

1 **Supplementary Figures and Tables**

2 **Supplementary Figure 1**

3 **A. PCA of genes affected by smoking in bronchial and nasal samples in an**

4 **independent U133A dataset.** Probesets for 85 of the 119 genes found to be

5 differentially expressed by smoking independently of collection site in the matched

6 bronchus and nose Exon array dataset were found on the U133A array. The expression

7 profiles of these probesets distinguish smokers from nonsmokers in both bronchus and

8 nose in an independent U133A dataset. Each point represents a subject (red: smokers,

9 blue: nonsmokers). There are 23 nonsmokers in blue and 34 smokers in red from U133A

10 bronchial dataset, and 8 non-smokers and 7 smokers from U133A nasal dataset.

11

12 **B. PCA of genes previously identified to be affected by smoking in bronchial airway**

13 **in the current Exon array dataset.** Probesets for 71 of the 97 genes found to be

14 differentially expressed by smoking in our previous U133A bronchial epithelium dataset

15 (Spira et al 2004), were found on the Exon array. The expression profiles of these

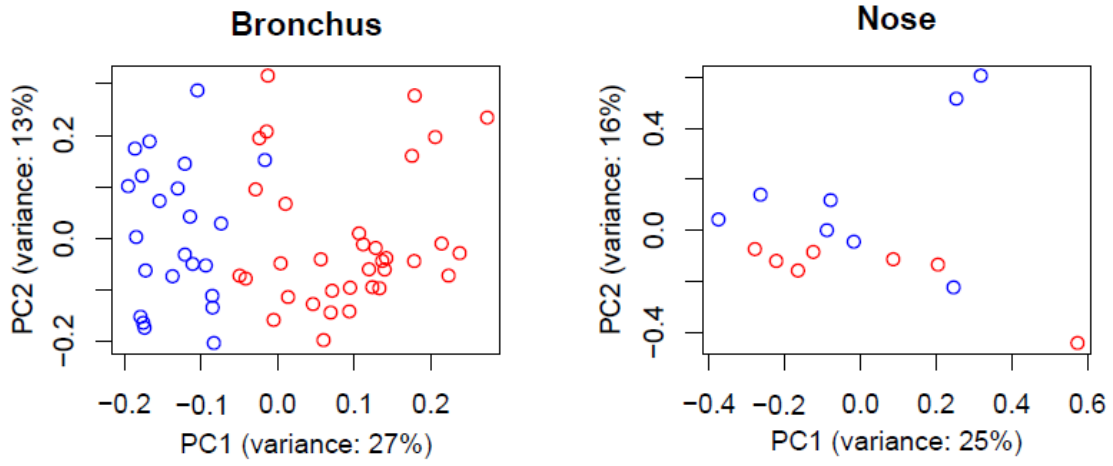
16 probesets distinguish smokers from nonsmokers in both bronchus and nose in the Exon

17 dataset. Each point represents a subject (red: smokers, blue: nonsmokers). There are 13

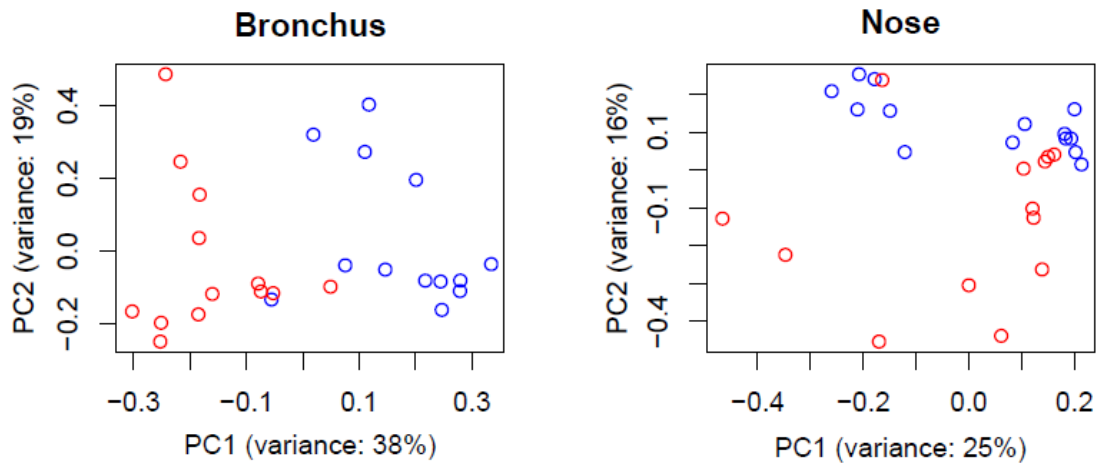
18 nonsmokers and 13 smokers in bronchus, and 14 non-smokers and 13 smokers in nose.

