

RESEARCH ARTICLE

Bioinformatics Analysis of the Effects of Tobacco Smoke on Gene Expression

Chunhua Cao¹, Jianhua Chen^{2*}, Chengqi Lyu¹, Jia Yu¹, Wei Zhao¹, Yi Wang¹, Derong Zou^{1*}

1 Department of Stomatology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, 200233, China, **2** Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, 200030, China

* derongzou@qq.com (DZ); chenjhv@hotmail.com (JC)



OPEN ACCESS

Citation: Cao C, Chen J, Lyu C, Yu J, Zhao W, Wang Y, et al. (2015) Bioinformatics Analysis of the Effects of Tobacco Smoke on Gene Expression. PLoS ONE 10(12): e0143377. doi:10.1371/journal.pone.0143377

Editor: Marinus F.W. te Pas, Wageningen UR Livestock Research, NETHERLANDS

Received: July 30, 2015

Accepted: November 4, 2015

Published: December 2, 2015

Copyright: © 2015 Cao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Youth Research Project of Shanghai Health and Family Planning Commission (20134Y192), the foundation of Shanghai Jiao Tong University Affiliated Sixth People's Hospital (1517), Shanghai Key Laboratory of Psychotic Disorders (13dz2260500), the Science and Technology Commission of Shanghai Municipality (12JC1407301), Shanghai Natural Science Foundation (12ZR1447200) and the Natural Science Foundation of China (81470714, 31400825).

Abstract

This study was designed to explore the effects of tobacco smoke on gene expression through bioinformatics analyses. Gene expression profile GSE17913 was downloaded from the Gene Expression Omnibus database. The differentially expressed genes (DEGs) in buccal mucosa tissues between 39 active smokers and 40 never smokers were identified. Gene Ontology (GO) and pathway enrichment analyses of DEGs were performed, followed by protein-protein interaction (PPI) network, transcriptional regulatory network as well as miRNA-target regulatory network construction. In total, 88 up-regulated DEGs and 106 down-regulated DEGs were identified. Among these DEGs, cytochrome P450, family 1, subfamily A, polypeptide 1 (*CYP1A1*) and *CYP1B1* were enriched in the Metabolism of xenobiotics by cytochrome P450 pathway. In the PPI network, tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*), and *CYP1A1* were hub genes. In the transcriptional regulatory network, transcription factors of MYC associated factor X (*MAX*) and upstream transcription factor 1 (*USF1*) regulated many overlapped DEGs. In addition, protein tyrosine phosphatase, receptor type, D (*PTPRD*) was regulated by multiple miRNAs in the miRNA-DEG regulatory network. *CYP1A1*, *CYP1B1*, *YWHAZ* and *PTPRD*, and TF of *MAX* and *USF1* may have the potential to be used as biomarkers and therapeutic targets in tobacco smoke-related pathological changes.

Introduction

There are approximately 1.3 billion people who smoke cigarettes in the world. Tobacco is an important risk factor for multiple human malignancies, which causes almost 5 million preventable deaths every year [1, 2]. More than 100 tumor promoters and carcinogens have been identified in tobacco [3]. Tobacco combustion products contain polycyclic aromatic hydrocarbons (PAHs), which have been suggested to be carcinogenic [4]. A variety of cancers such as lung cancer, oral cavity cancer, esophageal cancer and liver cancer are attributed to cigarette smoke [5]. In addition, smoking can reduce the efficacy of targeted anticancer therapies by stimulating

Competing Interests: The authors have declared that no competing interests exist.

metabolic clearance [6, 7]. A better understanding of the mechanisms underlying these diseases may lead to more effective treatments for cancer patients.

Previous studies have suggested that tobacco smoke causes a field of injury in the oral mucosa epithelial cells, which are among the most important physiological barriers against several exogenous stimuli [8]. Tobacco smoke has been found to affect gene expression in many tissues and cells, including epithelial cells [9]. Some studies have reported that histologically normal large airway epithelial cells of smokers display allelic loss [10] and p53 mutations [11]. An in vitro study by Pickett et al. [12] found that when a single source of human airway epithelial cells was exposed to the same dose of cigarette smoke condensate from 10 different cigarettes, 21 genes were altered by 9 of the 10 cigarettes. In addition, based on the transcriptome profiling, Spira et al. [13] have indicated that smoking induces the expression of genes involved in redox stress and xenobiotic metabolism in the large airway epithelial cells. Theoretically, the development of a transcriptome-based biomarker to identify high-risk smokers may provide a basis to protect against the carcinogenic effects of cigarette smoke. However, related studies are far from being enough.

In the current study, we downloaded the microarray data GSE17913 and identified the differentially expressed genes (DEGs) in buccal mucosa tissues between 39 active smokers and 40 never smokers. We performed pathway enrichment analysis and protein-protein interaction (PPI) network construction. The transcriptional regulatory network and the miRNA-target regulatory network were constructed. We aimed to define the effects of smoking on the oral epithelial cells. The findings of this study may provide new insights into the carcinogenic effects of smoking and then develop useful preventive strategies.

