Published in final edited form as:

Immunogenetics. 2010 April ; 62(4): 237–251. doi:10.1007/s00251-010-0431-6.

Impact of female cigarette smoking on circulating B cells in vivo: the suppressed *ICOSLG*, *TCF3*, and *VCAM1* gene functional network may inhibit normal cell function

Feng Pan

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Tie-Lin Yang

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Orthopedic Surgery and Basic Medical Science, School of Medicine, University of Missouri-Kansas City, 2411 Holmes Street. Room M3-C03, Kansas City, MO 64108-2792, USA

Xiang-Ding Chen

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Laboratory of Molecular and Statistical Genetics, College of Life Sciences Hunan Normal University, Changsha, Hunan 410081, People's Republic of China

Yuan Chen

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Ge Gao

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Yao-Zhong Liu

Department of Orthopedic Surgery and Basic Medical Science, School of Medicine, University of Missouri-Kansas City, 2411 Holmes Street. Room M3-C03, Kansas City, MO 64108-2792, USA

Yu-Fang Pei

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Orthopedic Surgery and Basic Medical Science, School of Medicine, University of Missouri-Kansas City, 2411 Holmes Street. Room M3-C03, Kansas City, MO 64108-2792, USA

Bao-Yong Sha

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Yan Jiang

Institute of Systems Science, School of Management, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China

Chao Xu

Institute of Systems Science, School of Management, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China

Robert R. Recker

Department of Biomedical Sciences and Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE 68131, USA

Hong-Wen Deng

The Key Laboratory of Biomedical Information Engineering of Ministry of Education, and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

Department of Orthopedic Surgery and Basic Medical Science, School of Medicine, University of Missouri-Kansas City, 2411 Holmes Street. Room M3-C03, Kansas City, MO 64108-2792, USA

Laboratory of Molecular and Statistical Genetics, College of Life Sciences Hunan Normal University, Changsha, Hunan 410081, People's Republic of China

Center of Systematic Biomedical Research, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China

Hong-Wen Deng: dengh@umkc.edu

Abstract

As pivotal immune guardians, B cells were found to be directly associated with the onset and development of many smoking-induced diseases. However, the *in vivo* molecular response of B cells underlying the female cigarette smoking remains unknown. Using the genome-wide Affymetrix HG-133A GeneChip[®] microarray, we firstly compared the gene expression profiles of peripheral circulating B cells between 39 smoking and 40 non-smoking healthy US white women. A total of 125 differential expressed genes were identified in our study, and 75.2% of them were down-regulated in smokers. We further obtained genotypes of 702 single nucleotide polymorphisms in those promising genes and assessed their associations with smoking status. Using a novel multicriteria evaluation model integrating information from microarray and the association studies, several genes were further revealed to play important roles in the response of smoking, including ICOSLG (CD275, inducible T-cell co-stimulator ligand), TCF3 (E2A immunoglobulin enhancer binding factors E12/E47), VCAM1 (CD106, vascular cell adhesion molecule 1), CCR1 (CD191, chemokine C-C motif receptor 1) and IL13 (interleukin 13). The differential expression of ICOSLG ($p = 0.0130$) and TCF3 ($p = 0.0125$) genes between the two groups were confirmed by realtime reverse transcription PCR experiment. Our findings support the functional importance of the identified genes in response to the smoking stimulus. This is the first *in vivo* genome-wide expression study on B cells at today's context of high prevalence rate of

smoking for women. Our results highlight the potential usage of integrated analyses for unveiling the novel pathogenesis mechanism and emphasized the significance of B cells in the etiology of smoking-induced disease.

Keywords

Cigarette smoking; B cells; Microarray; Genome-wide; Association

Introduction

Cigarette smoking, the leading cause of death in USA, accounts for ~440,000 deaths and costs more than 167 billion dollars, annually (Marshall et al. 2006). Smoking is the predominant cause of lung cancer (Alberg et al. 2007; Gandini et al. 2008) and highly associated with the onset and development of cancers, stroke, heart attack, vascular disease, respiratory diseases, and other severe chronic disease, such as chronic obstructive pulmonary disease (COPD) (Bergen and Caporaso 1999).

The components of inhalational smoke enter the body through the pulmonary alveoli and are distributed throughout all tissues by blood via capillary vessel. Thus, the smoking components travel within the blood, which contains immune cells such as B cells, T cells, monocytes, and other active components. These immune guardians, such as B cells, may react to smoking stimulus and exert on other organs and tissues with systematical effects (Buttner et al. 2007). Molecular studies showed that several nicotinic receptors up-regulated in the B cells in the presence of smoking components (Cloeze-Tayarani and Changeux 2007; Skok et al. 2003). Long-term exposure to nicotine can induce B cells to decrease the antibody secretion, inhibit the cell proliferation and development, and finally suppress the normal immune function (Skok et al. 2003, 2005, 2007).

As a result, smoking may induce abnormal changes in the immune system and cause diseases via the dysregulation of impaired B cells. Particularly, many studies have demonstrated that smoking can suppress the normal function of B cells (Chan et al. 1990; Geng et al. 1995; Moszczynski et al. 2001; Savage et al. 1991; Yang et al. 1999) and hence increase the risk of COPD and emphysema (Brandsma et al. 2008; van der Strate et al. 2006), asthma (Tsoumakidou et al. 2007), chronic periodontitis (Al-Ghamdi and Anil 2007; Barbour et al. 1997), and inflammatory reaction (Bosken et al. 1992). The healthy consequences of smoking for women are worse than men (Mackay and Amos 2003). Cohort studies showed that female smokers appear to have increased susceptibility to tobacco carcinogens (Henschke et al. 2006). Women are more vulnerable to the cigarette-smoking-induced respiratory symptoms (asthma, chronic bronchitis, wheezing, and/or breathlessness) than men (Langhammer et al. 2000). Smoking affects the fertility of women and causes early menopause (Langhammer et al. 2000). In the USA and the UK, cigarette smoking becomes the single most important preventable cause of premature death in adult women, accounting for at least a third of all deaths in the age from 35 to 69 (Amos 1996).

Microarray has been employed to provide novel clues of the pathogenesis of smoking-induced disease in given specific samples. The peripheral mononuclear white cells (PMWCs), which contain mixed immune cells, had already been used for gene expression studies in response to smoking in vitro (Ryder et al. 2004; van Leeuwen et al. 2005, 2007). A pilot in vivo study further purified the fresh T cells from PMWCs and applied microarray to investigate the role of T cells in male smokers (Buttner et al. 2007). The differentially expressed genes (DEGs) identified by those studies contributed to illuminating the involvement of immune cells in smoking-induced reactions. Although many clinical and

cohort studies indicated that women are more sensitive to smoking than man, studies seldom focused on the in vivo molecular responses and interactions of genes for the immune cells underlying the female smoking behavior. This may be because, in many countries (especially in the developing countries), cigarette smoking still tends to be regarded as a major male healthy issue (Mackay and Amos 2003; Amos 1996). One other clinical reason could be because women have a lower rate of death of lung cancer compared with men after diagnosis (Henschke et al. 2006), which appeared to relieve the serious results of smoking in women to a certain extent.

In this study, using the genome-wide Affymetrix HG-133A GeneChip[®] microarray, we compared the gene expression profiles of peripheral circulating B cells between smoking and non-smoking white women in the USA. Knowledge about the identified DEGs may fill the gaps of existing female smoking studies and further contribute to our understanding of pathogenesis of disease induced by smoking in the context of an increasing prevalence rate of cigarette smoking among female population (Bergen and Caporaso 1999). This is the first genome-wide expression study on in vivo human B cells relating to smoking in females. A novel ICOSLG, TCF3, and vascular cell adhesion molecule 1 (VCAM1) gene network was suggested in B cells for the response of smoking-induced reactions. The results can also be used to accelerate the progress of gender-specific treatment or approaches for effective tobacco control.

