Supplemental Methods

Microarray Data Sets

The airway BC signature was previously characterized in our laboratory by genome-wide comparison of the BC transcriptome *vs* transcriptome of the intact airway epithelium derived from healthy nonsmokers, using Affymetrix HG-U133 Plus 2.0 (Affymetrix, Santa Clara, CA) [1]. The 199 lung adenoCa data set originally described by Chitale et al (http://cbio.mskcc.org/Public/lung_array_data/) [2] was used as the primary adenoCa cohort, which was re-reviewed histologically and updated with regard to clinical information (survival data). Of the 199 individuals, there were 182 with available microarray data and detailed clinical information. Those excluded individuals with unspecified gender (n=6), pathological stage IV (n=3) and unknown (n=2) or tumor pathology other than adenoCa (n=6). The transcriptome profiling of this data set was with the Affymetrix HG-U133A (n=87) and HG-U133A 2.0 (n=95) microarrays. Of the 1161 airway BC signature genes [1], 862 were on the microarrays; for the purpose of the present study, these 862 genes are referred to as the "airway BC signature." The 544 non-BC signature genes (i.e., genes up-regulated in the intact airway epithelium *vs.* airway BC) on the microarray were also analyzed.

To visualize similarity and differences between various carcinoma subtypes, lung adenoCa, lung SqCa and airway BC were compared using principal component analysis (PCA; GeneSpring software) based on expression of the airway BC signature [1]. All cancer data sets used for this analysis were based on the Affymetrix HG-U133 Plus 2.0 array and are publically available at the Gene Expression Omnibus (GEO), including: 68 lung adenoCa (GSE12667) [3], 40 lung adenoCa (GSE19188) [4], 40 lung adenoCa and 18 lung SqCa (GSE10245) [5], 55 colorectal cancer (GSE17537) [6], 129 breast cancer (GSE5460) [7], 91 hepatocellular carcinoma (GSE9843) [8] and 39 pancreatic cancer (GSE15471) [9].

To validate the relationship of the expression of airway BC signature genes to survival for adenoCa, two independent lung adenoCa data sets were assessed, including 58 adenoCa (validation cohort 1, http://data.cgt.duke.edu/oncogene.php and GSE3141) [10] and 327 of 442 adenoCa (validation cohort 2, https://caarraydb.nci.nih.gov/caarray/publicExperimentDetailAction.do?ex) [11], excluding 104 subjects analyzed in Memorial Sloan Kettering Cancer Center, the majority of which are present in the primary cohort, and 11 large cell neuroendocrine carcinoma samples identified based on pathologic re-evaluation [12].

Calculation of the BC index (I_{BC}) for each individual subject

For each airway BC signature gene, the median expression level was determined with respect to each array. Then the I_{BC} was calculated for each subject as a number of the airway BC genes having expression levels higher than median level in the analyzed cohort using the formula:

$$I_{BC} = \sum_{n=1}^{862} E_n$$

where E1 had a value of 1 if the expression level for gene 1 was >median level of adenoCa subjects or had a value of 0 if the expression level is ≤median level of adenoCa subjects; E2 is the index for gene 2, etc.

Comparison of Differentiation-associated Molecular Patterns in BC-high *vs* BC-low AdenoCa

To determine differences in the expression of differentiation-associated molecular features in BC-high *vs* BC-low adenoCa, expression of genes associated with the major cell types of the human airway epithelium were compared in the lung adenoCa subtypes of the primary cohort. The genes assessed included those associated with ciliated cells [forkhead box J1

(FOXJ1) and dynein axonemal intermediate chain 1 (DNAI1)]; mucus-secreting cells [mucin 5AC (MUC5AC) and trefoil factor 3 (TFF3)] [13]; Clara cells [NK2 homeobox 1 (NKX2-1) and secretoglobin 1A1 (SCGB1A)] [14], and neuroendocrine cells [synaptophysin (SYP) and chromogranin A (CHGA)] [15]. In addition, expression of genes related to epithelial-mesenchymal transition (EMT), including snail homolog 1(SNAI1), snail homolog 2 (SNAI2), twist homolog 1 (TWIST1), and N-cadherin (CDH2) [16] were also analyzed.

