

Coordinate Control of Expression of Nrf2-Modulated Genes in the Human Small Airway Epithelium Is Highly Responsive to Cigarette Smoking

Ralf-Harto Hübner,¹ Jamie D Schwartz,¹ Bishnu P De,¹ Barbara Ferris,¹ Larsson Omberg,² Jason G Mezey,² Neil R Hackett,¹ and Ronald G Crystal¹

¹Department of Genetic Medicine, Weill Cornell Medical College, New York, New York; ²Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York, United States of America

Nuclear factor erythroid 2-related factor 2 (Nrf2) is an oxidant-responsive transcription factor known to induce detoxifying and antioxidant genes. Cigarette smoke, with its large oxidant content, is a major stress on the cells of small airway epithelium, which are vulnerable to oxidant damage. We assessed the role of cigarette smoke in activation of Nrf2 in the human small airway epithelium *in vivo*. Fiberoptic bronchoscopy was used to sample the small airway epithelium in healthy-nonsmoker and healthy-smoker, and gene expression was assessed using microarrays. Relative to nonsmokers, Nrf2 protein in the small airway epithelium of smokers was activated and localized in the nucleus. The human homologs of 201 known murine Nrf2-modulated genes were identified, and 13 highly smoking-responsive Nrf2-modulated genes were identified. Construction of an Nrf2 index to assess the expression levels of these 13 genes in the airway epithelium of smokers showed coordinate control, an observation confirmed by quantitative PCR. This coordinate level of expression of the 13 Nrf2-modulated genes was independent of smoking history or demographic parameters. The Nrf2 index was used to identify two novel Nrf2-modulated, smoking-responsive genes, *pirin* (*PIR*) and *UDP glucuronosyltransferase 1-family polypeptide A4* (*UGT1A4*). Both genes were demonstrated to contain functional antioxidant response elements in the promoter region. These observations suggest that Nrf2 plays an important role in regulating cellular defenses against smoking in the highly vulnerable small airway epithelium cells, and that there is variability within the human population in the Nrf2 responsiveness to oxidant burden.

© 2009 The Feinstein Institute for Medical Research, www.feinsteininstitute.org

Online address: <http://www.molmed.org>

doi: 10.2119/molmed.2008.00130

INTRODUCTION

In cigarette smokers, phase II detoxifying and antioxidant enzymes are an important mechanism in the protection of the small airway epithelium from the more than 10¹⁴ oxidants present in each puff of cigarette smoke (1–4). Defense against oxidants is a critical cellular function, and oxidant-mediated somatic mutations play an important role in the mechanisms of malignant transformation of the airway epithelium in association with cigarette smoking (2,5–7). Central to the cellular oxidant defense

mechanism is the nuclear factor erythroid 2-related factor 2 (NFE2L2, Nrf2), a member of the “Cap ‘n’ Collar” family of the basic leucine transcription factors, known to coordinate the induction of phase II detoxifying and antioxidant enzymes (8–11). In the resting state, Nrf2 is bound to the cytoplasmic inhibitor Kelch-like epichlorohydrin-associated protein 1 (KEAP1) (12,13). Upon activation, Nrf2 dissociates from KEAP1, translocates into the nucleus, and binds to antioxidant response element sequences, resulting in transcriptional activation

of genes that help to protect the cell from oxidants (8,12,13).

Understanding of the role of Nrf2 in the lung has been advanced by studies of Nrf2 knockout mice exposed to a variety of mediators of lung injury, including bleomycin, elastase, mechanical ventilation, ovalbumin sensitization, lipopolysaccharide, and environmental stresses including hyperoxia, butylated hydroxytoluene, benz(a)pyrene, diesel exhaust particles, and cigarette smoke (14–26). Microarray analyses of lungs from Nrf2 knockout mice and wild-type controls have helped to identify Nrf2-modulated genes, for example, genes that help protect mice from the development of cigarette smoke-induced lung injuries (23).

On the basis of the knowledge that Nrf2-modulated genes respond to cigarette

Address correspondence and reprint requests to Ronald G. Crystal, Department of Genetic Medicine, Weill Cornell Medical College, 1300 York Avenue, Box 96, New York, New York 10065. Phone: 646-962-4363; Fax: 646-962-0220; E-mail: geneticmedicine@med.cornell.edu. Submitted March 10, 2009; Accepted for publication March 20, 2009; Epub (www.molmed.org) ahead of print March 20, 2009.

rette smoke and other toxic reagents, and in the context that cigarette smoking places a major oxidant stress on the lung, with the initial damage in the small airway epithelium (1–4,20,23,27–30), we asked: does cigarette smoking induce activation of Nrf2 in the human small airway epithelium with coordinate control of genes with antioxidant response elements, and are these genes the same or different from those observed in mice? To address these questions, we sampled the small airway epithelium of 83 individuals (38 healthy nonsmokers and 45 healthy smokers) to determine if Nrf2 was activated in healthy smokers, and if so, to identify which of the known murine Nrf2-modulated genes are up-regulated by smoking in this cell population. The data demonstrate that Nrf2 is activated in the small airway epithelium of healthy smokers. Using the known murine Nrf2-modulated genes in the lung as a starting point (14,16–26), and Affymetrix Human Genome U133 Plus 2.0 microarrays to assess genome-wide gene expression, we observed that 13 human homologues of 201 known murine Nrf2-modulated genes were highly responsive to cigarette smoke in the small airway epithelium of healthy smokers. To identify additional genes that may also be regulated by Nrf2 in the human small airway epithelium, an index was created to quantify the extent of Nrf2-responsive gene expression. We then searched among all genes expressed in the small airway epithelium of healthy individuals for genes whose expression, with TaqMan quantitative confirmation, correlated with the Nrf2 index. Using this strategy, we identified two genes not previously recognized to be controlled by Nrf2, *PIR* and *UGT1A4*, both of which contained several antioxidant response elements 5' to the gene, and both of which responded *in vitro* to Nrf2.

MATERIALS AND METHODS

Study Population

Healthy nonsmokers ($n = 38$) and healthy current cigarette smokers ($n = 45$)

were recruited through advertisements in local newspapers and on electronic bulletin boards, and through an ongoing program of free spirometry screening in the Department of Genetic Medicine and Division of Pulmonary and Critical Care Medicine. The evaluation of all individuals was performed at the Weill Cornell NIH Clinical and Translational Science Center and Department of Genetic Medicine Clinical Research Facility, using institutional review board–approved clinical protocols. Nonsmokers and smokers were determined to be healthy on the basis of standard history, physical exam, complete blood count, coagulation studies, liver function tests, HIV-1 serology, urine studies, chest x-ray, electrocardiogram, and pulmonary function tests. Current smoking status was evaluated on the basis of history (pack-year), venous carboxyhemoglobin levels, and urine analysis for nicotine metabolites. The inclusion criteria for healthy nonsmokers were “never smoking history” and normal physical exam, lung function, and chest x-ray, with smoking-related blood and urine within the nonsmoker range. The criteria for healthy smokers were current smoking history, normal physical exam, lung function, chest x-ray, and smoking-related urine and blood parameters consistent with that of a current smoker (31,32).

Sampling Small Airway Epithelium

We collected small airway epithelium brushes by using fiberoptic bronchoscopy, as previously described (33,34). Briefly, after the study participant had received mild sedation with meperidine and midazolam and routine anesthesia of the vocal cords and bronchial airways with topical lidocaine, a fiberoptic bronchoscope (Pentax, EB-1530T3) was directed to proximal to the opening of a desired lobar bronchus. Small airway epithelium cells were collected from the 10th- to 12th-order bronchi of the right lower lobe. A 2.0-mm disposable brush was advanced 7 to 10 cm further distally from the third-order bronchial branching, and the distal end of the brush was wedged

into the 10th- to 12th-generation branching of the right lower lobe. Small airway epithelium was collected by gently gliding the brush back and forth 5 to 10 times in 8 to 10 different locations in the same general area. Cells were detached from the brush by flicking and were immediately transferred into aliquots of ice-cold LHC8 medium (Gibco, Grand Island, NY, USA). Total cell number was counted on a hemocytometer, and cell viability was estimated by Trypan Blue exclusion and expressed as a percentage of the total cells recovered. About $1\text{--}2 \times 10^4$ cells were processed immediately for RNA extraction, and $2\text{--}5 \times 10^6$ cells were frozen at -80°C in cell-freezing medium (Sigma-Aldrich, St. Louis, MO, USA) for further processing for Western analysis and electrophoretic mobility shift assays. Additionally, 2×10^4 cells per slide were used for a differential cell count and for immunohistochemistry. Slides for a differential cell count and for immunohistochemistry were prepared by centrifugation (Cytospin 11; Shandon Instruments, Pittsburgh, PA, USA), and stored at 4°C until further processing. For differential cell count, cells were stained with Dif-Quik (Baxter Healthcare, Miami, FL, USA). All samples were confirmed as bonafide small airway epithelial samples by expression of genes encoding surfactant proteins and Clara secretory protein (33).

RNA Extraction, Microarray Processing, and Data Analysis

Analyses were performed by using Affymetrix (Santa Clara, CA, USA) Human Genome U133 Plus 2.0 microarrays (54,675 probe sets representing approximately 47,000 full-length human gene transcripts) and associated protocols. Total RNA was extracted by using a modified version of the TRIzol method (Invitrogen, Carlsbad, CA, USA), followed by RNeasy (Qiagen, Valencia, CA, USA) to remove residual DNA. RNA samples were stored in RNA Secure (Ambion, Austin, TX, USA) at -80°C until further processing. An aliquot of each RNA sample was run on

an Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) to visualize and quantify the degree of RNA integrity. The concentration was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Double-stranded cDNA was synthesized from 1–2 µg of total RNA using the GeneChip One-Cycle cDNA Synthesis Kit, followed by cleanup with a GeneChip Sample Cleanup Module, *in vitro* transcription reaction performed by using the GeneChip IVT Labeling Kit, and clean-up and quantification of the biotin-labeled cRNA yield by spectrophotometric analysis. All kits were from Affymetrix. Hybridizations to test chips and to the microarrays were performed according to Affymetrix protocols, and microarrays were processed by the Affymetrix fluidics station and scanned with the Affymetrix GeneChip Array Scanner 3000 7G. Overall microarray quality was verified by the following criteria: (a) RNA integrity number ≥ 7.0 ; (b) 3'/5' ratio for GAPDH ≤ 3 ; and (c) scaling factor ≤ 10.0 . Captured images were processed by using the MAS 5.0 algorithm (Affymetrix Microarray Suite Version 5.0 software), which takes into account the perfect match and mismatch probes. The data were normalized by using GeneSpring version 7.3 software (Agilent Technologies) per array, by dividing the raw data by the 50th percentile of all measurements and by gene.

Nrf2 Activation in Small Airway Epithelium of Healthy Nonsmokers and Smokers

On the basis of the knowledge that Nrf2 translocates to the nucleus to protect cells from oxidative stress, such as that caused by smoking (11,12,20,23,30, 35), we applied immunohistochemical and Western analysis of Nrf2 to assess the activation status of Nrf2 in small airway epithelial cells obtained from healthy nonsmokers and healthy smokers.

For immunohistochemistry, slides with cytopins of the small airway epithelium

were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS), pH 7.4, at 23°C for 20 min. To enhance staining, an antigen retrieval step was carried out by incubating the slide at 90°C for 20 min in a water steamer (Black & Decker, Hunt Valley, MD, USA), in citrate solution (BD Biosciences Pharmingen, San Diego, CA, USA), followed by cooling at 23°C for 20 min. The slides were blocked in 5% donkey serum for 45 min to reduce background staining, then incubated overnight at 4°C with an affinity-purified rabbit polyclonal antibody, mapped to the C-terminus of human Nrf2 (Santa Cruz Biotechnology, Santa Cruz, CA, USA), diluted 1:1000 in PBS with 5% donkey serum. To ensure the specificity of the Nrf2 antibody, blocking Nrf2 peptide (Santa Cruz Biotechnology) was used as a control. The Vectastain Elite ABC kit (Vector Laboratories, Burlingame, CA, USA) and the 3-amino-9-ethyl carbazole substrate kit (Vector Laboratories) were used to detect antibody binding, and the slides were counterstained with hematoxylin (Sigma-Aldrich) and mounted using GVA mounting medium (Zymed, San Francisco, CA, USA). Brightfield microscopy was performed using a Nikon Microphot microscope equipped with a Plan 40 × N.A. 0.70 objective lens. Images were captured with an Olympus DP70 CCD camera.