Materials and methods

Subjects

The study was approved by the Institutional Review Board, and all the subjects signed informed consent documents before entering the project. All the study subjects were US whites of European origin living in Omaha, NE and its surrounding areas.

Microarray and RT-PCR study sample—We recruited 80 unrelated healthy women age 40–60, comprising 40 smokers and 40 non-smokers. The 40 subjects in each group can be also equally categorized to 20 premenopause and 20 postmenopause women. This recruitment strategy was adopted to avoid the potential effects on B cell gene expression profiling due to different menopausal status. A total number of 79 subjects were finally used for this study. We dropped one individual because the corresponding microarray chip was failed or damaged with abnormal background signal and thus cannot pass the strict quality control of dChip software. Smoking-related data were recorded using a nurse-administered questionnaire, which also included a detailed medication and disease history. Subjects were categorized as “smokers” based on recorded answers of smoking history in the questionnaire. The 39 current smokers were actively smoking 1.0 ± 0.5 pack/day and had a average 31.89 ± 9.37 years (mean $\pm$ SD) smoking history with the range from 14 to 48 years. In the nonsmoking group, only one individual (Nsm-52, 56 years old) smoked before when she is 19 and quit at her age of 29 years. After carefully inspecting, the other 39 women never smoked before. All of these 40 subjects were thus defined as “non-smokers.” There were no differences in age ($p > 0.7$) between the smokers (49.37 ± 8.05 years) and non-smokers (49.98 ± 7.92 years). Additional strict exclusion criteria elaborated in Appendix 2 of the Electronic supplementary material were used to minimize other potential effects of non-smoking factors on B cell gene expression changes.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-010-0431-6) contains supplementary material, which is available to authorized users.

Association study sample—A random sample containing 840 unrelated whites (417 men and 423 women) was identified from our established and expanding genetic repertoire currently containing more than ~10,000 subjects. Similarly, subjects were categorized as “smokers” based on their answer in the questionnaire. A subject who had never smoked was defined as a “non-smoker.” For non-smokers, we excluded those subjects younger than the age of 25. There are 224 and 154 qualified smokers in the male and female subgroup, respectively. We documented the detail subject characteristics in our recent genome-wide smoking association study published elsewhere (Liu et al. 2009a).

Experimental procedures

B cell isolation, total RNA extraction, and gene chip preparation—Seventy milliliters of blood were drawn from each recruited female. B cell isolation from 70 ml whole blood was performed using a positive isolation method with Dynabeads® CD19 (Pan B) and DETACHaBEAD® CD19 (DynaL Biotech, Lake Success, NY, USA) following the manufacturer’s protocols. B cell purity was assessed by flow cytometry (BD Biosciences, San Jose, CA, USA) with fluorescence-labeled antibodies, PE-CD19 and FITC-CD45. The average purity was 96.3% with <1% deviation.

Total RNA from B cells was extracted using Qiagen RNeasy Mini Kit (Qiagen, Inc., Valencia, CA, USA). Total RNA concentration and integrity were determined by an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA, USA). Each RNA sample has an excellent integrity number >9.0 in this study, indicating that RNA degradation due to processing was minimal and negligible. Finally, a 4 µg total RNA/sample was used for the production of complementary RNA (cRNA). The production of cRNA, hybridization, and scanning of the HG-133A GeneChip® were performed according to the manufacturer’s protocol (Affymetrix, Santa Clara, CA, USA), as we did before (Lei et al. 2009; Liu et al. 2005).

Real-time RT-PCR—Two-step real-time reverse transcription polymerase chain reaction (RT-PCR) was used to verify the differentially expressed genes identified from the analyses of chip experiments. All reactions were run in triplicates for each gene. The first step is RT for synthesis of complementary DNA from total RNA, and the second step is real-time quantitative PCR. All the RT reagents were supplied by Taqman® Reverse Transcription Reagents (Applied Biosystems, Foster City, CA, USA). Real-time quantitative PCR was performed in a 25-µl reaction volume using standard protocols on an Applied Biosystems 7900HT. The experimental procedures were followed as the manufacture’s protocol and elaborated in Appendix 2 of the Electronic supplementary material.

Genotyping—Genomic DNA was extracted from whole human blood. We used the genotyping platform Affymetrix Mapping 250K Nsp and 250K Sty arrays to conduct the single nucleotide polymorphism (SNP) genotyping according to the standard protocol recommended by the manufacture. The SNP density of this platform can cover the whole genome. However, we only choose 702 SNPs, which covered most of our identified DEGs for further candidate gene association (some of the DEGs were not covered by the Affymetrix 500K chips). All the genotyping results has been strictly tested, and only those that passed the quality control will be involved in our later association analysis (Liu et al. 2009a,b,c; Xiong et al. 2009).

Data analyses

Differential expression analyses—We submitted our gene expression profiling to Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>), and the access number was GSE18723. Microarray Suite 5.0 (MAS 5.0, Affymetrix, Santa Clara, CA, USA) software was used to generate array raw data in CEL files. A preselecting procedure adopted by our

previous study (Liu et al. 2005) was also used here. We selected a subset of 7,215 probe sets from a total of 22,283 for ~14,500 genes in the HG-133A array. The selection criteria firstly remove those unreliable probes that did not express in B cells (MAS 5.0 generated fluorescence intensities value <150) or overexpress due to non-specific hybridization (>10,000). Then, the second process minimizes the multiple-testing problem by excluding the genes with relatively constant expression value across the samples from further analysis (Liu et al. 2005; Golub et al. 1999). Thus, we actually tested those probes with more biological information (large variance). The approaches employed here can improve the statistical power. Subsequent data analyses were performed on these selected probe sets. The fold change of each probe set was calculated by comparing the mean expression value given by MAS 5.0 in all smokers with the mean value in non-smokers.

We applied multiple analytic approaches including mas5.0 algorithm (Affymetrix, Santa Clara, CA, USA), robust multiarray algorithm (RMA) (Irizarry et al. 2003), the improved GC-content adjusted RMA (GCRMA) (Wu et al. 2004), and dChip (Li and Hung Wong 2001; Li and Wong 2001) for converting and normalizing our raw probe data to gene expression values, respectively. The expression data generated by mas5.0, RMA, GCRMA, and dChip were analyzed respectively by Bioconductor's Multitest package to identify DEGs between the smoking and the nonsmoking groups using *t* test. The Benjamini and Hochberg (BH) procedure (Benjamini and Hochberg 1995) was used for multiple-testing adjustment, and the adjusted $p < 0.05$ was used as the significant criterion.

Finally, we set our criteria for identifying the significant candidate genes (probe sets). If a gene was commonly suggested by all four methods (mas5.0, RMA, GCRMA, and dChip) (BH adjusted $p < 0.05$) and whose expression value was changed by more than ± 1.5 -fold, we selected this one as the identified DEG. This combination of multiple analytic approaches was used in order to minimize the number of false positives due to potential bias of different algorithms and non-specific hybridization.

Clustering and gene ontology analyses—According to the similarity in our samples and gene expression, the identified DEGs were further clustered and visualized by Cluster&TreeView software (Eisen et al. 1998), using hierarchical method with mas5.0 normalized data in two dimensions at both the gene and sample levels (Getz et al. 2000; Wu and Dewey 2006). To gain an overall picture of potential functions of the DEGs, we classified the genes according to four organizing principles (biological process, molecular function, cellular component, and the involved pathways) of the gene ontology (GO) database (<http://www.geneontology.org/>) by Onto-Express (<http://vortex.cs.wayne.edu/ontoexpress/>) and Pathway-Express (<http://vortex.cs.wayne.edu/ontoexpress/>).

Multicriteria evaluation model analyses—Briefly, the multicriteria evaluation method can integrate multiple criteria (information) and finally get an optimal decision. We used this model in order to get the optimal and reliable functional gene network, which associate with female smoking. The results of both candidate gene SNP association analysis (p values) and gene functional database network analysis (database scores) were served as multiple criteria and combined into the model. The whole procedures of applying the model were elaborated in Appendix 2 of the Electronic supplementary material.

