1B1* catalyzes the hydroxylation of estradiol to 4-hydroxy estradiol (49). 4-Hydroxycatechol estrogen is a highly reactive catechol estrogen, which is further oxidized to estrogen-3,4-quinone that can react with DNA to form unstable adducts, leading to depurination and mutations. Although the link between *CYP1B1*,

estrogen metabolism, and breast carcinogenesis has been intensively investigated (49), much less attention has been given to aerodigestive malignancies. Marked increases in levels of CYP1B1 mRNA were found in the oral mucosa of both male and female smokers. Because menstruating females produce higher levels of estrogen than males, it is possible that increased CYP1B1-mediated catabolism of estradiol occurs in the aerodigestive tracts of female smokers, resulting in enhanced mutagenesis and elevated cancer risk. As shown in Table 3, the magnitude of induction of AKR and UGT family members was greater in the oral mucosa of female than male smokers. By contrast, there was greater suppression of insulin-like growth factor-like family member 1 in the oral mucosa of female than male smokers. Although these findings need to be validated in larger studies, these differences could also help to explain gender-dependent differences in the risk of cancer. For example, as detailed above, activation of PAH-*trans*-dihydrodiols by AKRs leads to reactive oxygen species-mediated genotoxicity (22).

Our results also suggest that smoking induces similar changes in gene expression in the oral and bronchial epithelium (Table 4; Supplementary Table S5). For example, smoking is associated with increased expression of several genes (*CYP1A1*, *CYP1B1*, *NQO1*, *ALDH3A1*, and *UGTs*) involved in xenobiotic metabolism in both oral and bronchial epithelium. In addition to being important for understanding carcinogenesis, smoking-related changes in xenobiotic metabolism may alter the activity of selected chemopreventive agents (4, 5) and targeted anticancer therapies (6), resulting in reduced efficacy. Increased levels of CEACAM family members and GPX2 were found in both the oral and bronchial epithelium of smokers. These findings agree with other recent studies (21) and suggest that easily accessible oral epithelial cells provide insights into tobacco-induced molecular changes not only in the oral cavity but also in the bronchial epithelium. Use of oral epithelium should be considered as a surrogate tissue in future lung cancer prevention trials.

A powerful tool in computational biology is the ability to compare existing sets of expression data for patterns. The expression profile data from the current study were compared with expression profiles of drugs and small-

molecule inhibitors. This computational analysis suggested that geldanamycin, a Hsp90 inhibitor, might suppress the changes in the transcriptome induced by cigarette smoke. Consistent with this prediction, we showed that geldanamycin blocked tobacco smoke-mediated induction of CYP1A1 and CYP1B1 *in vitro*. These results are consistent with other evidence that Hsp90 inhibitors suppress AHR-mediated activation of *CYP1A1* and *CYP1B1* transcription (50). In addition to suppressing PAH-mediated induction of CYP1A1 and CYP1B1, inhibitors of Hsp90 have multiple other effects. It is predictable, for example, that Hsp90 inhibitors will downregulate levels of multiple other client proteins and suppress the induction of other AHR-regulated genes. AHR-dependent genes play a role in both the activation and the detoxification of tobacco carcinogens. Given the overall complexity of these effects, it is uncertain whether systemic or topical treatment with a Hsp90 inhibitor will suppress the mutagenic effects of tobacco smoke or have a chemopreventive effect. Additional studies will be needed to address these questions. More importantly, our findings illustrate the potential use of computational biology as a strategy to identify chemopreventive agents.

#### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

#### Acknowledgments

We thank Jenny Xiang (Microarray Core, Weill Cornell Medical College) for expert help, Professor Harel Weinstein and Piali Mukherjee for helpful discussions, and Kevin C. Dorff (HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Medical College of Cornell University) for web page design.

#### Grant Support

NIH grants R25 CA105012, T32 CA09685, P01 CA106451, and CTSC ULI-RR024996.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 09/09/2009; revised 10/06/2009; accepted 10/16/2009; published OnlineFirst 02/23/2010.