To quantify the nuclear accumulation of Nrf2 in small airway epithelium in response to smoking, cytoplasmic and nuclear extracts from small airway epithelium cells of healthy nonsmokers and healthy smokers were assessed by Western analysis. Cytoplasmic and nuclear extracts were isolated stepwise by using a commercially available nuclear and cytoplasmic extraction kit (Pierce Biotechnology, Rockford, IL, USA). Briefly, after thawing and one washing with PBS (pH 7.4), small airway epithelium cells were lysed in 200 µL of ice cold cytoplasmic extraction reagent 1 for 10 min, and 11 µL of ice-cold cytoplasmic extraction reagent 2 for 1 min, according to the manufacturer's protocol. After centrifugation at

16,000g for 5 min, the supernatant (cytoplasmic extract) was stored on ice, while the insoluble pellet fraction was resuspended in 25 µL of nuclear-extraction reagent. After a 40-min incubation period, with vortexing every 10 min, and another centrifugation at 16,000g for 10 min, the supernatant (nuclear extract) was collected. Protein concentration was assessed in nuclear and cytoplasmic extracts by using a BCA protein concentration kit (Pierce). Samples were adjusted to equal amounts of protein (50 µg protein/lane for cytoplasmic protein and 30 µg protein/lane for nuclear protein). Samples were mixed with Nupage LDS sample buffer and reducing agent (Invitrogen) and loaded onto a Novex Tris-Glycine 4%–12% gel (Invitrogen). Proteins were transferred (25 V, 90 min, 23°C) to a 0.45-µm pore size polyvinylidene fluoride membrane (BioRad) in Tris-glycine transfer buffer (Invitrogen). After transfer, the membrane was blocked in PBS containing 5% milk powder (Bio-Rad) for 1 h, then the membrane was incubated overnight at 4°C with a 1:1000 dilution of antihuman Nrf2 monoclonal IgG (Abnova, Taipei, Taiwan). Detection was performed by using horseradish peroxidase-conjugated secondary antimouse antibody (Santa Cruz), at a 1:2000 dilution, and the enhanced chemiluminescence reagent system (GE Healthcare, Pittsburgh, PA, USA), using Kodak BioMax light film. To evaluate the levels of nuclear protein in the nuclear extracts, the blots were stripped in Western-blot stripping buffer (Pierce) for 15 min, at 23°C. After being blocked in PBS containing 5% milk powder for 1 h at 23°C, the blots with nuclear extracts were incubated with a polyclonal goat-antihuman lamin B antibody (Santa Cruz Biotechnology) at a 1:300 dilution, and with a secondary antigoat horseradish peroxidase-conjugated antibody (Santa Cruz Biotechnology) at 1:2000, for 90 min each at 23°C, with detection as described above. The blots with cytoplasmic extracts were incubated with a polyclonal rabbit anti-β-tubulin antibody (Imgenex, San Diego, CA, USA) at a 1:1000 dilution,

and with a secondary antirabbit horse-radish peroxidase-conjugated antibody (Santa Cruz Biotechnology) at 1:2000, for 90 min each at 23°C. To quantitatively assess the amount of Nrf2 in the nucleus and cytoplasm, the films were digitally imaged, maintaining exposure within the linear range of detection. The contrast was inverted, the pixel intensity of each band determined, and the background pixel intensity for a negative area of film of identical size was subtracted by using MetaMorph image analysis software (Universal Imaging, Downingtown, PA, USA). Bands on the lamin B and β -tubulin Western blots were scanned and analyzed by densitometry to ensure equal protein loading.

Expression Levels of Nrf2-Modulated Genes in the Small Airway Epithelium

We compiled a list of known murine Nrf2-modulated genes consisting of all murine genes reported to be responsive to different oxidative stress conditions in the lung, as identified in a total of 12 reports by several groups using Nrf2 knockout mice (14,16–26). This list included 201 genes that were upregulated in wild-type mice compared with Nrf2 knockout mice (Supplemental Data, Table 1). Among these 201 genes, 187 human homologs were identified by using the NCBI (National Center for Bioinformatics) HomoloGene database (<http://www.ncbi.nlm.nih.gov/homologene>) and the Information Hyperlinked Over Proteins database (<http://www.ihop-net.org/UniPub/iHOP/>). The impact of smoking on these 187 human genes was assessed by using the gene expression microarray data from the 45 smokers and 38 nonsmokers with following criteria: (a) Affymetrix detection call of Present ("P call") in > 20 % of samples, (b) magnitude of fold-change in average expression value for healthy smokers versus nonsmokers > 1.5, (c) significance level in expression value for healthy smokers versus nonsmokers $P < 0.01$. A final list was established consisting of putative, highly smoking-responsive, Nrf2-modulated human genes.

Index for Nrf2-Modulated Gene Expression in the Small Airway Epithelium of Healthy Smokers

If Nrf2-modulated genes are coordinately controlled in the small airway epithelium of smokers, then all of the Nrf2 human homologs that are significantly modulated by smoking should be modulated to a similar extent (for example, low, medium, or high expression level) in each individual. To assess this thesis, an Nrf2 index was created that ranked healthy smokers based on their expression levels of the 13 Nrf2-modulated genes. For each gene, the individuals were divided into quartiles based on the level of gene expression; each individual was assigned a quartile designation of 1, 2, 3, or 4 for each of the putative Nrf2-responsive genes. To avoid biasing the Nrf2 index toward genes that were represented by multiple probe sets, the quartile designations were averaged and rounded to the nearest single-decimal unit for genes represented by more than 1 probe set. To calculate the Nrf2 index for each individual, the quartile designations of that individual were averaged across all Nrf2-modulated genes and reported as mean \pm SD. A z score Nrf2 index was used to validate the quartile-based index. This validation was accomplished by determining for each subject the z score for each expression value and summing the z scores for the 13 genes in the initial gene list for each of the 45 smokers.

To visualize whether Nrf2-modulated genes were concordantly regulated among the healthy smokers, Nrf2 genes were ordered alphabetically, and all healthy smokers were ranked according to the Nrf2 index. Gene clusters were visualized based on the level of gene expression, with a different color assigned to each quartile. The distribution of the Nrf2 index was compared with that of indices generated in the same way using control genes. Two other gene lists were compiled as controls, both compromised of randomly chosen genes ("P call" in >20% of samples) generated by using Excel 2003 software (Microsoft

Corporation, Redmond, WA, USA): (a) genes randomly chosen from all probe sets corresponding to unique genes and (b) genes randomly chosen from a total of 476 smoking-responsive genes identified from our dataset (defined as > 1.5-fold, $P < 0.01$ after Benjamini-Hochberg correction, for healthy smokers versus nonsmokers; Supplemental Data, Table 2). To rule out the possibility that the Nrf2 index might reflect differing demographic characteristic or differing exposure to smoking, we compared the demographics and smoking history of individuals from each quartile. In addition, the overall coordination of the expression levels of the Nrf2-related genes was determined by averaging the r^2 values for all pairs of genes from all healthy smokers and comparing that mean r^2 value to the mean value from the set of random genes and set of random smoking-related genes (Supplementary Table 2).

TaqMan Confirmation of Microarray Data

TaqMan real-time reverse-transcription-polymerase chain reaction (RT-PCR) was performed on RNA samples from the small airway epithelium of healthy smokers with low and high Nrf2 index values, and the results compared with those from the Human Genome U133 Plus 2.0 microarray. cDNA was synthesized from 2 μ g RNA in a 100 μ L reaction volume, using the TaqMan Reverse Transcriptase Reaction Kit (Applied Biosystems, Foster City, CA, USA), with random hexamers as primers. Two dilutions, 1:10 and 1:100, were made from each sample, and duplicate wells were run for each dilution. TaqMan PCR reactions were carried out using pre-made gene-expression kits from Applied Biosystems for Nrf2-modulated genes, and 2 μ L of cDNA was used in each 25 μ L reaction. The endogenous control was human β -actin (Applied Biosystems). Relative expression levels were determined using the $\Delta\Delta$ Ct method, with the average value of the low-ranked smokers as the calibrator. The PCR reactions were run in

an Applied Biosystems Sequence Detection System 7500, and the relative quantity was determined using the algorithm provided by the manufacturer.

Identification of Potential New Human Nrf2-Modulated Genes in the Small Airway Epithelium

To identify unknown genes that may also be modulated by Nrf2 in the small airway epithelium, the Nrf2 index was correlated with the expression values of all genes expressed in the lung of healthy smokers ("P call" in >20% of samples), using the r^2 value to identify known and putative new Nrf2-modulated genes. For the 25 most highly correlated genes, the presence of the primary core sequence of the antioxidant response element (RTGAYNNNGCR) (23,36) in the promoter region of each gene (5' untranslated region and 10,000-bp upstream of the transcriptional start site) was determined by searching with Genamics Expression 1.1 Pattern Finder Tool Software (Genamics, Hamilton, New Zealand).

Validation of Newly Identified Nrf2-Modulated Genes

Newly identified Nrf2-modulated genes were validated with electrophoretic mobility shift assay (37,38). We incubated 10 µg of nuclear extract from small airway epithelium with a 32 P-labeled antioxidant response element oligonucleotide from *NAD(P)H dehydrogenase, quinone 1* (NQO1), a known Nrf2-modulated gene (10,37). A fifty-fold excess of unlabeled NQO1 oligonucleotide was used as a specific competitor to determine the specificity of binding. A fifty-fold excess of nonspecific oligonucleotide (a random sequence from the *von Willebrand factor* gene) was used as a negative control for the competition. To demonstrate that putative antioxidant response elements from the newly identified Nrf2-modulated genes bind to Nrf2 in a similar manner as does the NQO1-antioxidant response element, unlabeled antioxidant response element oligonucleotides from each newly identified Nrf2-modulated gene were used as spe-

cific competitors, in a fifty-fold excess, against the labeled NQO1 oligonucleotide. To confirm the results, one of the antioxidant response element oligonucleotides from each gene proven to be functional was radioactively labeled and competed with an unlabeled oligonucleotide as a specific competitor, and with a nonspecific oligonucleotide as a negative control. The sequences of the oligonucleotides (Operon Biotechnologies, Huntsville, AL, USA) are listed in Supplemental Data Table 3. The oligonucleotides were annealed by heating to 95°C in an annealing buffer of 10 mmol/L Tris, pH 8.0, 50 mmol/L NaCl, and 1 mmol/L EDTA (ethylenediaminetetraacetic acid). These oligonucleotides (200 ng) were end-labeled with 300 µCi 32 P ATP by a T4 polynucleotide kinase (New England Biolab, Ipswich, MA, USA), and afterward purified, according to the protocol, using the QIAquick Nucleotide Removal kit (Qiagen, Valencia, CA, USA). Approximately 3 ng (300,000 cpm) of labeled probe was used in each reaction. The probes were incubated with nuclear protein, extracted as described above, for 30 min at 23°C in a binding buffer of 0.02 mol/L Hepes pH 8.0, 0.05 mol/L KCl, 0.05 µmol/L EDTA, 1 µmol/L MgCl₂, and 5% glycerol, and with a fifty-fold excess of an unrelated, noncompetitive oligonucleotide to reduce nonspecific binding. For super shift analysis (39), labeled oligonucleotides were incubated with 1.65 µg of an affinity purified antihuman Nrf2 rabbit polyclonal IgG (Atlas Antibodies, Stockholm, Sweden) for 30 min. Pure rabbit IgG (Jackson ImmunoResearch, West Grove, PA, USA) was used as a control at the same concentration. The mixture was separated on a 6% native polyacrylamide gel (Invitrogen). The gels were subsequently dried, and exposed to a BioMax X-Ray film for 5 to 24 h.

Statistical Analysis

Human Genome U133 Plus 2.0 microarrays were analyzed using GeneSpring software. Average expression values in small airway samples were calculated

from normalized expression levels for healthy nonsmokers and healthy smokers. Statistical comparisons between nominal variables were calculated using an unpaired, two-tailed *t* test with unequal variance. The comparison of SDs in the index of Nrf2-modulated genes against indices of random genes and random smoking-responsive genes was calculated with the nonparametric Wilcoxon rank test. A chi-square test was used for statistical comparison for categorical variables. To rule out the possibility that the Nrf2 index might reflect differing demographic characteristics or differing exposure to smoking, individuals of each quartile were compared by ANOVA. Correlation analyses were all performed using the Spearman Correlation. All tests other than the two-tailed *t* test were calculated using StatView version 5.0 (SAS Institute). Data are presented as mean ± standard error. A *P* value < 0.05 was considered to be significant.

Web Deposition of Data

All data have been deposited in the Gene Expression Omnibus (GEO) site (<http://www.ncbi.nlm.nih.gov/geo/>), which is curated by the NCBI under accession number GSE11952.

All supplementary materials are available online at www.molmed.org.

