Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

Impact of Smoking Cessation on Global Gene Expression in the Bronchial Epithelium of Chronic Smokers

Li Zhang,[1] J. Jack Lee,[1] Hongli Tang,[2] You-Hong Fan,[2] Lianchun Xiao,[1] Hening Ren,[2] Jonathan Kurie,[2] Rodolfo C. Morice,[2] Waun Ki Hong[2] and Li Mao[2]

Abstract Cigarette smoke is the major cause of lung cancer and can interact in complex ways with drugs for lung cancer prevention or therapy. Molecular genetic research promises to elucidate the biological mechanisms underlying divergent drug effects in smokers versus nonsmokers and to help in developing new approaches for controlling lung cancer. The present study compared global gene expression profiles (determined via Affymetrix microarray measurements in bronchial epithelial cells) between chronic smokers, former smokers, and never smokers. Smoking effects on global gene expression were determined from a combined analysis of three independent data sets. Differential expression between current and never smokers occurred in 591 of 13,902 measured genes ($P < 0.01$ and >2-fold change; pooled data)—a profound effect. In contrast, differential expression between current and former smokers occurred in only 145 of the measured genes ($P < 0.01$ and >2-fold change; pooled data). Nine of these 145 genes showed consistent and significant changes in each of the three data sets ($P < 0.01$ and >2-fold change), with eight being down-regulated in former smokers. Seven of the eight down-regulated genes, including CYP1B1 and three AKR genes, influence the metabolism of carcinogens and/or therapeutic/chemopreventive agents. Our data comparing former and current smokers allowed us to pinpoint the genes involved in smoking-drug interactions in lung cancer prevention and therapy. These findings have important implications for developing new targeted and dosing approaches for prevention and therapy in the lung and other sites, highlighting the importance of monitoring smoking status in patients receiving oncologic drug interventions.

Chronic cigarette smoking is the major cause of lung cancer and remains so for years even after smoking cessation (1, 2). Therefore, the development of agents for controlling lung cancer generally targets, virtually by default, current and former heavy smokers. Smoking status, however, seems to influence response to various chemopreventive and chemotherapeutic agents and clinical outcomes of their use (3, 4). Three large randomized clinical trials to prevent lung cancer—the Alpha-Tocopherol, Beta-Carotene Prevention Study (5), Carotene and Retinol Efficacy Trial (6), and Lung Intergroup Trial (7)—showed that current heavy smokers had harmful interactions (higher lung cancer mortality, incidence, and recurrence) with preventive agents (versus control arms); agent effects in former smokers were generally neutral and were not readily interpretable in never smokers because of the exclusion or very limited number of these patients in these trials. Certain lung cancer therapy regimens have been shown to be less effective in current smokers than in former and never smokers (8, 9). Smoking can stimulate the metabolic clearance of targeted anticancer therapies, undoubtedly diminishing therapeutic benefit (9, 10). These data highlight the importance of understanding the biological effect of chronic smoking on lung tissue.

To understand why smokers and former smokers have differential responses to agents for preventing or treating lung cancer, we analyzed and compared global gene expression profiles in three independent cancer-free cohorts comprising current, former, and never smokers.

Materials and Methods

Study population

This study included current smokers, former smokers, and never smokers with no evidence of cancer and collected from separate, independent studies conducted at The University of Texas M. D. Anderson Cancer Center (MDACC; two studies) and the Boston Medical Center (BMC; one study). The three data sets associated with the three studies are called MDACC-1, MDACC-2, and BMC throughout this article. Former smoking was defined as having quit smoking for at least

Authors' Affiliations: Departments of [1]Bioinformatics and Computational Biology and [2]Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas
Received 12/22/2007; revised 02/18/2008; accepted 02/21/2008.
Grant support: National Cancer Institute grant P01 CA91844 and Department of Defense grant W81XWH-04-1-0142.
Note: Supplementary data for this article are available at Cancer Prevention Research Online (http://cancerprevres.aacrjournals.org/).
Requests for reprints: Li Zhang, M. D. Anderson Cancer Center, Houston, TX 77030. E-mail: Lzhangli@mdanderson.org, or Li Mao, M. D. Anderson Cancer Center. E-mail: lmao@mdanderson.org.
©2008 American Association for Cancer Research.
doi:10.1158/1940-6207.CAPR-07-0017

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

12 months before study entry. Participants included in the MDACC-1 and MDACC-2 data sets came from the placebo arm of an ongoing chemoprevention trial done at M. D. Anderson Cancer Center. All MDACC subjects were clinically free of cancer at enrollment and underwent a bronchoscopy at baseline. Bronchial brushes were done at six predetermined sites including the entry area at each of the five main lobes and the carina, as previously described (11). The study was approved by the MDACC Institutional Review Board, and all MDACC participants gave signed informed consent. BMC data set participants had a bronchoscopy at the BMC and were analyzed in a previously reported study (12) as well as in the current study. Potential subjects for the MDACC or BMC data sets in the current study were excluded if their specimen images (produced as discussed below in "cRNA preparation and microarray hybridization") had defects, evidence of blood contamination, or other problems that did not meet image quality criteria applied consistently across all three data sets. All MDACC patients had smoking history, with average pack-years of 40.6 (±13). Their average age is 58 (±8) years. Sixty-one percent of them are male and 78% of them are White. More details about the demographic data of MDACC data sets can be found in Supplementary Table S2.