#### References

- Gritz ER, Dresler C, Sarna L. Smoking, the missing drug interaction in clinical trials: ignoring the obvious. *Cancer Epidemiol Biomarkers Prev* 2005;14:2287–93.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum* 2004;83:1–1438.
- Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer* 2003;3:733–44.
- Mayne ST, Lippman SM. Cigarettes: a smoking gun in cancer chemoprevention. *J Natl Cancer Inst* 2005;97:1319–21.
- The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study Group. The effect of vitamin E and  $\beta$  carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med* 1994;330:1029–35.
- Hamilton M, Wolf JL, Rusk J, et al. Effects of smoking on the pharmacokinetics of erlotinib. *Clin Cancer Res* 2006;12:2166–71.
- Fox JL, Rosenzweig KE, Ostroff JS. The effect of smoking status on survival following radiation therapy for non-small cell lung cancer. *Lung Cancer* 2004;44:287–93.
- Pantarotto J, Malone S, Dahrour S, Gallant V, Eapen L. Smoking is associated with worse outcomes in patients with prostate cancer treated by radical radiotherapy. *BJU Int* 2007;99:564–9.
- Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005;353:123–32.
- Browman GP, Wong G, Hodson I, et al. Influence of cigarette smoking on the efficacy of radiation therapy in head and neck cancer. *N Engl J Med* 1993;328:159–63.

11. Khuri FR, Lee JJ, Lippman SM, et al. Randomized phase III trial of low-dose isotretinoin for prevention of second primary tumors in stage I and II head and neck cancer patients. *J Natl Cancer Inst* 2006;98:441–50.
12. Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type. *Am J Epidemiol* 1993;138:281–93.
13. International Early Lung Cancer Action Program Investigators. Women's susceptibility to tobacco carcinogens and survival after diagnosis of lung cancer. *JAMA* 2006;296:180–4.
14. Neugut AI, Jacobson JS. Women and lung cancer: gender equality at a crossroad? *JAMA* 2006;296:218–9.
15. Powell CA, Klares S, O'Connor G, Brody JS. Loss of heterozygosity in epithelial cells obtained by bronchial brushing: clinical utility in lung cancer. *Clin Cancer Res* 1999;5:2025–34.
16. Franklin WA, Gazdar AF, Haney J, et al. Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis. *J Clin Invest* 1997;100:2133–7.
17. Wistuba II, Mao L, Gazdar AF. Smoking molecular damage in bronchial epithelium. *Oncogene* 2002;21:7298–306.
18. Guo M, House MG, Hooker C, et al. Promoter hypermethylation of resected bronchial margins: a field defect of changes? *Clin Cancer Res* 2004;10:5131–6.
19. Spira A, Beane J, Shah V, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A* 2004;101:10143–8.
20. Spira A, Beane JE, Shah V, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007;13:361–6.
21. Sridhar S, Schembri F, Zeskind J, et al. Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics* 2008;9:259.
22. Park JH, Mangal D, Tacka KA, et al. Evidence for the aldo-keto reductase pathway of polycyclic aromatic *trans*-dihydrodiol activation in human lung A549 cells. *Proc Natl Acad Sci U S A* 2008;105:6846–51.
23. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31:e15.
24. Cope LM, Irizarry RA, Jaffee HA, Wu Z, Speed TP. A benchmark for Affymetrix GeneChip expression measures. *Bioinformatics* 2004;20:323–31.
25. Lindros KO, Oinonen T, Kettunen E, Sippel H, Muro-Lupori C, Koivusalo M. Aryl hydrocarbon receptor-associated genes in rat liver: regional coincidence of aldehyde dehydrogenase 3 and glutathione transferase Ya. *Biochem Pharmacol* 1998;55:413–21.
26. Erichsen TJ, Ehmer U, Kalthoff S, et al. Genetic variability of aryl hydrocarbon receptor (AhR)-mediated regulation of the human UDP glucuronosyltransferase (UGT) 1A4 gene. *Toxicol Appl Pharmacol* 2008;230:252–60.
27. Landkisch TO, Gillman TC, Erichsen TJ, et al. Aryl hydrocarbon receptor-mediated regulation of the human estrogen and bile acid UDP-glucuronosyltransferase 1A3 gene. *Arch Toxicol* 2008;82:573–82.
28. Olinga P, Elferink MG, Draaisma AL, et al. Coordinated induction of drug transporters and phase I and II metabolism in human liver slices. *Eur J Pharm Sci* 2008;33:380–9.
29. Hosack DA, Dennis G, Jr., Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. *Genome Biol* 2003;4:R70.
30. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
31. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;313:1929–35.
32. Sacks PG. Cell, tissue and organ culture as *in vitro* models to study the biology of squamous cell carcinomas of the head and neck. *Cancer Metastasis Rev* 1996;15:27–51.
33. Gümüş ZH, Du B, Kacker A, et al. Effects of tobacco smoke on gene expression and cellular pathways in a cellular model of oral leukoplakia. *Cancer Prev Res* 2008;1:100–11.
34. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. *J Biol Chem* 1951;193:265–75.
35. Stevens EA, Mezrich JD, Bradfield CA. The aryl hydrocarbon receptor: a perspective on potential roles in the immune system. *Immunology* 2009;127:299–311.
36. Lou H, Du S, Ji Q, et al. Induction of AKR1C2 by phase II inducers: identification of a distal consensus antioxidant response element regulated by NRF2. *Mol Pharmacol* 2006;69:1662–72.
37. Penning TM, Drury JE. Human aldo-keto reductases: function, gene regulation and single nucleotide polymorphisms. *Arch Biochem Biophys* 2007;464:241–50.
38. Sreerama L, Sladek NE. Three different stable human breast adenocarcinoma sublines that overexpress ALDH3A1 and certain other enzymes, apparently as a consequence of constitutively up-regulated gene transcription mediated by transactivated EpREs (electrophile responsive elements) present in the 5'-upstream regions of these genes. *Chem Biol Interact* 2001;130–132:247–60.
39. Kekatpure VD, Dannenberg AJ, Subbaramaiah K. HDAC6 modulates Hsp90 chaperone activity and regulates activation of aryl hydrocarbon receptor signaling. *J Biol Chem* 2009;284:7436–45.
40. Conney AH, Miller EC, Miller JA. Substrate-induced synthesis and other properties of benzopyrene hydroxylase in rat liver. *J Biol Chem* 1957;228:753–66.
41. Volk DE, Thivyanathan V, Rice JS, et al. Solution structure of a *cis*-opened (10*R*)-*N*6-deoxyadenosine adduct of (9*S*,10*R*)-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene in a DNA duplex. *Biochemistry* 2003;42:1410–20.
42. Dannenberg AJ, Subbaramaiah K. Targeting cyclooxygenase-2 in human neoplasia: rationale and promise. *Cancer Cell* 2003;4:431–6.
43. Jayaprakash V, Rigual NR, Moysich KB, et al. Chemoprevention of head and neck cancer with aspirin: a case-control study. *Arch Otolaryngol Head Neck Surg* 2006;132:1231–6.
44. Barrett AW, Williams DM, Scott J. Effect of tobacco and alcohol consumption on the Langerhans cell population of human lingual epithelium determined using a monoclonal antibody against HLADR. *J Oral Pathol Med* 1991;20:49–52.
45. Andrews FJ, Halliday GM, Narkowicz CK, Muller HK. Indomethacin inhibits the chemical carcinogen benzo[*a*]pyrene but not dimethylbenz[*a*]anthracene from altering Langerhans cell distribution and morphology. *Br J Dermatol* 1991;124:29–36.
46. Arenberger P, Kemeny L, Rupec R, Bieber T, Ruzicka T. Langerhans cells of the human skin possess high-affinity 12(*S*)-hydroxyicosate-traenoic acid receptors. *Eur J Immunol* 1992;22:2469–72.
47. West KA, Brognard J, Clark AS, et al. Rapid Akt activation by nicotine and a tobacco carcinogen modulates the phenotype of normal human airway epithelial cells. *J Clin Invest* 2003;111:81–90.
48. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;40:616–22.
49. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. *N Engl J Med* 2006;354:270–82.
50. Hughes D, Guttenplan JB, Marcus CB, Subbaramaiah K, Dannenberg AJ. Heat shock protein 90 inhibitors suppress aryl hydrocarbon receptor-mediated activation of CYP1A1 and CYP1B1 transcription and DNA adduct formation. *Cancer Prev Res* 2008;1:485–93.