RESULTS

Study Population of Smokers and Nonsmokers

Small airway epithelial samples were collected from 83 individuals comprising 38 healthy nonsmokers and 45 healthy smokers (Table 1). The nonsmoker and smoker groups were similar with regard to age (*P* > 0.4), sex (*P* > 0.8), and ethnicity (*P* > 0.2). Healthy smokers had a mean smoking history of 27 ± 2 pack-years. Compared with healthy nonsmokers, healthy smokers demonstrated a mildly higher total number of cells recovered in the brushings (*P* < 0.04, Table 1). The differentials of the epithelium cells mostly did not differ between healthy

Table 1. Study population demographics and biologic sample data^a

Parameter	Healthy nonsmokers	Healthy smokers	<i>P</i> value ^b
n	38	45	
Sex, male/female	28/10	32/13	<i>P</i> > 0.8
Age, y	43 ± 2	43 ± 1	<i>P</i> > 0.4
Ancestry, B/W/H/A/O ^c	16/16/4/1/1	28/11/6/0/0	<i>P</i> > 0.2
Smoking history, pack-year	0 ± 0	27 ± 2	<i>P</i> < 1 × 10 ⁻¹⁴
Urine nicotine, ng/mL	0 ± 0	967 ± 161	<i>P</i> < 4 × 10 ⁻⁷
Urine cotinine, ng/mL	0 ± 0	1055 ± 119	<i>P</i> < 4 × 10 ⁻¹¹
Venous carboxy hemoglobin, %	0 ± 0	2.0 ± 0.3	<i>P</i> < 2 × 10 ⁻⁶
Pulmonary function parameters ^d			
FVC	109 ± 2	109 ± 2	<i>P</i> > 0.6
FEV1	105 ± 3	108 ± 2	<i>P</i> > 0.5
FEV1/FVC	81 ± 1	81 ± 1	<i>P</i> > 0.7
TLC	103 ± 2	100 ± 2	<i>P</i> > 0.4
DLCO	99 ± 2	95 ± 2	<i>P</i> > 0.3
Epithelial cells recovered, n × 10 ⁶	5.8 ± 0.3	6.5 ± 0.4	<i>P</i> < 0.04
Epithelial cells, %	99.7 ± 0.2	99.8 ± 0.1	<i>P</i> > 0.4
Inflammatory cells, %	0.2 ± 0.1	0.2 ± 0.1	<i>P</i> > 0.8
Differential cell count			
Ciliated, %	75.3 ± 1.1	72.3 ± 1.3	<i>P</i> > 0.06
Secretory, %	7.1 ± 0.6	7.1 ± 0.5	<i>P</i> > 0.4
Basal, %	10.1 ± 0.8	10.6 ± 0.7	<i>P</i> > 0.5
Undifferentiated, %	7.3 ± 0.6	10.0 ± 0.9	<i>P</i> < 0.01

^aData are presented as mean ± standard error.^bGroup comparison with unpaired, two-tailed *t* test with unequal variance and chi-square test.^cB, black; W, white; H, Hispanic; A, Asian; O, other.^dPulmonary function testing parameters are given as percentage of predicted value, with the exception of FEV1/FVC, which is reported as percentage observed; FVC, forced vital capacity; FEV1, forced expiratory volume in 1 s; TLC, total lung capacity; DLCO, diffusing capacity.

nonsmokers and smokers (*P* > 0.07), except that undifferentiated cells were more frequently found in smokers than in nonsmokers (*P* < 0.01).

Localization of Nrf2 in Healthy Small Airway Epithelium of Nonsmokers and Smokers

To confirm airway epithelium expression, and on the basis of knowledge that Nrf2 translocates to the nucleus to protect cells from oxidative stresses such as smoking (11,12,20,23), we carried out immunohistochemical analysis by using small airway epithelial cells obtained by brushing. With an antibody dilution of 1:1000, we stained about 20%–30% of brushed cells for Nrf2 (Figure 1 A–B, D–F), which were specifically blocked by using a blocking peptide as a negative control (Figure 1C, F). The intensity of

Nrf2 staining was variable among cells. However, for nonsmokers there was minimal Nrf2 detected in the nucleus (Figure 1A, B). In contrast, the epithelial cells of smokers showed faint cytoplasmic Nrf2 staining with prominent strong areas of stain in the nucleus (Figure 1D, E). There were typically 2 or 3 spots of Nrf2 staining per cell consistent with location at the nucleoli (arrows).

To quantify the nuclear accumulation of Nrf2 in small airway epithelium in response to smoking, nuclear and cytoplasmic extracts from healthy nonsmokers and smokers were assessed by Western analysis. Using multiple brushes from the airway, we obtained approximately 30 μg of nuclear and 50 μg of cytoplasmic protein from each individual. Equal protein loading among the samples was verified with antibodies

that identify the nucleus (lamin B) and cytoplasm (β-tubulin). Western analysis demonstrated increased Nrf2 protein concentration in the cytoplasm of healthy nonsmokers as compared with healthy smokers (Figure 1G). In contrast, the nuclear extract showed a higher amount of Nrf2 protein in smokers compared with nonsmokers (Figure 1H). Quantification of single bands by densitometry revealed a 3.6-fold increase in levels of Nrf2 in the cytoplasm of healthy nonsmokers compared with healthy smokers (*P* < 0.02; Figure 1I), and a 3.7-fold increase in levels of Nrf2 in the nucleus of healthy smokers compared with healthy nonsmokers (*P* < 0.05, Figure 1J). Because of limitations in biological materials from human subjects, it was not possible to correlate gene expression at the mRNA level with protein levels measured by Western analysis in the same subject.

Expression Levels of Nrf2-Modulated Genes in Small Airway Epithelium

A list of murine Nrf2-modulated genes was compiled based on reports of genes responsive to different oxidative stress conditions in the lung of Nrf2 knockout mice (14,16–26). In total, 201 murine oxidative stress responsive genes in the lung were identified, and 187 human homologs were determined to be expressed in the small airway epithelium (Supplemental Data, Table 1). The impact of smoking on these 187 human genes was assessed by using the gene expression microarray data from the 45 smokers and 38 nonsmokers described above. From these 187 genes, we established a final list consisting of 13 putative human, Nrf2-modulated, highly smoking-responsive genes (Table 2). All 13 genes were significantly upregulated in smokers compared with nonsmokers (all comparisons, *P* < 10⁻⁴; Figure 2), with a mean increase in expression level of 4.5 ± 1.7-fold in response to smoking. Among the other known murine Nrf2-modulated genes, an additional 16 genes were found to be significantly modulated by smoking, with smoker-to-

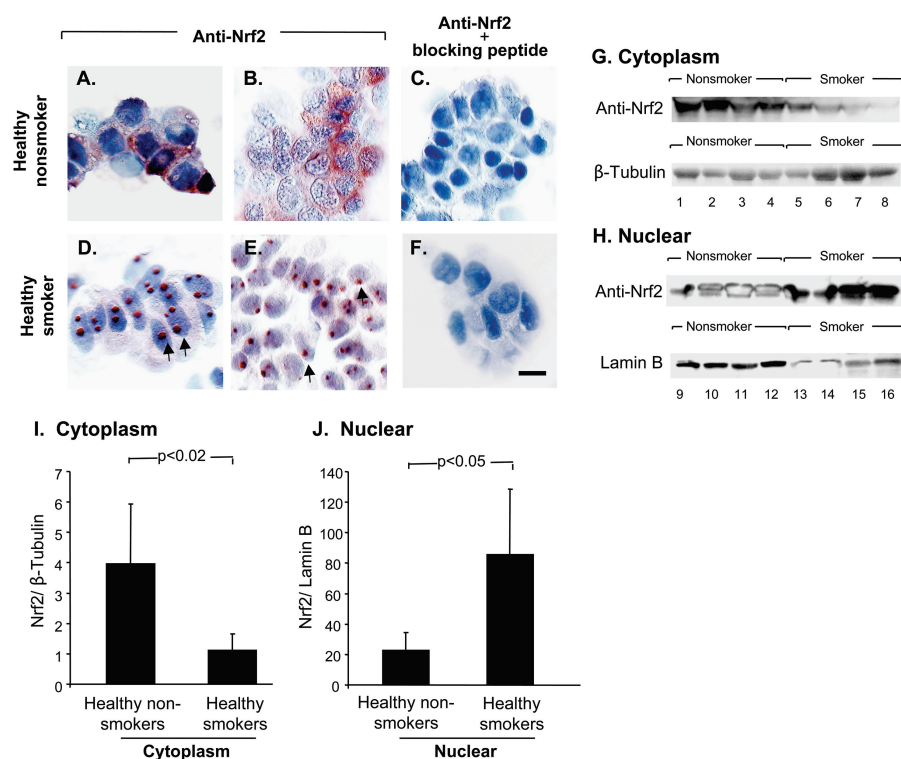


Figure 1. Localization of Nrf2 in small airway epithelium of healthy nonsmokers and healthy smokers. Cytospin preparations of small airway epithelium were assessed by immunohistochemical analysis and Western analysis. (A–F) Immunohistochemistry. (A, B) Healthy nonsmokers, + Nrf2 antibody; (C) healthy nonsmoker, + Nrf2 antibody + blocking peptide; (D–E) healthy smokers, + Nrf2 antibody; (F) healthy smoker, + Nrf2 antibody + blocking peptide. For (D–E), arrows indicate nucleoli positive for Nrf2 staining. For (A–F), bar = 10 μ m. (G–J) Western analyses. For the Western analyses, small airway epithelial cells of four healthy nonsmokers and four healthy smokers provided nuclear and cytoplasmic extracts, which were normalized to equal protein concentration. (G) Western analysis of Nrf2 in cytoplasmic extracts from small airway epithelial cells of healthy nonsmoker and smoker (upper panel) and β -tubulin as an equal loading control (lower panel). Lane 1–4, examples of nonsmokers; lanes 5–8, examples of smokers. (H) Western analysis of Nrf2 in nuclear extract from small airway epithelium from healthy nonsmoker and healthy smoker (upper panel) and lamin B as loading control (lower panel). Lanes 9–12, examples of nonsmokers; lanes 13–16, examples of smokers. (I) Quantitative analyses of band densities for Nrf2 in cytoplasm (normalized to β -tubulin). (J) Quantitative analysis of band densities for Nrf2 in nucleus (normalized to lamin B). Data shown are mean \pm standard error.

nonsmoker P values of 0.05 to 10^{-4} . However, because the fold-change of smoker to nonsmoker was <1.5 in all cases, these genes were not considered further (Supplemental Data, Table 4).

Index for Nrf2-Modulated Genes in the Small Airway Epithelium of Healthy Smokers

If the human Nrf2-modulated, smoking responsive genes have been correctly

identified, then it follows that there should be coordinate regulation of the 13 genes among the smoking subjects (that is, a subject who has a high expression level of 1 of the 13 genes should also have high expression of the other 12). To assess the average small airway epithelium gene expression of the Nrf2-modulated smoking responsive genes, an Nrf2 index was created that ranks healthy smokers on the basis of their overall ex-

pression levels of the 13 Nrf2-modulated genes. For each gene, individuals were divided into quartiles on the basis of the level of gene expression, so each individual was assigned a quartile designation of 1 to 4 for each of the 13 genes. This analysis demonstrated that among smokers, there were clearly individuals who had expression levels for most or all genes in the upper quartile, whereas other subjects had expression levels for most or all genes in the lower quartile. The Nrf2 index for each individual is the average of the 13 quartile designations (Figure 3A). The use of a discontinuous score for determining the Nrf2 index was validated by calculating the index by using an alternate strategy in which the sum of the z scores for each gene was calculated for each subject. This method gave an index that was strongly correlated ($r^2 = 0.95$) with the quartile-based score (see Supplementary Figure S1).