Data and Methods

Affymetrix microarray data

The microarray data GSE17913 [14] was downloaded from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database based on the platform of the Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix Inc., Santa Clara, California, USA). The dataset consists of 79 samples, including 40 samples of buccal mucosa from never smokers (< 100 cigarettes per lifetime; 20 never smoker females; 20 never smoker males) and 39 samples of buccal mucosa from active smokers (≥ 15 pack year exposure; 19 smoker females; 20 smoker males). The demographic characteristics of never smokers and smokers were shown in Table 1.

Data preprocessing and differential expression analysis

The original array data were preprocessed with background correction and quartile data normalization by robust multiarray average (RMA) [15]. Then, the probes were mapped to the

Table 1. Demographic characteristics of never smokers and smokers.

	Female (N = 20)	Male (N = 20)	P value
Never smokers			
Age (years) median	45	45	0.51
Age (years) range	26–66	30–55	
Smokers			
Age (years) median	43	45.5	0.87
Age (years) range	27–63	30–54	
Pack-year median	25	32.5	0.37
Pack-year range	15–66	15–60	

doi:10.1371/journal.pone.0143377.t001

gene symbols based on the microarray annotation information in R Bioconductor package `hgu133a2.db`, and the expression values of multiple probes for a given gene were reduced to a single value by taking their median expression value.

The paired t-test based on the `limma` package [16] in Bioconductor was used to identify the DEGs between active smokers and never smokers. The Benjamini & Hochberg method was used to decrease the false positive rate of the p-value, and the false discovery rate (FDR) was calculated. $FDR < 0.05$ was considered the cutoff value.

Gene ontology and pathway enrichment analyses

The `clusterProfiler` (available at <http://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>) [17] package in R is used to automate the process of biological term classification and the enrichment analysis of gene clusters. To analyze the DEGs at the functional level, Gene Ontology Biological Process (GO BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the `clusterProfiler` tool. $P\text{-value} < 0.05$ was set as the threshold value. In addition, the pathway figure was described using `pathview` (available at <http://bioconductor.org/packages/release/bioc/html/pathview.html> and at <http://Pathview.r-forge.r-project.org/>) [18] in the R Bioconductor package.

PPI network construction

The Search Tool for the Retrieval of Interacting Genes (STRING) database (<http://string-db.org/>) [19] is a precomputed global resource used to evaluate PPI information. In this paper, the STRING online tool was applied to analyze the PPI of DEGs, and only experimentally validated interactions with a combined score > 0.4 were selected as significant. The PPI network was constructed using `cytoscape` [20].

Transcriptional regulatory network construction

The Encyclopedia Of DNA Elements (ENCODE, <http://genome.ucsc.edu/ENCODE/>) Project aimed to provide comprehensive annotations of candidate functional elements in the human genome for scientific and medical communities [21]. In this study, we extracted all the human transcription factor (TF) binding sites were extracted from the ENCODE database. Next, the repeatability of each TF binding site was analyzed, and the TF binding site that appeared in at least 2 independent samples was selected for further analysis. Then, combining with the annotation data of genetic transcription area, we screened out all the TFs located in the gene promoter regions (1 kb upstream of the transcription start site (TSS) to 0.5 kb downstream of TSS). The TF-gene pairs were then constructed. Finally, we extracted the TF-DEG pairs from the TF-gene pairs and constructed the DEG-associated transcriptional regulatory network.

miRNA-target regulatory network construction

The `starBase v2.0` (<http://starbase.sysu.edu.cn/>) database provides the RNA-RNA and protein-RNA interaction networks from 108 CLIP-Seq (PAR-CLIP, HITS-CLIP, iCLIP, CLASH) data sets generated by 37 independent studies [22]. It also contains the miRNA-target regulatory networks which are verified by experiment and predicted by 5 algorithms (TargetScan, miRanda, Pictar2, PITA and RNA22). In our study, we screened out the miRNA-target pairs that were not only verified by at least 1 experiment but also predicted by at least 3 algorithms. Then the miRNA-target regulatory network was constructed.

Results

Identification of DEGs

A total of 394 DEGs were identified between active smokers and never smokers, of which 288 were up-regulated and 106 were down-regulated ([S1 Table](#)). The results are shown in the heat-map ([Fig 1](#)).

GO and pathway enrichment analyses of DEGs

After pathway enrichment analysis, 3 pathways were obtained, namely the Metabolism of xenobiotics by cytochrome P450, Osteoclast differentiation, and Chemokine signaling pathway ([Table 2](#)). The Metabolism of xenobiotics by cytochrome P450 pathway was enriched by several DEGs including cytochrome P450, family 1, subfamily A, polypeptide 1 (*CYP1A1*) and *CYP1B1*. Specifically, the DEG distribution in the pathway of Metabolism of xenobiotics by cytochrome P45 was shown in [Fig 2](#). Additionally, the DEGs were mainly enriched in GO BP terms related to biological processes, cellular processes and single-organism processes.

All GO BP and KEGG pathways for the DEGs were shown in [S2 Table](#).

PPI network construction

The constructed PPI network was shown in [Fig 3](#). The PPI network consisted of 139 nodes and 183 edges ([S3 Table](#)). Among these gene nodes, tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*), NCK adaptor protein 1 (*NCK1*) and cytochrome P450, family 1, subfamily A, polypeptide 1 (*CYP1A1*) had the highest degrees of 9. Moreover, *YWHAZ* and *NCK1* were slightly down-regulated, and *CYP1A1* was significantly up-regulated.