Comparative Analysis of Airway BC Signature Gene Expression in Lung AdenoCa and Lung Squamous Cell Carcinoma (SqCa)

To compare the expression of the airway BC signature in lung adenoCa to SqCa, the 58 adenoCa and 53 SqCa described by Bild et al [10] was analyzed. To compare the overall airway BC signature expression between adenoCa and SqCa, the BC index was calculated based on the median level of each airway BC signature gene in the adenoCa subjects, and the I_{BC} values were assessed using Mann-Whitney test. For identification of the airway BC signature genes differentially expressed in BC-high adenoCa *vs* SqCa, the criteria used was $p<0.05$ with a Benjamini-Hochberg correction to limit the false positive rate. Differential expression of selected airway BC signature genes in the analyzed lung cancer subtypes (BC-low adenoCa, BC-high adenoCa and SqCa) was additionally analyzed using Mann-Whitney test.

References

1. Hackett NR, Shaykhiev R, Walters MS, et al. The Human Airway Epithelial Basal Cell Transcriptome. *PLoS One* 2011; 6: e18378.

2. Chitale D, Gong Y, Taylor BS, et al. An Integrated Genomic Analysis of Lung Cancer Reveals Loss of DUSP4 in EGFR-Mutant Tumors. *Oncogene* 2009; 28: 2773-2783.

3. Ding L, Getz G, Wheeler DA, et al. Somatic Mutations Affect Key Pathways in Lung Adenocarcinoma. *Nature* 2008; 455: 1069-1075.

4. Hou J, Aerts J, den HB, et al. Gene Expression-Based Classification of Non-Small Cell Lung Carcinomas and Survival Prediction. *PLoS One* 2010; 5: e10312.

5. Kuner R, Muley T, Meister M, et al. Global Gene Expression Analysis Reveals Specific Patterns of Cell Junctions in Non-Small Cell Lung Cancer Subtypes. *Lung Cancer* 2009; 63: 32-38.

6. Smith JJ, Deane NG, Wu F, et al. Experimentally Derived Metastasis Gene Expression Profile Predicts Recurrence and Death in Patients With Colon Cancer. *Gastroenterology* 2010; 138: 958-968.

7. Lu X, Lu X, Wang ZC, et al. Predicting Features of Breast Cancer With Gene Expression Patterns. *Breast Cancer Res Treat* 2008; 108: 191-201.

8. Chiang DY, Villanueva A, Hoshida Y, et al. Focal Gains of VEGFA and Molecular Classification of Hepatocellular Carcinoma. *Cancer Res* 2008; 68: 6779-6788.

9. Badea L, Herlea V, Dima SO, et al. Combined Gene Expression Analysis of Whole-Tissue and Microdissected Pancreatic Ductal Adenocarcinoma Identifies Genes Specifically Overexpressed in Tumor Epithelia. *Hepatogastroenterology* 2008; 55: 2016-2027.

10. Bild AH, Yao G, Chang JT, et al. Oncogenic Pathway Signatures in Human Cancers As a Guide to Targeted Therapies. *Nature* 2006; 439: 353-357.

11. Shedden K, Taylor JM, Enkemann SA, et al. Gene Expression-Based Survival Prediction in Lung Adenocarcinoma: a Multi-Site, Blinded Validation Study. *Nat Med* 2008; 14: 822-827.

12. Bryant CM, Albertus DL, Kim S, et al. Clinically Relevant Characterization of Lung Adenocarcinoma Subtypes Based on Cellular Pathways: an International Validation Study. *PLoS One* 2010; 5: e11712.