The association statistical analyses were carried out with SAS software (SAS Institute Inc., Cary, NC, USA). Genotypic association analysis was conducted with logistic regression for the association with smoking status (dependent variables) after testing the potential population stratification. Two methods were employed to crosscheck the population stratification: the Structure 2.2 software (<http://pritch.bsd.uchicago.edu/software.html>;

Pritchard et al. 2000), which used Markov chain Monte Carlo algorithm to cluster individuals into different cryptic subpopulations on the basis of 200 random unlinked SNPs data. Then, the genomic control method was used to generate the inflation factor (Devlin and Roeder 1999). Ideally, for a homogeneous population with no stratification, the inflation factor should be equal to or near 1.0. For our total sample, the estimated value was 1.009. Both of the two methods indicated essentially no substructure in our sample.

The effects of covariates like age and/or sex were adjusted by comparing with the corresponding restricted model. In the first original model, both covariates and genotype were included, and in the restricted model, only the effects of the covariates were estimated. We compared these two models to assess the effects of the genotype. The significance of the covariate-adjusted association, which is the difference in likelihood of the two models, was tested using a chi-squared test (Liu et al. 2009a). The detailed whole procedures were elaborated in our smoking association study (Liu et al. 2009a).

To investigate the direct (physical) and/or indirect (functional) interactions among the genes identified, we use the STRING database (<http://string.embl.de/>) for functional network analysis. The STRING database provided a score for each gene–gene interaction, which is computed as the joint probability of the probabilities from the different evidence channels (protein interaction, fusion, co-expression, text mining, etc.), correcting for the probability of randomly observing an interaction (von Mering et al. 2005). The high database score means that there are more experimental or predicted evidences for gene–gene functional relationship (interaction).

Real-time RT-PCR analysis—Briefly, the Student's *t* test was employed to compare the relative gene expression [i.e. $2^{-\Delta CT}$, where ΔCT (amplified cycle threshold) = ($CT_{\text{Target Gene}} - CT_{\text{GAPDH control}}$)] between the smoking and non-smoking groups (Liu et al. 2005). For the ICOSLG gene, the subjects Nsm-29, Nsm-58, and Sm-75 were discarded for *t* test due to the amplification failed. For the TCF3 gene, the subjects Nsm-10, Sm-20, Nsm-29, Nsm-58, Sm-63, Sm-66, Sm-67, Sm-75, and Sm-89 were discarded for the same reason.

Results

Differential expression analyses

On average, $35.16 \pm 1.96\%$ of the total of 22,283 probe sets in the array were called “present” for our samples based on the analysis with the MAS 5.0 suite software. According to our criteria (Appendix 1, Table A1 of the Electronic supplementary material), we identified 125 genes differentially expressed between the smoking and non-smoking groups, among which 94 genes (75.2%) were down-regulated (fold change ≤ -1.5) in the smoking sample, and other 31 genes were up-regulated (fold change $\geq +1.5$) in smokers.

Clustering and gene ontology analyses

Figure 1 shows the results of the clustering analyses of the DEGs. Seventy-nine studied subjects were separated into two distinct major clusters by the unsupervised clustering analysis. Only six individuals (7.6%) were incorrectly categorized. Some important genes discussed latter in our study were also marked by rectangle in Fig. 1. We can see similar gene expression pattern among these genes. Particularly, the ICOSLG and VCAM1, the IL13 and MUC5AC, the ARF6 and CCR1 paired with each other and presented a very close distance. Other genes were distributed in many clusters. The TCF3 and the GATA2 genes remained relatively independent.

Figure 2 represents the results of the GO enrichment and pathway enrichment analyses. In the “molecular function” principle, functions of the 125 genes were focused on the protein binding and nucleotide binding. The most significant ontological term was “oxygen transporter activity” with a $p < 1.00 \times 10^{-4}$, which indicated that smoking may cause potential oxidative stress on the B cells. The p value was provided to determine the significance of probability that the ontological category of genes in the data set is explained by chance alone. In the “biological process” principle, the functions were mainly on immune response. In the “cellular component” principle, the products of those genes were primarily located on plasma membrane and cytoplasm. The pathway enrichment result showed the identified gene enriched in the cytokine–cytokine receptor interaction pathway.

Multicriteria evaluation model analyses

A total of 702 SNPs of our identified DEGs were queried for association analysis. There are 30 and 24 DEGs associated with smoking in total sample and female subgroup, respectively. Nineteen genes among them were significant in both total and female samples (overlapped). Using STRING to further analyze the 125 DEGs, a functional associated network was constructed (Fig. 3). This network includes 40 genes, and 29 of them were DEGs. Information regarding the involved genes with the ontology results was summarized in Table 1 and Appendix 1, Table A2 of the Electronic supplementary material.

The multicriteria evaluation model generated a weighted sum score C for each DEG, based on the association p value and the STRING database score. This score simultaneously measured the significance of DEGs in both association and STRING pathway analysis. Generally, the higher this score, the more significant of DEGs in our combined analyses (Appendix 2 of the Electronic supplementary material). We ranked their C scores from the largest to zero and provided the top 30 results in Appendix 2 of the Electronic supplementary material. Obviously, the genes that associated with smoking and/or actively interacted with others had a high C score, such as IL13, TCF3, GATA2, MYH11, VCAM1, ARF6, GIT2, CCR1, and ICOSLG (Appendix 2, Table A2 of the Electronic supplementary material).

Interestingly, the results showed that VCAM1, CCR1, and GATA2 were highly associated with smoking status in both total and female sample (Table 2). The association of ICOSLG gene was detected only in the female subgroup (Table 2). Other SNPs with known function were marked in Table 2. Many functionally important genes are included in both the network and the meaningful ontology categories (Table 1).

Based on the results of evaluation model and prior known functional relevance of DEGs, we built the potential biological relationships and network for the most promising genes with their immune functions in B cells (Fig. 4). In this network, there are 12 DEGs induced by the smoking stimulation in women, which influence broad aspects including the cell development, the cytokine secretion, the signaling transduction, the chemotactic response, and the B cell lineage commitment.

Gender-specific comparison

We also provided the association signals of male subgroup in Table 2 in order to investigate whether the identified DEG can effect on both sexes or not. Interestingly, the significant SNPs ($p < 0.05$) effect on men cannot be detected in women, which indicated that our study probably mainly identified the female-specific results. For each gene, based on only the association signals, the VCAM, CCR1, and CPM may associate with male smoking.

Real-time RT-PCR analyses

Based on the limited total RNA quantity and primer assay availability, their highly statistical significance of differential expression, their functional importance in our multicriteria evaluation analyses, and their prior known function relevance to smoking response, we selected the ICOSLG and TCF3 genes for replication. The relative expression level was 0.130 ± 0.096 (mean $\pm$ SD) in smoking group and 0.219 ± 0.176 in non-smoking group for the ICOSLG gene. For the TCF3 gene, it was 0.425 ± 0.238 and 0.581 ± 0.245 , respectively. These two genes were significantly down-regulated in the smoking group with $p = 0.0130$ (ICOSLG) and 0.0125 (TCF3), respectively.

Discussion

B cells, as major producers of antibodies in the human body (Parkin and Cohen 2001), may be involved in the development of diseases induced by cigarette smoking (Chan et al. 1990; Tsoumakidou et al. 2007; Kidney et al. 1996; Regius et al. 1990; Vignes et al. 2000). A common hypothesis for the onset of these diseases is that smoking influences the functions of B cells and thus inhibits the production of protective immunoglobulins and suppresses normal immune response. In this study, we identified 125 DEGs, 75.3% of them were down-regulated in the smokers, which indicated a broad suppression of smoking stimulating on B cells. Based on the functional and replication analyses, we suggested a potential novel B cell-mediated pathophysiological mechanism for the etiology of smoking-induced diseases. The down-regulation of genes in the functional network may systematically inhibit different immune functions in B cells (Fig. 4). The potential oxidative stress on the B cells may also play roles in the development of microvascular and small airway complications (Brandsma et al. 2008; Hasnis et al. 2007).