Transcriptional regulatory network of DEGs

The DEG-associated transcriptional regulatory network was shown in [Fig 4](#). The network consisted of 2 TFs, 101 target DEGs and 131 edges ([S4 Table](#)). [Fig 4](#) showed that the TF MYC associated factor X (MAX) was down-regulated and upstream transcription factor 1 (USF1) was up-regulated, with many overlapped DEGs regulated by the 2 TFs.

miRNA-DEG regulatory network

In total, 210 miRNAs, 2 TFs and 118 genes were included in the miRNA-DEG regulatory network ([S5 Table](#)). Almost every miRNA regulated two DEGs, but several DEGs were also regulated by multiple miRNAs, such as mesoderm induction early response 1, family member 3 (*MIER3*), cysteine-rich hydrophobic domain 1 (*CHIC1*) and protein tyrosine phosphatase, receptor type, D (*PTPRD*).

Discussion

The identification of DEGs between smokers and never smokers is important to understanding the mechanisms underlying the carcinogenic effects of smoking. In this study, 288 up-regulated DEGs and 106 down-regulated DEGs were identified. Among these DEGs, *CYP1A1* and *CYP1B1* were mainly enriched in the pathway of Metabolism of xenobiotics by cytochrome P450. In the PPI network, *YWHAZ*, *NCK1* and *CYP1A1* were hub genes. In the transcriptional regulatory network, the TFs MAX and USF1 regulated many overlapping DEGs. In addition, *PTPRD* was regulated by multiple miRNAs in the miRNA-DEG regulatory network.

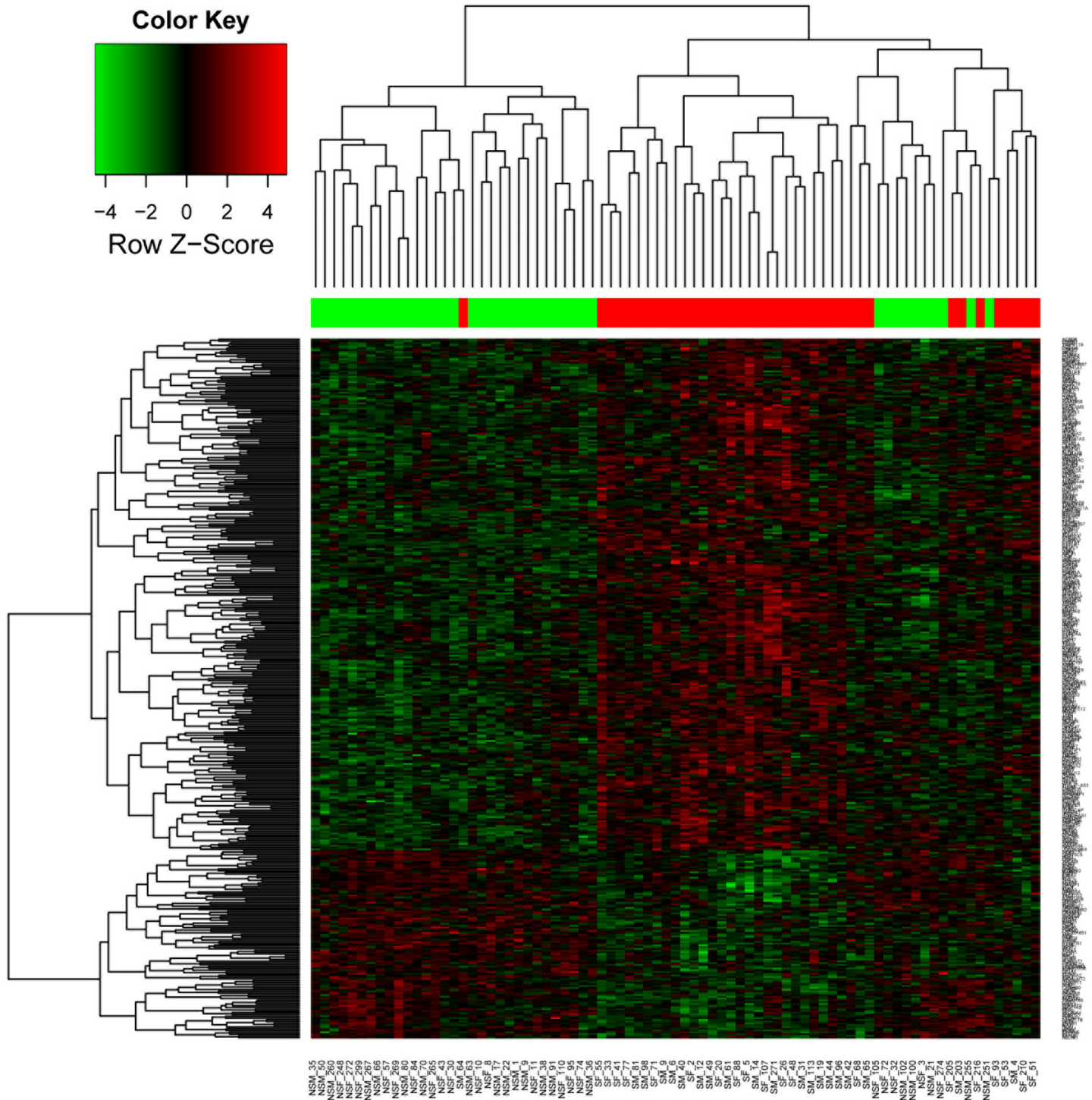


Fig 1. Heatmap plot of differentially expressed genes across all samples. Upper color bar represents sample classes; red represents smoker group; green represents non-smoker group.

doi:10.1371/journal.pone.0143377.g001

Cytochrome P450 enzymes can catalyze the biotransformation of various xenobiotic compounds to form ultimate toxicants [23]. Mutations in certain *CYP* genes contribute to clinically

Table 2. Pathway enrichment result for differentially expressed genes (DEGs).