19

A



B



20

21

22

23

24

25

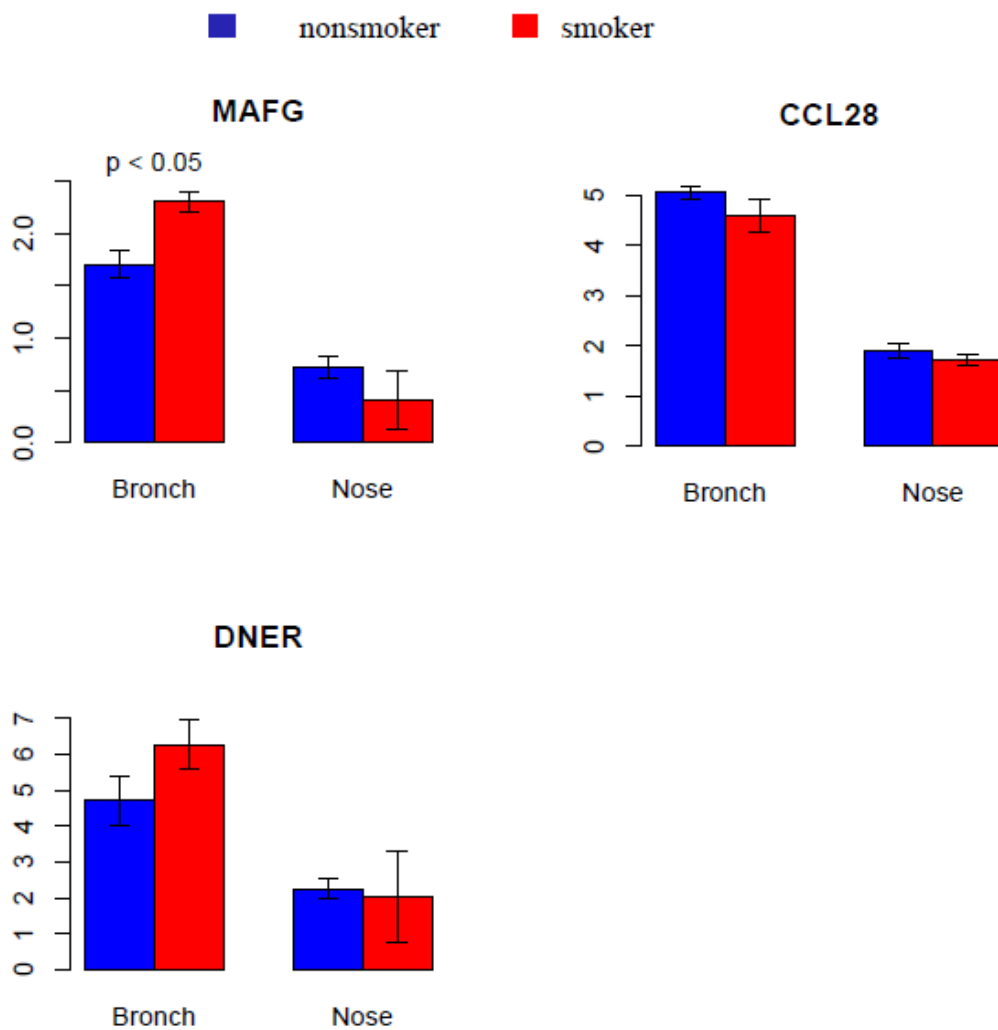
26

27

28 **Supplementary Figure 2**

29 RT-PCR validation of genes that are more dramatically affected by smoking in nose or
30 bronchus. Error bars indicate standard error. P-values indicate the Student's t-test
31 between the relative expression differences in 3 smokers versus 3 nonsmokers
32 respectively.

33



34

35

36

37 **Supplementary Table 1**

38 Study information including the present study and other two studies.

	Present Study		Spira et al. 2004 ⁽¹⁾	Sridhar et al. 2008 ⁽²⁾ (unmatched subjects)	
	Bronchus	Nose	Bronchus	Bronchus	Nose
nonsmokers	14 [#]	14	23*	23*	8
smokers	13	13	34*	34*	7

39 # One bronchial sample from one never smoker was identified as an outlier and excluded from
40 the analysis

41 *These samples represent the same data.

42

43 **Supplementary Table 2**

44 Primer sequences for candidate genes and a control gene (GAPDH).

gene	Forward	Reverse
CYP1B1	TATGCAGGAGCTTCTGGGAGA	TGGGCCCTTTAAGTCTTTGACTC
CYP1A1	CATGCTGACCCTGGGAAAGA	TGGCTCATCCTTGACAGTGC
AKR1B10	AGATCACAGTGAACCTTAGTCCTGTTATAGAC	AAAACAGCACCTCGATTCTCGT
TMEM45A	AGCGAACCTGCTATCTTGGTTC	CCCTCCAAAATTTCCAATCGAT
SEC14L3	CCACTTCCCTTGTTGCACATAG	TGCTTCCAGGAAAGTCCCTAAA
MAFG	CACCAGCGTCATCACAATAGTAAAG	CCCTACGATCGGGCATCC
CCL28	CACACCTTTTATCCCAGCACTTC	TCAAGTGATCCTCCTGCCTCA
DNER	GTTTGAGTGCCGGCTTTCTG	TGCTACGTTACGTGGTTTTTCCTAAC
GAPDH	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG

45

46

47 **Supplementary Table 3**