The healthy smokers showed a wide range of Nrf2 index values, with a 25th-to-75th percentile range of 1.7 to 3.2, and individuals were almost evenly distributed among the index groups. Of all 45 smokers, 13 smokers ranked in the Nrf2 index between 1 and 2, 17 smokers between 2 and 3, and 15 smokers between 3 and 4 (Figure 3B). To demonstrate that this result represents coordinate control, the distribution of the smokers according to the 13 Nrf2-modulated gene index was compared with that of an index generated in the same way by using 13 random genes (Supplemental Table 2). This comparison gave a narrower range with a 25th to 75th percentile range of 2.2 to 2.7. Similarly, for 13 smoking-responsive genes randomly chosen from a total of 476 smoking-responsive genes identified from our data set (corrected $P < 0.05$, fold-change > 1.5), the 25th and 75th percentile of the index was 2.2 to 2.8, narrower than the range of the putative Nrf2-modulated, smoking-responsive genes. For both control gene indices, smokers were predominately distributed between the index values two and three, suggesting a random distribution of these control genes, without coordinate

Table 2. Human homologues of murine Nrf2-modulated genes highly smoking-responsive in the human small airway epithelium.^a

Mouse gene symbol	Mouse gene title	Human homolog symbol	Fold-change, smoker/nonsmoker	P value (smoker/nonsmoker) ^b	References
<i>Adh7</i>	<i>Alcohol dehydrogenase 7</i>	<i>ADH7</i>	5.7	$P < 4 \times 10^{-16}$	(19,23)
<i>Fgrp, Fr-1</i>	<i>Aldo-keto reductase family 1, member B10</i>	<i>AKR1B10</i>	23.5	$P < 2 \times 10^{-8}$	(23)
<i>Aldh3a1</i>	<i>Aldehyde dehydrogenase 3 family, member A1</i>	<i>ALDH3A1</i>	4.8	$P < 6 \times 10^{-15}$	(19,71)
<i>G6pdx</i>	<i>Glucose-6-phosphate dehydrogenase</i>	<i>G6PD</i>	1.9	$P < 6 \times 10^{-6}$	(19,23,71)
<i>Gclc</i>	<i>Glutamate-cysteine ligase, catalytic subunit</i>	<i>GCLC</i>	1.6	$P < 2 \times 10^{-4}$	(16,18,19,22-24,26)
<i>Gpx2</i>	<i>Glutathione peroxidase 2</i>	<i>GPX2</i>	6.4	$P < 2 \times 10^{-12}$	(10,17,18,23,25,26)
<i>Gsr</i>	<i>Glutathione reductase</i>	<i>GSR</i>	1.6	$P < 1 \times 10^{-6}$	(10,23,24)
<i>Mdh-1, Mod-1</i>	<i>Malic enzyme 1, NADP(+) -dependent, cytosolic</i>	<i>ME1</i>	3.8	$P < 2 \times 10^{-14}$	(10,19,23,26)
<i>Nqo1</i>	<i>NAD(P)H dehydrogenase, quinone 1</i>	<i>NQO1</i>	2.4	$P < 2 \times 10^{-14}$	(10,16-21,23,41)
<i>Pgd</i>	<i>Phosphogluconate dehydrogenase</i>	<i>PGD</i>	1.6	$P < 8 \times 10^{-6}$	(23)
<i>Tkt, P68</i>	<i>Transketolase (Wernicke-Korsakoff syndrome)</i>	<i>TKT</i>	1.8	$P < 3 \times 10^{-8}$	(19)
<i>Txnrd1</i>	<i>Thioredoxin reductase 1</i>	<i>TXNRD1</i>	1.9	$P < 6 \times 10^{-11}$	(18,19,22,23,26)
<i>Ugt1-06</i>	<i>UDP glucuronosyltransferase 1 family, polypeptide A6</i>	<i>UGT1A6</i>	2.0	$P < 3 \times 10^{-7}$	(16-19)

^aCriteria: (a) Affymetrix Detection Call of Present ("P call") in > 20 % of samples; (b) magnitude of fold-change in average expression value for healthy smokers versus nonsmokers > 1.5; (c) significance level in expression value for healthy smokers versus nonsmokers $P < 0.01$.

^bGroup comparison with unpaired, two-tailed *t* test with unequal variance.

regulation (Figure 3B). Of the 45 smokers 37 (82%) were ranked between the index values two and three in the random gene index, and 33 of 45 smokers (73%) were ranked between two and three in the random smoking-responsive gene index ($P < 0.0001$ versus Nrf2-modulated genes, chi-square test). Other evidence of the coordinate regulation of the Nrf2 genes is the comparison of SDs of the Nrf2 index for each individual, which should be small if all genes are coordinately regulated. For Nrf2-modulated genes, the mean \pm SD of all smokers was 0.67 ± 0.03 , a value significantly lower than for random genes (1.08 ± 0.02 , $P < 10^{-4}$) or that of random smoking-responsive genes (1.06 ± 0.03 , $P < 10^{-4}$, Figure 3C).

To confirm the relatedness of the expression profile of these genes, we used pairwise Pearson correlations between the 13 genes across all subjects and then compared the average of these correlations to correlations derived from random sets of genes on the array or smoking-related genes. The mean r^2 for the Nrf2-modulated genes was 0.36 ± 0.02 compared with 0.10 ± 0.01 for random genes and 0.05 ± 0.01 for random smoking-regulated genes.

To assess whether the Nrf2 index reflected differing demographic character-

istics or differing exposure to smoking among the subjects examined, the healthy smokers were divided into three ranges (1–2, low expression of Nrf2-modulated genes; 2–3, moderate expression; 3–4, high expression). The demographic characteristics (sex, age, and ethnicity) were not different among the three groups ($P > 0.05$ by chi-square test for sex and ethnicity, ANOVA for age, Table 3). Similarly, the smoking exposure as assessed by history in pack-years, urine nicotine, urine cotinine, and serum carboxyhemoglobin was not different among the three groups (all comparisons $P > 0.05$ by ANOVA, Table 3). Thus, the Nrf2 index is not influenced by known demographic characteristics. Similarly, with analysis using the *z* score method, the *P* value for correlation of Nrf2 index with smoking history ($P = 0.10$), urine nicotine ($P = 0.50$), urine cotinine ($P = 0.19$), and carboxyhemoglobin ($P = 0.24$) were all insignificant.

Correlation of Nrf2 Index to Expression of all Genes in Small Airway Epithelium

If Nrf2-modulated genes have coordinate regulation in the human small airway epithelium, then other potential Nrf2-modulated genes might be identi-

fied by searching among all genes expressed in the lung of healthy smokers for genes whose expression correlates with the Nrf2 index. Of the top 25 probe sets whose small airway expression level was correlated with the Nrf2 index (Table 4), 24 were smoking responsive, with a mean increase of expression level of 2.9 ± 0.2 -fold in response to smoking (P values all $< 6 \times 10^{-5}$). The 25 probe sets corresponded to 18 uniquely identified genes. Overall, these 25 probe sets were highly correlated with the Nrf2 index, ($r^2 > 0.51$, $P < 0.0001$), suggesting that all these genes might have the same coordinate regulation and control. The gene expression that was most strongly associated with the Nrf2 index was that of *transaldolase 1* (*TALDO1*, Figure 4A), a key enzyme of the pentose phosphate pathway, which protects cellular integrity from oxygen injury. Other strongly correlated genes included: *ALDH3A1* (Figure 4B), an enzyme that plays an important role in the detoxification of alcohol-derived acetaldehydes; *NQO1* (Figure 4C), an enzyme that prevents the reduction of quinones that results in the production of radical species; and *TXNRD1* (Figure 4D), a protein that reduces thioredoxins and plays a role in selenium metabolism and protection from oxidative stress.

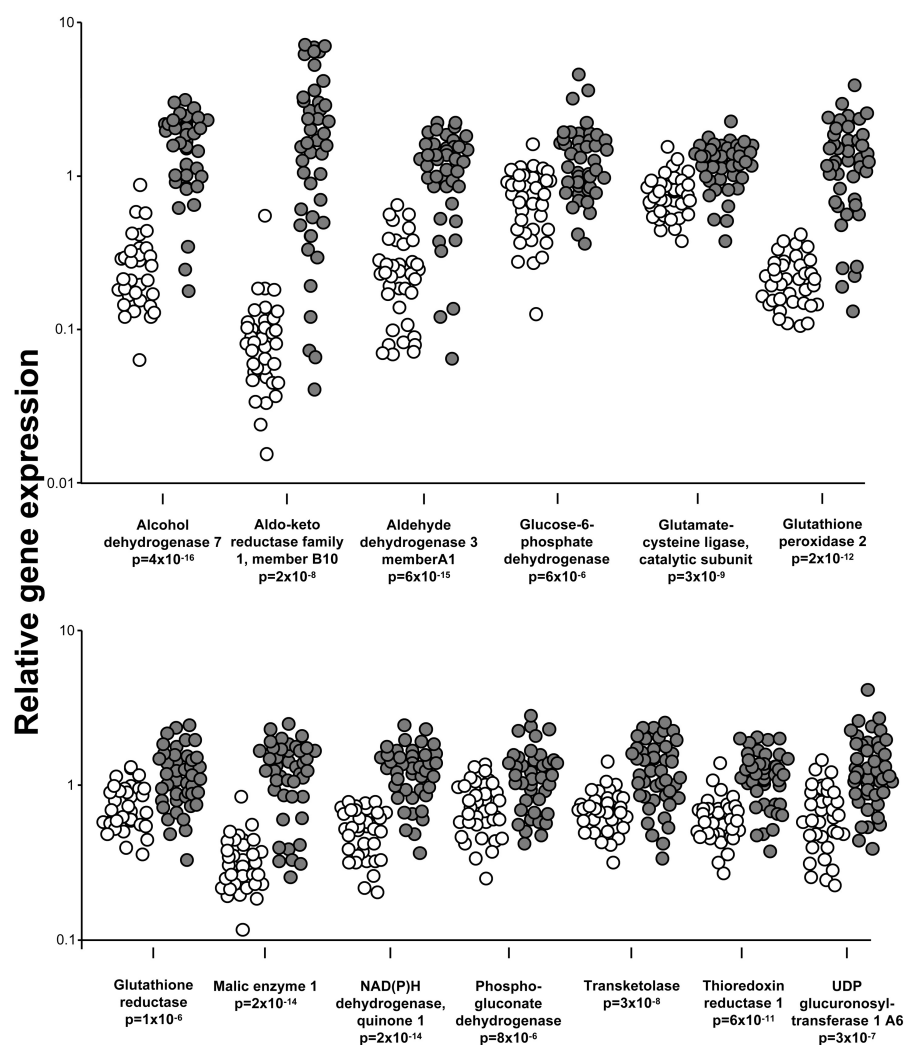


Figure 2. Expression levels of Nrf2-modulated genes in small airway epithelium. A list of murine Nrf2-modulated genes was compiled and used to identify the human homologs. The gene expression in small airway epithelium was examined to generate a list of putative human Nrf2-modulated genes that are highly responsive to smoking (Supplemental Table 1). Shown is the normalized expression level for each individual (38 healthy non-smokers (○), 45 healthy smokers (●)). The gene names and the *P* values are located on the abscissa, the relative gene expression levels are on the ordinate as a log scale.

The coordinate regulation of these genes was confirmed by unsupervised clustering. The complete set of smoking-responsive genes was clustered for a set of healthy smokers, and the locations of the NRF2-modulated genes within the cluster was determined. A small subsection of the cluster containing 52 genes (Supplementary Figure S2) included 12 of the genes listed in Table 4, a finding

that suggests a strong coordinate regulation. Other genes found in this part of the cluster, for example, *carbonyl reductases 1 and 3* (*CBR1*, *CBR3*) may also be modulated by Nrf2.

TaqMan Confirmation of Microarray Data

To validate the results obtained from the microarrays, TaqMan real-time RT-

PCR was carried out to assess several Nrf2-modulated genes by use of RNA samples from six healthy smokers classified as low expressers (Nrf2 index 1–2) and six healthy smokers categorized as high expressers (Nrf2 index 3–4). The TaqMan-PCR data were consistent with the microarray data for all three genes (Figure 5A). In the microarray analyses, *TALDO1* (3.0-fold, $P < 10^{-6}$), *ME1* (4.3-fold, $P < 10^{-5}$) and *TXNRD1* (2.8-fold, $P < 10^{-4}$) were significantly upregulated in the high expressers compared with the low expressers, in small airway epithelium. Use of TaqMan analysis on the same subjects confirmed that *TALDO1*, *ME1*, and *TXNRD1*, were significantly upregulated in the high expressers (2.4-fold, 4.3-fold and 2.8-fold, respectively; all $P < 0.05$; Figure 5B).

Identification of Potential New Human Nrf2-Modulated Genes in Small Airway Epithelium

If correlation with the Nrf2 index predicts regulation of small airway gene expression level by Nrf2, then some of the genes identified by this method may have previously been reported to be Nrf2 modulated. Of the 18 genes that were identified as having a strong association to the Nrf2 index, 14 have previously been reported as Nrf2-modulated on the basis of a transcription induction by Nrf2 in studies using Nrf2 knockout mice and/or by gel-shift assays specific for Nrf2: *TALDO1* (10,40); *ALDH3A1* (23), *NQO1* (10,16–21,23,41); *TXNRD1* (18,19,22,23,26); *ME1* (10,19,23,26); *aldo-keto reductase family 1, member C1 and member C2* (*AKR1C1*, *AKR1C2*) (42); *glutathione peroxidase 2* (*GPX2*) (10,17,18,23,25,26); *transketolase* (*TKT*) (19); *UGT1A6* (16–19,43); *glutamate-cysteine ligase, modifier subunit* (*GCLM*) (10,19,23,25); *sequestome 1* (*SQSTM1*) (23); *ADH7* (19,23); and *GSR* (10,23,24). We identified a very similar (22 of 25) list of potential Nrf2-modulated genes by using the continuous *z* score method of index calculation.