ID	Description	P-value	Gene Symbol
hsa00980	Metabolism of xenobiotics by cytochrome P450	2.80E-05	CYP1B1/CYP1A1/ALDH3A1/GSTA3/AKR1C2/AKR1C4/GSTM4/AKR1C3/GSTM3
hsa04380	Osteoclast differentiation	2.61E-02	NCF4/PPARG/SOCS3/TNFRSF11A/BLNK/AKT3/TYK2
hsa04062	Chemokine signaling pathway	2.76E-02	CCL5/CCL26/CCL18/CCR2/AKT3/JAK2/CCR7/CCL22/JAK3

doi:10.1371/journal.pone.0143377.t002

relevant diseases including malignancy [24]. In this study, Metabolism of xenobiotics by cytochrome P450 was a significant pathway and was enriched by several DEGs, including *CYP1B1* and *CYP1A1*. As mentioned above, the combustion products of tobacco contain PAHs, which can be converted into reactive metabolites via *CYP1A1*. These reactive metabolites may be involved in the initiation of carcinogenesis via the formation of bulky PAH-DNA adducts [25, 26]. In the human lung, high expression of *CYP1A1* has been associated with increased lung cancer risk [27]. Additionally, *CYP1B1* can also activate various carcinogens: for instance, *CYP1B1* can catalyze the formation of dihydrodiols of specific PAHs and their subsequent oxidation into carcinogenic dihydrodiol epoxides [28]. *CYP1B1* is also commonly overexpressed in human malignancies [29]. Our result was consistent with the findings above, and therefore, *CYP1A1* and *CYP1B1* may be potential targets in smoking-mediated malignancies.

In the PPI network, *YWHAZ* was one of the hub genes with the highest degree. Its encoded proteins are involved in many vital cellular processes such as signal transduction, metabolism, cell cycle regulation and apoptosis. *YWHAZ* protein expression is well known to be related to advanced disease grade and poor clinical outcome in lung cancer patients [30]. Research has found that *YWHAZ* is a potential regulator of the function of β -catenin, which is a central effector of Wnt signaling in tumorigenesis and metastasis [31]. In particular, tobacco smoke exposure may lead to the translocation of β -catenin via cooperation with interleukin-1 β [32]. Taken together, *YWHAZ* may be a marker gene in tobacco smoke-related pathological changes.

Furthermore, in the transcriptional regulatory network, the TFs *MAX* and *USF1* regulated many overlapped DEGs. *MAX* is a member of the basic helix-loop-helix leucine zipper (bHLHZ) family of transcription factors. It can form heterodimers with other family members, including *MYC*, which is an oncoprotein implicated in cell growth, proliferation, differentiation and apoptosis [33]. The dysregulated expression of *MYC* has been reported in a wide range of human malignancies [34]. On the other hand, *USF1* also encodes a member of the bHLHZ family and functions as a cellular transcription factor that regulates the expression of numerous genes involved in cellular proliferation and the cell cycle [35, 36]. Importantly, Wu et al. [37] have demonstrated that nicotine, a component of tobacco smoke, can enhance *USF1* translocation from the cytoplasm to the nucleus. As a result, we speculated that tobacco smoke might induce cancer by targeting genes such as *MAX* and *USF1*.

In the miRNA-DEG regulatory network, several DEGs were regulated by multiple miRNAs, such as *PTPRD*. The protein encoded by *PTPRD* is a member of the protein tyrosine phosphatase (PTP) family, signaling molecules that regulate cell growth, differentiation, and oncogenic transformation [38]. *PTPRD* mutations have been found in lung cancer and other malignancies [39]. Currently, there are few reports about the relationship between tobacco smoke and the dysregulated expression of *PTPRD*. Thus, we speculated that *PTPRD* might be a potential biomarker associated with the carcinogenic effect of tobacco smoke.