Bronchial brush processing and RNA extraction

For samples in MDACC-1 and MDACC-2, brushes were placed on bronchoscopy in 3 mL of plain DMEM culture (Life Technologies, Inc.) in sterile tissue culture tubes and stored at 4°C for processing the same day. The tubes were vortexed lightly to detach cells from the brushes. After removal of the brush from the tube, the cell suspension was centrifuged at 2,500 rpm for 5 min. Cell pellets were then washed with 2 mL of PBS twice, and an aliquot of material was saved at –80°C until RNA extraction. For the microarray analysis, cells from the six brushing sites of the same individual were pooled together for RNA extraction. We used TRIzol reagent (Invitrogen) for total RNA extraction according to the manufacturer's protocol, with a yield of 1 to 4 μg of total RNA per sample. Integrity of the RNA was confirmed by running it on an RNA 6000 Neno LabChip (Agilent Technologies). The samples in BMC data set were processed similarly as described (12), except that a single-round amplification protocol was used.

cRNA preparation and microarray hybridization

The first and second cDNA strands were synthesized as previously described (13). The first reverse transcription was done in the absence of biotin-labeled ribonucleotides, resulting in unlabeled cRNA, which was then used as starting material for the second cycle. In the second cycle, the first and second cDNA strands were synthesized. The second transcription was done in the presence of biotin-labeled-ribonucleotides, resulting in labeled cRNA. The cRNA was fragmented and checked by gel electrophoresis, as reported earlier (13). The Affymetrix GeneChip system was used for hybridization, staining, and imaging of the probe arrays. Hybridization cocktails of 300 μL, each containing 15 μg of cRNA and exogenous hybridization controls, were prepared as previously described and hybridized to U133A or U133A plus GeneChips (Affymetrix) overnight at 42°C. Hybridized fragments were detected with streptavidin linked to phycoerythrin (Molecular Probes). GeneChips were scanned and imaged using Affymetrix Microarray Analysis Suite version 5.0.

Microarray data normalization

There were two array types used in this study: U133A and U133 Plus 2. The U133A array contains ~500,000 distinct probe features interrogating 18,400 human transcripts and variants, including 13,902 well-characterized genes. The U133 plus 2.0 array contains all probe features that are on the U133A array. In addition, there are 9,921 new probe sets representing 6,500 new genes. To facilitate straightforward comparison of the data, we used only the probes that are common to both array types. We also ignored data of MM probes. We used PM probes common to U133A and U133 Plus 2 arrays to perform quantile normalization on the probe level data (14). The procedure was done so that the distributions of the probe signal intensities of a sample are identical for all samples within a data set. Then, we used PDNN model (15) to quantify the gene expression values from the normalized probe signal intensity data. We then applied median-centering normalization on the probe set level data so that the median of expression values of a sample was made to be the same for all the samples in all of the data sets.

Identification of differentially expressed genes

Differential expression was identified to be similar to that described by Wang et al. (16). We used Z values (defined below) to assess differentially expressed genes between current and never or former smokers, in whom the magnitude of the Z values is assumed to represent the effect of smoking cessation. For a given data set containing n_A and n_B samples in groups A and B, respectively, we compute the following test statistic Z for each probe set:

$$Z = D/\sigma \qquad (A)$$

where D is average difference between the log expression values between A and B groups. σ is the estimated standard deviation (SD) of D:

$$\sigma = \sqrt{\sigma_A{}^2/n_A + \sigma_B^2/n_B} \qquad (B)$$

where σ_A^2 and σ_B^2 are estimated variances of log expression values in groups A and B, respectively. These variances were estimated using Loess fit between the mean log expression values and the SD of the log expression values. The underlying assumption is that the mean and the SD are related by a smooth function, which allows the analysis method to treat the SD as if it were known.

Combining Z values from different data sets

The Z values obtained from the three data sets can be combined using the following formula:

$$Z = (Z_1/\sigma_1^2 + Z_2/\sigma_2^2 + Z_3/\sigma_3^2)/\omega^2 \qquad (C)$$

$$\omega^2 = 1/\sigma_1^2 + 1/\sigma_2^2 + 1/\sigma_3^2 \qquad (D)$$

where Z_1, Z_2, and Z_3 were calculated using Eq. A from MDACC-1, MDACC-2, and BMC data sets, respectively; σ_1, σ_2, and σ_3 were calculated using Eq. B from MDACC-1, MDACC-2, and BMC data sets, respectively.

The test statistic Z is supposed to form a T distribution if the log expression values are normally distributed. However, the observed data slightly deviate from the normal distribution because they contain more extreme values. Consequently, the significance of Z can be overestimated.

