

Variability of Antioxidant-Related Gene Expression in the Airway Epithelium of Cigarette Smokers

Neil R. Hackett, Adriana Heguy, Ben-Gary Harvey, Timothy P. O'Connor, Karsta Luettich, Douglas B. Flieder, Rana Kaplan, and Ronald G. Crystal

Belfer Gene Therapy Core Facility, Department of Genetic Medicine, Division of Pulmonary and Critical Care Medicine, and Department of Pathology, Weill Medical College of Cornell University, New York, New York

Cigarette smoking is the major risk factor for developing chronic bronchitis, yet only 15–20% of smokers develop this disorder. Because oxidants are the major mechanism of smoking-induced airway damage, we hypothesized that smoking is associated with upregulation of various antioxidant-related genes in the airway epithelium, but the magnitude of the response shows high inter-individual variability. Microarray analysis was used to assess levels of expression of 44 antioxidant-related genes in four categories (catalase/superoxide dismutase family; glutathione metabolism; redox balance; and pentose phosphate cycle) in bronchoscopy-obtained airway epithelium of matched cohorts (13 current smokers, 9 nonsmokers), none of whom had lung disease. There was minimal variation in gene expression levels within the same individual (right versus left lung or over time), but significant upregulation of 16/44 antioxidant-related genes in smoker epithelium compared with nonsmokers. Subgroups of smokers were identified with clusters of expression levels of antioxidant-related genes. We propose that the antioxidant-related genes demonstrating the most variability in the level of expression in smokers may be useful genetic markers in epidemiologic studies assessing susceptibility to smoking-induced chronic bronchitis.

Chronic bronchitis, a common form of chronic obstructive pulmonary disease (COPD) associated with persistent inflammation and pathogen colonization of the airway epithelium, is defined clinically by the presence of a chronic productive cough for at least 3 mo in two consecutive years in an individual in whom other causes of chronic cough have been excluded (1). Cigarette smoking is the major risk factor for the development of chronic bronchitis, and there is a dose-dependent link of the cumulative amount of smoking and the incidence of the disease (2, 3). Compelling evidence supports the concept that the pathogenesis of chronic bronchitis is associated with the chronic stress of oxidants in cigarette smoke overwhelming the antioxidant defenses of the airway epithelium, resulting in persistent injury to the airways (4, 5).

(Received in original form December 31, 2002 and in revised form April 2, 2003)

Address correspondence to: Ronald G. Crystal, Department of Genetic Medicine, Weill Medical College of Cornell University, 515 East 71st Street, Suite 1000, New York, New York 10021. E-mail: geneticmedicine@med.cornell.edu

Abbreviations: α 1-antitrypsin, AAT; chronic obstructive pulmonary disease, COPD; coefficient of variation, CV; epithelial lining fluid, ELF; γ -glutamate cysteine ligase, GCL; superoxide dismutase, SOD.

Am. J. Respir. Cell Mol. Biol. Vol. 29 pp. 331–343, 2003

Originally Published in Press as DOI: 10.1165/rcmb.2002-0321OC on April 17, 2003
Internet address: www.atsjournals.org

Despite the clear link of smoking to the risk for chronic bronchitis, only 15–20% of cigarette smokers develop COPD (3, 6), suggesting that there must be risk factors other than smoking that contribute to the susceptibility to this disease. Clues that genetic variability plays an important role come from studies demonstrating familial clustering of impaired lung function in response to smoking, and studies documenting susceptibility to smoking-related airflow obstruction in twins (3, 7, 8). In this context, the focus of this study is to identify candidate genes linked to the risk for chronic bronchitis in association with cigarette smoking. Because it is the airway epithelium that takes the brunt of the oxidant burden in cigarette smoke, we hypothesize that the airway epithelium responds to smoking by upregulating the expression of genes related to antioxidant protection, but the response to cigarette smoking for each gene is variable among the population, i.e., different individuals upregulate specific antioxidant-related genes differently in the airway epithelium in response to the stress of smoking. If these hypotheses are correct, then current smokers who have smoked an average of 20 pack-years, but are still phenotypically normal (i.e., have not yet developed airway disease), should demonstrate upregulation of antioxidant-related genes in the airway epithelium, but there should be variability in the relative amount of upregulation of these genes among the smokers.

Our approach to identify candidate genes relevant to the risk for chronic bronchitis is based on the ability to use fiberoptic bronchoscopy with airway brushing to obtain fresh samples of airway epithelium of smokers and nonsmoker normal volunteers, and to use microarray analysis to assess the expression of antioxidant-related genes in these fresh samples of airway epithelium (9–11). We have evaluated the expression of airway epithelial antioxidant-related genes in matched cohorts of 13 current cigarette smokers and 9 nonsmokers, none of whom had a history or evidence of lung disease. The mRNA levels in the airway epithelium of 44 antioxidant-related genes in 4 categories were quantified, including: (i) catalase/superoxide dismutase family; (ii) glutathione metabolism; (iii) redox balance; and (iv) pentose phosphate cycle. Of the 44 antioxidant-related genes assessed, 16 were significantly upregulated in the smokers compared with the nonsmokers, many more than have been previously linked to smoking. There was minimal intra-individual variability in the level of expression of these genes between the right and left lung or over time, indicating that the expression level of the genes assessed in this study represents a repeatable characteristic of the individual. However, there was considerable inter-individual variability among

the smokers in the expression levels of the 16 upregulated genes. Further, there was evidence of clustering of the levels of antioxidant-related gene expression among the smokers, suggesting that there are subgroups of individuals whose airway epithelium responds differently to cigarette smoking. Based on these observations, we propose that there is genetic variability in the antioxidant defenses of the airway epithelium in response to the stress of cigarette smoking, and that those antioxidant-related genes that demonstrate the most marked variability should be considered as candidates for future epidemiologic studies assessing the role of genetic variation among cigarette smokers in the risk for the development of chronic bronchitis.

Materials and Methods

Study Individuals

Two groups of individuals were recruited using advertisements in local newspapers: phenotypically normal nonsmokers and phenotypically normal chronic smokers with an approximate history of 20 pack-years of smoking, but no history of current respiratory tract infection, chronic bronchitis, or lung cancer. The study was approved by the Weill Cornell Medical College Institutional Review Board, and written informed consent was obtained from each individual before enrollment in the study.

To ensure that both groups fit the enrollment criteria, individuals underwent an initial screening evaluation, which included a history of smoking habits, respiratory tract symptoms and prior illnesses, a complete physical exam, blood studies, urine analysis, chest roentgenogram, electrocardiogram, and lung function tests. The blood studies included blood cell counts, coagulation parameters, serum electrolytes, liver and kidney function tests, serum evaluation for human immunodeficiency virus antibodies, hepatitis profile (A, B, and C), anti-nuclear antibodies, sedimentation rate, and rheumatoid factor. Screening evaluation relevant to smoking habits included the urinary levels of nicotine and its derivative cotinine, and serum levels of carboxyhemoglobin. Upon completion of the baseline evaluation, all individuals who met the inclusion/exclusion criteria underwent fiberoptic bronchoscopy to obtain airway epithelial cells. Smokers were instructed not to smoke following the evening before undergoing bronchoscopy.

Collection of Airway Epithelial Cells

Fiberoptic bronchoscopy was performed to collect airway epithelial cells using methods developed in our laboratory to ensure the extraction of high quality RNA for gene expression analysis (9–11). A 1-mm disposable brush (Wiltek Medical, Inc., Winston-Salem, NC) was advanced through the working channel of the bronchoscope. The airway epithelial cells were obtained by gently gliding the brush back and forth on the airway epithelium 5–10 times in 10 different locations in the same general area. Two independent samples of airway epithelium were obtained from each individual, from the third branching of the bronchi in the left and right lower lobes. Four individuals, including one nonsmoker and three smokers, were reevaluated at intervals of 3–7 mo to assess variability in gene expression over time. The cells were detached from the brush by flicking into 5 ml of ice-cold LHC8 medium (Invitrogen, Carlsbad, CA). An aliquot of 0.5 ml was kept for differential cell count. The remainder (4.5 ml) was immediately processed for RNA extraction. Total cell number was determined by counting on a hemocytometer. Differential cell count (epithelial versus inflammatory cells) was assessed on sedimented cells prepared by cyto-centrifugation (Cytospin 11; Shandon Instruments, Pittsburgh, PA)

stained with DiffQuik (Baxter Healthcare, Miami, FL). The samples from the right and left lungs were kept independent of each other throughout subsequent analyses. For a subset of three smokers and three nonsmokers, lung biopsy specimens were collected and processed for histologic examination using hematoxylin and eosin staining. The sections were evaluated for the presence of inflammatory cells, percentage of basal cells, goblet cells, and intermediate and ciliated cells.

Preparation of cDNA and Hybridization of cRNA

We chose to use microarray analysis to assess the level of expression of genes encoding antioxidant-related proteins rather than Northern analysis, RNase protection, or quantitative PCR-based methods because microarrays offer a strategy to assess the expression level of a large number of genes simultaneously using a small RNA sample. This is of critical importance for this study because it rests on assessing mRNA levels of a large number of antioxidant-related genes in freshly obtained samples of pure airway epithelium from volunteer normal non-smokers and smokers. The airway brushing methodology via fiberoptic bronchoscopy typically yields only 10 to 15 μg RNA from normal individuals (10). While this is adequate for microarray methodology, other methods of quantification of mRNA levels require far more RNA to assess this number of genes.