# Cancer Prevention Research

## Effects of Cigarette Smoke on the Human Oral Mucosal Transcriptome

Jay O. Boyle, Zeynep H. Gümüş, Ashutosh Kacker, et al.

*Cancer Prev Res* 2010;3:266-278. Published OnlineFirst February 23, 2010.

**Updated version** Access the most recent version of this article at:  
doi:[10.1158/1940-6207.CAPR-09-0192](https://doi.org/10.1158/1940-6207.CAPR-09-0192)

**Supplementary Material** Access the most recent supplemental material at:  
<http://cancerpreventionresearch.aacrjournals.org/content/suppl/2010/02/23/1940-6207.CAPR-09-0192.DC1>  
<http://cancerpreventionresearch.aacrjournals.org/content/suppl/2010/03/04/1940-6207.CAPR-09-0192.DC2>

**Cited articles** This article cites 50 articles, 14 of which you can access for free at:  
<http://cancerpreventionresearch.aacrjournals.org/content/3/3/266.full#ref-list-1>

**Citing articles** This article has been cited by 17 HighWire-hosted articles. Access the articles at:  
<http://cancerpreventionresearch.aacrjournals.org/content/3/3/266.full#related-urls>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, use this link  
<http://cancerpreventionresearch.aacrjournals.org/content/3/3/266>.  
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.