To alleviate the bias due to the assumption of normal distribution, we used permuted data to compute Z^*. The expression values are randomly permuted for each probe set within each data set. The permutation was done 10 times to construct an empirical cumulative distribution function of Z^*. This distribution was assumed to be

Table 1. Sample sizes and array types of the microarray data sets

Data set	FS	CS	NS	Array type
MDACC-1	7	11	0	U133A
MDACC-2	8	15	0	U133 Plus 2
BMC	9	30	19	U133A

Abbreviations: FS, former smoker; CS, current smoker; NS, never smoker.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

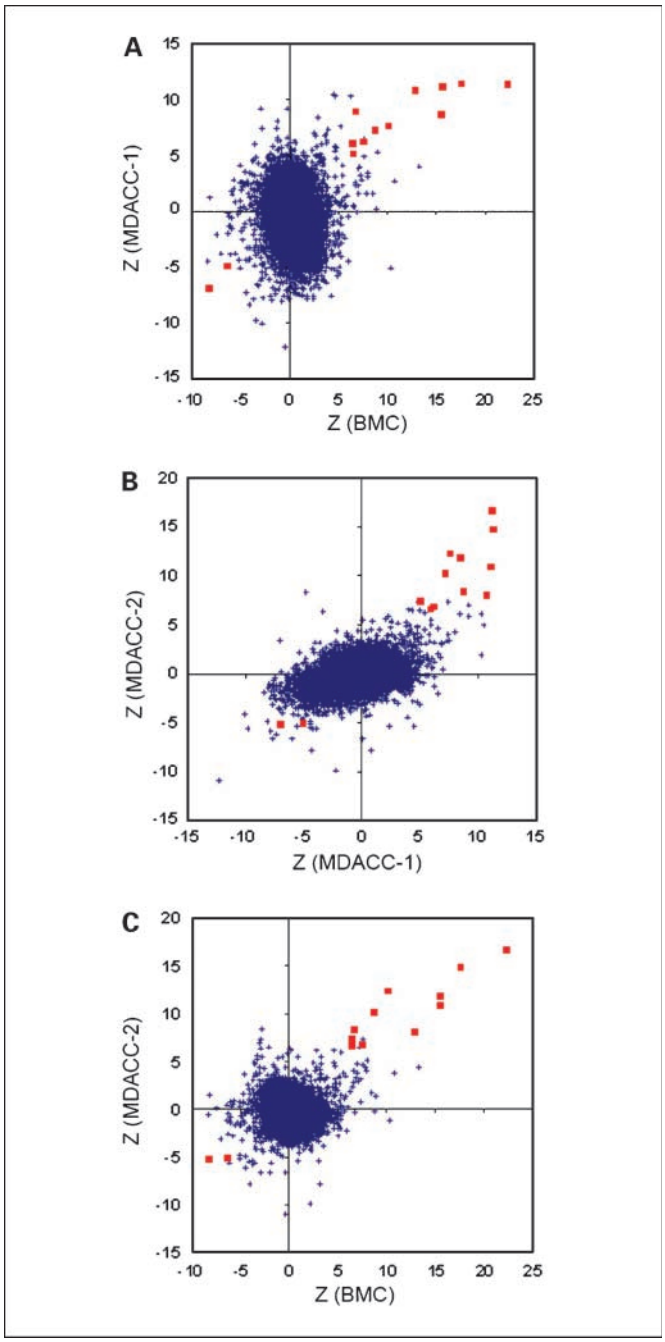

Fig. 1. Comparison of *Z* values obtained from the three data sets (BMC, MDACC-1, and MDACC-2). Each point in these scatter plots represents a probe set. The probe sets with absolute *Z* values >5 in all three data sets are shown in red. Detailed data on these probe sets are in Table 2.

the distribution of *Z* values under null hypothesis (i.e., no differential expression), and it was used to estimate the *P* values and the false discovery rate associated with *Z* values. The permutation was done within each of the data sets, but never across the data sets. Note that other than the permutation step, our method is the same as that described by Wang et al. (16).

Results

Our overall study population numbered 99 individuals, composed of 56 current smokers, 24 former smokers, and 19 never smokers from three independent data sets (Table 1). The MDACC-1 and MDACC-2 data sets included 41 chronic smokers (26 current, 15 former) enrolled in an ongoing chemoprevention trial at M. D. Anderson Cancer Center. All 41 of these subjects had at least a 20-pack-year smoking history. Demographic characteristics of the MDACC cohorts are included in Supplementary Table S1. The BMC data set was composed of 75 current, former, and never smokers. Never smokers with significant environmental cigarette exposure and subjects with respiratory symptoms or who regularly use inhaled medications were excluded. We selected 58 members of the BMC cohort for the present analysis (Table 1) and excluded 17 subjects. Exclusions from either the BMC or MDACC data sets were based on image quality criteria applied consistently across all three data sets.