However, some genes identified had not previously been reported as transcriptionally induced by Nrf2 but have

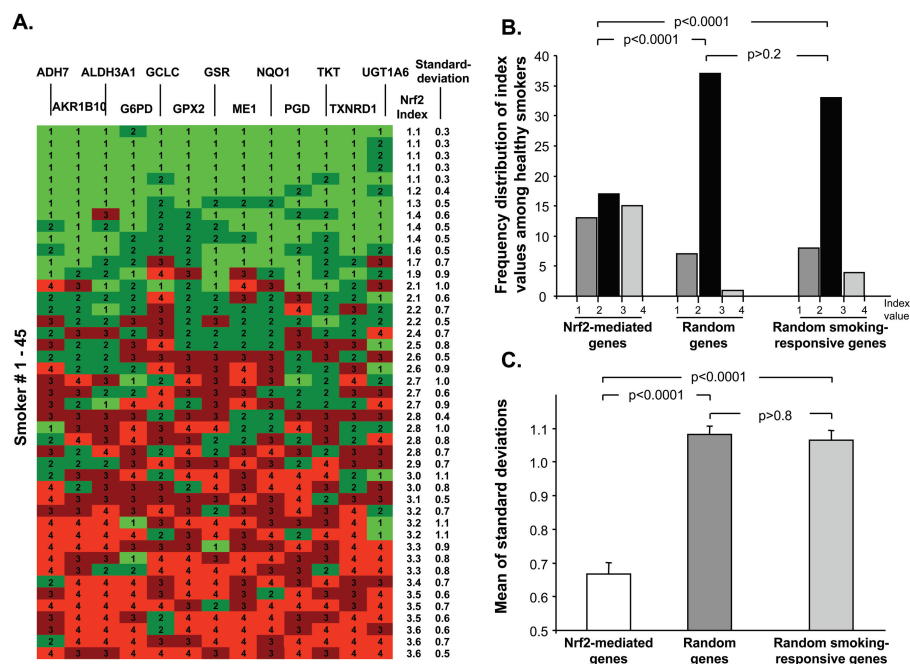


Figure 3. Index for Nrf2-modulated genes in the small airway epithelium of healthy smokers. We created an Nrf2 index that ranks the expression levels of the 13 Nrf2-modulated genes significantly modulated by smoking for all 45 healthy smokers. For each gene, individuals were divided into quartiles on the basis of the level of gene expression. Each individual was assigned a quartile designation of one, two, three, or four for each of the 13 genes. To calculate the Nrf2 index, for each individual the quartile designations for the individual were averaged across all genes. (A) Cluster analysis of Nrf2-modulated genes (from left to right) for 45 healthy smokers (from top to the bottom). The number in each object represents the quartile based on the gene expression, first quartile bright green, second dark green, third dark red, fourth bright red. The mean \pm SD of the Nrf2 index is presented at the right. Note that all of the healthy smokers tend to have the same quartile for all genes, with some smokers showing low expression (that is, all green bars) of the 13 genes, some smokers intermediate expression (mixed colors), and some smokers with all 13 genes showing high Nrf2 expression (that is, all red bars). (B,C) The Nrf2 index distribution for the 45 healthy smokers was compared with an index generated in the same way by using 13 “random genes” or 13 “random smoking-responsive genes” (see Supplemental Table 2). (B) Comparison of frequency distribution of smokers among index values. Shown are the indices for the 13 Nrf2-modulated genes, the 13 random genes and the 13 “random smoking-responsive genes.” For each, the data is grouped by index values of 1–2, 2–3, and 3–4. Note that smokers ranked by random genes or random smoking-responsive genes have indices mostly distributed between 2 and 3, but smokers ranked by the Nrf2 genes are almost evenly distributed. (C) Comparison of SDs of individual smokers of the 13 Nrf2-modulated genes, 13 random genes, and 13 random smoking-responsive genes. Data shown are mean \pm standard error. Note that the Nrf2-modulated genes have significantly less variation.

a strong association to the Nrf2 index. These included *PIR*, a gene of the cubin family (Figure 6A); *ATB binding cassette, sub-family B, member 6* (*ABCB6*), encoding a transporter protein of the superfamily of the ATP-binding cassette (Figure 6B); *UDP glucosyltransferase 1*

family, polypeptide A4 (*UGT1A4*), part of the complex locus which encodes several UDP-glucosyltransferases (Figure 6C); and *aldo-keto reductase family 1, member C3* (*AKR1C3*), a member of the aldo/ketoreductase superfamily (Figure 6D).

To evaluate whether *PIR*, *ABCB6*, *UGT1A4*, and *AKR1C3* are potential candidate genes modulated by Nrf2, the presence of the primary core sequence of the antioxidant response element (RTGAYNNNGCR) (10,37) was assessed by searching the promoter region of each gene with Genamics Expression 1.1 Pattern Finder Tool software. This analysis showed that *PIR*, *ABCB6*, and *UGT1A4*, but not *AKR1C3*, contain one or more antioxidant response elements in the genomic sequence upstream of their transcription start sites, suggesting these genes as novel Nrf2-modulated genes (Table 4).

Validation of Newly Identified Nrf2-Modulated Genes

If *PIR*, *ABCB6*, and *UGT1A4* are Nrf2-modulated genes, then the DNA binding activity of Nrf2 to the antioxidant response element of these genes should be detectable by electrophoretic mobility-shift assay. A strong DNA-binding activity of nuclear protein from the small airway epithelium of healthy smokers for NQO1-antioxidant response element was observed (Figure 7A, lanes 2, 6, 10, 18, 23). Nrf2 binding was demonstrated by the disappearance of DNA–protein complexes by competition with excess unlabeled NQO1 oligonucleotides (lanes 3, 11, 19, 24), but not with nonspecific unlabeled oligonucleotides (lanes 4, 16, 21, 29). Specificity was further confirmed by the disappearance of the gel shift band in the presence of an anti-Nrf2 affinity-purified polyclonal antibody (lane 7), whereas no disappearance was seen with the use of a corresponding IgG control (lane 8). Instead of a supershift, the disappearance of the gel shift band suggests a competition for binding between the Nrf2 antibody and NQO1 oligonucleotide, similar to that reported for Nrf2 binding to an antioxidant response element of *UGT1A1* with nuclear extracts from HepG2 cells after oxidative stress (38).

PIR contains four antioxidant response elements located 33 bp (lane 12) downstream of the transcription site and –3209

Table 3. Comparison of demographics of healthy smokers with different ranges of the Nrf2 index.^a

Parameter	Ranges of Nrf2 index ^b			P value ^c
	1–2	2–3	3–4	
n	13	17	15	
Age, years	40 ± 2	42 ± 1	46 ± 2	P > 0.2
Sex, male/female	7/6	11/6	14/1	P > 0.8
Ancestry, B/W/H/A/O	7/2/4/0/0	12/4/1/0/0	9/5/1/0/0	P > 0.2
Smoking history, pack-year	22 ± 3	27 ± 4	32 ± 5	P > 0.4
Urine nicotine, ng/mL	1205 ± 412	801 ± 217	935 ± 212	P > 0.2
Urine cotinine, ng/mL	894 ± 206	913 ± 192	1367 ± 212	P > 0.3
Venous carboxyhemoglobin, %	1.6 ± 0.4	1.8 ± 0.6	2.0 ± 0.3	P > 0.1

^aData are presented as mean ± standard error.

^b1–2 = low expression of Nrf2-modulated genes; 2–3 = moderate expression level of Nrf2-modulated genes; 3–4 = high expression of Nrf2-modulated genes.

^cGroup comparison with ANOVA or chi-square test.

^dB, black; W, white; H, Hispanic; A, Asian; O, other.

bp (lane 13), –3480 bp (lane 14), and –5566 bp (lane 15) upstream of the transcription start site. If these are functional Nrf2-regulated antioxidant response elements, then an excess of oligonucleotides consisting of these antioxidant response elements should compete specifically against the labeled NQO1 antioxidant response element. For *PIR*, functional antioxidant response elements were found at the locations of –3209 bp (lane 13) and –5566 bp (lane 15) upstream of the transcription start site, but not at 33 bp (lane 12) or –3480 bp (lane 14) from the transcription site.

For *ABCB6*, one antioxidant response element was identified at –7575 bp upstream of the transcription site. An excess of oligonucleotides consisting of this antioxidant response element did not compete against the labeled NQO1 antioxidant response (lane 20).

For *UGT1A4*, which contains 4 antioxidant response elements at –3043 bp (lane 25), –3159 bp (lane 26), –3885 bp (lane 27), and –5523 bp (lane 28) upstream of the transcriptional start site, two antioxidants response elements, at –3885 bp (lane 27) and –5523 bp (lane 28), competed against the NQO1-antioxidant response element—though not as strongly as either the NQO1 oligonucleotide itself or as the functional pirin antioxidant response elements—

suggesting that these two are probably functional antioxidant response elements.

To confirm these results, one of the antioxidant response element oligonucleotides of *PIR* and *UGT1A4* that was proven to be functional was radioactively labeled and competed with an unlabeled oligonucleotide as a specific competitor, and with a nonspecific oligonucleotide as a negative control. A strong DNA binding activity of nuclear protein from small airway epithelium was observed for labeled *PIR* antioxidant response element –3209 bp (Figure 7B, lane 31, 35) and for labeled *UGT1A4* antioxidant response element –3385 bp (Figure 7C, lane 39, 43). Nrf2 binding was demonstrated by the disappearance of DNA–protein complexes by competition with excess unlabeled *PIR* oligonucleotides (Figure 7B, lane 32) and *UGT1A4* oligonucleotide (Figure 7C, lane 40), but not with nonspecific vWF unlabeled oligonucleotides (Figure 7B and 7C, lanes 33 and 41). Specificity was further confirmed by the disappearance of the gel shift band in the presence of an anti-Nrf2 affinity purified polyclonal antibody (Figure 7B and 7C, lanes 36 and 44), whereas no disappearance was seen using a corresponding IgG control (Figure 7B and 7C, lanes 37 and 45).

DISCUSSION

Nuclear factor erythroid 2–related factor 2 (NFE2L2, Nrf2), an oxidant responding transcription factor, is known to induce phase 2 detoxifying and antioxidant genes to protect cells from oxidative stress (8–11). Based on the knowledge that cigarette smoke with its large oxidant content is a major stressor of the small airway epithelium, with cells that are vulnerable to oxidant damage and its associated malignant transformation (2,5–7), we asked the questions: does cigarette smoking induce activation of Nrf2 in the human small airway epithelium, is there an associated coordinate control of Nrf2-modulated genes, and are the highly-responsive Nrf2-modulated genes the same or different from those observed in Nrf2 murine models? Fiberoptic bronchoscopy was used to sample pure populations of small airway epithelium in 38 healthy nonsmokers and 45 healthy smokers, and gene expression of human homologs of 201 known murine Nrf2-modulated genes were assessed using Affymetrix Human Genome U133 Plus 2.0 microarrays. In the human small airway epithelium of healthy normal smokers compared with healthy nonsmokers, Nrf2 protein was significantly activated in the cells, and 13 Nrf2-modulated genes were identified as highly smoking responsive. Comparing different indices created based on the extent of gene expression of Nrf2-modulated genes and other control genes, we observed a significant concordant regulation for Nrf2-modulated genes. To identify new human Nrf2-modulated genes, the index of Nrf2-modulated genes was correlated to all genes expressed in the small airway epithelium. There was a concordant regulation of gene expression among some highly associated genes, which was confirmed with TaqMan quantitative PCR and was independent of smoking history. Two additional highly smoking-responsive Nrf2-modulated genes in the small airway epithelium were identified, *PIR* and *UGT1A4*, both of which contain several antioxidant response elements 5' to the

Table 4. Identification of genes not previously associated with Nrf-2 control identified by the correlation of the Nrf2 index with the level of expression of all genes in the small airway epithelium of healthy smokers.^a