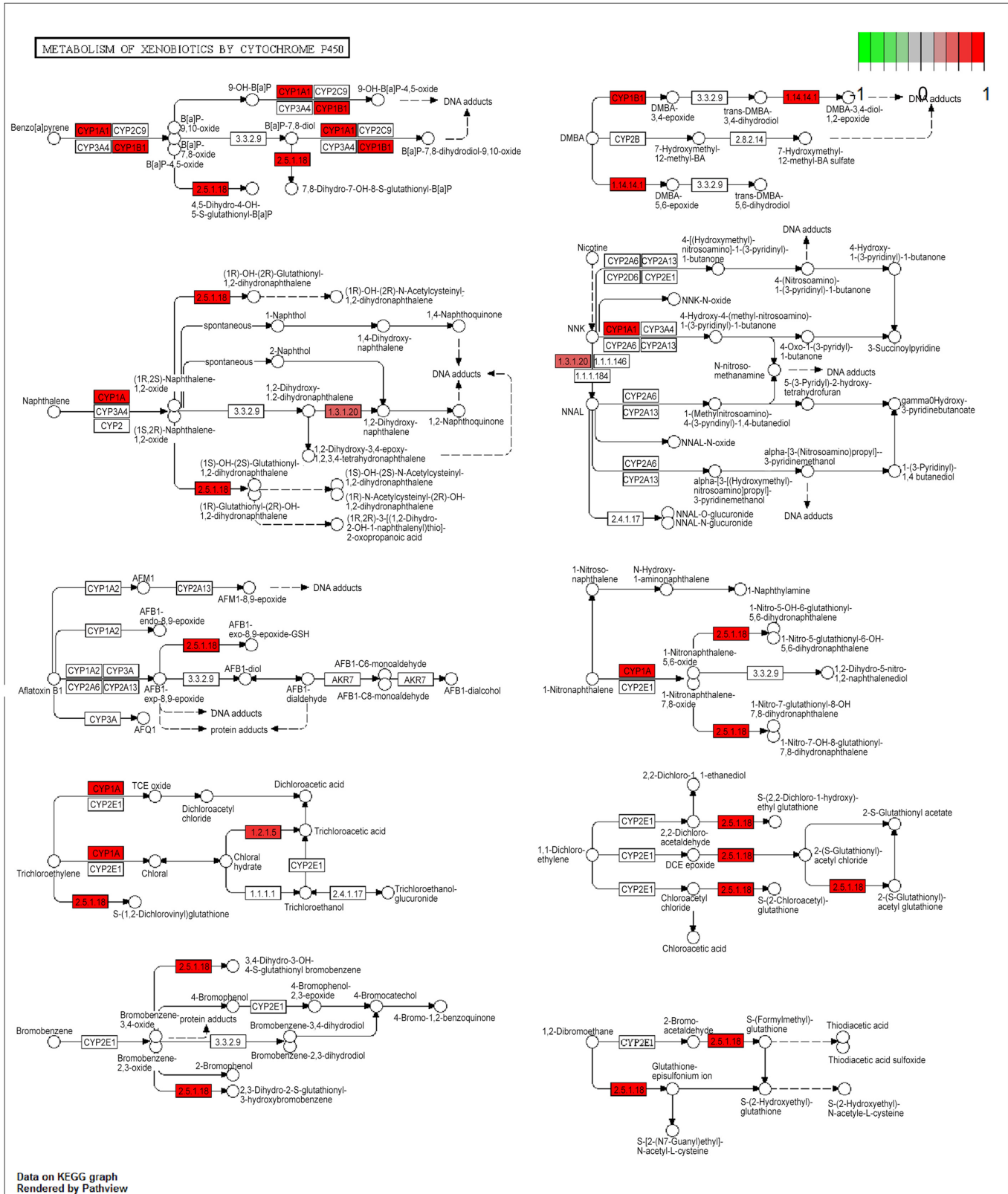


Fig 2. Differentially expressed gene (DEG) distribution in metabolism of xenobiotics by cytochrome P450 pathway. Red represents DEGs, and color shade represents log FC variation.