First, we determined *Z* values (defined in Materials and Methods) in the three data sets separately. Then, we compared the *Z* values from each data set and, as shown in Fig. 1, we found that the genes with the most significant differential expressions (shown in red) are similar among the three data sets. The largest *Z* values mostly are located in the first and third quadrants in the scatter plots of Fig. 1, indicating that these

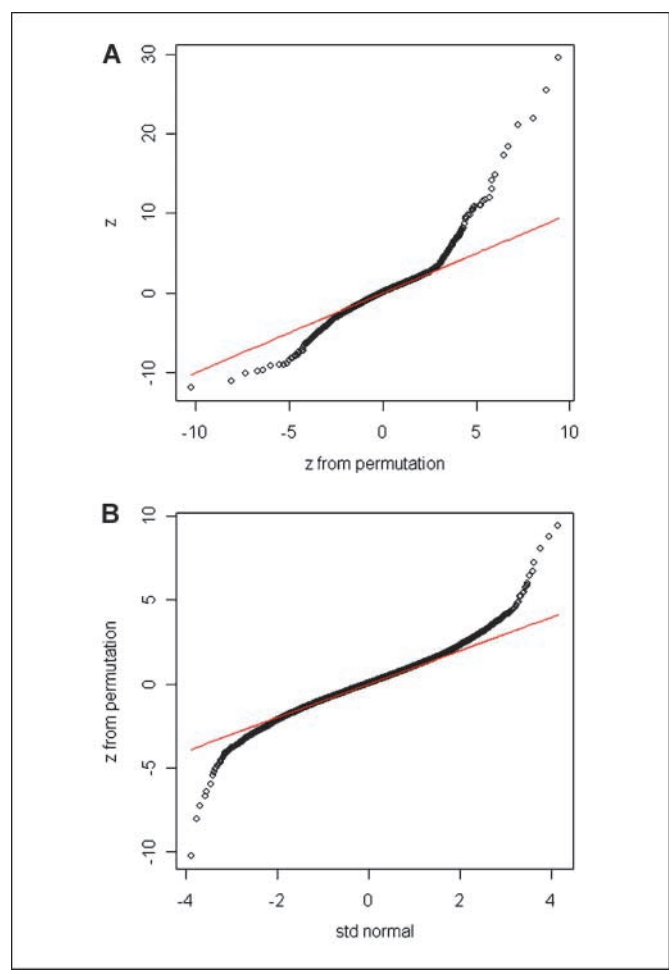

Fig. 2. Quantile-quantile plots of *Z* values. *A*, quantile of *Z* values versus quantile of *Z* values obtained from permuted data. *B*, quantile of *Z* values from permutation data versus quantile values of standard normal distribution.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

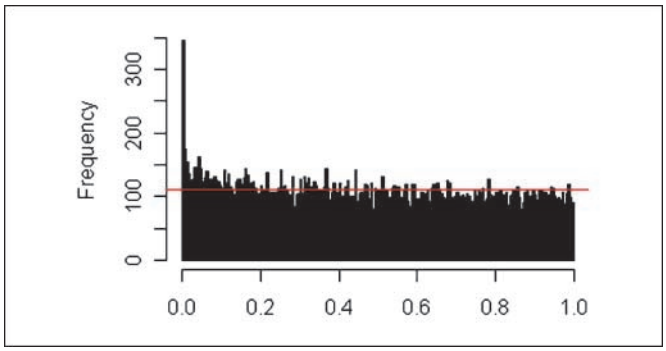

Fig. 3. Histogram of *P* values in search of differential expression between current and former smokers. Based on BUM estimate, 345 probe sets were identified as differentially expressed with a false discovery rate of 32%. One hundred seventy-six of the 345 probe sets have a fold change >2. Detailed gene information on the 176 probe sets (145 genes) is provided in Supplementary Table S2. The *P* values were evaluated on the basis of the combined *Z* values from the three data sets and the combined *Z* values from the permuted data.

changes in gene expression are consistent among the three data sets.

To further assess the statistical significance of the changes in gene expression between former and current smokers, we used quantile-quantile plots (Fig. 2) of *Z* values and BUM plots (Fig. 3) to evaluate the distribution of *P* values. Figure 2A compares the quantiles of *Z* values calculated from combining all three microarray data sets and the quantiles of Z_P (*Z* values calculated from permuted data). With permuted data, *Z* values are bounded between −10 and 10. The *Z* values from observed data contain clear outliers >10. Without differential expression, the data points in Fig. 2A should be close to the diagonal line (shown in red). Ideally, if the gene expression data obey normal distributions and are independent from each other, we would expect values of Z_P to form a standard normal distribution. However, Fig. 2B shows that Z_P's have wider ranges than that from standard normal distribution. Consequently, we used the distribution of Z_P as that from the null hypothesis (no differential expression between former and current smokers) to compute the *P* values of *Z* instead of using the standard normal distribution.