13. Dvorak A, Tilley AE, Shaykhiev R, et al. Do Airway Epithelium Air-Liquid Cultures Represent the in Vivo Airway Epithelium Transcriptome? *Am J Respir Cell Mol Biol* 2011; 44: 465-473.

14. Zhang L, Whitsett JA, Stripp BR. Regulation of Clara Cell Secretory Protein Gene Transcription by Thyroid Transcription Factor-1. *Biochim Biophys Acta* 1997; 1350: 359-367.

15. Wiedenmann B, Huttner WB. Synaptophysin and Chromogranins/Secretogranins--Widespread Constituents of Distinct Types of Neuroendocrine Vesicles and New Tools in Tumor Diagnosis. *Virchows Arch B Cell Pathol Incl Mol Pathol* 1989; 58: 95-121.

16. Kalluri R, Weinberg RA. The Basics of Epithelial-Mesenchymal Transition. *J Clin Invest* 2009; 119: 1420-1428.

Supplemental Tables

Supplemental Table I	Lung Adenocarcinoma Patient Characteristics
Supplemental Table II	Gene Expression Analysis of Airway Basal Cell (BC) Signature in Lung Adenocarcinoma (adenoCa)
Supplemental Table III	Multivariate Cox Regression Analyses Including the Category Associated with the Airway Basal Cell (BC) Signature

Supplemental Table I. Lung Adenocarcinoma Patient Characteristics

	Primary cohort (Chitalle et al.)	**Validation 1 (Bild et al.)**	**Validation 2 (Shedden et al.)**
Number of patients	182	58	327
Age, mean ± S.D.	66.7±10.7	N.A.[2]	64.3±10.2
Gender			
Male	78	N.A.	180
Female	104	N.A.	147
Smoking history			
Never	35	N.A.	28
Ever	146	N.A.	210
Unknown	1	N.A.	89
Pathological stage[1]			
I	63	N.A.	200
II	80	N.A.	65
III	39	N.A.	60
Unknown	0	N.A.	2
Median follow-up (month)	43.8	30.1	48.8
Number of death	90	32	191

[1] Pathological stage was based on 6th edition TNM staging.

[2] Abbreviations: N.A.: not available

Supplemental Table II. Characterization of the 10-Gene Basal Cell (BC)-high Lung Adenocarcinoma (adenoCa) Signature

Gene name	Gene symbol	Functions	Identification of BC-high adenoCa	
			Sensitivity (%)	Specificity (%)
Phosphoglycerate mutase 1 (brain)	PGAM1	Enzyme in the glycolytic pathway	93.5	93.5
Transmembrane protein 158 (gene/pseudogene)	TMEM158	Ras-induced senescence	91.3	82.6
Solute carrier family 16, member 3 (monocarboxylic acid transporters)	SLC16A3	Lactic acid and pyruvate transport	89.1	87
Desmoglein 2	DSG2	Cadherin cell adhesion	89.1	80.4
Serine/threonine kinase receptor-associated protein	STRAP	RNA splicing	89.1	80.4
Glutaredoxin 2	GLRX2	Induction of apoptosis by oxidative stress	87	89.1
ArfGAP with FG repeats 1	AGFG1	RNA trafficking or localization	87	84.8
Leucine rich repeat containing 42	LRRC42	(unknown)	87	84.8
Sec23 homolog A (S. cerevisiae)	SEC23A	Endoplasmic reticulum-Golgi protein trafficking	87	84.8
Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	PLOD2	Stability of intermolecular collagen cross-links	87	76.1

1 Based on the microarray analysis of the primary adenoCa data set.

Abbreviations: BC: basal cell, adenoCa: adenocarcinoma, AE: airway epithelium, N.S.: not significant.