All analyses were performed with the Affymetrix HuGeneFL chip using the protocols from Affymetrix (Santa Clara, CA). Total RNA was extracted from the brushed cells using TRIzol (Life Technologies) followed by RNeasy (Qiagen, Valencia, CA) to remove residual DNA, a procedure yielding $\sim 2 \mu\text{g}$ from 10^6 cells. First-strand cDNA was synthesized using the T7-(dT)₂₄ primer (sequence 5'-GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG-[dT]₂₄-3', HPLC purified, from Oligos Etc., Wilsonville, OR) and converted to double stranded cDNA using Superscript Choice system (Life Technologies). Double-stranded cDNA was purified by phenol chloroform extraction and precipitation, and the size distribution examined after agarose gel electrophoresis. This material was then used for synthesis of the biotinylated RNA transcript using the BioArray HighYield reagents (Enzo, New York, NY), purified by the RNeasy kit (Qiagen) and fragmented immediately before use. As specified by Affymetrix, the labeled cRNA was first hybridized to the test chip and then, when satisfactory, to the HuGeneFL GeneChip, for 16 h. The GeneChips were processed by the fluidics station under the control of the Microarray Suite software (Affymetrix) to receive the appropriate reagents and washes for detection of hybridized biotinylated cRNA and then manually transferred to the scanner for data acquisition.

Microarray Data Analysis

As recommended by Affymetrix, the image data on each individual microarray chip was scaled to an arbitrary target intensity, using the Microarray Suite software (version 5.1). Data analysis was performed on all 44 microarrays representing the airway epithelium from the right and left lungs of the 22 individuals (9 nonsmokers and 13 smokers, first visit), plus 8 microarrays representing the 4 individuals who were sampled during a second visit (1 nonsmoker and 3 smokers). The raw data was normalized using the GeneSpring software (Silicon Genetics, Redwood City, CA). All 52 microarrays passed GeneSpring quality acceptance criteria. Normalization was performed using the default normalization parameters recommended by the software, as follows: (i) per microarray sample, by dividing the raw data by the 50th percentile of all measurements; and (ii) per gene, by dividing the raw data by the

median of the expression level for the gene in all samples. Data from probe sets representing genes that failed the Affymetrix Detection criterion (labeled "Absent" or "A," or "Marginal" or "M") in all 52 microarrays were eliminated from further assessment. All further analyses were performed on the remaining 4,512 genes selected using this criteria.

Clustering and tree building programs (45) were used to compare the overall gene expression patterns among samples from smokers and nonsmokers. These analyses used only the samples from each individual's first visit, using one value per probe set (resulting from averaging the value obtained from the right and left lung samples). To further evaluate global expression patterns of gene expression in airway epithelium, global gene expression patterns were compared by location (right lung versus left lung in the same individual); among different individuals with similar smoking status (nonsmoker compared with nonsmoker; smoker compared with smoker); and nonsmoker compared with smoker. These analyses used the 921 genes that were identified as "Present" (Affymetrix Detection "P") in all of the microarray experiments. The data was normalized per microarray sample only, as described above. The r^2 value was used as an index of the overall similarity between two samples.

The intra-individual variability in global gene expression pattern over time was assessed by comparing the expression levels of the 921 genes expressed in samples obtained from a given individual from separate visits. Four individuals were assessed by repeated bronchoscopy and airway brushing, at intervals of 3–7 mo (1 nonsmoker, NS1, and 3 smokers, S2, S3, and S6).

The antioxidant-related genes were identified by biological pathway searches of the Affymetrix annotated databases (www.affymetrix.com) for the HuGeneFL chip, and also searching the gene descriptions in the annotated lists by key words. These genes were grouped into four subcategories: (i) catalase and superoxide dismutases; (ii) glutathione pathway; (iii) redox balance; and (iv) pentose phosphate pathway.

Statistical Analysis

The demographic data for nonsmokers and smokers were assessed for age, cell yield, % epithelial and nonepithelial cells using Student's two-tailed t test. Sex and race of the smokers and nonsmokers were compared by Chi-square test. Overall correlation of global gene expression pattern and expression patterns of the 44 antioxidant-related genes between two samples (for example, between left and right of the same individual, between two individuals, or between the same individual on two visits) was assessed using the r^2 of the simple linear regression plot of the expression values normalized by array only. Assessment of the effects of age, sex, and smoking on which genes were significantly changed was done using the GeneSpring software, by calculating the P values using the Welch t test and the Benjamini-Hochberg correction for False Discovery Rate (46). These analyses were done on the 4,512 genes that passed the Affymetrix Detection criterion (labeled "Present") in at least one microarray, and genes were assumed to be significantly upregulated or downregulated if the calculated P value was < 0.05 . The mean value of expression of a given gene was calculated for each individual as the average of expression levels in the left and right lung. The mean value of expression of a given gene in smokers and in nonsmokers was used to determine the fold change in gene expression between the two groups. The coefficient of variation (CV) of expression level of each antioxidant-related gene was calculated for smokers and nonsmokers as the standard deviation divided by the mean expression level, and reported as a percentage. The average CV for the 16 antioxidant-related genes that were significantly upregulated in the smokers was compared with

TABLE 1
*Study population and airway epithelial samples**

Parameter	Nonsmoker	Smoker
n	9	13
Sex, M/F	7/2	9/4
Age, yr	39 ± 13	38 ± 8
Race, B/W/H	5/4/0	6/6/1
Smoking history, pack-years	0	21 ± 9
Urine nicotine, ng/ml [†]	8 ± 7	3,493 ± 1,110
Urine cotinine, ng/ml [†]	Negative	1,118 ± 457
Venous carboxyhemoglobin, % [‡]	N.D.	4 ± 2
Pulmonary function parameters [§]		
FVC	102 ± 3	102 ± 3
FEV ₁	102 ± 2	102 ± 3
FEV ₁ /FVC	101 ± 3	99 ± 2
TLC	103 ± 3	99 ± 3
D _{LCO}	88 ± 2	85 ± 3
D/VA	96 ± 2	93 ± 4
Epithelial cells		
Total number recovered (×10 ⁶)	9.4 ± 3.5	8.5 ± 3.9
% epithelial cells	99 ± 1	99 ± 1
% inflammatory cells	1 ± 1	1 ± 1

Definition of abbreviations: B, Black; D_{LCO}, total diffusing capacity; D/VA, diffusing capacity corrected for alveolar volume; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; H, Hispanic; N.D., not determined; TLC, total lung capacity; W, White.

* Demographic characteristics of 22 individuals in the study. The epithelial cell data represent the samples used for the 44 microarrays, 2/individual, collected during first visit. Data are presented as mean ± SE.

[†] Urine nicotine and cotinine used as screen to insure current smoking: > 200 = active smoker, 50–200 = passive smoker; < 50 = nonsmoker. Data represent the mean of two determinations from the day of the initial screening and day of bronchoscopy.

[‡] Determined for 8 of 13 smokers; venous carboxyhemoglobin was used as a secondary marker of current smoking; nonsmokers $< 1.5\%$.

[§] Pulmonary tests are expressed as % of predicted.

the average CV for the 28 antioxidant-related genes that were not significantly altered by smoking using the two-tailed Student's t test, assuming unequal variance (due to the unequal sample sizes).

Results

Study Population and Airway Epithelium Samples

All individuals were determined to be phenotypically normal based on standard history, physical exam, routine blood and urine studies, chest X-ray, EKG, and pulmonary function tests (Table 1). The smoker and nonsmoker groups were similar in regard to sex, age, and race ($P > 0.4$, all comparisons). Urine nicotine, urine cotinine, and venous carboxyhemoglobin levels verified that the individuals who gave a history of current smoking were current smokers, and that those reporting nonsmoking were nonsmokers. The cells recovered by bronchial brushing contained four major epithelial cell types, including: ciliated, basal, secretory and undifferentiated cells, and a small number of inflammatory cells. In all cases, $> 10^6$ epithelial cells were obtained, which is more than sufficient for the microarray analysis. The total number of cells recovered and percentage of epithelial cells within the samples were similar among the smoking and nonsmoking groups ($P > 0.3$, all comparisons). Hematoxylin and eosin staining of biopsy sections examined