Signaling transduction and cytokine secretion

VCAM1 (CD106) is a member of the Ig superfamily and encodes a surface adhesion protein, which mediates cell adhesion and signal transduction (Mebius 2003). At the early stage of B cell development, VCAM1 adheres the pro-B cells to stromal cells, which is one of the important steps for pro-B cells evolving to pre-B cells (Funk et al. 1994). Thus, the down-regulation of VCAM1 in smokers' B cells revealed a suppression of immune signal transduction and may interrupt the cellular development.

Furthermore, it has been demonstrated that VCAM1 directly interacts with the IL13 gene (Fig. 3; Sironi et al. 1994). In our study, the down-regulation of VCAM1 (fold change = -1.50) in smokers was correlated with the down-regulation of IL13 (fold change = -1.56) as expected. IL13 is a well-known candidate gene for the smoking-exposed COPD, asthma, and other serious respiratory disease (Cozen et al. 2004; Feleszko et al. 2006; Lee et al. 2007; Sadeghnejad et al. 2007, 2008; Zheng et al. 2000). It encodes an immunoregulatory-stimulating cytokine, which involves in several stages of B-cell differentiation and maturation (Cocks et al. 1993; Defrance et al. 1994; McKenzie et al. 1993). The down-regulation of IL13 (fold change = -1.56) may cause dysregulation of major histocompatibility complex class II expression, inhibit the IgE class switching, and interrupt the inflammatory cytokine production (Wynn 2003). We speculated that the suppression would significantly reduce the normal cell number and activity of B cells under the stimulation of smoking.

T cell-driven B cell activation and plasma cells formation process

The down-regulation of ICOSLG in B cells (fold change = -1.68) is a signature of in vivo clusters of B cells and plasma cells in close contact with T cells (Hutloff et al. 2004). Down-regulation of ICOSLG in our study indicated the interaction of B cells with T cells, thus may

mediate the T cell-driven B cell activation process and force the formation of memory B cells and plasma cells (secreting the antibodies; Hutloff et al. 2004). In addition, both down-regulation of ICOSLG and IL13 may be critically relevant in inducing an acute immune response and perpetuating inflammation in chronically immune-mediated disorders (Hu et al. 2007).

Notably, another inhibited (fold change = -1.50) gene ARF6 associated with the formation of plasma cell was also identified in our study. ARF6 was responsible for vesicular transport and thus is important for the formation of plasma cell from B cells (Yang et al. 1998). The suppression of ARF6 may influence cell-mediated immune response and differentiation of other related immune cells (e.g., monocyte) (Yang et al. 1998).

Chemotactic responses and motility

Chemokines and their receptors perform the roles of signal transduction and are critical for the recruitment of immune cells to the site of inflammation. CCR1 as one of the chemokine receptor also enhanced the migration and circulation of B cells (Corcione et al. 2002). A high level expression of the CCR1 may improve the chemotactic responses for B cells, such as facilitating their preferential localization in lymphoid tissues (Ehrhardt et al. 2005) and infiltrating to the site of inflammation or infection (Noda et al. 2007). A significant low expression of CCR1 (fold change = -2.10) in our study thus suggested a limited ability of chemotactic responses in smokers. Therefore, one of the possible reasons for the high frequency of chronic inflammations or infections that occurred in smokers may be attributed to the ability of B cells migrating or localizing to the target site was weakened or even damaged.

On the other hand, one of the up-regulated gene (fold change = $+2.00$) GIT2 (G protein-coupled receptor kinase interacting ArfGAP 2) may impair the motility of B cells because only the down-regulation of GIT2 can enhance the cell motility (Frank et al. 2006).

B cell lineage commitment

TCF3 gene, which regulates the development of B cells (Bradney et al. 2003), was also identified to be suppressed in the smokers. TCF3 served as pivotal regulator for different sets of orchestrated and stage-specific genes in lineage commitment and differentiation of B cells (Gisler and Sigvardsson 2002; Greenbaum et al. 2004). The down-regulation of TCF3 gene usually happened in the old B cells or elderly people, which indicated the reduced protection to infectious agents (Frasca et al. 2008). More importantly, down-regulated TCF3 gene associated with the decline in the percentage of memory B cells but an increase in that of naive B cells (Frasca et al. 2008). Therefore, the suppressed TCF3 gene (fold change = -1.56) in smokers indicated a B cell defect or overaged in the immune system.

TCF3 also interacted with GATA binding protein 2 (GATA2) gene (Fig. 3). GATA2 gene is expressed in multipotent progenitors prior to many kinds of cell lineage commitment (Murrell and Green 1995). Induced expression of GATA2 promotes proliferation and differentiation (Briegel et al. 1993). Down-regulated GATA2 with TCF3 provided novel clues for factors regulating the B cell lineage commitment.

Novel clues for the pathogenesis of obstruction disease imbedded in B cells

Some risk genes of obstruction disease in the respiratory and circulating system were identified here, such as VCAM1 (Cavusoglu et al. 2004), IL13 (Lee et al. 2007), and MUC5AC (Kraft et al. 2008). Smoking can up-regulate the VCAM1 gene in endothelial cells and may cause atherosclerosis due to the vascular cellular overadhesion (Cavusoglu et al. 2004; Cirillo et al. 2007). In addition, smoking can up-regulate IL13 and induce the

overexpression of mucus (MUC5AC protein) in the airway epithelial cells of asthma and COPD patients (Nakao et al. 2008).

However, our in vivo study revealed the down-regulation of the VCAM1, IL13, and MUC5AC in smoking-stimulated B cells. This phenomenon observed in women may suggest a different mechanism in B cells, which was different from the one in epithelial cells. Presumably, smoking-stimulated overexpression of VCAM1, IL13, and MUC5AC in endothelial cells surface, which leads to over accumulation of mucus and adhesion protein. On the other hand, smoking inhibited expression of VCAM1, IL13, and MUC5AC in B cells, which interrupted the normal immune signal transduction and response. The effects of suppressed immune function in B cells combined with the accumulated mucus and adhesion proteins in endothelial cells surface may induce the obstruction in the respiratory and circulating system. Interestingly, supportive evidence showed that some smoking-induced diseases may involve interactions between immune cells and vascular endothelium (Cavusoglu et al. 2004). Our study further contributed useful clues for this pathogenesis.

In conclusion, we have characterized the genome-wide expression profile for the first time and, by extrapolation, defined the functional relationship of a specific set of genes from the B cells across a relatively large sample of females with/without smoking habit. Whether this mechanism is female-specific or can affect on both sex still needs further investigation.

Acknowledgments

We thank Dr. Peng Xiao for coordinating the RT-PCR experiment. This work was partially supported by the LB595 and LB692 grant from the State of Nebraska. The study was also benefited from support from Xi'an Jiaotong University, Shanghai Leading Academic Discipline Project (project number S30501) and the Ministry of Education of China. HWD was partially supported by grants from NIH (P50AR055081, R01AG026564, R01AR050496, RC2DE020756, R01AR057049, and R03TW008221) and Franklin D. Dickson/Missouri Endowment.