Probe set ID ^b	Gene symbol	Gene title	r ²	Location of ARE, bp ^c	References
201463_s_at	TALDO1	Transaldolase 1	0.81	-5578	(10,40)
205623_at	ALDH3A1	Aldehyde dehydrogenase 3 family, member A1	0.79	-7761	(23)
201467_s_at	NQO1	NAD(P)H dehydrogenase, quinone 1	0.79	-8447, -386	(16-21,41)
201266_at	TXNRD1	Thioredoxin reductase 1	0.77	-9409, -4225, -20, 20179, 21409, 21979, 2809	(18,19,22,23,26)
207469_s_at	PIR	Pirin	0.76	-5466, -3480, -3209, 33	None
204059_s_at	ME1	Malic enzyme 1, NADP(+) -dependent, cytosolic	0.76	-7400, -2164, -129	(10,23,26,37)
210519_s_at	NQO1	NAD(P)H dehydrogenase, quinone 1	0.74	-8447, -386	(10,16-23,41)
211653_x_at	AKR1C2	Aldo-keto reductase family 1, member C2	0.74	-8778, -6724, -5404, -2803, -1689, -1655	(42)
202831_at	GPX2	Glutathione peroxidase 2	0.73	-6897, -5842, -3600, -55	(10,17,23,25,26,37)
208700_s_at	TKT	Transketolase	0.70	-7278, -6731, -5713, -1261	(19)
203192_at	ABCB6	ATP-binding cassette, sub-family B, member 6	0.69	-7575	None
209699_x_at	AKR1C2	Aldo-keto reductase family 1, member C2	0.67	-8778, -6724, -5404, -2803, -1689, -1655	(42)
204532_x_at	UGT1A4	UDP glucuronosyltransferase 1 family, polypeptide A4	0.66	-5523, -3885, -3159, -3043	None
215125_s_at	UGT1A1-10	UDP glucuronosyltransferase 1 family, polypeptide A1-10 ^d	0.66	No unique gene ^d	
207126_x_at	UGT1A4	UDP glucuronosyltransferase 1 family, polypeptide A4	0.65	-5523, -3885, -3159, -3043	None
206094_x_at	UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6	0.65	-7826	(16-19,43)
209160_at	AKR1C3	Aldo-keto reductase family 1, member C3	0.65	No ARE ^e	
216594_x_at	AKR1C1	Aldo-keto reductase family 1, member C1	0.65	-9655, -6191, -4718, -3583, -1717	(42)
203925_at	GCLM	Glutamate-cysteine ligase, modifier subunit	0.63	-9	(10,19,23,25)
204151_x_at	AKR1C1	Aldo-keto reductase family 1, member C1	0.63	-9655, -6191, -4718, -3583, -1717	(42)
1555854_at	—	Transcribed locus	0.61	Not unique gene ^f	
208596_s_at	UGT1A1-10	UDP glucuronosyltransferase 1 family, polypeptide A1-10	0.61	Not unique gene ^d	
201471_s_at	SQSTM1	Sequestosome 1	0.60	-1313, -345	(23)
210505_at	ADH7	Alcohol dehydrogenase 7	0.60	-2859	(19,23)
225609_at	GSR	Glutathione reductase	0.55	-9339, -12	(10,23,24)

^aNrf2 index was correlated with the expression values of all genes expressed in the lung of healthy smokers ("P call" in > 20 % of samples), the 25 probe sets representing the highest correlation are listed.

^bProbe set identification number (ID) based on Affymetrix Human Genome U133 plus 2.0 Array

^cLocation of primary core sequence (RTGAYNNNGCR) of the antioxidant response element (ARE) in base pairs (bp) upstream (minus) or down stream (plus) of transcription site by searching the promoter area of each gene with Genamics Expression 1.1 Pattern Finder Tool software.

^dProbe set does not represent a unique gene. It includes several members of UDP-glucuronosyltransferase family, including *UGT1A1*, *UGT1A10*, *UGT1A3*, *UGT1A4*, *UGT1A5*, *UGT1A6*, *UGT1A7*, *UGT1A8*, *UGT1A9*.

^eNo antioxidant response element (ARE) identified with Genamics Expression 1.1 Pattern Finder Tool software

^fProbe set does not represent a unique gene. The transcribed locus lies between *AKRC1* and *AKRC2* genes.

gene. Using electrophoretic mobility shift assay, we found that some antioxidant response elements of *PIR* and *UGT1A4* responded *in vitro* to activated Nrf2. These observations are consistent with the concept that Nrf2 plays an important role in regulating cellular defenses against smoking in the highly vulnerable small airway epithelium cell population, and that there is variability among the

population in the relative Nrf2 responsiveness to a similar oxidant burden.

Nuclear Factor Erythroid 2-Related Factor 2

Oxidants and related radicals are known inducers of phase II detoxification and antioxidant pathways (1-4). The ability of cells to activate these pathways is an important cellular defense

system. Studies of the "Cap 'n' Collar" family of transcription factors identified Nrf2 as the master regulator of antioxidant protection (8-11). The cellular mechanism of Nrf2 activation is not fully understood and may include several pathways. It is known that phosphorylation of Nrf2 through protein kinases controls the binding of Nrf2 to Keap1 in the cytoplasm, where it rests in

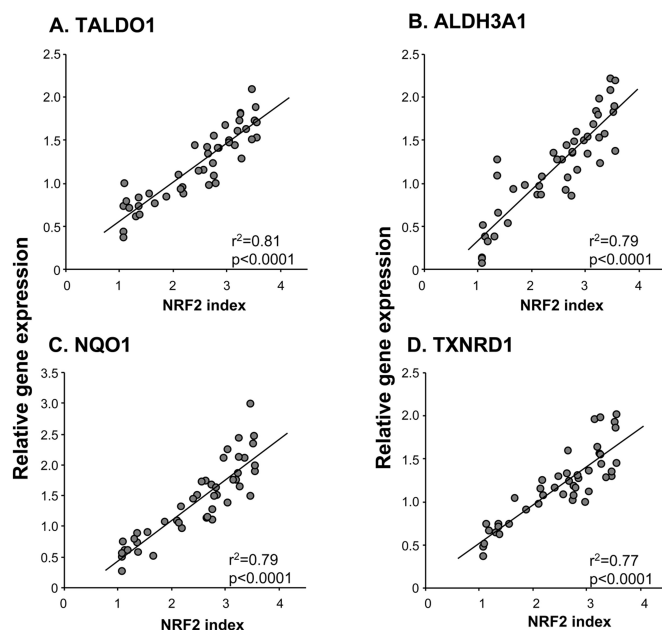


Figure 4. Correlation of expression of the Nrf2-modulated genes in small airway epithelium with the Nrf2 index. Shown are four examples of highly correlated genes. Each point in each panel represents a healthy smoker. (A) *Transaldolase 1* (TALDO1); (B) *aldehyde dehydrogenase 3 family, member A1* (ALDH3A1); (C) *NAD(P)H dehydrogenase, quinone 1* (NQO1); (D) *thioredoxin reductase 1* (TXNRD1).

the inactivated form and is targeted for ubiquitin-dependent degradation in the absence of antioxidant stress (12,13). Keap1 constitutively targets Nrf2, and this complex serves as a biological sensor in the cytosol (12,13). During oxidative stress, oxidants react with thiol groups of Keap1, eliciting a conformational change that causes a dissociation of Nrf2, allowing Nrf2 to translocate into the nucleus (12). An alternative model proposes that electrophilic stresses impair the Keap1-modulated proteosomal degradation of Nrf2, which enables Nrf2 synthesized *de novo* after exposure to the stress to accumulate directly in the nucleus, bypassing the Keap1 gate (13). Once in the nucleus, Nrf2 forms a heterodimer with muscle aponeurotic fibrosarcoma protein, and the heterodimer binds with high affinity to antioxidant response elements upstream of the antioxidant genes (8,12,13). In addition, mitogen-activated protein kinases also regulate antioxidant response element binding activity of Nrf2 (8).

The role of Nrf2 in the lung has been significantly advanced by studies of Nrf2 knockout mice exposed to a variety of mediators of lung injury. After stress, lungs of Nrf2-deficient mice are characterized by inflammation, evidence of oxidative stress, and increased apoptosis compared with the wild-type littermates (14–26). Microarray analyses of gene expression in lungs from Nrf2 knockout mice and wild-type controls have helped to identify Nrf2-modulated genes, for example, genes that help protect mice from development of cigarette smoke-induced lung injuries (23). Comprehensive studies of changes in gene expression using microarrays have helped to identify novel Nrf2-modulated genes (19,22,23). For example, Rangasamy *et al.* (23) have shown that after exposure to cigarette smoke, Nrf2 induces expression of 45 genes in the lung, mainly related to the glutathione, thioredoxin, and NADPH regenerating systems. Similarly, oxidative lung injury upregulated 125 Nrf2-

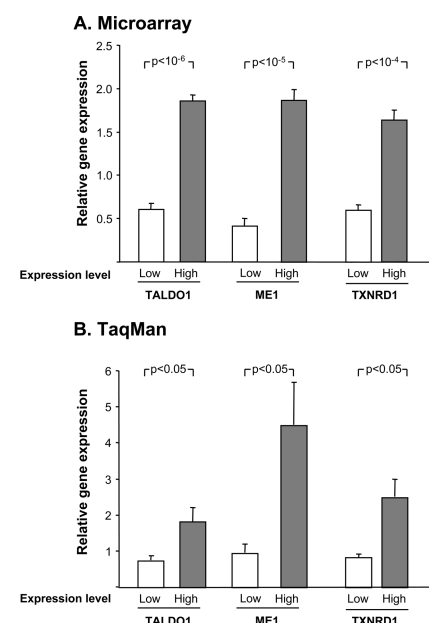


Figure 5. Quantitative TaqMan RT-PCR confirmation of the microarray data. TaqMan RT-PCR was used to assess three Nrf2-modulated genes by use of RNA samples from six healthy smokers classified as low expressors (Nrf2 index 1–2) and six healthy smokers who were categorized as high expressors (Nrf2 index 3–4). The normalized average gene expression level is represented on the ordinate, and the individual genes are represented on the abscissa. (A) Microarray data; (B) TaqMan RT-PCR data. Data shown are mean \pm standard error. TALDO1, transaldolase 1; ME1, malic enzyme 1; TXNRD1, thioredoxin reductase 1.

modulated genes, including many known antioxidant genes, but also novel genes, such as *protein kinase C α* (19), a gene linked to asthma and pulmonary adenocarcinoma (44). Aoki *et al.* (15) demonstrated that Nrf2 knockout mice bred with transgenic mice that are susceptible to *in vivo* mutations (guanine phosphoribosyltransferase delta mice) have enhanced spontaneous and benz(a)pyrene-induced mutation rates in the lungs, suggesting that Nrf2 also protects genomic DNA from certain types of cancer. These observations are consistent with the concept that Nrf2-modulated genes confer protection

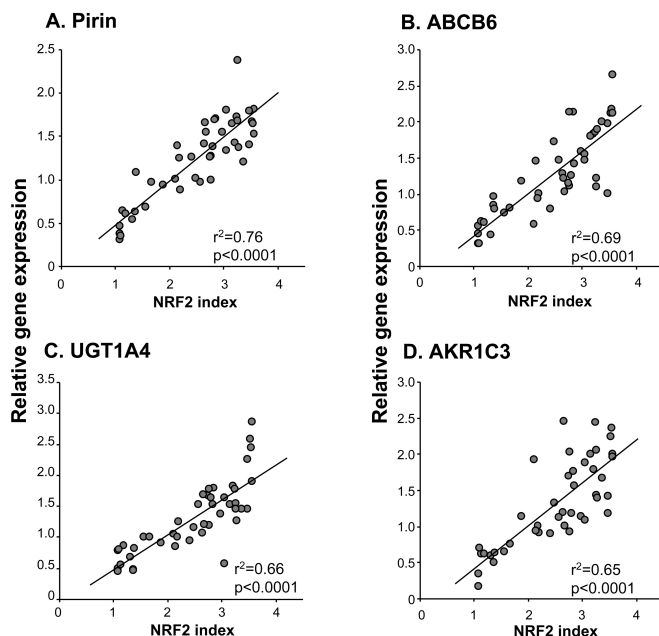


Figure 6. Identification of human Nrf2-modulated genes not previously linked to Nrf2 control. Shown are genes not previously reported as transcriptionally induced by Nrf2, but whose expression in the small airway epithelium is strongly associated with the Nrf2 index. Each point in each panel represents a healthy smoker with Nrf2 index plotted against expression level for: (A) *Pirin* (*PIR*); (B) *ATP-binding cassette, sub-family B, member 6* (*ABCB6*); (C) *UDP gluconosyltransferase 1 family, polypeptide A4* (*UGT1A4*); (D) *Aldo-keto reductase family 1, member C3* (*AKR1C3*).

against carcinogenesis. For example, the *NQO1* gene catalyzes the detoxification of quinones and stabilizes the tumor suppressor p53 protein (45,46). Additionally, *NQO1*-deficient individuals have a higher risk of developing leukemia (47,48) and other type of cancers (49,50). Another example is the family members of glutathione-S-transferases (GSTs) that catalyze the conjugation of reduced glutathione. Humans lacking GST have a increased risk of developing cancer in many organs, including breast, colorectal, thyroid, lung, stomach, prostate, and bladder (51–57). Some cancer chemopreventive agents, such as isothiocyanates, induce Nrf2-modulated genes and are protective against carcinogen-induced tumorigenesis (10,58,59). Some cancer cell lines are also characterized by high induction of Nrf2-modulated genes. For example, mutations in Keap1 of lung cancer cell lines lead to an increased constitutive

expression of Nrf2-modulated genes, with associated increased resistance to chemotherapeutic drugs (60–62).