The BUM plot (17) presented a histogram of the *P* values. Under the null hypothesis, the *P* values should form a uniform distribution. The sharp spike at the left side of Fig. 3 represents the effects of differential expression contradicting the null hypothesis. The uniform part of the histogram is indicated by the red line in Fig. 3. The area above the red line contains ~1,200 probe sets, which is our estimated number of genes that are differentially expressed between the former smokers and current smokers. Only a subset of these genes is identifiable, however. According to the BUM method (17), we found 345 probe sets that were differentially expressed at a *P* value of <0.01, for which the false discovery rate was estimated to be 32%. Of the 345 probe sets, 176 have a >2-fold difference in expression (details of these 176 probe sets are shown in Supplementary Table S2). These 176 probe sets represent 145 nonredundant significantly differentially expressed genes (>2-fold change; $P < 0.01$). These 145 genes include 9 genes (Table 2) with consistent and significant changes in each of the three data sets ($P < 0.01$; >2-fold change). Eight of the nine genes are down-regulated after smoking cessation; one is up-regulated. To test the general accuracy of our microarray measurements, we compared them with reverse transcription-PCR measurements of a selected panel of genes, finding that the reverse transcription-PCR and microarray measurements were highly correlated, 96% [e.g., in the case of *ALDH3A1* (Supplementary Fig. S1), which is the gene with the largest change between former and current smokers (Table 2)]. Furthermore, although not calculated, the false discovery rate for the subset of 176 probe sets

Table 2. Genes with consistent fold changes >2 in each ($P < 0.01$) and across ($P \leq 0.0001$) the three data sets

Gene	Fold changes				*P* (comb.)	Full name	RefSeq	Probe set
	BMC	MDACC-1	MDACC-1	Comb.				
ALDH3A1	6.9	9.4	4.0	6.2	0.0000	aldehyde dehydrogenase 3 family, member A1	NM_000691	205623_at
CYP1B1	4.2	5.7	6.7	4.9	0.0000	cytochrome *P*450, member 1B1	NM_000104	202436_s_at
MUC5AC	2.2	9.6	3.0	3.5	0.0000	mucin 5AC, oligomeric mucus/gel-forming	XM_001130382	214385_s_at
AKR1C2	3.3	4.2	3.5	3.5	0.0000	aldo-keto reductase family 1, member C2	NM_001354	209699_x_at
AKR1B10	3.2	4.2	3.8	3.5	0.0000	aldo-keto reductase family 1, member B10	NM_020299	206561_s_at
AKR1C1	2.8	4.0	3.3	3.2	0.0000	aldo-keto reductase family 1, member C1	NM_001353	204151_x_at
NQO1	2.8	2.3	2.4	2.6	0.0001	NAD(P)H dehydrogenase, quinone 1	NM_000903	210519_s_at
AKR1C3	2.5	2.1	3.1	2.5	0.0000	aldo-keto reductase family 1, member C3	NM_003739	209160_at
SCGB1A1	−2.0	−2.4	−2.6	−2.4	0.0001	secretoglobin, family 1A, member 1 (uteroglobin)	NM_003357	205725_at

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

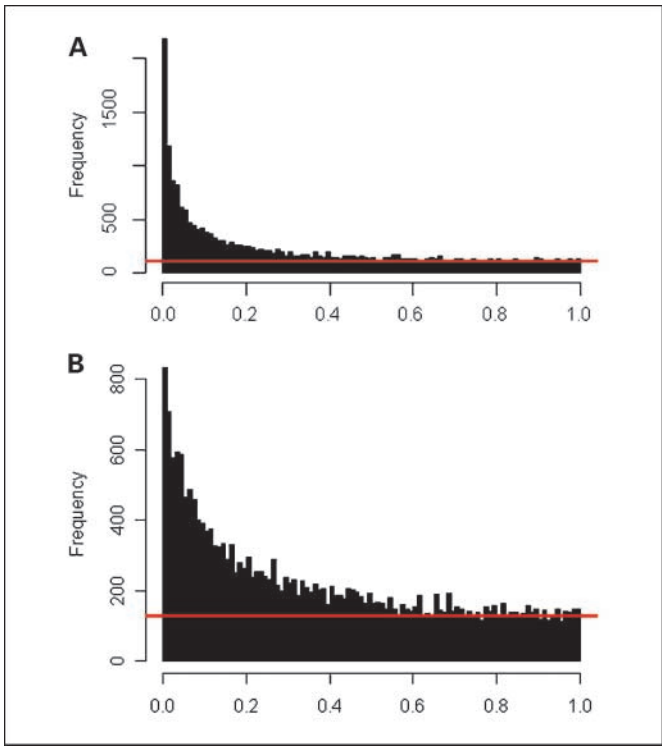

Fig. 4. Histograms of *P* values in search of differential expression between never smokers and current smokers (*A*) and between former smokers and never smokers (*B*). Only data from BMC data set were used the plots. False discovery rates were estimated to be 5% and 16% for $P < 0.01$ in *A* and *B*, respectively.

should be lower than that (32%) estimated for the 345 probe sets, and the false discovery rate for the nine changed genes that were validated across three data sets should be lower still because each subset adds new criteria that increase reliability.

For comparison, we also examined differential gene expression between current and never smokers (Fig. 4A). Similar to that in Fig. 3, the peak volume above the red line represents the number of differentially expressed genes, which is ~11,000 probe sets. This number is >9 times greater than the number detected in the comparison between former and current smokers (Fig. 3). We found 591 nonredundant genes with statistically significant changes (>2-fold change and $P < 0.01$) in pooled data of the three data sets, a group that is >4 times larger than the group of such differentially expressed genes detected in the comparison between current and former smokers. Of the 145 significantly changed genes between current and former smokers, 77 are consistent with, and 68 are not consistent with, the 591 such genes between current and never smokers (Supplementary Table S2). The nine genes with consistent and significant changes between former and current smokers in each of the three data sets are in the subset of 77 common, significantly changed genes. Figure 4B compares former smokers with never smokers.