doi:10.1371/journal.pone.0143377.g002

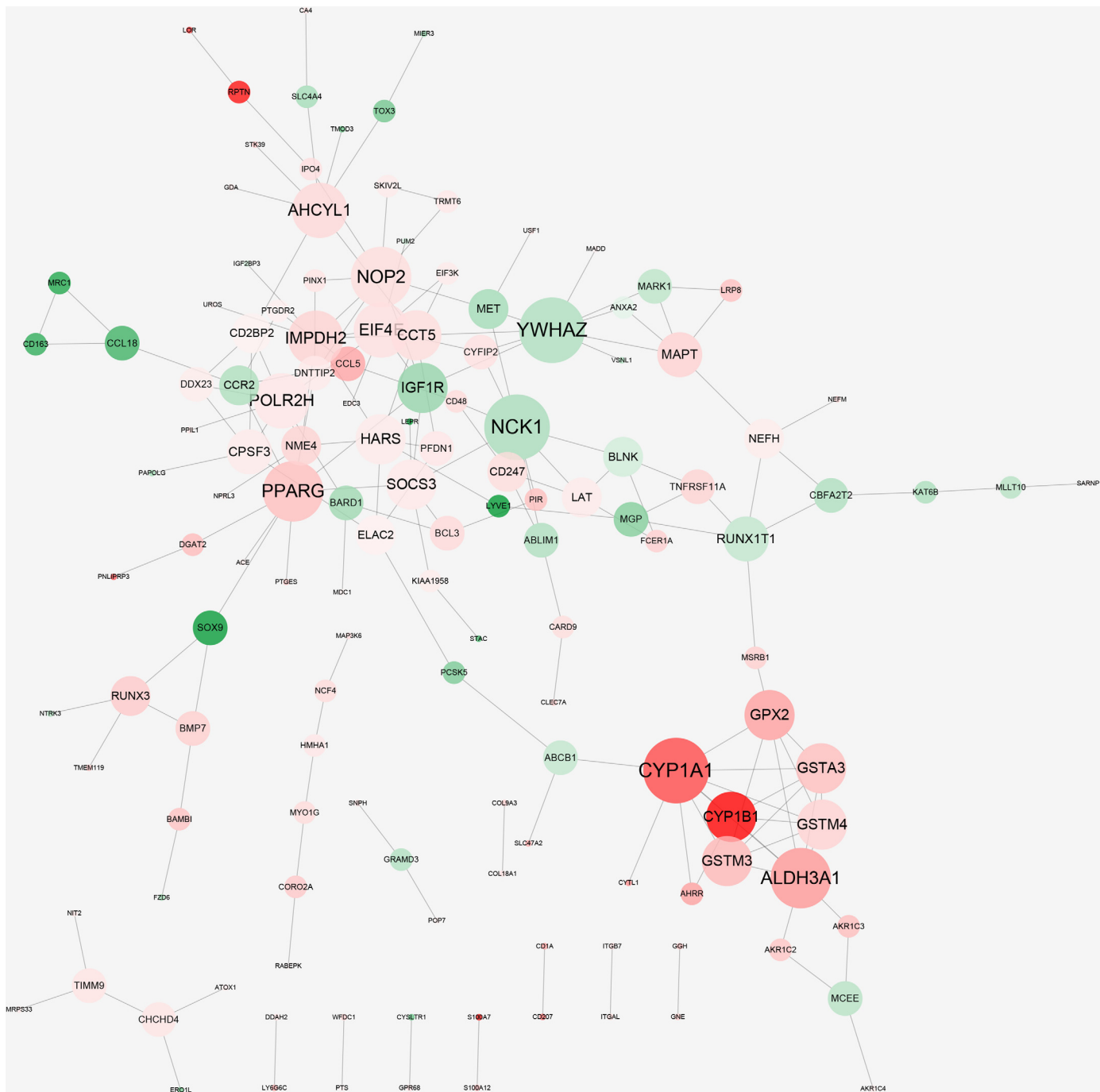


Fig 3. Protein-protein interaction (PPI) network constructed by the differentially expressed genes (DEGs). Node size represents node degree; a larger size indicates a larger degree. Red represents up-regulation, and green represents down-regulation.

doi:10.1371/journal.pone.0143377.g003

In conclusion, we have used a bioinformatics approach to identify the marker genes related to the carcinogenic role of tobacco smoke. These DEGs, such as *CYP1A1*, *CYP1B1*, *YWHAZ* and *PTPRD*, and TF of *MAX* and *USF1*, may have the potential to be used as biomarkers and therapeutic targets in tobacco smoke-related pathological changes. The findings in this study may contribute to our further understanding of the underlying molecular mechanisms that are

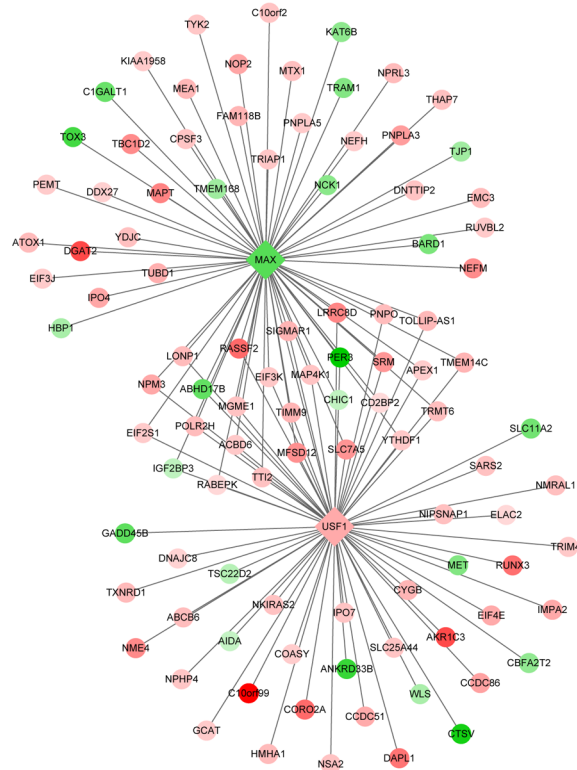


Fig 4. Transcriptional regulatory network for transcription factors of MAX and USF1. Diamond represents transcription factor, and circle represents differentially expressed genes (DEGs). Red represents up-regulated expression, and green represents down-regulated expression.

doi:10.1371/journal.pone.0143377.g004

modulated by tobacco smoke. Further genetic and experimental studies with larger sample sizes are still needed to confirm the results.

Supporting Information

S1 Table. All identified differentially expressed genes (DEGs) between active smokers and never smokers.

(XLS)

S2 Table. All Gene Ontology Biological Processes (GO BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for the DEGs.

(XLS)

S3 Table. All the nodes and edges in the protein-protein interaction (PPI) network.

(XLS)

S4 Table. All the nodes and edges in the transcriptional regulatory network.

(XLS)

S5 Table. All the nodes and edges in the microRNA regulatory network.

(XLS)

Acknowledgments

This work was supported by the Youth Research Project of Shanghai Health and Family Planning Commission (20134Y192), the foundation of Shanghai Jiao Tong University Affiliated Sixth People's Hospital (1517), Shanghai Key Laboratory of Psychotic Disorders (13dz2260500), the Science and Technology Commission of Shanghai Municipality (12JC1407301), Shanghai Natural Science Foundation (12ZR1447200) and the Natural Science Foundation of China (81470714, 31400825).