Supplemental Table III. Gene Expression Analysis of Airway Basal Cell (BC) Signature in Lung Adenocarcinoma (adenoCa)

Compared gene sets	**Cohorts**	**p value (t-test)**	
		< 0.5x median	**> 2x median**
Airway BC signature *vs* non-BC signature	1) Primary cohort (n=182, Chitale et al) 2) Validation cohort 1 (n=58, Bild et al) 3) Validation cohort 2 (n=327, Shedden et al.)	p<0.002	p<0.0006
Airway BC signature *vs* randomly selected gene sets 1-50	Airway BC signature 1) Primary cohort (n=182, Chitale et al) 2) Validation cohort 1 (n=58, Bild et al) 3) Validation cohort 2 (n=327, Shedden et al.) Ramdomly selected gene sets 1-50 1) Primary cohort (n=182, Chitale et al)	p<0.0000005	p<0.02

Supplemental Figure Legends

Supplemental Figure 1. Diagram representing experimental flow of the study.

Supplemental Figure 2. Correlation between BC index and NK2 homeobox 1 (NKX2-1) expression in primary lung adenocarcinoma cohort (n=182). Y-axis - BC index; x-axis – normalized NKX2-1 expression levels; Pearson correlation (r) and p value are indicated.

Supplemental Figure 3. Immunohistological analysis of BC-high lung adenoCa, BC-low lung adenoCa, lung squamous cell carcinoma (SqCa) and normal lung tissue for the expression of the thyroid transcription factor-1 (TTF-1) and tumor protein TP63. Representative BC-high and BC-low adenoCa biopsy samples were selected based on the gene expression data as described in Methods; scale bar - 50µm for original magnification x60 for normal lung tissue, 20 µm for original magnification x20 for lung cancer tissue.

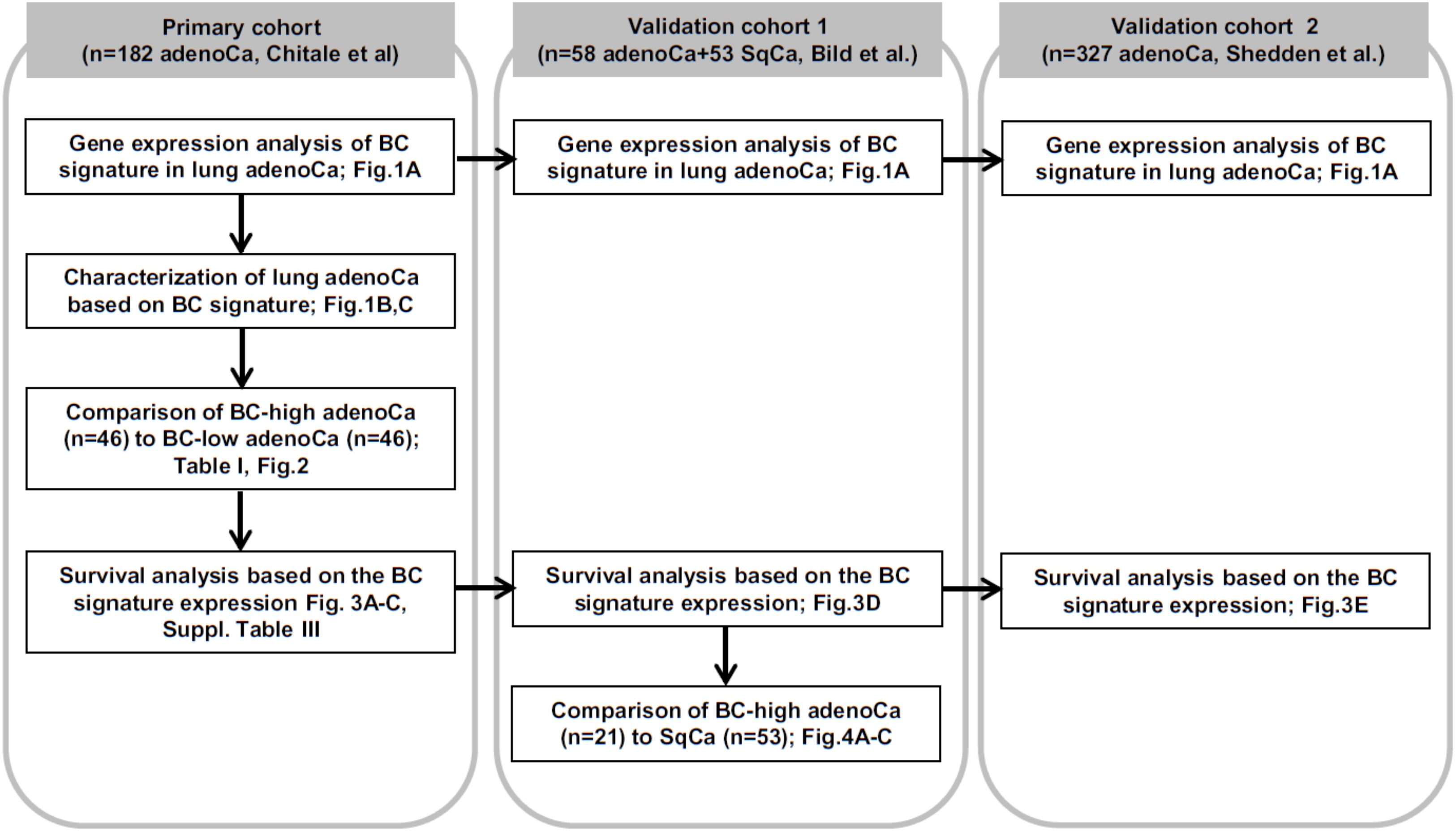
Supplemental Figure 4. Examples of expression of small cell lung carcinoma-related genes in BC-high adenoCa compared to BC-low adenoCa**:** tumor protein p53 (TP53); retinoblastoma 1 (RB1), V-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma derived (avian) (MYCL1, also known as L-MYC). In all panels, log2-transformed normalized gene expression levels based on the microarray analysis are shown; n=46 in each group. Outliers are indicated on the basis of interquartile range (IQR); ◦ -1.5 x IQR to 3 x IQR, * - more or less than 3 x IQR.