References

- Alberg AJ, Ford JG, Samet JM. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* 2007;132(3 Suppl):29S–55S. [PubMed: 17873159]
- Al-Ghamdi HS, Anil S. Serum antibody levels in smoker and non-smoker Saudi subjects with chronic periodontitis. *J Periodontol* 2007;78(6):1043–1050. [PubMed: 17539718]
- Amos A. Women and smoking. *Br Med Bull* 1996;52(1):74–89. [PubMed: 8746298]
- Barbour SE, Nakashima K, Zhang JB, Tangada S, Hahn CL, Schenkein HA, Tew JG. Tobacco and smoking: environmental factors that modify the host response (immune system) and have an impact on periodontal health. *Crit Rev Oral Biol Med* 1997;8(4):437–460. [PubMed: 9391754]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 1995;57(1):289–300.
- Bergen AW, Caporaso N. Cigarette smoking. *J Natl Cancer Inst* 1999;91(16):1365–1375. [PubMed: 10451441]
- Bosken CH, Hards J, Gatter K, Hogg JC. Characterization of the inflammatory reaction in the peripheral airways of cigarette smokers using immunocytochemistry. *Am Rev Respir Dis* 1992;145(4 Pt 1):911–917. [PubMed: 1554220]
- Bradney C, Hjelmeland M, Komatsu Y, Yoshida M, Yao TP, Zhuang Y. Regulation of E2A activities by histone acetyltransferases in B lymphocyte development. *J Biol Chem* 2003;278(4):2370–2376. [PubMed: 12435739]
- Brandsma CA, Hylkema MN, van der Strate BW, Slebos DJ, Luinge MA, Geerlings M, Timens W, Postma DS, Kerstjens HA. Heme oxygenase-1 prevents smoke induced B-cell infiltrates: a role for regulatory T cells? *Respir Res* 2008;9:17. [PubMed: 18252008]
- Briegleb K, Lim KC, Plank C, Beug H, Engel JD, Zenke M. Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. *Genes Dev* 1993;7(6):1097–1109. [PubMed: 8504932]

- Buttner P, Mosig S, Funke H. Gene expression profiles of T lymphocytes are sensitive to the influence of heavy smoking: a pilot study. *Immunogenetics* 2007;59(1):37–43. [PubMed: 17149619]
- Cavusoglu Y, Timuralp B, Us T, Akgun Y, Kudaiberdieva G, Gorenek B, Unalir A, Goktekin O, Ata N. Cigarette smoking increases plasma concentrations of vascular cell adhesion molecule-1 in patients with coronary artery disease. *Angiology* 2004;55(4):397–402. [PubMed: 15258685]
- Chan MA, Benedict SH, Carstairs KC, Francombe WH, Gelfand EW. Expansion of B lymphocytes with an unusual immunoglobulin rearrangement associated with atypical lymphocytosis and cigarette smoking. *Am J Respir Cell Mol Biol* 1990;2(6):549–552. [PubMed: 2346660]
- Cirillo P, Pacileo M, De Rosa S, Calabro P, Gargiulo A, Angri V, Prevete N, Fiorentino I, Ucci G, Sasso L, Petrillo G, D'Amore SM, Chiariello M. HMG-CoA reductase inhibitors reduce nicotine-induced expression of cellular adhesion molecules in cultured human coronary endothelial cells. *J Vasc Res* 2007;44(6):460–470. [PubMed: 17657162]
- Cloez-Tayarani I, Changeux JP. Nicotine and serotonin in immune regulation and inflammatory processes: a perspective. *J Leukoc Biol* 2007;81(3):599–606. [PubMed: 17108054]
- Cocks BG, de Waal MR, Galizzi JP, de Vries JE, Aversa G. IL-13 induces proliferation and differentiation of human B cells activated by the CD40 ligand. *Int Immunol* 1993;5(6):657–663. [PubMed: 7688562]
- Corcione A, Tortolina G, Bonecchi R, Battilana N, Tadorelli G, Malavasi F, Sozzani S, Ottonello L, Dallegri F, Pistoia V. Chemotaxis of human tonsil B lymphocytes to CC chemokine receptor (CCR) 1, CCR2 and CCR4 ligands is restricted to non-germinal center cells. *Int Immunol* 2002;14(8):883–892. [PubMed: 12147625]
- Cozen W, Diaz-Sanchez D, James Gauderman W, Zadnick J, Cockburn MG, Gill PS, Masood R, Hamilton AS, Jyrala M, Mack TM. Th1 and Th2 cytokines and IgE levels in identical twins with varying levels of cigarette consumption. *J Clin Immunol* 2004;24(6):617–622. [PubMed: 15622446]
- Defrance T, Carayon P, Billian G, Guillemot JC, Minty A, Caput D, Ferrara P. Interleukin 13 is a B cell stimulating factor. *J Exp Med* 1994;179(1):135–143. [PubMed: 7903680]
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;55(4):997–1004. [PubMed: 11315092]
- Ehrhardt GR, Hsu JT, Gartland L, Leu CM, Zhang S, Davis RS, Cooper MD. Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells. *J Exp Med* 2005;202(6):783–791. [PubMed: 16157685]
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95(25):14863–14868. [PubMed: 9843981]
- Felieszko W, Zawadzka-Krajewska A, Matysiak K, Lewandowska D, Peradzynska J, Dinh QT, Hamelmann E, Groneberg DA, Kulus M. Parental tobacco smoking is associated with augmented IL-13 secretion in children with allergic asthma. *J Allergy Clin Immunol* 2006;117(1):97–102. [PubMed: 16387591]
- Frank SR, Adelstein MR, Hansen SH. GIT2 represses Crk- and Rac1-regulated cell spreading and Cdc42-mediated focal adhesion turnover. *Embo J* 2006;25(9):1848–1859. [PubMed: 16628223]
- Frasca D, Landin AM, Lechner SC, Ryan JG, Schwartz R, Riley RL, Blomberg BB. Aging down-regulates the transcription factor E2A, activation-induced cytidine deaminase, and Ig class switch in human B cells. *J Immunol* 2008;180(8):5283–5290. [PubMed: 18390709]
- Funk PE, Kincade PW, Witte PL. Native associations of early hematopoietic stem cells and stromal cells isolated in bone marrow cell aggregates. *Blood* 1994;83(2):361–369. [PubMed: 7506947]
- Gandini S, Botteri E, Iodice S, Boniol M, Lowenfels AB, Maisonneuve P, Boyle P. Tobacco smoking and cancer: a meta-analysis. *Int J Cancer* 2008;122(1):155–164. [PubMed: 17893872]
- Geng Y, Savage SM, Johnson LJ, Seagrave J, Sopor ML. Effects of nicotine on the immune response. I. Chronic exposure to nicotine impairs antigen receptor-mediated signal transduction in lymphocytes. *Toxicol Appl Pharmacol* 1995;135(2):268–278. [PubMed: 8545837]
- Getz G, Levine E, Domany E. Coupled two-way clustering analysis of gene microarray data. *Proc Natl Acad Sci U S A* 2000;97(22):12079–12084. [PubMed: 11035779]
- Gisler R, Sigvardsson M. The human V-preB promoter is a target for coordinated activation by early B cell factor and E47. *J Immunol* 2002;168(10):5130–5138. [PubMed: 11994467]

- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286(5439):531–537. [PubMed: 10521349]
- Greenbaum S, Lazorchak AS, Zhuang Y. Differential functions for the transcription factor E2A in positive and negative gene regulation in pre-B lymphocytes. *J Biol Chem* 2004;279(43):45028–45035. [PubMed: 15310760]
- Hasnis E, Bar-Shai M, Burbea Z, Reznick AZ. Cigarette smoke-induced NF-kappaB activation in human lymphocytes: the effect of low and high exposure to gas phase of cigarette smoke. *J Physiol Pharmacol* 2007;58(Suppl 5, Pt 1):263–274. [PubMed: 18204136]
- Henschke CI, Yip R, Miettinen OS. Women's susceptibility to tobacco carcinogens and survival after diagnosis of lung cancer. *Jama* 2006;296(2):180–184. [PubMed: 16835423]
- Hu W, Janke A, Ortler S, Hartung HP, Leder C, Kieseier BC, Wiendl H. Expression of CD28-related costimulatory molecule and its ligand in inflammatory neuropathies. *Neurology* 2007;68(4):277–282. [PubMed: 17242332]
- Hutloff A, Buchner K, Reiter K, Baelde HJ, Odendahl M, Jacobi A, Dorner T, Kroczeck RA. Involvement of inducible costimulator in the exaggerated memory B cell and plasma cell generation in systemic lupus erythematosus. *Arthritis Rheum* 2004;50(10):3211–3220. [PubMed: 15476242]
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4(2):249–264. [PubMed: 12925520]
- Kidney JC, Wong AG, Efthimiadis A, Morris MM, Sears MR, Dolovich J, Hargreave FE. Elevated B cells in sputum of asthmatics. Close correlation with eosinophils. *Am J Respir Crit Care Med* 1996;153(2):540–544. [PubMed: 8564094]
- Kraft M, Adler KB, Ingram JL, Crews AL, Atkinson TP, Cairns CB, Krause DC, Chu HW. *Mycoplasma pneumoniae* induces airway epithelial cell expression of MUC5AC in asthma. *Eur Respir J* 2008;J31(1):43–46.
- Langhammer A, Johnsen R, Holmen J, Gulsvik A, Bjørmer L. Cigarette smoking gives more respiratory symptoms among women than among men. The Nord-Trøndelag Health Study (HUNT). *J Epidemiol Community Health* 2000;54(12):917–922. [PubMed: 11076988]
- Lee JS, Rosengart MR, Kondragunta V, Zhang Y, McMurray J, Branch RA, Choi AM, Sciruba FC. Inverse association of plasma IL-13 and inflammatory chemokines with lung function impairment in stable COPD: a cross-sectional cohort study. *Respir Res* 2007;8:64. [PubMed: 17868461]
- Lei SF, Wu S, Li LM, Deng FY, Xiao SM, Jiang C, Chen Y, Jiang H, Yang F, Tan LJ, Sun X, Zhu XZ, Liu MY, Liu YZ, Chen XD, Deng HW. An in vivo genome wide gene expression study of circulating monocytes suggested GBP1, STAT1 and CXCL10 as novel risk genes for the differentiation of peak bone mass. *Bone* 2009;44(5):1010–1014. [PubMed: 19223260]
- Li C, Hung Wong W. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol* 2001;2(8) RESEARCH0032.
- Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001;98(1):31–36. [PubMed: 11134512]
- Liu YZ, Dvornyk V, Lu Y, Shen H, Lappe JM, Recker RR, Deng HW. A novel pathophysiological mechanism for osteoporosis suggested by an in vivo gene expression study of circulating monocytes. *J Biol Chem* 2005;280(32):29011–29016. [PubMed: 15965235]
- Liu YZ, Pei YF, Guo YF, Wang L, Liu XG, Yan H, Xiong DH, Zhang YP, Levy S, Li J, Haddock CK, Papasian CJ, Xu Q, Ma JZ, Payne TJ, Recker RR, Li MD, Deng HW. Genome-wide association analyses suggested a novel mechanism for smoking behavior regulated by IL15. *Mol Psychiatry* 2009a;14(7):668–680. [PubMed: 19188921]
- Liu XG, Tan LJ, Lei SF, Liu YJ, Shen H, Wang L, Yan H, Guo YF, Xiong DH, Chen XD, Pan F, Yang TL, Zhang YP, Guo Y, Tang NL, Zhu XZ, Deng HY, Levy S, Recker RR, Papasian CJ, Deng HW. Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. *Am J Hum Genet* 2009b;84(3):418–423. [PubMed: 19268274]

- Liu YZ, Guo YF, Wang L, Tan LJ, Liu XG, Pei YF, Yan H, Xiong DH, Deng FY, Yu N, Zhang YP, Zhang L, Lei SF, Chen XD, Liu HB, Zhu XZ, Levy S, Papasian CJ, Drees BM, Hamilton JJ, Recker RR, Deng HW. Genome-wide association analyses identify SPOCK as a key novel gene underlying age at menarche. *PLoS Genet* 2009c;5(3):e1000420. [PubMed: 19282985]
- Mackay J, Amos A. Women and tobacco. *Respirology* 2003;8(2):123–130. [PubMed: 12753525]
- Marshall L, Schooley M, Ryan H, Cox P, Easton A, Heaton C, Jackson K, Davis KC, Homs G. Youth tobacco surveillance—United States, 2001–2002. *MMWR Surveill Summ* 2006;55(3):1–56. [PubMed: 16708059]
- McKenzie AN, Culpepper JA, de Waal MR, Briere F, Punnonen J, Aversa G, Sato A, Dang W, Cocks BG, Menon S, et al. Interleukin 13, a T-cell-derived cytokine that regulates human monocyte and B-cell function. *Proc Natl Acad Sci U S A* 1993;90(8):3735–3739. [PubMed: 8097324]
- Mebius RE. Organogenesis of lymphoid tissues. *Nat Rev Immunol* 2003;3(4):292–303. [PubMed: 12669020]
- Moszczynski P, Zabinski Z, Moszczynski P Jr, Rutowski J, Slowinski S, Tabarowski Z. Immunological findings in cigarette smokers. *Toxicol Lett* 2001;118(3):121–127. [PubMed: 11137318]
- Murrell AM, Green AR. Regulation of lineage restricted haemopoietic transcription factors in cell hybrids. *Oncogene* 1995;10(4):631–639. [PubMed: 7862440]
- Nakao I, Kanaji S, Ohta S, Matsushita H, Arima K, Yuyama N, Yamaya M, Nakayama K, Kubo H, Watanabe M, Sagara H, Sugiyama K, Tanaka H, Toda S, Hayashi H, Inoue H, Hoshino T, Shiraki A, Inoue M, Suzuki K, Aizawa H, Okinami S, Nagai H, Hasegawa M, Fukuda T, Green ED, Izuhara K. Identification of pendrin as a common mediator for mucus production in bronchial asthma and chronic obstructive pulmonary disease. *J Immunol* 2008;180(9):6262–6269. [PubMed: 18424749]
- Noda D, Hamachi T, Inoue K, Maeda K. Relationship between the presence of periodontopathic bacteria and the expression of chemokine receptor mRNA in inflamed gingival tissues. *J Periodontol* 2007;78(6):566–571. [PubMed: 17956471]
- Parkin J, Cohen B. An overview of the immune system. *Lancet* 2001;357(9270):1777–1789. [PubMed: 11403834]
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000;155(2):945–959. [PubMed: 10835412]
- Regius O, Rajczy K, Gergely I, Borzsonyi L, Lengyel E, Vargha P. The effect of smoking on peripheral blood lymphocytes and on some immunological parameters in old age. *Z Gerontol* 1990;23(3):163–167. [PubMed: 2392875]
- Ryder MI, Hyun W, Loomer P, Haqq C. Alteration of gene expression profiles of peripheral mononuclear blood cells by tobacco smoke: implications for periodontal diseases. *Oral Microbiol Immunol* 2004;19(1):39–49. [PubMed: 14678473]
- Sadeghnejad A, Meyers DA, Bottai M, Sterling DA, Bleecker ER, Ohar JA. IL13 promoter polymorphism 1112C/T modulates the adverse effect of tobacco smoking on lung function. *Am J Respir Crit Care Med* 2007;176(8):748–752. [PubMed: 17615386]
- Sadeghnejad A, Karmaus W, Arshad SH, Kurukulaaratchy R, Huebner M, Ewart S. IL13 gene polymorphisms modify the effect of exposure to tobacco smoke on persistent wheeze and asthma in childhood, a longitudinal study. *Respir Res* 2008;9:2. [PubMed: 18186920]
- Savage SM, Donaldson LA, Cherian S, Chilukuri R, White VA, Soper ML. Effects of cigarette smoke on the immune response. II. Chronic exposure to cigarette smoke inhibits surface immunoglobulin-mediated responses in B cells. *Toxicol Appl Pharmacol* 1991;111(3):523–529. [PubMed: 1746027]
- Sironi M, Sciacca FL, Matteucci C, Conni M, Vecchi A, Bernasconi S, Minty A, Caput D, Ferrara P, Colotta F, et al. Regulation of endothelial and mesothelial cell function by interleukin-13: selective induction of vascular cell adhesion molecule-1 and amplification of interleukin-6 production. *Blood* 1994;84(6):1913–1921. [PubMed: 7521694]
- Skok MV, Kalashnik EN, Koval LN, Tsetlin VI, Utkin YN, Changeux JP, Grailhe R. Functional nicotinic acetylcholine receptors are expressed in B lymphocyte-derived cell lines. *Mol Pharmacol* 2003;64(4):885–889. [PubMed: 14500745]