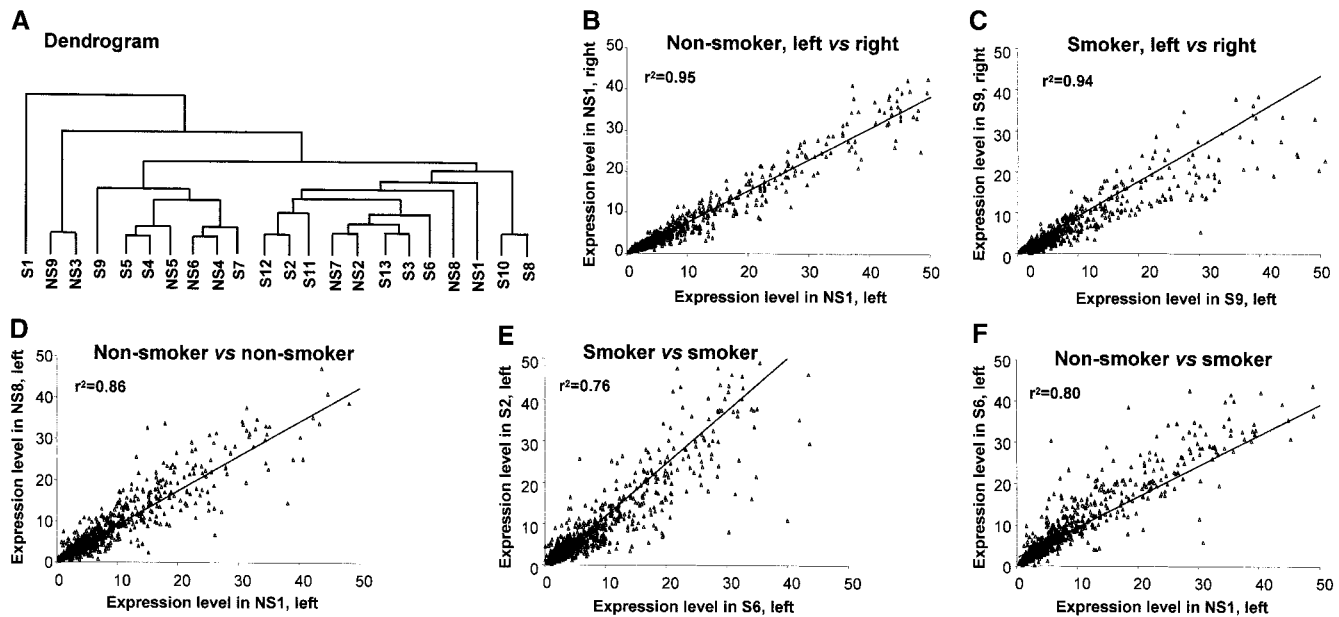


Figure 1. Assessment of the relationship of global airway epithelium gene expression pattern among different individuals. (A) Dendrogram. (B–F) Variability of airway gene epithelium gene expression level within and between individuals. (A) Normalized expression levels for each gene of the 4,512 expressed genes were analyzed using the Cluster program (45) by hierarchical clustering with the complete linkage clustering method. The clustered data was then exported to TreeView (45) for visualization. Nonsmokers (1–9) are indicated as “NSX,” and smokers (1–13) as “SX,” where “X” is the number assigned to the individual. For B–F, expression level, normalized by array, for the 921 genes scored “Present” on all microarrays are plotted for representative comparisons. The lines represent the linear least squares best fit. (B) Comparison of gene expression levels in the airway epithelium from the left and right lower lobe of the same nonsmoker (NS1). (C) Comparison of gene expression levels in the airway epithelium for the left and right lower lobe of the same smoker (S9). (D) Comparison of gene expression levels in the airway epithelium from the lower left lobe of two different nonsmokers (NS8 versus NS1). (E) Comparison of gene expression levels in the airway epithelium from the lower left lobe of two different smokers (S2 versus S6). (F) Comparison of gene expression levels in the airway epithelium from the lower left lobe of two different individuals, a nonsmoker (NS1) and a smoker (S6).

at low and high power did not reveal any differences in terms of number of inflammatory cells, percentage of basal cells, goblet cells, or intermediate or ciliated cells between smokers and nonsmokers.

Global Assessment of Gene Expression

The HuGeneFL chip includes probes representing ~ 5,000 full-length human genes. In preliminary analyses of the 22 study individuals, 921 genes were identified as expressed (“Present” in the Affymetrix definition) on all chips for all individuals, 4,512 were identified as expressed on at least 1 chip among the 22 individuals, and 44 genes identified to be in the category of antioxidant-related genes were expressed on at least 1 chip. For each of the analyses of the data, the number of genes analyzed is specifically noted.

Global analysis of gene expression by cluster analysis did not segregate the smokers from the nonsmokers (Figure 1A).

Assessments of global airway epithelial gene expression comparing the right to left lung in the same individual (nonsmokers or smokers) and among different individuals (nonsmoker versus nonsmoker, smoker versus smoker, and nonsmoker versus smoker) revealed highly significant ($P < 0.001$) positive correlations for all comparisons (Figures 1B–1F). For example, expression levels of airway epithelium sampled on the same day from the right and left lower lobe

of a representative nonsmoker (Figure 1B) and right and left lower lobe of a smoker (Figure 1C) showed high correlations, with r^2 values of 0.95 and 0.94, respectively. Correlations of global gene expression levels between two individuals were significant ($P < 0.001$, all comparisons) when comparing two nonsmokers ($r^2 = 0.86$; Figure 1D), two smokers ($r^2 = 0.76$; Figure 1E), or a smoker versus a nonsmoker ($r^2 = 0.80$; Figure 1F). Not surprisingly, the correlation coefficients were greater in intra-individual comparisons than in inter-individual comparisons. Importantly, comparison of the expression levels (average of right and left lung) assessed at the first visit and second visit for the same individual showed a high degree of correlation between the two samples in all cases (Figure 2). Interestingly, the r^2 value for the only nonsmoker that was sampled in two separate visits was 0.78, compared with r^2 values ≥ 0.88 for all three smokers. This individual was also the one with the longest time elapsed between visits (7 mo as opposed to 5, 4, and 3 mo for the three smokers). Because only four individuals were available to sample on two separate occasions, separated by variable periods of time, and because all r^2 values show a high degree of correlations between visits, the biological significance of this observation (if any) cannot be determined at this point.

Assessment of the effect of age, sex, and cigarette smok-

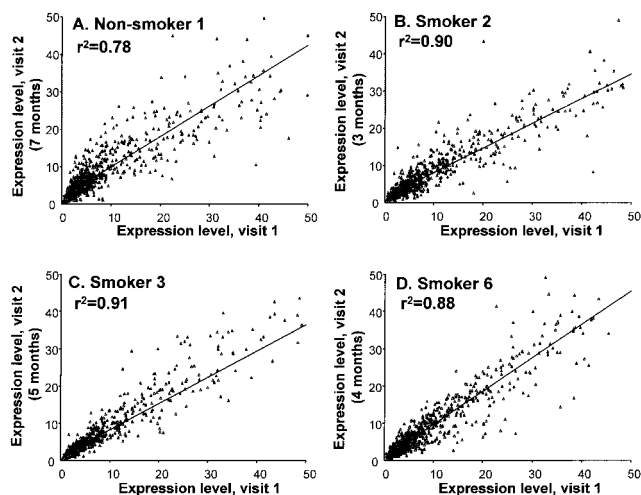


Figure 2. Variability in individual global gene expression pattern as a function of time. The gene expression levels of the 921 genes present in all microarrays in visit 1 were plotted against gene expression in visit 2 of the same individual. The expression levels are average from the left and right lung samples. (A) Comparison of gene expression levels in the airway epithelium from two visits, 7 mo apart, of nonsmoker NS1. (B) Comparison of gene expression levels in the airway epithelium from two visits, 3 mo apart, of smoker S2. (C) Comparison of gene expression levels in the airway epithelium from two visits, 5 mo apart, of smoker S3. (D) Comparison of gene expression levels in the airway epithelium from two visits, 4 mo apart, of smoker S6. The linear least squares best fit line and correlation levels expressed as r^2 are indicated in each panel.

ing on the expression level of the 4,512 expressed genes demonstrated no significant correlations with age (Figure 3A) or sex (except for a ribosomal Y-linked protein; Figure 3B). In contrast, when current smoking status was used as the variable for comparison of gene expression patterns, a total of 98 genes were identified that were differentially expressed between the smokers and the nonsmokers at a significant level (Figure 3C). The 98 genes differentially expressed in the airway epithelium of smokers compared with nonsmokers fell into various categories of biological function, including genes involved in the detoxification of xenobiotics (some of which were previously described to be upregulated by cigarette smoke, such as cytochrome P450 1A1 and 1B1) (12, 13), genes potentially involved in carcinogenesis, and the antioxidant-related genes. Of the 98 genes that were differentially expressed at a statistically significant level, 55 were upregulated in smokers by at least 1.5-fold, and 11 were downregulated by at least 1.5-fold in smokers compared with nonsmokers.

Expression of Antioxidant-Related Genes

The 44 antioxidant-related genes on the HuGeneFL chip, identified by searching the Affymetrix annotated databases for that chip, were subgrouped into 4 categories, including: (i) catalase and superoxide dismutases; (ii) glutathione metabolism; (iii) redox balance; and (iv) pentose phosphate pathway (Table 2). Among these 44 genes, 16 were upregu-

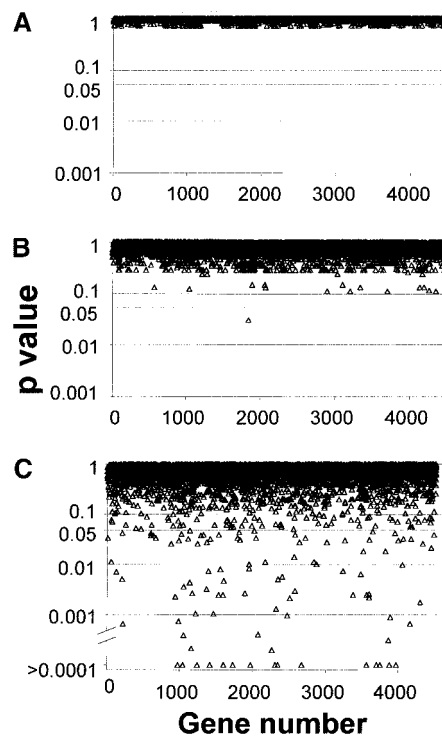


Figure 3. Significance of age (A; ≥ 40 yr old versus < 40 yr old), sex (B), and smoking (C) on expression of genes in the airway epithelium. For the 4,512 genes called “Present” in at least 1 of the microarrays, P values were calculated using the Welch t test, with Benjamini-Hochberg correction for false discovery rate (46). Ordinate, P value for each gene; abscissa, arbitrary gene number from 1 to 4,512, derived from the alphabetical order of the GenBank or TIGR identifiers, left to right.

lated in the airway epithelium of smokers, with a P value of < 0.05 (Table 2), and none were significantly downregulated.