The scope of differential expressions in Fig. 4A is much larger than that in Fig. 3, which may be due to differences in sample size. A principal component analysis (Fig. 5), however, supported the conclusion that the larger differential expression in Fig. 4A compared with that in Fig. 3 is not simply due to sample size. The gene expression profile of each patient is represented by its two main principal components. Two distinct clusters emerge in Fig. 5, and the cluster to the left (Comp1 <–10) contains mostly never smokers. The right-side cluster is predominated by a mixture of current and former smokers, which supports the conclusion that former smokers are more similar to current smokers than to never smokers.

Discussion

In probing 13,902 genes, we found that 591 were differentially expressed in current versus never smokers and that only 145 of these 591 (25%) were also differentially expressed in current versus former smokers. Among these 145 genes, 9 were significantly differentially expressed (8 overexpressed, 1 underexpressed; Table 2) by >2-fold in current versus former smokers in each ($P < 0.01$) and in the pooled data ($P < 0.0001$) of the three data sets (two MDACC, one BMC) included in this study. Therefore, our present study pinpoints and validates nine differentially expressed genes in former versus current smokers.

Seven of the eight validated genes overexpressed in current smokers—CYP1B1, four AKRs, ALDH3A1, and NQO1 (Table 2)—are involved in drug and/or carcinogen metabolism (9, 10, 18–27). Polycyclic aromatic hydrocarbons in tobacco smoke are known to bind to and activate the aryl hydrocarbon receptor and thus induce CYP1B1 (10). CYP1B1 expression is of special interest because it may contribute both to increased drug metabolism and to carcinogenesis of the aerodigestive tract (1, 18–20). The metabolic clearance of docetaxel, tamoxifen, gefitinib, erlotinib, and other cancer prevention and therapy drugs is enhanced by YP1B1 (9, 21–23). Up-regulation of CYP1B1 and the six other validated overexpressed metabolizing genes by smoking is likely involved in the adverse interactions between smoking and drugs for lung cancer prevention and therapy; smoking cessation down-regulates these gene expressions and thus may reduce or eliminate the adverse drug interactions.

Four of the eight most-up-regulated genes we detected in current smokers (Table 2) are members of the AKR

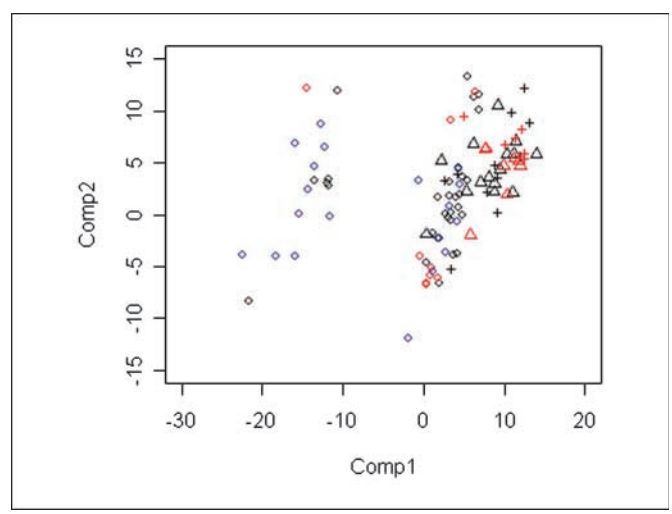

Fig. 5. Principal component analysis. The two main principal components were used to visualize the relationships among patients with different smoking status. Each point represents a patient. Current smokers were shown in black, former smokers in red, and never smokers in blue. BMC data were shown in circles, MDACC-1 data in pluses, and MDACC-2 data in triangles.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

family (AKR1C1, AKR1C2, AKR1C3, and AKR1B10; ref. 28). AKR1B10 is overexpressed in non–small-cell lung cancer and squamous metaplasia in association with smoking (24, 26). AKR1C1, AKR1C2, and AKR1C3 are known to be involved in tobacco carcinogen and/or drug metabolism. AKR1C1 overexpression is correlated with a poor prognosis of non–small-cell lung cancer and is associated with chemotherapeutic drug resistance (25, 29). Data suggest that a potential role of AKR1B10 in retinoic acid signaling (30) may be a factor in the negative effects of retinoic acid (retinoids) and its relative β-carotene in smokers in chemoprevention trials (3–7). Several studies have shown that overexpression of AKR1C1, AKR1C2, or AKR1C3 contributes to the resistance of various tumor types, including lung cancer, to cisplatin-based chemotherapy (25, 31–33).

Various biases can produce inconsistencies between similar data sets. These biases can stem from differences in age, race, sex, smoking history, and sample processing. Regarding sample processing, for example, MDACC-1 and MDACC-2 samples involved two rounds of RNA amplification versus a single round in the BMC set. Two rounds of amplification are known to cause loss of signals for probes that target far away from the 3′ end of mRNA sequences. The consistent changes in smoking cessation–related genes in all three independent data sets support the robustness of our present findings.