Oxidants in Cigarette Smoke and the Response of Small Airway Epithelium to Smoking

Cigarette smoke can be divided into particulate matter and a gas phase, both rich sources of oxidants (1–4,63). The gas phase of cigarette smoke contains a large number of various species of oxidants and free radicals, including nitrogen-, carbon-, and oxygen-centered radicals, and produces esters and peroxyesters of nitrous and nitric acids (63,64). The radicals in tar include a mixture of semi-quinones, hydroquinones, and quinones, which can penetrate mammalian cells and bind to and nick DNA (63,65). Other polyphenolic agents, such as catechol, also generate hydrogen peroxide and superoxide free radicals (63). It is estimated that more than 10^{14} free radicals are in-

haled per puff in the respiratory tract, putting an enormous oxidative stress on the small airway epithelium (1,27–30). The epithelial surface of the respiratory tract has an antioxidant defense system that includes glutathione, vitamins C and E, β -carotene, uric acid, and the enzymatic antioxidants, such as superoxide dismutases, catalases, enzymes associated with glutathione metabolism, and other proteins such as peroxiredoxins, thioredoxins, glutaredoxins, heme oxygenases and reductases (2,4,11). The homeostasis of lung cells depends on these antioxidants, and lung injury results where oxidative stress exceeds antioxidant defenses. In smokers, cigarette smoking induces a broad spectrum of oxidant-related genes in the airway epithelium not seen in non-smokers (11,12,20,23,28–30,33–35,66,67). Interestingly, Spira *et al.* (67) have evaluated gene expression in the large airway epithelium of smokers with possible lung cancer and found that oxidant related defense genes were downregulated in smokers with proven lung cancers as compared with smokers without lung cancer.

Nrf2-Modulated Genes in the Human Small Airway Epithelium in Response to Cigarette Smoking

Based on all reports of genes responsive to different oxidative stress conditions in the lungs of Nrf2 knockout mice (14,16–26), we compiled a list of 201 murine oxidative stress-responsive genes in the lung, and 187 human homologs were identified as expressed in small airway epithelium in humans. Using the gene expression data from 45 smokers and 38 nonsmokers, we identified a final list of 13 putative human, Nrf2-modulated, highly smoking-responsive genes. In agreement with the murine data (14,16–26), these genes showed a coordinate regulation in the small airway epithelium with certain smokers showing a consistently high level or consistently low level of Nrf2-modulated genes by microarray analysis with TaqMan-quantitative confirmation. Interestingly, the Nrf2 index was not dependent on current smoking, smoking history, or other demographic

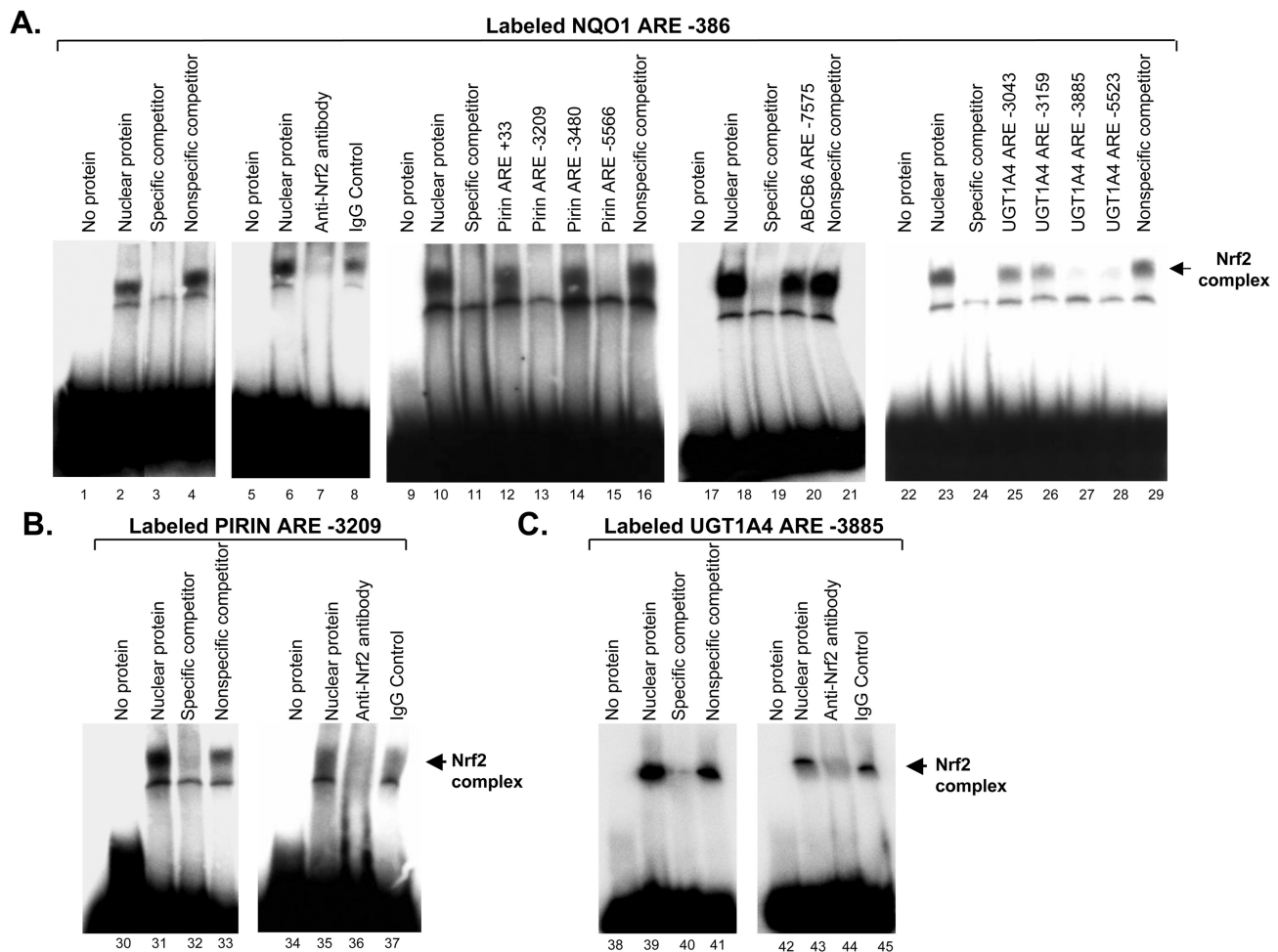


Figure 7. Validation of newly identified Nrf2-responsive genes by electrophoretic mobility shift assay. Nuclear extract from small airway epithelium was incubated with 32 P-labeled antioxidant response element (ARE) oligonucleotides and specifically competed with a fifty-fold excess of ARE oligonucleotides from the same gene or with ARE oligonucleotides from putative Nrf2-modulated genes. (A) *NAD(P)H dehydrogenase, quinone 1 (NQO1)* ARE, located at -386 bp upstream of the transcription site. Lane 1, no protein; lane 2, nuclear protein alone; lane 3, nuclear protein + specific competitor (NQO1 unlabeled ARE); lane 4, nuclear protein + nonspecific competitor (von Willebrand (vWF) oligonucleotides); lane 5, no protein; lane 6, nuclear protein alone; lane 7, nuclear protein + anti-Nrf2 antibody; lane 8, nuclear protein + IgG control; lane 9, no protein; lane 10, nuclear protein alone; lane 11, nuclear protein + specific competitor (NQO1 unlabeled ARE); lane 12, nuclear protein + pirin (PIR) unlabeled ARE at position +33 bp (downstream of the transcription start site); lane 13, nuclear protein + PIR unlabeled ARE at position -3209 bp (upstream of the transcription start site); lane 14, nuclear protein + PIR unlabeled ARE at position -3480 bp; lane 15, nuclear protein + PIR unlabeled ARE at position -5566 bp; lane 16, nuclear protein + nonspecific competitor (vWF oligonucleotides); lane 17, no protein; lane 18, nuclear protein alone; lane 19, nuclear protein + specific competitor (NQO1 unlabeled ARE); lane 20, nuclear protein + ABCB6 unlabeled ARE at position -7575 bp; lane 21, nuclear protein + nonspecific competitor (vWF oligonucleotides); lane 22, no protein; lane 23, nuclear protein alone; lane 24, nuclear protein + specific competitor (NQO1 unlabeled ARE); lane 25, nuclear protein + UGT1A4 unlabeled ARE at -3043 bp upstream of the transcription start site; lane 26, nuclear protein + UGT1A4 unlabeled ARE at position -3159; lane 27, nuclear protein + UGT1A4 unlabeled ARE at position -3885 bp; lane 28, nuclear protein + UGT1A4 unlabeled ARE at position -5523 bp; lane 29, nuclear protein + nonspecific competitor (von Willebrand factor (vWF) oligonucleotides). For lane 13, lane 15, lane 27, and lane 28, disappearance of gel-shift band indicates functional ARE. (B) Pirin (PIR) ARE. Radioactively labeled PIR-ARE at position -3209 bp (lane 30-37). Lane 30, no protein; lane 31, nuclear protein alone; lane 32, nuclear protein + specific competitor (PIR unlabeled ARE); lane 33, nuclear protein + nonspecific competitor (vWF oligonucleotides); lane 34, no protein; lane 35, nuclear protein alone; lane 36, nuclear protein + anti-Nrf2 antibody; lane 37, nuclear protein + IgG control; (C) radioactively labeled UGT1A4-ARE at position -3885 bp (lane 38-45). Lane 38, no protein; lane 39, nuclear protein alone; lane 40, nuclear protein + specific competitor (UGT1A4 unlabeled ARE); lane 41, nuclear protein + nonspecific competitor (vWF oligonucleotides); lane 42, no protein; lane 43, nuclear protein alone; lane 44, nuclear protein + anti-Nrf2 antibody; lane 45, nuclear protein + IgG control.

characteristics such as sex, age, and ethnicity, raising the possibility that the variability of Nrf2-modulated gene expression is, at least in part, genetically determined. This variability among smokers might lead to the hypothesis for future studies that healthy smokers with low Nrf2-modulated gene expression may represent a group of individuals at higher risk for smoking-induced lung injuries. A recent study identified a single polymorphism in the promoter of Nrf2 significantly associated with a higher risk of acute lung injury after major trauma in humans, and proposed Nrf2 as a candidate susceptibility gene for oxidative related illnesses (68). This is in agreement with a report published by Malhotra *et al.* (69), that the level of expression of Nrf2-modulated genes in the lung parenchyma were negatively correlated with severity of chronic obstructive lung disease.

The Nrf2 index, which is based on the expression level of the 13 putative human highly smoking-responsive genes, was used to identify further Nrf2-responsive genes expressed in human small airway epithelium. Most of the top genes which were identified as having a strong association with the Nrf2 index have been previously reported as Nrf2-modulated on the basis of transcription induction by Nrf2 in studies using Nrf2 knockout mice and/or by gel shift assays specific for Nrf2. Two additional highly smoking-responsive genes were identified, *PIR* and *UGT1A4*, both of which contain several functional antioxidant response elements. *PIR*, a gene of the cubin family, has been recently described to be highly upregulated by chronic cigarette smoking, and it was associated with bronchial epithelium cell apoptosis (70). Interestingly, the structure of pirin comprises two β -barrel domains, with an iron-binding cofactor, suggesting an enzymatic role in biological redox reactions (71). *UGT1A4* is part of the complex locus encoding several UDP-glucuronosyltransferases, which are enzymes of the glucuronidation pathway that transforms small lipophilic molecules, such as steroids, bilirubin, hormones, and drugs, into water-soluble,

excretable metabolites (43). Although other UDP-glucuronosyltransferases have already been connected with Nrf2 (16–19), *UGT1A4* has not specifically been shown to have a functional antioxidant response element.

In conclusion, these observations demonstrate that Nrf2 plays an important role in regulating cellular defenses in highly vulnerable cell populations such as the small airway epithelium after cigarette smoking. Nrf2-modulated genes have been found to be upregulated in healthy cigarette smokers in a concordant pattern. The variability among the population in the relative Nrf2 responsiveness to a similar oxidant burden might suggest a genetic determination, which would have implications in the risk for lung disease related to cigarette smoking.

ACKNOWLEDGMENTS

We thank the Fellows of the Division of Pulmonary and Critical Care Medicine for assistance with screening of subjects and bronchoscopies, T O'Connor and P Leopold for helpful discussions; M Teater, H Webb, T Raman, and D Dang for excellent technical assistance; H Lou for providing differential counts; and T Virgin-Bryan and N Mohamed for help in preparing this manuscript. These studies were supported, in part, by R01 HL074326, P50 HL084936 and UL1RR024996 (the WMC CTSC); and the Will Rogers Memorial Fund, Los Angeles, CA.

DISCLOSURE

We declare that the authors have no competing interests as defined by *Molecular Medicine*, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES

1. Pryor WA, Prier DG, Church DF. (1983) Electron-spin resonance study of mainstream and side-stream cigarette smoke: nature of the free radicals in gas-phase smoke and in cigarette tar. *Environ. Health Perspect.* 47:345–55.
2. Rahman I, MacNee W. (1996) Role of oxidants/antioxidants in smoking-induced lung diseases. *Free Radic. Biol. Med.* 21:669–81.
3. MacNee W. (2000) Oxidants/antioxidants and COPD. *Chest* 117:303S–17S.