In agreement with previously published observations indicating that neither catalase nor the superoxide dismutases are upregulated in the human lung in response to oxidative stress (14, 15), the expression levels of the genes coding for these enzymes in the study population were not elevated in smokers versus nonsmokers (0 of 4 genes in this category; Table 2, Figure 4A). In contrast to catalase and the superoxide dismutases (SODs), 33% (7 of 21) of the genes encoding enzymes of the glutathione pathway were upregulated in the airway epithelium of smokers compared with nonsmokers. The most dramatic example was that observed for the glutathione peroxidase 2 gene (Table 2, Figure 4B). In addition to genes of the glutathione pathway, 57% (4 of 7) of the genes involved in redox balance were significantly upregulated in the airway epithelium of smokers, including alcohol dehydrogenase 7, an enzyme involved in the synthesis of retinoic acid (vitamin A), an antioxidant (16); two members of the aldo-keto reductase family 1, including C3, which catalyzes the reduction of dehydroascorbate into ascorbic acid, and B1, which metabolizes toxic aldehydes generated by lipid peroxidation, thus serving as an antioxi-

TABLE 2
*Antioxidant-related gene expression in phenotypically normal smokers**

Category	Gene ID	Description	Ratio Smokers/Nonsmokers	P value [†]
Catalase, superoxide dismutase	X04085	Catalase	1.36	0.26520
	X02317	Superoxide dismutase 1	1.26	0.28280
	J02947	Superoxide dismutase 3	1.17	0.52690
	X65965	Superoxide dismutase 2	0.80	0.30020
Glutathione metabolism	X68314	Glutathione peroxidase 2	5.09	< 0.00001
	L35546	Glutamate-cysteine ligase (regulatory subunit)	2.50	0.03952
	X15722	Glutathione reductase	2.31	0.03122
	M90656	Glutamate-cysteine ligase (catalytic subunit)	2.29	0.00040
	D00632	Glutathione peroxidase 3	1.86	0.00710
	X69433	NADP⁺ mitochondrial isocitrate dehydrogenase 2	1.48	0.00089
	M16594	Glutathione S-transferase A2	1.48	0.00247
	M96233	Glutathione S-transferase M4	1.54	0.20900
	Z84718	Glutathione S-transferase theta 1	1.60	0.25400
	U46499	Microsomal glutathione S-transferase 1	1.48	0.16310
	J05459	Glutathione S-transferase M3	1.44	0.38330
	X65727	Glutathione S-transferase A2, alternative transcript	1.30	0.17720
	U34683	Glutathione synthetase	1.27	0.08468
	U90313	Glutathione-S-transferase like	1.17	0.54650
	Y00433	Glutathione peroxidase 1	1.15	0.61390
	X71973	Glutathione peroxidase 4	1.10	0.60840
	X76648	Glutaredoxin	1.08	0.82590
	L38503	Glutathione S-transferase theta 2	1.06	0.88580
	U77604	Microsomal glutathione S-transferase 2	1.03	0.91830
	U86529	Glutathione S-transferase zeta 1	1.01	0.91830
M24485	Glutathione S-transferase pi	0.97	0.98250	
Redox balance	X76342	Alcohol dehydrogenase 7	4.35	0.00001
	D17793	Aldo-keto reductase 1C3	4.11	0.00006
	X15414	Aldo-keto reductase 1B1	2.78	0.00217
	X91247	Thioredoxin reductase 1	2.20	0.00065
	X77584	Thioredoxin	1.39	0.07244
	U25182	Peroxiredoxin 4	1.16	0.54650
	J02888	NADH menadione oxidoreductase 2, dioxin-inducible	1.10	0.65780
	L19437	Transaldolase 1	2.89	0.00103
Pentose phosphate cycle	U30255	Phosphogluconate dehydrogenase	2.00	0.00002
	L12711	transketolase	1.99	0.00022
	X55448	Glucose 6 phosphate dehydrogenase	1.85	< 0.00001
	X12447	Fructose-bisphosphate aldolase A	1.64	0.00247
	D25328	Phosphofructokinase, platelet	1.46	0.41430
	K03515	Glucose phosphate isomerase	1.27	0.19570
	M83088	Phosphoglucomutase 1	1.12	0.53720
	Y00971	Phosphoribosyl pyrophosphate synthetase 2	1.06	0.86100
	J04794	Aldo-keto reductase 1A1	1.04	0.88330
	D00860	Phosphoribosyl pyrophosphate synthetase 1	0.99	0.94040
U24183	Phosphofructokinase, muscle	0.97	0.97270	

* For each gene in the antioxidant-related categories, shown is the ratio of gene expression for smokers to nonsmokers (average for the 13 smokers, divided by an average of 9 nonsmokers).

[†] P value based on Welch *t* test, using the left and right lungs of an individual as replicate samples; those genes with *P* < 0.05 difference among smokers and nonsmokers are in bold.

dant defense (17); and thioredoxin reductase 1, a member of the pyridine nucleotide-disulfide oxidoreductase family involved in maintaining redox balance (18) (Table 2, Figure 4C). Finally, 42% (5 of 12) of the genes coding for enzymes of the pentose phosphate pathway were upregulated in the airway epithelium of smokers, including glucose-6-phosphate dehydrogenase, phosphogluconate dehydrogenase, transketolase, transaldolase 1, and fructose-bisphosphate aldolase A (Table 2, Figure 4D).

Variability in Expression of Antioxidant-Related Genes Within and Among Individuals

One of the central objectives of this study is to examine not only the extent to which smokers show changes in antioxidant-related gene expression in airway epithelium, but also to assess the variability of the smoking-induced changes in gene expression within and among individuals. For many of the antioxidant-related genes observed to be upregulated in the smokers, there was marked inter-individual variability

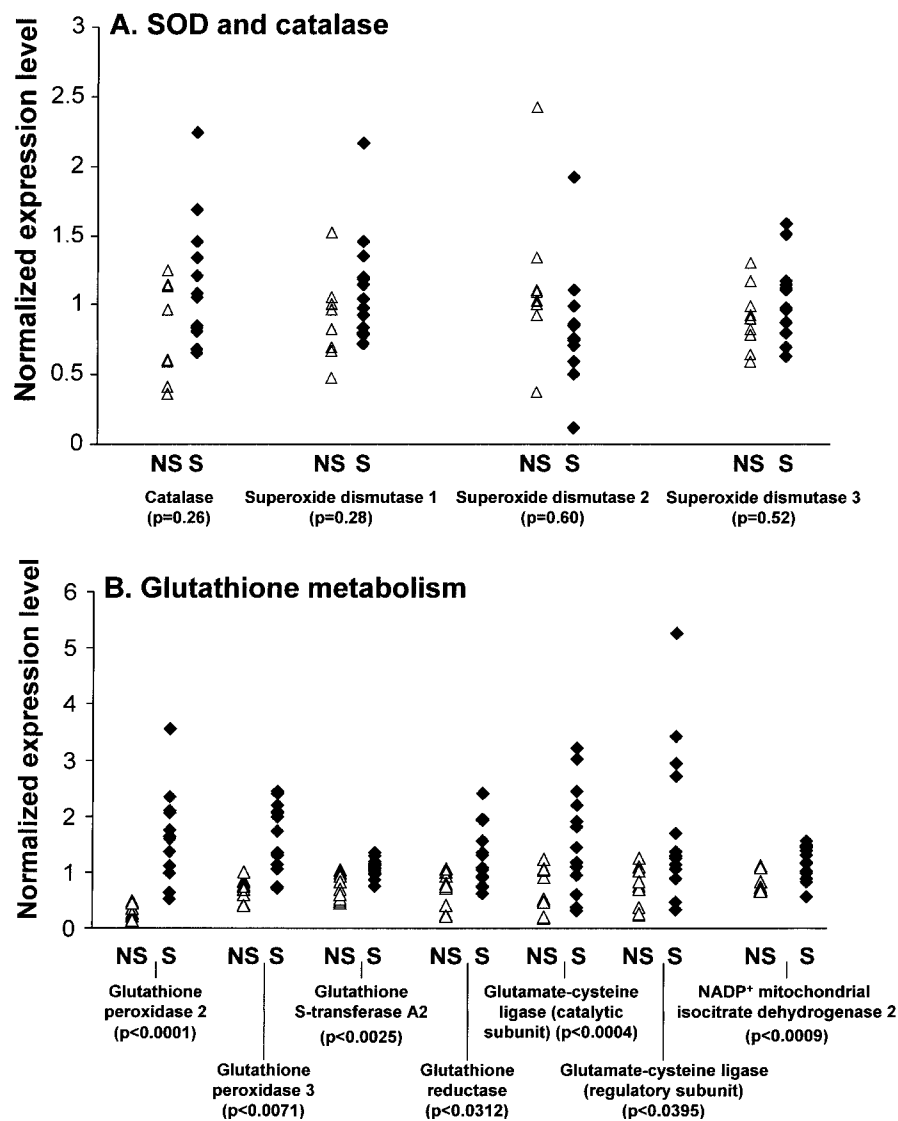


Figure 4. Examples of relative expression levels of antioxidant-related genes in non-smokers and smokers. (A) Catalase and superoxide dismutases. (B) Glutathione pathway. (C) Redox balance. (D) Pentose phosphate pathway. For all panels, the values plotted represent the average of the normalized expression level in the right and left lung for each individual, including 9 nonsmokers (*open triangles*, NS) and 13 smokers (*filled diamonds*, S). The identity of the gene and *P* value are indicated below the abscissa. (A) Relative expression levels of genes encoding catalase and superoxide dismutases 1–3 in the airway epithelium of smokers and nonsmokers. Note that although these genes are not upregulated in a statistically significant fashion, there is inter-individual variability in expression levels among nonsmokers and smokers, and four smokers have levels of catalase expression above the highest level observed in the nonsmokers. (B) Relative expression levels of genes encoding enzymes of the glutathione pathway observed to be significantly upregulated in the airway epithelium of smokers compared with nonsmokers. Only genes with $P < 0.05$ are represented in this figure. Note the broad inter-individual variability in the expression levels of these genes among different smokers. (C) Relative expression levels of genes encoding enzymes involved in redox balance of antioxidant systems. Note that smokers have significantly greater expression levels, and show greater variability in expression levels for several of these genes. One of the genes, thioredoxin, is not significantly different between smokers and nonsmokers, although it approaches significance. (D) Relative expression levels of genes encoding enzymes involved in the pentose phosphate pathway. Note that smokers show significantly greater expression than nonsmokers for five of the enzymes.

in the expression levels of these genes among different smokers, with some smokers expressing levels similar to that of the nonsmokers, and others expressing variable degrees of higher levels.