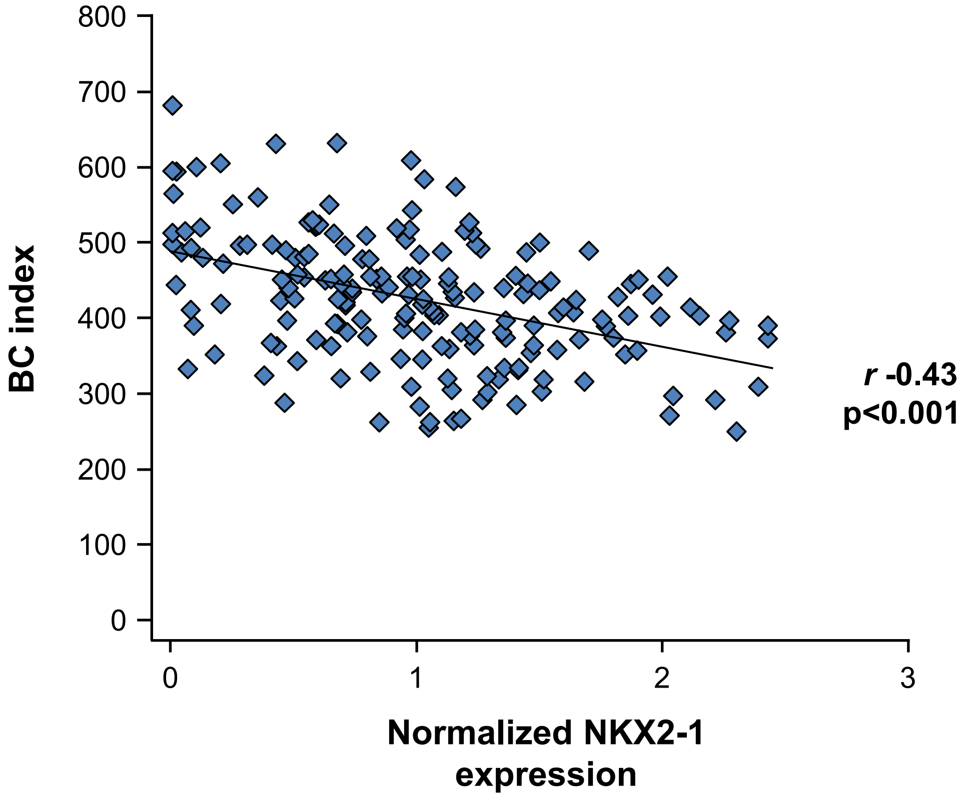
Supplemental Figure 5. Disease free survival after surgery of BC-high adenoCa (red) vs BC-low adenoCa (blue) in primary lung adenocarcinoma cohort (n=182). p values were determined by the log-rank test; the number of individuals in each group are indicated.