Author Contributions

Conceived and designed the experiments: CC WZ. Analyzed the data: JC. Contributed reagents/materials/analysis tools: JY. Wrote the paper: CC JC DZ WZ. Collected important background information and acquisition of data: CL JY YW.

References

1. Organization WH (2006) The facts about smoking and health. Fact Sheet May. 30.
2. Gritz ER, Dresler C, Sarna L (2005) Smoking, the missing drug interaction in clinical trials: ignoring the obvious. *Cancer Epidemiol Biomarkers Prev.* 14:2287–2293. PMID: [16214906](#)
3. Mayne ST, Lippman SM (2005) Cigarettes: a smoking gun in cancer chemoprevention. *J Natl Cancer Inst.* 97:1319–1321. PMID: [16174848](#)
4. Godschalk RW, Van Schooten F-J, Bartsch H (2003) A critical evaluation of DNA adducts as biological markers for human exposure to polycyclic aromatic compounds. *BMB Reports.* 36:1–11.
5. Humans IWGotEoCRt, Organization WH, Cancer IAfRo, *Tobacco Smoke and Involuntary Smoking: This Publication Represents the Views and Expert Opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which Met in Lyon, 11–18 June 2002.* 2004: IARC.
6. Hamilton M, Wolf JL, Rusk J, Beard SE, Clark GM, Witt K, et al. (2006) Effects of smoking on the pharmacokinetics of erlotinib. *Clin Cancer Res.* 12:2166–2171. PMID: [16609030](#)
7. Pantarotto J, Malone S, Dahrrouge S, Gallant V, Eapen L (2007) Smoking is associated with worse outcomes in patients with prostate cancer treated by radical radiotherapy. *BJU Int.* 99:564–569. PMID: [17166241](#)
8. Donetti E, Gualerzi A, Bedoni M, Volpari T, Sciarabba M, Tartaglia G, et al. (2010) Desmoglein 3 and keratin 10 expressions are reduced by chronic exposure to cigarette smoke in human keratinised oral mucosa explants. *Arch Oral Biol.* 55:815–823. doi: [10.1016/j.archoralbio.2010.07.001](#) PMID: [20667405](#)
9. Brody JS (2012) Transcriptome alterations induced by cigarette smoke. *Int J Cancer.* 131:2754–2762. doi: [10.1002/ijc.27829](#) PMID: [22961494](#)
10. Wistuba II, Virmani AK, Gazdar AF, Lam S, LeRiche J, Behrens C, et al. (1997) Molecular damage in the bronchial epithelium of current and former smokers. *J Natl Cancer Inst.* 89:1366–1373. PMID: [9308707](#)
11. Powell CA, Klares S, O'Connor G, Brody JS (1999) Loss of heterozygosity in epithelial cells obtained by bronchial brushing: clinical utility in lung cancer. *Clin Cancer Res.* 5:2025–2034. PMID: [10473082](#)
12. Pickett G, Seagrave J, Boggs S, Polzin G, Richter P, and Tesfaigzi Y (2009) Effects of 10 cigarette smoke condensates on primary human airway epithelial cells by comparative gene and cytokine expression studies. *Toxicol Sci.* kfp298.
13. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, et al. (2004) Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A.* 101:10143–10148. PMID: [15210990](#)
14. Boyle JO, Gümüş ZH, Kacker A, Choksi VL, Bocker JM, Zhou XK, et al. (2010) Effects of cigarette smoke on the human oral mucosal transcriptome. *Cancer Prev Res.* 3:266–278.
15. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* 20:307–315. PMID: [14960456](#)
16. Smyth GK, Limma: linear models for microarray data, in *Bioinformatics and computational biology solutions using R and Bioconductor.* 2005, Springer. p. 397–420.
17. Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 16:284–287. doi: [10.1089/omi.2011.0118](#) PMID: [22455463](#)