As a control, to assess whether the expression level of smoking-induced antioxidant-related genes varied from visit to visit within one single individual, microarray data was assessed from one nonsmoker (nonsmoker 1) and from three smokers (smokers 2, 3, and 6) who had repeat fiberoptic bronchoscopy and brushing of the airway epithelium at intervals of 3–7 mo. Global gene expression showed excellent intra-individual correlations from visit 1 to visit 2 (Figure 2). Assessment of individual repeatability for antioxidant-related genes (first visit data for all 44 antioxidant-related genes compared with second visit data) demonstrated good correlation ($P < 0.001$, all comparisons; Figures 5A–5D).

These results suggest that the expression levels of antioxidant-related genes did not vary greatly with time within the same individual. To examine intra-individual variability in expression levels of the 44 antioxidant-related genes at one given time, expression levels in the left lung were compared with expression levels found in the right lung in the same 4 individuals (Figures 5E–5H). The high correlations ($P < 0.001$, all cases) demonstrate remarkable consistency in antioxidant-related gene expression levels in different locations within the lungs of the same individuals.

Having identified the specific antioxidant-related genes that were upregulated in smokers, and having confirmed the consistency in expression levels within individuals over time or between two airway locations, the extent of inter-individual variability in expression level among smokers was determined. To quantify this inter-individual variability

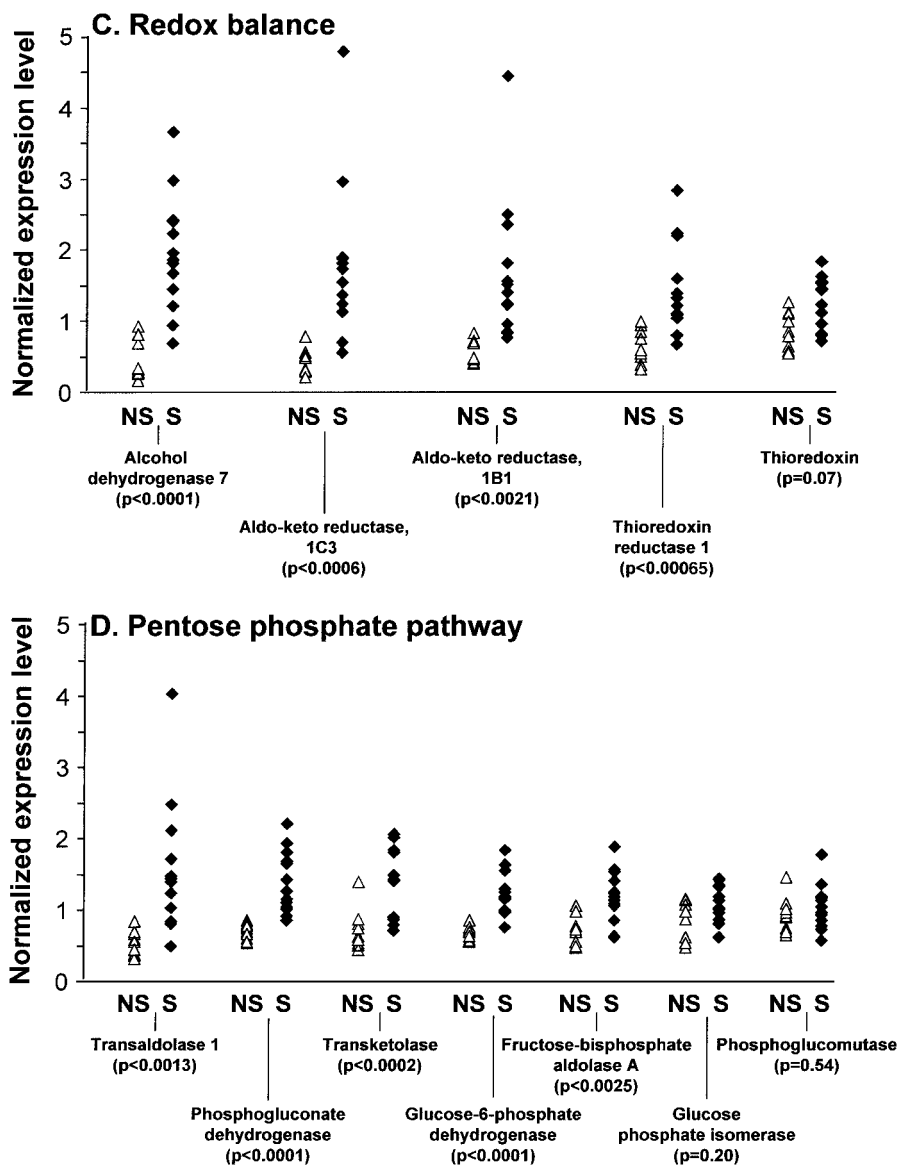


Figure 4. Continued.

in the smokers, the coefficient of variation (CV, expressed as a percentage) of normalized expression levels was assessed for each of the 44 antioxidant-related genes. For the 16 genes that were significantly upregulated in smokers, the average CV was 43.4%, a value that was significantly greater ($P < 0.025$) than the average CV in the 28 genes that were not significantly upregulated in the smokers (31.5%). These data suggest that smokers not only show altered expression levels for these genes compared with nonsmokers, but that the extent of inter-individual variability in the smokers is greater in these genes than in genes that are not upregulated by smoking.

For the 16 of the 44 antioxidant-related genes that were upregulated in airway epithelium in association with cigarette smoking, the CV ranged from 15–75%, i.e., some genes showed greater variability in the levels of expression than others (Table 3). Five genes had CV of $> 50\%$, including glutamate-cysteine ligase (regulatory subunit) (75.3%),

aldo-keto reductase 1C3 (65.5%), aldo-keto reductase 1B1 (60.8%), glutathione reductase (59.8%), and transaldolase 1 (58.7%). The observed variability in gene expression levels is not likely to be due entirely to inherited genetic differences among individuals, and could be affected by variables such as smoking history. While it is not possible to control for all potentially confounding variables, none of the 16 antioxidant genes that were upregulated in the airway epithelium in association with cigarette smoking showed a significant correlation with the age of the smoker, and none of the r^2 values exceeded 0.15. Among the same 16 genes, only glutathione peroxidase 3 showed a significant correlation ($r^2 = 0.32$, $P = 0.024$) between gene expression level and smoking history (in pack-years). Hence, little of the observed variability is explained by these potentially confounding variables.

Analysis of the pattern of expression of the 16 statistically significant antioxidant-related genes by hierarchical clustering

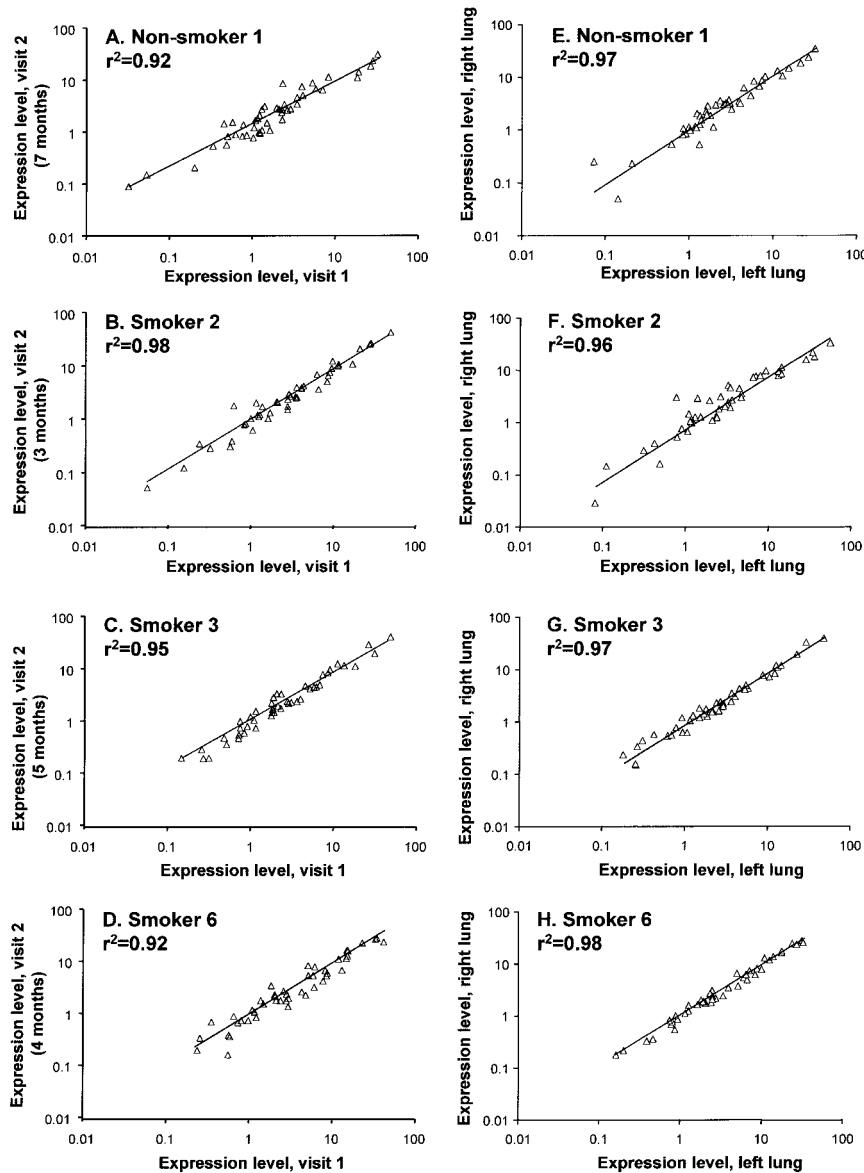


Figure 5. Correlation between first and second visit, and between right and left lung, in the expression levels of the 44 antioxidant-related genes in four individuals followed over time. Each point in each panel represents a single gene, and the data is plotted on log axes for clarity. In the comparisons of two visits (A–D), values from right and left lungs were averaged for each visit. (A) Nonsmoker 1, 7 mo between evaluations. (B) Smoker 2, 3 mo between evaluations. (C) Smoker 3, 5 mo between evaluations. (D) Smoker 6, 4 mo between evaluations. (E–H) Values from the left and right lung samples from the same four individuals, collected on the second visit. The high r^2 values for individuals at different times (A–D; range, 0.92–0.98) and from the two lungs of the same individual at the same time (E–H; range, 0.96–0.98) indicate the consistency in expression levels for these genes within individuals over time and between left and right lungs.