- Skok M, Grailhe R, Changeux JP. Nicotinic receptors regulate B lymphocyte activation and immune response. *Eur J Pharmacol* 2005;517(3):246–251. [PubMed: 15963492]
- Skok MV, Grailhe R, Agenes F, Changeux JP. The role of nicotinic receptors in B-lymphocyte development and activation. *Life Sci* 2007;80(24–25):2334–2336. [PubMed: 17363009]
- Tsoumakidou M, Elston W, Zhu J, Wang Z, Gamble E, Siafakas NM, Barnes NC, Jeffery PK. Cigarette smoking alters bronchial mucosal immunity in asthma. *Am J Respir Crit Care Med* 2007;175(9):919–925. [PubMed: 17303795]
- van der Strate BW, Postma DS, Brandsma CA, Melgert BN, Luinge MA, Geerlings M, Hylkema MN, van den Berg A, Timens W, Kerstjens HA. Cigarette smoke-induced emphysema: a role for the B cell? *Am J Respir Crit Care Med* 2006;173(7):751–758. [PubMed: 16399994]
- van Leeuwen DM, Gottschalk RW, van Herwijnen MH, Moonen EJ, Kleinjans JC, van Delft JH. Differential gene expression in human peripheral blood mononuclear cells induced by cigarette smoke and its constituents. *Toxicol Sci* 2005;86(1):200–210. [PubMed: 15829617]
- van Leeuwen DM, van Agen E, Gottschalk RW, Vlietinck R, Gielen M, van Herwijnen MH, Maas LM, Kleinjans JC, van Delft JH. Cigarette smoke-induced differential gene expression in blood cells from monozygotic twin pairs. *Carcinogenesis* 2007;28(3):691–697. [PubMed: 17056606]
- Vignes S, Oksenhendler E, Quint L, Daniel MT, Mariette X, Clauvel JP. Polyclonal B lymphocytosis and hyper-IgM: immunodeficiency and/or benign lymphoid proliferation associated with tobacco? *Rev Med Interne* 2000;21(3):236–241. [PubMed: 10763184]
- von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 2005;33:D433–D437. (Database issue). [PubMed: 15608232]
- Wu X, Dewey TG. From microarray to biological networks: analysis of gene expression profiles. *Methods Mol Biol* 2006;316:35–48. [PubMed: 16671399]
- Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F. A model-based background adjustment for oligonucleotide expression arrays. *J Am Stat Assoc* 2004;99(468):909–917.
- Wynn TA. IL-13 effector functions. *Annu Rev Immunol* 2003;21:425–456. [PubMed: 12615888]
- Xiong DH, Liu XG, Guo YF, Tan LJ, Wang L, Sha BY, Tang ZH, Pan F, Yang TL, Chen XD, Lei SF, Yerges LM, Zhu XZ, Wheeler VW, Patrick AL, Bunker CH, Guo Y, Yan H, Pei YF, Zhang YP, Levy S, Papasian CJ, Xiao P, Lundberg YW, Recker RR, Liu YZ, Liu YJ, Zmuda JM, Deng HW. Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. *Am J Hum Genet* 2009;84(3):388–398. [PubMed: 19249006]
- Yang CZ, Heimberg H, D'Souza-Schorey C, Mueckler MM, Stahl PD. Subcellular distribution and differential expression of endogenous ADP-ribosylation factor 6 in mammalian cells. *J Biol Chem* 1998;273(7):4006–4011. [PubMed: 9461590]
- Yang Q, Hergenbahn M, Weninger A, Bartsch H. Cigarette smoke induces direct DNA damage in the human B-lymphoid cell line Raji. *Carcinogenesis* 1999;20(9):1769–1775. [PubMed: 10469623]
- Zheng T, Zhu Z, Wang Z, Homer RJ, Ma B, Riese RJ Jr, Chapman HA Jr, Shapiro SD, Elias JA. Inducible targeting of IL-13 to the adult lung causes matrix metalloproteinase- and cathepsin-dependent emphysema. *J Clin Invest* 2000;106(9):1081–1093. [PubMed: 11067861]

Fig. 1. Two-dimensional hierarchical clustering picture. The horizontal axis shows the clustering of 79 subjects within the two major clusters, and the vertical axis represents the identified 125 genes clustered based on similar expression patterns. *Sm* smoking, *Nsm* non-smoking, and numbers for subject codes. The rectangle-marked genes were identified by our multicriteria model and discussed in detail

Fig. 2. Classifications of gene ontology (GO) and pathway enrichment analyses. **a–c** Classifications of the DEGs according to the gene ontology (GO) principles. The *p* values (≤ 0.01 , which indicated a significant biological meaning for the results) and the numbers with percentage of the genes in each category are indicated on the *left*. **d** The pathway enrichment of the DEGs. The “input genes” indicates the numbers of the matched genes in our results. The “pathway genes” represents the total numbers of the genes involved in the pathway indexed by KEGG database (<http://www.genome.jp/kegg/>)

Fig. 3.

A network diagram constructed by STRING. Each ball represents the corresponding gene. Gene names with the underscore are focus genes included in the input 125 DEGs. Stronger functional associations are represented by *thicker lines*. The *p* value is used to rank how relevant the network is to the total input genes. It takes into account the number of focus genes in the network and the size of the network to approximate how relevant this network is to the original list of genes. The larger ball means the gene has detail protein structure information in the STRING database

Fig. 4. Relationships of discussed genes with their immune functions in B cells. Each rectangle represents the corresponding gene. The *small dotted rectangle* represents genes (ICOS, PIK3CG, PDPK1, and AKT1) that may mediate the relationships but not belong to our identified gene list. Genes with similar functional relationship are *circled using dotted arch rectangles*. The *arrows* indicate the tendency of up-regulation or down-regulation in our study

Table 1

Classification of genes in the network by ontological and pathway groups

Gene Symbol	AffyID	Genomic Location	Fold (S/N)	Molecular Function	Biological Process	Cellular Component	KEGG Pathway
HBA1/HBA2	217414_x_at	16p13.3	-2.97	Oxygen, heme binding and transporter activity	Oxygen transport	Hemoglobin complex	
MYH11	207961_x_at	16p13.11	-2.35	Nucleotide binding Calmodulin binding	Muscle thick filament assembly	Melanosome	
HBA1/HBA2	211699_x_at	16p13.3	-2.23	Oxygen, heme binding and transporter activity	Oxygen transport	Hemoglobin complex	
CCR1	205099_s_at	3p21	-2.10	Rhodopsin-like receptor activity	Immune response	Plasma membrane	Cytokine-cytokine receptor interaction
				Protein binding			
HBA1	211745_x_at	16p13.3	-2.05	Oxygen, heme binding and transporter activity	Oxygen transport	Hemoglobin complex	
HBB	211696_x_at	11p15.5	-2.03	Oxygen, heme binding and transporter activity	Oxygen transport Regulation of blood vessel size	Hemoglobin complex	
					Nitric oxide transport		
HBA1/HBA2	204018_x_at	16p13.3	-2.01	Oxygen, heme binding and transporter activity	Oxygen transport	Hemoglobin complex	
HBA1/HBA2	209458_x_at	16p13.3	-1.95	Oxygen, heme binding and transporter activity	Oxygen transport	Hemoglobin complex	
GRIN1	210782_x_at	9q34.3	-1.83	Protein binding		Plasma membrane Cell junction	Calcium signaling Neuroactive ligand-receptor interaction
MUC5AC	217182_at	11p15.5	-1.75				
HOXB1	208224_at	17q21.3	-1.74				
B7H2 (ICOSLG)	211197_s_at	21q22.3	-1.68		Immune response		Cell adhesion molecules (CAMs)
NPPA	209957_s_at	1p36.21	-1.61		Regulation of blood vessel size		
EPHB3	204600_at	3q21	-1.61	Nucleotide binding Ephrin receptor activity			Axon guidance
CIB2	205008_s_at	15q24	-1.59	Protein binding			
FANCA	215530_at	16q24.3	-1.57	Protein binding		cytoplasm	
IL13	207844_at	5q31	-1.56	Interleukin-13 receptor binding	Immune Response		Cytokine-cytokine receptor interaction Jak-STAT signaling Fc epsilon RI signaling Jak-STAT signaling