48 GO molecular function categories, human KEGG pathways, and GenMAPP pathways
 49 over-represented among 119 genes commonly changed by smoking in bronchus and nose
 50 compared to all annotated genes on the human exon arrays (FDR < 0.05).

Category	Term	Count	PValue	Benjamini
SP_PIR_KEYWORDS	microsome	9	2.10E-07	1.12E-04
SP_PIR_KEYWORDS	oxidoreductase	17	1.59E-07	1.70E-04
GOTERM_MF_ALL	GO:0050381~unspecific monooxygenase activity	6	2.96E-07	8.53E-04
GOTERM_MF_ALL	GO:0016712~oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen	6	1.04E-06	0.001491
SP_PIR_KEYWORDS	iron	10	1.48E-05	0.003151
INTERPRO	IPR002401:Cytochrome P450, E-class, group I	7	5.53E-07	0.003268
SP_PIR_KEYWORDS	nadp	8	1.35E-05	0.003583
SP_PIR_KEYWORDS	monooxygenase	7	1.13E-05	0.004004
GOTERM_CC_ALL	GO:0042598~vesicular fraction GO:0016705~oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	9	1.44E-05	0.006241
GOTERM_MF_ALL	GO:0016491~oxidoreductase activity	18	1.49E-05	0.008559
GOTERM_MF_ALL	GO:0005506~iron ion binding	11	1.31E-05	0.009402
GOTERM_CC_ALL	GO:0005792~microsome	9	1.18E-05	0.010148
INTERPRO	IPR001128:Cytochrome P450	7	3.62E-06	0.010668
SP_PIR_KEYWORDS	heme	7	1.00E-04	0.017638
GOTERM_CC_ALL	GO:0000267~cell fraction	18	7.74E-05	0.022134
GOTERM_MF_ALL	GO:0004497~monooxygenase activity	7	6.01E-05	0.028437
SP_PIR_KEYWORDS	pentose phosphate pathway	3	1.98E-04	0.029716
GOTERM_MF_ALL	GO:0020037~heme binding	7	1.83E-04	0.043602
GOTERM_MF_ALL	GO:0046906~tetrapyrrole binding	7	1.83E-04	0.049602