analysis also suggested inter-individual variability, and revealed subgroups of individuals with similar patterns of expression, distinct from other subgroups (Figure 6). The hierarchical clustering analysis included data from both visits for those individuals subjected to repeated sampling and the first and second visit data clustered very close to each other for three of the four individuals, providing another illustration of good repeatability of gene expression levels within individuals. In the context of comparison among different individuals, some smokers had high expression levels of a subset of genes, whereas other smokers had low expression levels for that particular subset. For example, smokers 6, 8, and 10 upregulated a set of these 16 genes, including GPX3, TXNRD1, GCLR, AKR1C3, TALDO1, AKR1B1, GPX2, and ADH7 (abbreviations are described in Figure 6 legend), whereas smokers 7 and 11 failed to upregulate many of these genes. These data suggest that the pattern of upregulation of antioxidant-related gene ex-

pression in the airway epithelium of healthy smokers is variable among individual smokers, with some individuals showing similar patterns to one another, for a specific group of genes.

Discussion

Although cigarette smoking is the major risk factor for developing chronic bronchitis, only 15–20% of smokers develop clinically significant airflow obstruction, implicating the involvement of genetic factors in determining susceptibility to this airway disease (1–8). The focus of this study is on airway epithelial expression of antioxidant-related genes, the extent to which the expression levels of these genes are altered in smokers compared with nonsmokers, and the magnitude of inter-individual variability in the expression levels among smokers. The hypothesis is that there is inter-individual variability in airway epithelial

TABLE 3

*Genes that show significantly greater expression levels in smokers than nonsmokers, ranked by the coefficient of variation in smokers**

Gene ID	Description	Coefficient of Variation in Smokers [†]
L35546	Glutamate-cysteine ligase (regulatory subunit)	75.3
D17793	Aldo-keto reductase 1C3	65.5
X15414	Aldo-keto reductase 1B1	60.8
X15722	Glutathione reductase	59.8
L19437	Transaldolase 1	58.7
X68314	Glutathione peroxidase 2	48.0
X91247	Thioredoxin reductase 1	44.1
D00632	Glutathione peroxidase 3	42.4
X76342	Alcohol dehydrogenase 7	42.0
M90656	Glutamate-cysteine ligase (catalytic subunit)	37.1
L12711	Transketolase	34.1
U30255	Phosphogluconate dehydrogenase	30.7
X12447	Fructose-bisphosphate aldolase A	30.5
X69433	NADP ⁺ mitochondrial isocitrate dehydrogenase 2	25.7
X55448	Glucose-6-phosphate dehydrogenase	24.7
M16594	Glutathione S-transferase A2	15.1

* See Table 2 for total list of antioxidant-related genes.

[†] Calculated as standard deviation, divided by the mean, and expressed as a percentage.

upregulation of expression of antioxidant-related genes in response to smoking, and that the gene expression pattern of the antioxidant-related genes in the airway epithelium is a factor in susceptibility to bronchitis. Assessment of global gene expression patterns in the airway epithelium of 22 phenotypically normal nonsmokers and smokers did not segregate the nonsmokers from the smokers, suggesting that the effects of smoking on airway epithelium gene expression in healthy individuals are modest compared with the overall variability in gene expression in the airway epithelium of normal healthy humans, as measured by microarray experiments. Because this method generates a vast number of data points simultaneously for thousands of genes, it is also possible that part of this overall variability is due to the nature of the assay. In contrast, assessment of the level of expression of 44 antioxidant-related genes in these samples, demonstrated significant upregulation of 16 antioxidant-related genes in the airway epithelium of smokers compared with nonsmokers. Many of these upregulated genes have not been previously linked to smoking. Importantly, although there was minimal variation in the levels of antioxidant-related gene expression within the same individual (right lung versus left lung or within the same individual over time), there was significant variation among individuals in the extent of upregulation of the antioxidant genes in the smokers. In the context that the oxidant-antioxidant balance in the airway epithelium plays a central role in cigarette smoking-induced chronic bronchitis, we propose that among the 16 significantly upregulated antioxidant-related genes in the smoker airway epithelium, those with the most intra-individual variability in expression levels may be useful as genetic markers for future epidemiologic studies regarding susceptibility to smoking-induced chronic bronchitis.

Oxidants, Antioxidants, and the Pathogenesis of Chronic Bronchitis

Cigarette smoke, the result of tobacco combustion, is an aerosol composed of 8% particulates and 92% vaporized

chemical suspended in a gaseous medium. Cigarette smoke contains massive amounts of oxidants; it is estimated that there are 10^{14} free radicals per "puff," many of which have a half-life sufficient to reach the lower respiratory tract (19). Both the gas phase and particulate matter of cigarette smoke contain various species of oxidants and free radicals, including nitric oxide in the gas phase and a mixture of semiquinones, hydroquinones, and quinones embedded in tar (19–22). Other polyphenolic agents, such as catechol, also generate hydrogen peroxide and superoxide free radicals (20–23). This oxidative stress can overwhelm the antioxidant defenses of the lung, leading to lung injury by various mechanisms, including lipid peroxidation of epithelial cell membranes (21, 24–26). Systemic and local increases in oxidants and decreases in antioxidants have been observed in smokers and individuals with COPD (24, 27–29). For example, the forced expiratory volume in 1 s in smokers with COPD correlates with glutathione levels in lung epithelial lining fluid (ELF) (30). Increased oxidative DNA damage, measured by 8-hydroxyguanine formation in human blood leukocytes and in lung tissue, has been observed in smokers compared with nonsmokers (31, 32).

Upregulation of Antioxidant-Related Genes in the Airway Epithelium of Smokers

Of the 44 antioxidant-related genes represented on the HuGeneFL chip, 16 (36%) showed significant upregulation in the airway epithelium of smokers compared with nonsmokers. Among the 4 subgroups within the 44 antioxidant-related genes, significant upregulation was observed in 0% of the catalase/SOD genes, 33% of the glutathione pathway genes, 57% of the redox balance genes, and 42% of the pentose phosphate cycle genes represented in the HuGeneFL chip. Other than a few of the genes associated with the glutathione pathway, most of the 16 antioxidant-related genes noted to be upregulated in the airway epithelium in association with smoking in healthy individuals represent novel observations.

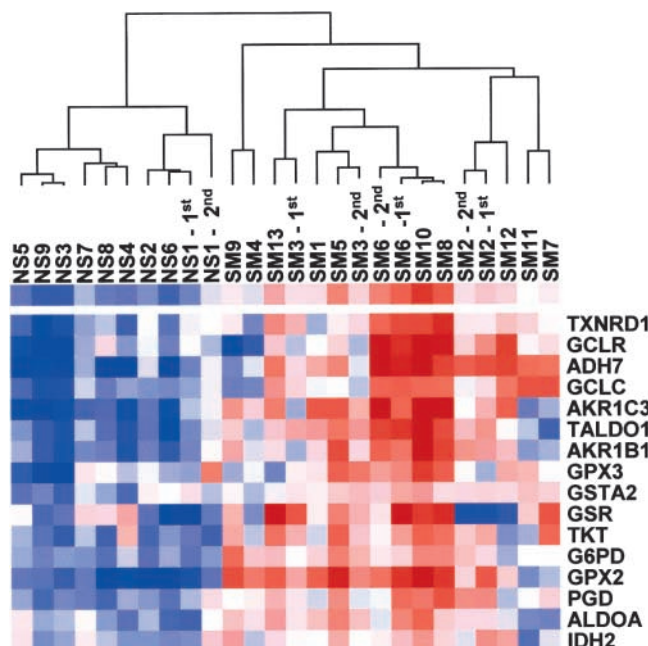


Figure 6. Cluster analysis of antioxidant-related gene expression levels in smokers and nonsmokers. The normalized, log-transformed expression levels (right and left lung average), for the 16 genes represented in *bold* in Table 2 ($P < 0.05$) were evaluated using the Cluster program (45) and subjected to hierarchical complete linkage clustering by both individual and gene. The resulting cluster was visualized with the TreeView (45) program. Genes expressed above average level are represented in *red*, genes expressed below average level are represented in *blue*. Average levels are represented in *white*. The intensity of the *red* and *blue* squares indicates the degree of up- or downregulation with respect to the average. Nonsmokers are labeled as NS followed by an individual number, and smokers are labeled as S followed by an individual number. Those individuals who were subjected to bronchoscopy and brushing on two separate visits are labeled “1st” and “2nd” and indicate each visit. As expected, since these genes were selected because their expression levels were altered in smokers compared with nonsmokers, the smokers and nonsmokers clearly segregate. Genes are labeled using the gene symbols, as follows: ADH7, alcohol dehydrogenase 7; GCLC, glutamate-cysteine ligase, catalytic subunit; GSTA2, glutathione S-transferase A2; GPX3, glutathione peroxidase 3; TXNRD1, thioredoxin reductase 1; GCLR, glutamate-cysteine ligase, regulatory subunit; AKR1C3, aldo-keto reductase 1C3; TALDO1, transaldolase 1; AKR1B1, aldo-keto reductase 1B1; IDH2, NADP⁺ mitochondrial isocitrate dehydrogenase 2; ALDOA, fructose-biphosphate aldolase A; GPX2, glutathione peroxidase 2; G6PD, glucose-6-phosphate dehydrogenase; TKT, transketolase; PGD, phosphogluconate dehydrogenase; and GSR, glutathione reductase.