18. Luo W, Brouwer C (2013) Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*. 29:1830–1831. doi: [10.1093/bioinformatics/btt285](https://doi.org/10.1093/bioinformatics/btt285) PMID: [23740750](https://pubmed.ncbi.nlm.nih.gov/23740750/)
19. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. (2014) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* gku1003.
20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Re.* 13:2498–2504.
21. Consortium EP (2004) The ENCODE (ENCyclopedia of DNA elements) project. *science*. 306:636–640. PMID: [15499007](https://pubmed.ncbi.nlm.nih.gov/15499007/)
22. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H (2013) starBase v2. 0: decoding miRNA–ceRNA, miRNA–ncRNA and protein–RNA interaction networks from large-scale CLIP–Seq data. *Nucleic Acids Res.* gkt1248.
23. Ding X, Kaminsky LS (2003) HUMAN EXTRAHEPATIC CYTOCHROMES P450: Function in Xenobiotic Metabolism and Tissue-Selective Chemical Toxicity in the Respiratory and Gastrointestinal Tracts*. *Annu Rev Pharmacol Toxicol.* 43:149–173. PMID: [12171978](https://pubmed.ncbi.nlm.nih.gov/12171978/)
24. Nebert DW, Russell DW (2002) Clinical importance of the cytochromes P450. *The Lancet.* 360:1155–1162.
25. Shimada T, Martin MV, Pruess-Schwartz D, Marnett LJ, Guengerich FP (1989) Roles of individual human cytochrome P-450 enzymes in the bioactivation of benzo (a) pyrene, 7, 8-dihydroxy-7, 8-dihydrobenzo (a) pyrene, and other dihydrodiol derivatives of polycyclic aromatic hydrocarbons. *Cancer Res.* 49:6304–6312. PMID: [2509067](https://pubmed.ncbi.nlm.nih.gov/2509067/)
26. Uppstad H, Øvrebo S, Haugen A, Møllerup S (2010) Importance of CYP1A1 and CYP1B1 in bioactivation of benzo [a] pyrene in human lung cell lines. *Toxicol Lett.* 192:221–228. doi: [10.1016/j.toxlet.2009.10.025](https://doi.org/10.1016/j.toxlet.2009.10.025) PMID: [19879933](https://pubmed.ncbi.nlm.nih.gov/19879933/)
27. Stücker I, Jacquet M, De Waziers I, Cenee S, Beaune P, Kremers P, et al. (2000) Relation between inducibility of CYP1A1, GSTM1 and lung cancer in a French population. *Pharmacogenet Genomics.* 10:617–627.
28. Kim JH, Stansbury KH, Walker NJ, Trush MA, Strickland PT, and Sutter TR (1998) Metabolism of benzo [a] pyrene and benzo [a] pyrene-7, 8-diol by human cytochrome P450 1B1. *Carcinogenesis.* 19:1847–1853. PMID: [9806168](https://pubmed.ncbi.nlm.nih.gov/9806168/)
29. Lin P, Chang H, Ho WL, Wu M-H, Su J-M (2003) Association of aryl hydrocarbon receptor and cytochrome P4501B1 expressions in human non-small cell lung cancers. *Lung Cancer.* 42:255–261. PMID: [14644512](https://pubmed.ncbi.nlm.nih.gov/14644512/)
30. Chen C-H, Chuang S-M, Yang M-F, Liao J-W, Yu S-L, and Chen JJ (2012) A Novel Function of YWHAZ/β-Catenin Axis in Promoting Epithelial–Mesenchymal Transition and Lung Cancer Metastasis. *Mol Cancer Res.* 10:1319–1331. doi: [10.1158/1541-7786.MCR-12-0189](https://doi.org/10.1158/1541-7786.MCR-12-0189) PMID: [22912335](https://pubmed.ncbi.nlm.nih.gov/22912335/)
31. Nelson WJ, Nusse R (2004) Convergence of Wnt, β-catenin, and cadherin pathways. *science*. 303:1483–1487. PMID: [15001769](https://pubmed.ncbi.nlm.nih.gov/15001769/)
32. Barbieri SS, Weksler BB (2007) Tobacco smoke cooperates with interleukin-1β to alter β-catenin trafficking in vascular endothelium resulting in increased permeability and induction of cyclooxygenase-2 expression in vitro and in vivo. *FASEB J.* 21:1831–1843. PMID: [17317723](https://pubmed.ncbi.nlm.nih.gov/17317723/)
33. Grandori C, Cowley SM, James LP, Eisenman RN (2000) The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu Rev Cell Dev Biol.* 16:653–699. PMID: [11031250](https://pubmed.ncbi.nlm.nih.gov/11031250/)
34. Lee JEA, Parsons LM, Quinn LM (2014) MYC function and regulation in flies: how *Drosophila* has enlightened MYC cancer biology. *AIMS Genetics.* 1:81–98.
35. Sanchez AP, Zhao J, You Y, Declèves A- E, Diamond-Stanic M, and Sharma K (2011) Role of the USF1 transcription factor in diabetic kidney disease. *Am J Physiol Renal Physiol.* 301:F271–F279. doi: [10.1152/ajprenal.00221.2011](https://doi.org/10.1152/ajprenal.00221.2011) PMID: [21543418](https://pubmed.ncbi.nlm.nih.gov/21543418/)
36. Niemiec P, Nowak T, Iwanicki T, Gorczynska-Kosiorz S, Balcerzyk A, Krauze J, et al. (2015) The rs2516839 Polymorphism of the USF1 Gene May Modulate Serum Triglyceride Levels in Response to Cigarette Smoking. *International journal of molecular sciences.* 16:13203–13216. doi: [10.3390/ijms160613203](https://doi.org/10.3390/ijms160613203) PMID: [26068452](https://pubmed.ncbi.nlm.nih.gov/26068452/)
37. Wu L, Wen L, Yang F, Wang X, Fang J (2007) [Expression, subcellular localization and nuclear translocation of transcription factor up stream stimulatory factor-1 in odontoblasts]. *Zhonghua kou qiang yi xue za zhi.* 42:559–560. PMID: [18070436](https://pubmed.ncbi.nlm.nih.gov/18070436/)
38. Östman A, Hellberg C, Böhmer FD (2006) Protein-tyrosine phosphatases and cancer. *Nat Rev Cancer.* 6:307–320. PMID: [16557282](https://pubmed.ncbi.nlm.nih.gov/16557282/)
39. Sato M, Takahashi K, Nagayama K, Arai Y, Ito N, Okada M, et al. (2005) Identification of chromosome arm 9p as the most frequent target of homozygous deletions in lung cancer. *Genes, Chromosomes and Cancer.* 44:405–414. PMID: [16114034](https://pubmed.ncbi.nlm.nih.gov/16114034/)