51

52

53

54 **Supplementary Table 4**

55 At FDR <0.05 and |fold-change| > 0.5, 119 genes that vary in expression between

56 smokers and nonsmokers were identified using a linear mixed-effects model including

57 bronchus and nose samples (see Figure 2). The log2 fold change was measured by the
 58 regression coefficient β_{status} between non-smokers and smokers when the site effect is
 59 fixed.

Transcript_Id	GeneSymbol	FoldChange(nonsmoker/smoker)	rawPvalue
3833948	CYP2A13	1.005	7.36E-11
3333488	SCGB1A1	0.922	7.75E-10
3625052	WDR72	-0.606	8.11E-10
2548699	CYP1B1	-3.204	9.59E-10
2319802	PGD	-0.653	5.40E-09
3614774	OCA2	-0.738	8.21E-09
2800711	ADCY2	0.817	1.49E-08
4000560	PIR	-1.012	3.15E-08
3316208	TALDO1	-0.633	1.01E-07
3649890	ABCC1	-0.528	1.54E-07
2599993	ABCB6	-0.607	3.70E-07
2711034	C3orf59	-0.596	5.49E-07
2786322	SLC7A11	-1.578	6.98E-07
2973995	EPB41L2	0.681	1.39E-06
3662201	MT1H	0.715	1.48E-06
2633691	TMEM45A	1.341	2.02E-06
3518418	KCTD12	0.739	2.04E-06
4025339	IDS	-0.562	2.53E-06
3652867	SCNN1G	1.094	3.46E-06
2696040	RAB6B	0.74	3.52E-06
3722129	CNTD1	0.772	3.57E-06
3025433	AKR1B10	-1.806	3.86E-06
2487918	ATP6V1B1	0.896	4.03E-06
3726298	TMEM92	-0.52	4.42E-06
2913694	CD109	-0.877	4.76E-06
3894545	SDCBP2	-0.647	5.40E-06
3540155	C14orf50	0.655	5.43E-06
3138464	PDE7A	0.521	5.46E-06
4013549	ITM2A	0.98	5.75E-06
2734047	AGPAT9	-0.54	6.29E-06
2517588	OSBPL6	0.864	6.78E-06
3894322	SRXN1	-0.643	6.93E-06
2908423	SLC29A1	0.634	6.96E-06
3957341	SEC14L3	1.792	7.02E-06
3884830	PPP1R16B	0.823	8.30E-06
3497195	CLDN10	-1.008	9.23E-06
2587790	GPR155	0.598	9.67E-06
3429460	TXNRD1	-0.661	1.05E-05
2814756	MAP1B	-0.67	1.26E-05
3818091	TMEM146	0.804	1.28E-05
3692999	MT1G	1.073	1.31E-05
3219621	CTNNA1	0.858	1.55E-05