Gene expression profiling in bronchoscopy specimens offers a direct assessment of the effects of cigarette smoking in the lungs. Gene expression patterns vary greatly between individuals, however, because of genetic variations and different environmental influences. A report by Spira et al. (on a relatively broad array of differentially expressed metabolizing and antioxidant genes in current versus former smokers; ref. 12) provided us with the opportunity to increase the robustness of our gene expression analyses by adding the BMC data set to our MDACC data sets. As we prepared our present results for publication, the Spira group published another report (34) that extended their earlier study, as do the complementary and confirmatory findings we report here. We were able to identify the specific drug-metabolizing genes involved in smoking-drug interactions, including the overexpressed genes in current versus former smokers, because of the cross validation and increased statistical power provided by adding the BMC data set (12) to our MDACC-1 and MDACC-2 data sets. The combined effect of these reports is to increase the robustness of their interrelated findings and thus their appeal for hypothesis generation.

Our results also show that the scope of genetic changes following smoking cessation is much smaller than that associated with chronic smoking (Figs. 2 and 4), possibly explaining the persistent high lung cancer risk in former smokers (35). Surprisingly at the time (~10 years ago), we and others previously found in assessments limited to specific genetic alterations that smoking-related genetic changes persisted after smoking cessation in a population similar to those of MDACC-1, MDACC-2, and BMC (36, 37). Showing similar genetic alterations in current and former smokers, results of the more sophisticated global genomic profiling approach of our present and other studies are consistent with the earlier findings (12, 34).

Our findings underscore the importance of smoking status in clinical trials, showing that smoking effects on metabolizing genes potentially can interfere with drugs in standard or investigational chemoprevention or therapy not only in the lung but in other sites as well. Future research directions should include (*a*) increased monitoring of smoking status and increased smoking cessation efforts in any trial setting because of adverse smoking effects on drug uptake and metabolism, and (*b*) the development of new dosing and targeted approaches to counteract adverse smoking-drug interactions in the lung. New targeted approaches should consider the signaling pathways of drug-metabolizing genes that were validated in this study.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733–44.
2. Cancer facts and figures, 2006 [article on the Internet]. American Cancer Society 2006. Available from: http://www.cancer.org/downloads/STT/CAFF2006PWSecured.pdf].
3. Mayne ST, Lippman SM. Cigarettes: a smoking gun in cancer chemoprevention. J Natl Cancer Inst 2005;97:1319–21.
4. 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.
5. The α-Tocopherol, β Carotene Cancer Prevention Study Group. The effect of vitamin E and β carotene on the incidence of lung cancer and other cancers in male smokers. N Engl J Med 1994;330: 1029–35.
6. Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of β carotene and vitamin A on lung cancer and cardiovascular disease. N Engl J Med 1996;334:1150–5.
7. Lippman SM, Lee JJ, Karp DD, et al. Randomized phase III intergroup trial of isotretinoin to prevent second primary tumors in stage I non-small-cell lung cancer. J Natl Cancer Inst 2001;93:605–18.
8. Zhang Z, Xu F, Wang S, et al. Influence of smoking on histologic type and the efficacy of adjuvant chemotherapy in resected non-small cell lung cancer. Lung CancerEpub ahead of print.
9. Hamilton M, Wolf JL, Rusk J, et al. Effects of smoking on the pharmacokinetics of erlotinib. Clin Cancer Res 2006;12:2166–71.
10. Port JL, Yamaguchi K, Du B, et al. Tobacco smoke induces CYP1B1 in the aerodigestive tract. Carcinogenesis 2004;25:2275–81.
11. Lee JS, Lippman SM, Benner SE, et al. Randomized placebo-controlled trial of isotretinoin in chemoprevention of bronchial squamous metaplasia. J Clin Oncol 1994;12:937–45.
12. 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.
13. Gold D, Coombes K, Medhane D, et al. A comparative analysis of data generated using two different target preparation methods for hybridization to high-density oligonucleotide microarrays. BMC Genomics 2004;5:2.
14. Bolstad BM, Irizzary RA, Astrand M, et al. A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. Bioinformatics 2003;19:185–93.
15. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. Nat Biotechnol 2003;21:818–21.
16. Wang J, Coombes KR, Highsmith WE, et al. Differences in gene expression between B-cell chronic lymphocytic leukemia and normal B cells: a meta-analysis of three microarray studies. Bioinformatics 2004;20:3166–78.
17. Pounds S, Morris SW. Estimating the occurrence of false positives and false negatives in microarray studies b approximating and partitioning the empirical distribution of *P* values. Bioinformatics 2003;19:1236–42.
18. Mahadevan B, Luch A, Atkin J, et al. Inhibition of human cytochrome *P*450 1B1 further clarifies its role in the activation of dibenzo[*a,l*]pyrene in cells in culture. J Biochem Mol Toxicol 2007;21:101–9.
19. Roos PH, Bolt HM. Cytochrome *P*450 interactions in human cancers: new aspects considering CYP1B1. Expert Opin Drug Metab Toxicol 2005;1: 187–202.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

20. Purnapatre K, Khattar SK, Saini KS. Cytochrome *P*450s in the development of target-based anticancer drugs. Cancer Lett 2008;259:1–15.