Catalase/SOD genes. Although catalase and the SODs play a central role in preventing oxidants from damaging the airway epithelium (33, 34), the available evidence suggests that humans do not upregulate the catalase and SOD genes in the airway epithelium in response to oxidative stress (14, 15). The data in the present study support this concept.

Glutathione pathway genes. Glutathione is a ubiquitous tripeptide (L- γ -glutamyl-L-cysteinyl-glycine) with a sulfhy-

dryl group that enables it to protect cells against oxidants and xenobiotics (5). Together with ATP, Mg²⁺, glycine, cysteine, and glutamate, glutathione synthesis requires the presence of two enzymes: γ -glutamate cysteine ligase (GCL, also known as γ -glutamylcysteine synthase) and glutathione synthase. GCL is the rate limiting enzyme in glutathione synthesis and cysteine is the limiting substrate (35). Glutathione homeostasis is regulated by feedback inhibition exerted by glutathione on GCL and by the glutathione redox system. Glutathione converts hydrogen peroxide and lipid peroxides to H₂O and nontoxic fatty acids in a reaction catalyzed by glutathione peroxidase. In the process, glutathione is converted to its oxidized form, which is subsequently reduced by glutathione reductase. The glutathione reductase reaction requires NADPH generated from the pentose phosphate pathway (5). Hence, effective regulation of glutathione activity involves several enzymes directly involved in its synthesis and cycling, as well as other enzymes in associated pathways of intermediary metabolism, including NADP⁺ mitochondrial isocitrate dehydrogenase, phosphogluconate dehydrogenase, and glucose-6-phosphate dehydrogenase.

In contrast to catalase and SOD, the respiratory epithelial surface responds to the stress of cigarette smoking by enhancing the expression of genes of the glutathione system (5). We observed that smokers had significant upregulation of 7 of 21 of the genes related to the glutathione pathway on the HuGeneFL chip, including glutathione reductase, glutathione peroxidase 2, glutathione peroxidase 3, glutathione S-transferase A2, glutamate-cysteine ligase (regulatory subunit), glutamate-cysteine ligase (catalytic subunit), and NADP⁺ mitochondrial isocitrate dehydrogenase 2. The mRNA expression of GCL has been previously reported as elevated in the lung epithelium of smokers with COPD (36). We have previously observed that cigarette smoking is associated with enhanced levels of glutathione in the lung ELF, suggesting that activation of the elevated glutathione system in the respiratory epithelium may represent a response to the high oxidant burden in cigarette smoke (37). Studies in animal models of smoking have demonstrated elevated levels of glutathione peroxidase mRNA in the bronchial epithelium of rats exposed to cigarette smoke for up to 14 d (38), and expression of glutathione peroxidase 2 is upregulated by hyperoxia in mouse lung (39). Extracellular glutathione peroxidase (also known as glutathione peroxidase 3) has been observed to be elevated in smokers' respiratory ELF, and mRNA levels of this enzyme are elevated in airway epithelial cells recovered from smokers (40).

Redox balance genes. Among the seven redox balance-related genes on the HuGeneFL chip, four were observed to be significantly upregulated in the smokers, including alcohol dehydrogenase 7, aldo-keto reductase 1C3, aldo-keto reductase 1B1, and thioredoxin reductase 1. Although these enzymes are not classical antioxidants like the enzymes of the glutathione pathway, they are involved in the synthesis of other antioxidants such as retinoic acid and ascorbic acid, as well as in oxidative processes related to the metabolism of tobacco xenobiotics, and thus are involved in the balance between oxidation and antioxidation (16–18).

Pentose phosphate pathway genes. The pentose phosphate cycle is a pathway of glucose turnover alternative to

glycolysis, and its major functions include the production of reducing equivalents in the form of NADPH. Because regeneration of reduced glutathione requires NADPH, a product of the pentose phosphate pathway, upregulation of these enzymes in association with cigarette smoking likely represents a physiologic response to the oxidative stress of cigarette smoke. Among the 12 pentose phosphate cycle-related genes on the HuGeneFL chip, 5 were significantly upregulated in the airway epithelium of the smokers, including transaldolase 1, phosphogluconate dehydrogenase, transketolase, glucose-6-phosphate dehydrogenase, and fructose-biphosphate aldolase A. None of the genes coding for enzymes of the pentose phosphate pathway have previously been linked to the pathogenesis of the smoking-induced airway epithelial disorders, except glucose-6-phosphate dehydrogenase, which has been reported to be induced by oxidative stress in lung (41).

Genetic Susceptibility to COPD

Although the marked variability in the development of COPD in response to cigarette smoking has been recognized for many decades, the only proven genetic factor predisposing to COPD is α 1-antitrypsin (AAT) deficiency, where airflow obstruction occurs predominantly in association with smoking (42). However, AAT deficiency is responsible for only 1–2% of all cases of COPD (43). Polymorphisms in other genes proposed to be linked to susceptibility to COPD include the promoter of the tumor necrosis factor- α , aromatic-inducible cytochrome P450 subfamily 1 polypeptide 1, cystic fibrosis transmembrane conductance regulator, α 2-macroglobulin, α 1-antichymotrypsin, and endothelial nitric oxide synthase, but definitive links with these genes and COPD have not been established (3, 8). The best evidence of non-AAT deficiency-related genetic susceptibility to COPD is the observation of familial clustering of impaired lung function in association with smoking, and twin studies documenting smoking-related genetic susceptibility to airflow obstruction (3, 7, 8). One study reported that first degree relatives (including smokers and ex-smokers) of 44 probands without AAT deficiency, but with accelerated early onset emphysema, had a 3-fold increased risk for chronic bronchitis relative to smokers from other families (7). Finally, a homozygous deletion of GSTM1, a member of the mu family of glutathione S-transferases, has been associated with severe emphysema in heavy smokers (44).

Variability in Airway Epithelial Gene Expression

The variability in gene expression observed in the airway epithelium in the same individual on the same day or over time represents a combination of the variability of the methodology (e.g., collection of samples, different chips) and the inherent biological variability of gene expression of airway epithelial cells. Assessment of global gene expression, as well as the subset of antioxidant-related genes, demonstrated remarkably little variation from the right to left lung in the same individual on the same day, or in the same individual over time. This observation is important in that it demonstrates that the expression level of the genes examined in this study represents a repeatable characteristic of the individual. This is a necessary condition for using the coefficient of variation to reflect legitimate inter-individual variability,

i.e., without reproducible intra-individual expression levels, high values for the coefficient of variation among the population could be attributed largely to intra-individual noise.

Based on the extensive data implicating oxidants in cigarette smoke as the major mechanism of smoking-induced airway damage, and the role of antioxidant-related pathways in protecting the airway epithelium from the stress of cigarette smoking, it is logical to hypothesize that variability in expression of genes coding for components of the antioxidant defense system of the airway epithelium plays a role in the variability in the responses of different individuals to the chronic oxidant stress of smoking. Among the 16 antioxidant-related genes significantly upregulated in response to smoking, there was variability in the level of upregulation, with some individuals responding differently to the stress of smoking. The variability among the 16 significantly upregulated antioxidant-related genes was observed in three categories: genes of the glutathione system, redox balance, and pentose phosphate pathway.

The observations in the present study that smoking is associated with significant, but variable, upregulation of 16 antioxidant genes in the airway epithelium of chronic smokers leads us to propose that there is a genetic basis for the variability observed in the expression of these genes, many of which have not been previously considered in regard to the pathogenesis of chronic bronchitis. Among these 16 antioxidant-related significantly upregulated genes, the 5 genes showing the most variability ($CV > 50\%$) included 2 in the glutathione pathway (glutathione reductase and the regulatory subunit of glutamate-cysteine ligase), 2 in the redox balance subgroup (aldo-keto reductase 1C3, and aldo-keto reductase 1B1), and 1 in the pentose phosphate cycle (transaldolase 1). There are no prior observations of variable upregulation of airway epithelial gene expression of genes in the redox balance or pentose phosphate pathway subgroups in response to smoking. However, relevant to the hypothesis of individual variability in airway epithelial antioxidant gene expression that is central to this study, we have previously observed a wide range of respiratory ELF glutathione levels in cigarette smokers, with some ELF levels 5-fold higher than the mean level in nonsmokers, whereas other glutathione ELF levels were within the normal range (37).

We propose that these 5 genes, with CV values greater than 50%, and perhaps others among the 16 antioxidant-related genes upregulated in the airway epithelium of smokers, are potential candidate genes related to the risk of susceptibility to smoking-induced chronic bronchitis. In regard to potential genetic variations in these genes, a search of the NCBI database (<http://www.ncbi.nlm.nih.gov/SNP/>) revealed a number of single nucleotide polymorphisms (SNPs) among these 5 genes. These SNPs could potentially have an effect on airway epithelial mRNA levels for these genes, resulting in differences in expression levels due to SNP genotype, and thus playing a role in the susceptibility to chronic bronchitis.