3247818	BICC1	0.711	1.58E-05
3611625	ALDH1A3	-0.629	1.92E-05
3335029	POLA2	0.547	2.22E-05
3791850	SERPINB13	-0.794	2.53E-05
2685304	PROS1	0.98	2.58E-05
2554018	EFEMP1	0.609	2.59E-05
2739308	EGF	-0.686	2.91E-05
3604147	KIAA1199	-0.584	2.93E-05
3013255	PEG10	0.581	3.10E-05
3061805	SGCE	0.535	3.16E-05
3633081	CYP1A1	-2.141	3.18E-05
3424218	ACSS3	0.829	3.34E-05
3623031	FBN1	0.676	3.46E-05
3012064	PFTK1	0.503	3.71E-05
3766013	10-Mar	0.607	3.92E-05
3943504	TIMP3	-0.848	4.14E-05
2890741	SCGB3A1	0.564	4.28E-05
4027416	G6PD	-0.636	4.29E-05
2678448	FAM107A	0.739	4.71E-05
2546795	CAPN13	0.696	5.01E-05
2533019	UGT1A9	-0.56	5.15E-05
2334986	CYP4X1	0.501	5.27E-05
2415266	CYP2J2	0.616	5.58E-05
2511603	GALNT5	-0.584	6.69E-05
3144346	RUNX1T1	0.575	7.68E-05
2414440	C1orf168	0.756	7.79E-05
2939593	PECI	0.718	8.07E-05
3089469	SORBS3	0.502	8.83E-05
3726691	ABCC3	-0.563	9.08E-05
3732092	CACNG4	-0.732	9.27E-05
3379597	MTL5	0.515	9.43E-05
3103293	RDH10	-0.521	9.95E-05
3356038	TMEM45B	-0.564	1.03E-04
3950602	PANX2	-0.562	1.09E-04
3136888	TOX	0.607	1.10E-04
3429555	EID3	-0.532	1.21E-04
3323413	HTATIP2	-0.529	1.22E-04
3375648	FTH1	-0.665	1.37E-04
3108648	C8orf47	0.573	1.40E-04
2772450	SULT1E1	0.654	1.42E-04
2838116	FABP6	0.86	1.47E-04
2866225	MEF2C	0.523	1.56E-04
3864375	LYPD3	-0.731	1.58E-04
3381879	P4HA3	0.507	1.72E-04
2811145	PART1	0.735	1.76E-04
2790324	RNF175	-0.642	1.98E-04
3782088	CABYR	-0.685	2.03E-04
2856044	EMB	0.544	2.04E-04
3823304	CYP4F3	-0.596	2.14E-04
3566495	C14orf37	0.619	2.14E-04

3484165	C13orf26	0.771	2.17E-04
3494629	SCEL	-0.844	2.19E-04
4000538	FIGF	-0.525	2.20E-04
3061621	TFPI2	0.624	2.21E-04
2431886	PDE4DIP	0.575	2.25E-04
3343293	CCDC81	0.648	2.61E-04
3596147	GCNT3	-0.655	2.72E-04
3450775	KIF21A	0.517	2.82E-04
3844781	PRG2	0.671	3.14E-04
3998766	KAL1	0.565	3.20E-04
3189311	PBX3	0.553	3.42E-04
3638607	ANPEP	0.634	3.89E-04
2898597	GMNN	0.53	3.92E-04
2828356	CSF2	-0.571	4.15E-04
3445741	MGP	0.596	4.36E-04
3662093	MT3	0.542	4.36E-04
3870895	CDC42EP5	-1.013	4.81E-04
3257246	IFIT1	0.73	4.88E-04
3190002	TTC16	0.507	5.32E-04
3282519	ARMC4	0.506	5.95E-04
3868857	KLK12	-1.132	6.17E-04
3853658	CYP4F11	-0.861	6.26E-04
2353988	FAM46C	0.54	6.43E-04
2343334	GIPC2	0.545	6.64E-04
3535780	PTGER2	0.616	6.64E-04
2881747	ANXA6	0.572	6.76E-04
3662696	CX3CL1	0.559	6.87E-04

60

61 References

- 62 1. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, Palma J and Brody JS. Effects of cigarette
63 smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A* 101: 10143-
64 10148, 2004.
- 65 2. Sridhar S, Schembri F, Zeskind J, Shah V, Gustafson AM, Steiling K, Liu G, Dumas YM, Zhang
66 X, Brody JS, Lenburg ME and Spira A. Smoking-induced gene expression changes in the
67 bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics* 9: 259, 2008.