Gene Symbol	AffyID	Genomic Location	Fold (S/N)	Molecular Function	Biological Process	Cellular Component	KEGG Pathway
E2A (TCF3)	216647_at	19p13.3	-1.56				Fc epsilon RI signaling
EFNB3	210883_x_at	17p13	-1.55	Transmembrane-ephrin receptor activity	B cell lineage commitment axon choice point recognition		Axon guidance
EIF4B	219599_at	12q13.13	-1.54				
GATA2	209710_at	3q21.3	-1.52	Protein binding	Phagocytosis		
VCAM1	203868_s_at	1p32-p31	-1.50	Protein binding		Plasma membrane	Cell adhesion molecules (CAMs) Leukocyte transendothelial migration
CPM	217557_s_at	12q14.3	-1.50			Plasma membrane	
ARF6	214182_at	14q21.3	-1.50	Nucleotide binding Protein Binding	Vesicle-mediated transport Regulation of Rac protein signal transduction	Plasma membrane Cytoplasm	
EPHB2	209588_at	1p36	-1.50	Transmembrane-ephrin receptor activity Nucleotide binding Ephrin receptor activity Protein binding		Endosome	Axon guidance
IL9R	208164_s_at	16p13.3	1.52				
DDX3X	212514_x_at	Xp11	1.53	ATP-dependent RNA helicase activity nucleotide binding		Cytoplasm Nuclear speck	
CBFB	206788_s_at	16q22.1	1.53	Protein binding			
IQGAP1	213446_s_at	15q26.1	1.57	Calmodulin binding		Plasma membrane Cytoplasm	Regulation of actin cytoskeleton
PDPK1	204524_at	16p13.3	1.70	Nucleotide binding Protein binding 3-Phosphoinositide dependent protein kinase activity		Plasma membrane Cytoplasm	mTOR signaling PPAR signaling
PIK3CG	206369_s_at	7q22.3	1.72	Protein binding			
SON	201085_s_at	21q22	1.77	Protein binding		Plasma membrane Cytoplasm	
GIT2	209876_at	12q24.1	2.00	Protein binding		Nuclear speck	

The fold change is ratio of the mean expression value given by Mas5.0 in the smoking to that in the non-smoking group. More detail information including the gene description and the identifier of Gene Ontology (GO) term is provided in Appendix 1, Table A2 of the Electronic supplementary material

Table 2
Association signals of candidate genes detected in the 840 unrelated US whites sample

ID	Affymetrix SNP ID	dbSNP ID	Position	Allele	MAF	P values in total sample	P values in female sample	P values in male sample
VCAM_#1	SNP_A-2157077	rs6702885	Chr1: 101006615	A/G	0.13	0.005	0.124	0.020
2	SNP_A-4284099	rs7545402	101093664	A/G	0.14	0.022	0.133	0.031
3	SNP_A-1791751	rs6673745	101047606	T/G	0.48	0.032	0.075	0.185
4	SNP_A-1868824	rs6662889	101047700	G/A	0.48	0.034	0.072	0.172
5	SNP_A-1914252	rs11578560	100943587	T/C	0.07	0.037	0.019	0.064
6	SNP_A-1900155	rs12132178	101044990	C/A	0.44	0.104	0.023	0.460
7	SNP_A-2232763	rs7548917	101011964	C/T	0.43	0.192	0.027	0.340
8	SNP_A-4300942	rs1409423	101008429	G/A	0.43	0.266	0.032	0.287
9	SNP_A-2115447	rs2027166	100849799	C/T	0.12	0.122	0.679	0.042
10	SNP_A-4224151	rs7415343	100849911	T/C	0.12	0.128	0.619	0.047
CCR1_#1	SNP_A-2208387	rs6767559	Chr3: 46102646	G/C	0.29	0.039	0.136	0.208
2	SNP_A-1863024	rs17215981	46099422	G/A	0.08	0.631	0.019	0.333
3	SNP_A-1825357	rs9859136	46070274	A/G	0.03	0.065	0.023	0.421
4	SNP_A-2064302	rs2131091	46132187	A/G	0.37	0.076	0.038	0.437
5	SNP_A-2283788	rs9879257	46133721	A/G	0.03	0.163	0.039	0.650
6	SNP_A-2227162	rs1546079	46125123	A/G	0.44	0.071	0.306	0.025
MYH11_#1	SNP_A-1805810	rs216164	Chr16: 15827119	T/C ^a	0.12	0.003	0.019	0.122
2	SNP_A-2054170	rs6498573	15786874	T/C	0.10	0.021	0.025	0.548
3	SNP_A-2096024	rs2272556	15737037	G/C	0.01	0.033	0.044	0.215
4	SNP_A-2102123	rs9922682	15854279	T/C ^a	0.21	0.196	0.015	0.093
5	SNP_A-4293845	rs3851702	15857063	G/A ^a	0.22	0.144	0.015	0.089
6	SNP_A-1821122	rs3851703	15857148	T/C ^a	0.23	0.146	0.016	0.087
IQGAP1_#1	SNP_A-1919370	rs16944523	Chr15: 88839298	C/T	0.11	0.021	0.023	0.243
2	SNP_A-1891732	rs8035980	88775394	C/G	0.35	0.025	0.050	0.273
3	SNP_A-1807318	rs8042861	88778337	T/G ^a	0.35	0.029	0.070	0.215
4	SNP_A-2264157	rs1978151	88838304	A/G ^a	0.36	0.036	0.063	0.276
5	SNP_A-4240454	rs10520686	88798433	C/G ^a	0.11	0.054	0.042	0.381

ID	Affymetrix SNP ID	dbSNP ID	Position	Allele	MAF	P values in total sample	P values in female sample	P values in male sample
CIB2_#1	SNP_A-2139504	rs16969507	Chr15: 76192672	C/T	0.03	0.082	0.050	0.573
CPM_#1	SNP_A-2164195	rs12425617	Chr12: 67814045	T/C	0.44	0.032	0.214	0.023
2	SNP_A-2044547	rs3741599	67536748	G/A ^a	0.23	0.096	0.844	0.014
3	SNP_A-1839367	rs2642556	67602131	A/G ^a	0.22	0.628	0.118	0.019
4	SNP_A-2192437	rs13486	67535354	G/A	0.23	0.054	0.710	0.020
5	SNP_A-2197875	rs2647952	67605530	G/A	0.21	0.602	0.151	0.022
6	SNP_A-1780436	rs1152934	67589531	T/G	0.28	0.153	0.144	0.042
GATA2_#1	SNP_A-2229022	rs2953128	Chr3: 129672262	A/G	0.42	0.009	0.037	0.260
ICOSLG_#1	SNP_A-4270705	rs2838519	Chr21: 44439451	A/G	0.49	0.080	0.039	0.693

SNP rs2953128 of GATA2 gene located in a promoter regulatory region, which may influence the transcription (<http://fastsnp.ibms.sinica.edu.tw>).

^a SNPs were intronic enhancers of the gene.