The presence of a large number of SNPs in these genes suggests the possibility to correlate some of these coding and noncoding SNPs with variability in the expression level of individual smokers. These genes, and perhaps others among the 16 smoking-induced significantly upregulated genes, are potential candidates for future epidemiologic

studies in older individuals with COPD, to identify genes that play a role in the susceptibility of smoking-induced chronic bronchitis.

Acknowledgments: The authors thank M. Harris for assistance in recruiting subjects, J. Xiang and H. Pan for advice on microarray data analysis, and N. Mohamed and T. Virgin-Bryan for help in preparing this manuscript. These studies were supported, in part, by NIH M01RR00047, and the Will Rogers Memorial Fund, Los Angeles, CA. R.K. is supported in part by an American Lung Association, New York Lung Fellowship.

References

- American Thoracic Society. 1995. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 152:S77–S121.
- Anthonsen, N. R., et al. 1994. Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The Lung Health Study. *JAMA* 272:1497–1505.
- Sethi, J. M., and C. L. Rochester. 2000. Smoking and chronic obstructive pulmonary disease. *Clin. Chest Med.* 21:67–86.
- Repine, J. E., A. Bast, and I. Lankhorst. 1997. Oxidative stress in chronic obstructive pulmonary disease. Oxidative Stress Study Group. *Am. J. Respir. Crit. Care Med.* 156:341–357.
- Rahman, I., and W. MacNee. 1999. Lung glutathione and oxidative stress: implications in cigarette smoke-induced airway disease. *Am. J. Physiol.* 277:L1067–L1088.
- Fletcher, C., R. Peto, and C. Tinker. 1976. The natural history of chronic bronchitis and emphysema: an eight-year study of early chronic obstructive lung disease in working men in London. Oxford University Press, London.
- Silverman, E. K., et al. 1998. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. *Am. J. Respir. Crit. Care Med.* 157:1770–1778.
- Sandford, A. J., and E. K. Silverman. 2002. Chronic obstructive pulmonary disease. 1: Susceptibility factors for COPD the genotype-environment interaction. *Thorax* 57:736–741.
- Danel, C., S. C. Erzurum, N. G. McElvaney, and R. G. Crystal. 1996. Quantitative assessment of the epithelial and inflammatory cell populations in large airways of normals and individuals with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* 153:362–368.
- Russi, T. J., and R. G. Crystal. 1997. Bronchoalveolar lavage and airway brushing to investigate the human lung. *In The Lung: Scientific Foundations.* Lippencott-Raven, Inc., Philadelphia. 371–382.
- Harvey, B. G., et al. 1999. Airway epithelial CFTR mRNA expression in cystic fibrosis patients after repetitive administration of a recombinant adenovirus. *J. Clin. Invest.* 104:1245–1255.
- Willey, J. C., et al. 1997. Quantitative RT-PCR measurement of cytochromes p450 1A1, 1B1, and 2B7, microsomal epoxide hydrolase, and NADPH oxidoreductase expression in lung cells of smokers and nonsmokers. *Am. J. Respir. Cell Mol. Biol.* 17:114–124.
- Spivack, S. D., et al. 2001. CYP1B1 expression in human lung. *Drug Metab. Dispos.* 29:916–922.
- Erzurum, S. C., et al. 1993. In vivo antioxidant gene expression in human airway epithelium of normal individuals exposed to 100% O₂. *J. Appl. Physiol.* 75:1256–1262.
- Yoo, J. H., S. C. Erzurum, J. G. Hay, P. Lemarchand, and R. G. Crystal. 1994. Vulnerability of the human airway epithelium to hyperoxia. Constitutive expression of the catalase gene in human bronchial epithelial cells despite oxidant stress. *J. Clin. Invest.* 93:297–302.
- Chou, C. F., C. L. Lai, Y. C. Chang, G. Duester, and S. J. Yin. 2002. Kinetic mechanism of human class IV alcohol dehydrogenase functioning as retinol dehydrogenase. *J. Biol. Chem.* 277:25209–25216.
- Burczynski, M. E., G. R. Sridhar, N. T. Palackal, and T. M. Penning. 2001. The reactive oxygen species- and Michael acceptor-inducible human aldo-keto reductase AKR1C1 reduces the alpha,beta-unsaturated aldehyde 4-hydroxy-2-nonenal to 1,4-dihydroxy-2-nonenal. *J. Biol. Chem.* 276:2890–2897.
- Arner, E. S., and A. Holmgren. 2000. Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.* 267:6102–6109.
- Pryor, W. A., and K. Stone. 1993. Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite. *Ann. N. Y. Acad. Sci.* 686:12–27.
- Deliconstantinos, G., V. Villiotou, and J. C. Stavrides. 1994. Scavenging effects of hemoglobin and related heme containing compounds on nitric oxide, reactive oxidants and carcinogenic volatile nitrosocompounds of cigarette smoke. A new method for protection against the dangerous cigarette constituents. *Anticancer Res.* 14:2717–2726.
- Spencer, J. P., et al. 1995. DNA damage in human respiratory tract epithelial cells: damage by gas phase cigarette smoke apparently involves attack by reactive nitrogen species in addition to oxygen radicals. *FEBS Lett.* 375:179–182.
- Reilly, M., N. Delanty, J. A. Lawson, and G. A. Fitz Gerald. 1996. Modulation of oxidant stress in vivo in chronic cigarette smokers. *Circulation* 94:19–25.
- Nakayama, T., M. Kodama, and C. Nagata. 1984. Generation of hydrogen peroxide and superoxide anion radical from cigarette smoke. *Gann* 75: 95–98.
- Morrow, J. D., et al. 1995. Increase in circulating products of lipid peroxidation (F2-isoprostanes) in smokers: smoking as a cause of oxidative damage. *N. Engl. J. Med.* 32:1198–1203.
- Altuntas, I., S. Dane, and K. Gumustekin. 2002. Effects of cigarette smoking on lipid peroxidation. *J. Basic Clin. Physiol. Pharmacol.* 13:69–72.
- Ho, Y. S., et al. 2002. Lipid peroxidation and cell death mechanisms in pulmonary epithelial cells induced by peroxynitrite and nitric oxide. *Arch. Toxicol.* 76:484–493.
- Rahman, I., D. Morrison, K. Donaldson, and W. MacNee. 1996. Systemic oxidative stress in asthma, COPD, and smokers. *Am. J. Respir. Crit. Care Med.* 154:1055–1060.
- Rennard, S. I. 2002. Overview of causes of COPD: new understanding of pathogenesis and mechanisms can guide future therapy. *Postgrad. Med.* 11:28–38.
- Wouters, E. F., E. C. Creutzberg, and A. M. Schols. 2002. Systemic effects in COPD. *Chest* 121:127S–130S.
- Linden, M., et al. 1993. Airway inflammation in smokers with nonobstructive and obstructive chronic bronchitis. *Am. Rev. Respir. Dis.* 148:1226–1232.
- Asami, S., et al. 1997. Cigarette smoking induces an increase in oxidative DNA damage, 8-hydroxydeoxyguanosine, in a central site of the human lung. *Carcinogenesis* 18:1763–1766.
- van Zeeland, A. A., A. J. de Groot, J. Hall, and F. Donato. 1999. 8-Hydroxydeoxyguanosine in DNA from leukocytes of healthy adults: relationship with cigarette smoking, environmental tobacco smoke, alcohol and coffee consumption. *Mutat. Res.* 439:249–257.
- MacNee, W. 2000. Oxidants/antioxidants and COPD. *Chest* 117:303S–317S.
- Comhair, S. A., and S. C. Erzurum. 2002. Antioxidant responses to oxidant-mediated lung diseases. *Am. J. Physiol.* 283:L246–L255.
- Luo, J. L., C. S. Huang, K. Babaoglu, and M. E. Anderson. 2000. Novel kinetics of mammalian glutathione synthetase: characterization of gamma-glutamyl substrate cooperative binding. *Biochem. Biophys. Res. Commun.* 275:577–581.
- Rahman, I., et al. 2000. Localization of gamma-glutamylcysteine synthetase messenger RNA expression in lungs of smokers and patients with chronic obstructive pulmonary disease. *Free Radic. Biol. Med.* 28:920–925.
- Cantin, A. M., S. L. North, R. C. Hubbard, and R. G. Crystal. 1987. Normal alveolar epithelial lining fluid contains high levels of glutathione. *J. Appl. Physiol.* 63:152–157.
- Gilks, C. B., K. Price, J. L. Wright, and A. Churg. 1998. Antioxidant gene expression in rat lung after exposure to cigarette smoke. *Am. J. Pathol.* 152:269–278.
- Cho, H. Y., et al. 2002. Role of NRF2 in protection against hyperoxic lung injury in mice. *Am. J. Respir. Cell Mol. Biol.* 26:175–182.
- Comhair, S. A., M. J. Thomassen, and S. C. Erzurum. 2000. Differential induction of extracellular glutathione peroxidase and nitric oxide synthase 2 in airways of healthy individuals exposed to 100% O₂ or cigarette smoke. *Am. J. Respir. Cell Mol. Biol.* 23:350–354.
- Kozar, R. A., et al. 2000. Antioxidant enzymes are induced during recovery from acute lung injury. *Crit. Care Med.* 28:2486–2491.
- Crystal, R. G. 1989. The α 1-antitrypsin gene and its deficiency states. *Trends Genet.* 5:411–417.
- Mahadeva, R., and D. A. Lomas. 1998. Genetics and respiratory disease: 2. Alpha 1-antitrypsin deficiency, cirrhosis and emphysema. *Thorax* 53:501–505.
- Harrison, D. J., A. M. Cantlay, F. Rae, D. Lamb, and C. A. Smith. 1997. Frequency of glutathione S-transferase M1 deletion in smokers with emphysema and lung cancer. *Hum. Exp. Toxicol.* 16:356–360.
- Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95:14863–14868.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and approach to multiple testing. *J. R. Statist. Soc. B* 57:289–300.