21. Sissung TM, Price DK, Sparreboom A, et al. Pharmacogenetics and regulation of human cytochrome *P*450 1B1: implications in hormone-mediated tumor metabolism and a novel target for therapeutic intervention. Mol Cancer Res 2006;4:135–50.

22. Li J, Zhao M, He P, et al. Differential metabolism of gefitinib and erlotinib by human cytochrome *P*450 enzymes. Clin Cancer Res 2007;13:3731–7.

23. Rochat B, Morsman JM, Murray GI, et al. Human CYP1B1 and anticancer agent metabolism: mechanism for tumor-specific drug inactivation? JPET 2001;296:537–41.

24. Fukumoto S, Yamauchi N, Moriguchi H, et al. Overexpression of the aldo-keto reductase family protein AKR1B10 is highly correlated with smokers' non-small cell lung carcinomas. Clin Cancer Res 2005;11:1776–85.

25. Penning TM. AKR1B10: a new diagnostic marker of non-small cell lung carcinoma in smokers. Clin Cancer Res 2005;11:1687–90.

26. Woenckhaus M, Klein-Hitpass L, Grepmeier U, et al. Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers. J Pathol 2006;210:192–204.

27. Sladek NE, Kollander R, Sreerama L, et al. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens. Cancer Chemother Pharmacol 2002;49:309–21.

28. Penning TM, Drury JE. Human aldo-keto reductases: function, gene regulation, and single nucleotide polymorphisms. Arch Biochem Biophys 2007;464:241–50.

29. Hsu NY, Ho HC, Chow KC, et al. Overexpression of dihydrodiol dehydrogenase as a prognostic marker of non-small cell lung cancer. Cancer Res 2001;61:2727–31.

30. Crosas B, Hyndman D, Gallego O, Martras S, Pares X, Flynn TG. Human aldose reductase and human small intestine aldose reductase are efficient retinal reductases: consequences for retinoid metabolism. Biochem J 2003;373:973–9.

31. Deng HB, Parekh HK, Chow KC, Simpkins H. Increased expression of dihydrodiol dehydrogenase induces resistance to cisplatin in human ovarian carcinoma cells. J Biol Chem 2002;277:15035–43.

32. Deng HB, Adikari M, Parekh HK, Simpkins H. Ubiquitous induction of resistance to platinum drugs in human ovarian cervical, germ-cell and lung carcinoma tumor cells overexpressing isoforms 1 and 2 of dihydrodiol dehydrogenase. Cancer Chemother Pharmacol 2004;54:301–7.

33. Chen J, Adikari M, Pallai R, Parekh HK, Simpkins H. Dihydrodiol dehydrogenases regulate the generation of reactive oxygen species and the development of cisplatin resistance in human ovarian carcinoma cells. Cancer Chemother Pharmacol 2007. Epub ahead of print 17661040.

34. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. Genome Biol 2007;8:R201.

35. Tong L, Spitz MR, Fueger JJ, et al. Lung carcinoma in former smokers. Cancer 1996;78:1004–10.

36. Mao L, Lee JS, Kurie JM, et al. Clonal genetic alterations in the lung of current and former smokers. J Natl Cancer Inst 1997;89:857–62.

37. Wistuba II, Lam S, Behrens C, et al. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89:1366–73.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.

Published OnlineFirst March 31, 2008; DOI: 10.1158/1940-6207.CAPR-07-0017

AACR American Association for Cancer Research

Cancer Prevention Research

Impact of Smoking Cessation on Global Gene Expression in the Bronchial Epithelium of Chronic Smokers

Li Zhang, J. Jack Lee, Hongli Tang, et al.

Cancer Prev Res 2008;1:112-118. Published OnlineFirst March 31, 2008.

Updated version	Access the most recent version of this article at: doi:10.1158/1940-6207.CAPR-07-0017
Supplementary Material	Access the most recent supplemental material at: http://cancerpreventionresearch.aacrjournals.org/content/suppl/2008/04/02/1940-6207.CAPR-07-0017.DC1 http://cancerpreventionresearch.aacrjournals.org/content/suppl/2009/05/08/1940-6207.CAPR-07-0017.DC2

Cited articles	This article cites 34 articles, 11 of which you can access for free at: http://cancerpreventionresearch.aacrjournals.org/content/1/2/112.full#ref-list-1
Citing articles	This article has been cited by 6 HighWire-hosted articles. Access the articles at: http://cancerpreventionresearch.aacrjournals.org/content/1/2/112.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.
Permissions	To request permission to re-use all or part of this article, use this link http://cancerpreventionresearch.aacrjournals.org/content/1/2/112. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.

Downloaded from cancerpreventionresearch.aacrjournals.org on October 23, 2021. © 2008 American Association for Cancer Research.