4. Rahman I, Biswas SK, Kode A. (2006) Oxidant and antioxidant balance in the airways and airway diseases. *Eur. J. Pharmacol.* 533:222–39.
5. Denissenko MF, Pao A, Tang M, Pfeifer GP. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* 274:430–2.
6. Hecht SS. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat. Rev. Cancer* 3:733–44.
7. Feng Z, Hu W, Hu Y, Tang MS. (2006) Acrolein is a major cigarette-related lung cancer agent: Preferential binding at p53 mutational hotspots and inhibition of DNA repair. *Proc. Natl. Acad. Sci. U. S. A.* 103:15404–9.
8. Igarashi K, *et al.* (1994) Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins. *Nature* 367:568–72.
9. Moi P, Chan K, Asunis I, Cao A, Kan YW. (1994) Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. *Proc. Natl. Acad. Sci. U. S. A.* 91:9926–30.
10. Thimmulappa RK, *et al.* (2002) Identification of Nrf2-regulated genes induced by the chemopreventive agent sulforaphane by oligonucleotide microarray. *Cancer Res.* 62:5196–203.
11. Cho HY, Reddy SP, Kleeberger SR. (2006) Nrf2 defends the lung from oxidative stress. *Antioxid. Redox. Signal.* 8:76–87.
12. Itoh K, *et al.* (1999) Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes. Dev.* 13:76–86.
13. Kobayashi A, *et al.* (2006) Oxidative and electrophilic stresses activate Nrf2 through inhibition of ubiquitination activity of Keap1. *Mol. Cell Biol.* 26:221–9.
14. Aoki Y, *et al.* (2001) Accelerated DNA adduct formation in the lung of the Nrf2 knockout mouse exposed to diesel exhaust. *Toxicol. Appl. Pharmacol.* 173:154–60.
15. Aoki Y, *et al.* (2007) Enhanced spontaneous and benzo(a)pyrene-induced mutations in the lung of Nrf2-deficient gpt delta mice. *Cancer Res.* 67:5643–8.
16. Chan K, Kan YW. (1999) Nrf2 is essential for protection against acute pulmonary injury in mice. *Proc. Natl. Acad. Sci. U. S. A.* 96:12731–6.
17. Cho HY, *et al.* (2002) Role of NRF2 in protection against hyperoxic lung injury in mice. *Am. J. Respir. Cell Mol. Biol.* 26:175–82.
18. Cho HY, Reddy SP, Yamamoto M, Kleeberger SR. (2004) The transcription factor NRF2 protects against pulmonary fibrosis. *FASEB J.* 18:1258–60.
19. Cho HY, Reddy SP, Debiase A, Yamamoto M, Kleeberger SR. (2005) Gene expression profiling of NRF2-mediated protection against oxidative injury. *Free Radic. Biol. Med.* 38:325–43.
20. Iizuka T, *et al.* (2005) Nrf2-deficient mice are highly susceptible to cigarette smoke-induced emphysema. *Genes Cells* 10:1113–25.
21. Ishii Y, *et al.* (2005) Transcription factor Nrf2 plays a pivotal role in protection against elastase-induced pulmonary inflammation and emphysema. *J. Immunol.* 175:6968–975.

22. Papaiahgari S, *et al.* (2007) Genetic and pharmacologic evidence links oxidative stress to ventilator-induced lung injury in mice. *Am. J. Respir. Crit. Care Med.* 176:1222–35.
23. Rangasamy T, *et al.* (2004) Genetic ablation of Nrf2 enhances susceptibility to cigarette smoke-induced emphysema in mice. *J. Clin. Invest.* 114:1248–59.
24. Rangasamy T, *et al.* (2005) Disruption of Nrf2 enhances susceptibility to severe airway inflammation and asthma in mice. *J. Exp. Med.* 202:47–59.
25. Singh A, *et al.* (2006) Glutathione peroxidase 2, the major cigarette smoke-inducible isoform of GPX in lungs, is regulated by Nrf2. *Am. J. Respir. Cell Mol. Biol.* 35:639–50.
26. Thimmulappa RK, *et al.* (2006) Nrf2 is a critical regulator of the innate immune response and survival during experimental sepsis. *J. Clin. Invest.* 116:984–95.
27. Hogg JC, Macklem PT, Thurlbeck WM. (1968) Site and nature of airway obstruction in chronic obstructive lung disease. *N. Engl. J. Med.* 278:1355–60.
28. Spira A, *et al.* (2004) Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc. Natl. Acad. Sci. U. S. A.* 101:10143–8.
29. Spira A, *et al.* (2004) Gene expression profiling of human lung tissue from smokers with severe emphysema. *Am. J. Respir. Cell Mol. Biol.* 31:601–10.
30. Adair-Kirk TL, *et al.* (2008) Distal airways in mice exposed to cigarette smoke: Nrf2-regulated genes are increased in Clara cells. *Am. J. Respir. Cell Mol. Biol.* 39:400–11.
31. Moyer TP, *et al.* (2002) Simultaneous analysis of nicotine, nicotine metabolites, and tobacco alkaloids in serum or urine by tandem mass spectrometry, with clinically relevant metabolic profiles. *Clin. Chem.* 48:1460–71.
32. Scherer G. (2006) Carboxyhemoglobin and thiocyanate as biomarkers of exposure to carbon monoxide and hydrogen cyanide in tobacco smoke. *Exp. Toxicol. Pathol.* 58:101–24.
33. Harvey BG, *et al.* (2007) Modification of gene expression of the small airway epithelium in response to cigarette smoking. *J. Mol. Med.* 85:39–53.
34. Ammous Z, *et al.* (2008) Variability in small airway epithelial gene expression among normal smokers. *Chest* 133:1344–53.
35. Kode A, *et al.* (2008) Resveratrol induces glutathione synthesis by activation of Nrf2 and protects against cigarette smoke-mediated oxidative stress in human lung epithelial cells. *Am. J. Physiol. Lung Cell Mol. Physiol.* 294:L478–88.
36. Wasserman WW, Fahl WE. (1997) Functional antioxidant responsive elements. *Proc. Natl. Acad. Sci. U. S. A.* 94:5361–6.
37. Tirumalai R, Rajesh KT, Mai KH, Biswal S. (2002) Acrolein causes transcriptional induction of phase II genes by activation of Nrf2 in human lung type II epithelial (A549) cells. *Toxicol. Lett.* 132:27–36.
38. Yueh MF, and Tukey RH. (2007) Nrf2-Keap1 signaling pathway regulates human UGT1A1 expression in vitro and in transgenic UGT1 mice. *J. Biol. Chem.* 282:8749–58.
39. Hellman LM, Fried MG. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* 2:1849–61.
40. Banki K, *et al.* (1997) The human transaldolase gene (TALDO1) is located on chromosome 11 at p15.4-p15.5. *Genomics* 45:233–8.
41. Ross D, *et al.* (2000) NAD(P)H:quinone oxidoreductase 1 (NQO1): chemoprotection, bioactivation, gene regulation and genetic polymorphisms. *Chem. Biol. Interact.* 129:77–97.
42. Lou H, Du S, Ji Q, Stolz A. (2006) Induction of AKR1C2 by phase II inducers: identification of a distal consensus antioxidant response element regulated by NRF2. *Mol. Pharmacol.* 69:1662–72.
43. Tukey RH, Strassburg CP. (2000) Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu. Rev. Pharmacol. Toxicol.* 40:581–616.
44. Dwyer-Nield LD, Paigen B, Porter SE, Malkinson AM. (2000) Quantitative trait locus mapping of genes regulating pulmonary PKC activity and PKC- α content. *Am. J. Physiol. Lung Cell Mol. Physiol.* 279:L326–32.
45. Asher G, Lotem J, Kama R, Sachs L, Shaul Y. (2002) NQO1 stabilizes p53 through a distinct pathway. *Proc. Natl. Acad. Sci. U. S. A.* 99:3099–104.
46. Gong X, Kole L, Iskander K, Jaiswal AK. (2007) NRH:quinone oxidoreductase 2 and NAD(P)H:quinone oxidoreductase 1 protect tumor suppressor p53 against 20s proteasomal degradation leading to stabilization and activation of p53. *Cancer Res.* 67:5380–8.
47. Rothman N, *et al.* (1997) Benzene poisoning, a risk factor for hematological malignancy, is associated with the NQO1 609C→T mutation and rapid fractional excretion of chlorzoxazone. *Cancer Res.* 57:2839–42.
48. Wiemels JL, *et al.* (1999) A lack of a functional NAD(P)H:quinone oxidoreductase allele is selectively associated with pediatric leukemias that have MLL fusions. United Kingdom Childhood Cancer Study Investigators. *Cancer Res.* 59:4095–9.
49. Clairmont A, *et al.* (1999) Association of NAD(P)H:quinone oxidoreductase (NQO1) null with numbers of basal cell carcinomas: use of a multivariate model to rank the relative importance of this polymorphism and those at other relevant loci. *Carcinogenesis* 20:1235–40.
50. Fagerholm R, *et al.* (2008) NAD(P)H:quinone oxidoreductase 1 NQO1(*2) genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat. Genet.* 40:844–53.
51. Garcia-Closas M, *et al.* (2005) NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 366:649–59.
52. Terry PD, Goodman M. (2006) Is the association between cigarette smoking and breast cancer modified by genotype? A review of epidemiologic studies and meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* 15:602–11.
53. Cotton SC, Sharp L, Little J, Brockton N. (2000) Glutathione S-transferase polymorphisms and colorectal cancer: a HuGE review. *Am. J. Epidemiol.* 151:7–32.
54. Granja F, *et al.* (2004) GST profiling may be useful in the screening for thyroid nodule malignancy. *Cancer Lett.* 209:129–37.
55. Carlsten C, Sagoo GS, Frodsham AJ, Burke W, Higgins JP. (2008) Glutathione S-transferase M1 (GSTM1) polymorphisms and lung cancer: a literature-based systematic HuGE review and meta-analysis. *Am. J. Epidemiol.* 167:759–74.
56. Boccia S, La TG, Gianfagna F, Mannocci A, Ricciardi G. (2006) Glutathione S-transferase T1 status and gastric cancer risk: a meta-analysis of the literature. *Mutagenesis* 21:115–23.
57. Agalliu I, Langeberg WJ, Lampe JW, Salinas CA, Stanford JL. (2006) Glutathione S-transferase M1, T1, and P1 polymorphisms and prostate cancer risk in middle-aged men. *Prostate* 66:146–56.
58. Fahey JW, Zhang Y, Talalay P. (1997) Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. *Proc. Natl. Acad. Sci. U. S. A.* 94:10367–72.
59. Munday R, *et al.* (2008) Inhibition of urinary bladder carcinogenesis by broccoli sprouts. *Cancer Res.* 68:1593–600.
60. Ohta T, *et al.* (2008) Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. *Cancer Res.* 68:1303–9.
61. Padmanabhan B, *et al.* (2006) Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. *Mol. Cell* 21:689–700.
62. Singh A, *et al.* (2006) Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer. *PLoS. Med.* 3:e420.
63. Pryor WA. (1997) Cigarette smoke radicals and the role of free radicals in chemical carcinogenicity. *Environ. Health Perspect.* 105 Suppl 4:875–82.
64. Pryor WA, and Stone K. (1993) Oxidants in cigarette smoke: radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite. *Ann. N. Y. Acad. Sci.* 686:12–27.
65. Stone KK, Bermudez E, Pryor WA. (1994) Aqueous extracts of cigarette tar containing the tar free radical cause DNA nicks in mammalian cells. *Environ. Health Perspect.* 102 Suppl 10:173–8.
66. Hackett NR, *et al.* (2003) Variability of antioxidant-related gene expression in the airway epithelium of cigarette smokers. *Am. J. Respir. Cell Mol. Biol.* 29:331–43.
67. Spira A, *et al.* (2007) Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat. Med.* 13:361–6.
68. Marzec JM, *et al.* (2007) Functional polymorphisms in the transcription factor NRF2 in humans increase the risk of acute lung injury. *FASEB J.* 21:2237–46.
69. Malhotra D, *et al.* (2008) Decline in NRF2 regulated antioxidants in COPD lungs due to loss of its positive regulator DJ-1. *Am. J. Respir. Crit. Care Med.*
70. Gelbman BD, Heguy A, O'Connor TP, Zabner J, Crystal RG. (2007) Upregulation of pirin expression by chronic cigarette smoking is associated with bronchial epithelial cell apoptosis. *Respir. Res.* 8:10.
71. Pang H, *et al.* (2004) Crystal structure of human pirin: an iron-binding nuclear protein and transcription cofactor. *J. Biol. Chem.* 279:1491–8.