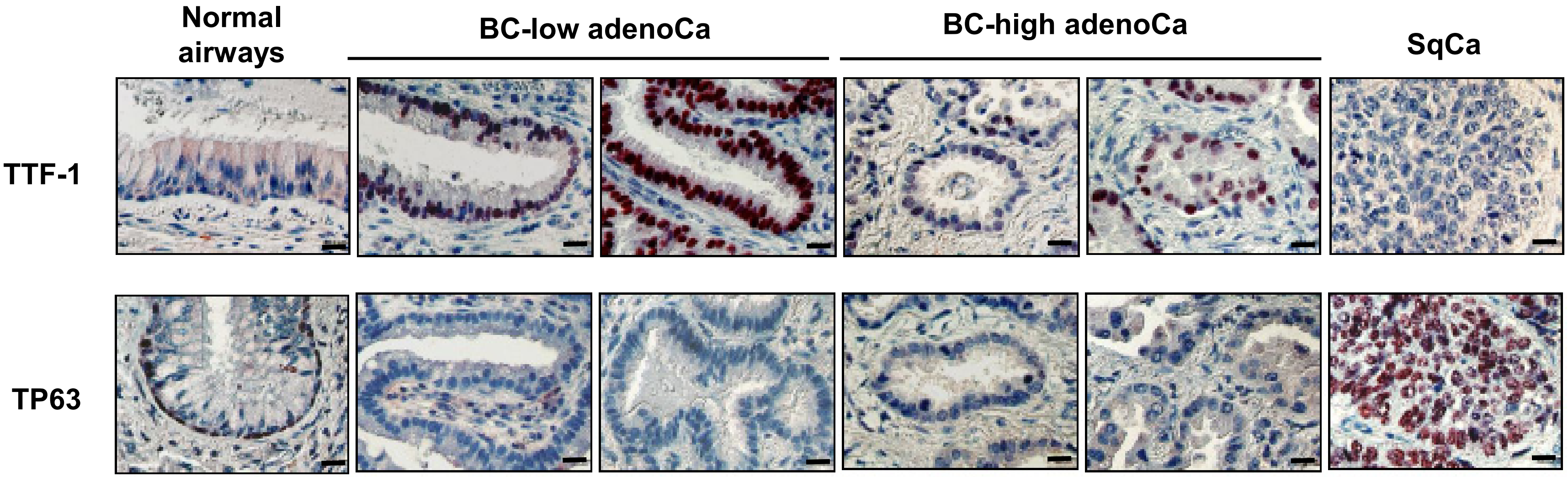
Supplemental Figure 6. Survival analysis of lung squamous cell carcinoma (SqCa) patients based on the BC signature expression in the lung cancer cohort (Bild et al.) compared to the adenoCa. **A.** Categorization of BC-high (red) and BC-low (blue) patients with adenocarcinoma (adenoCa) and SqCa. SqCa samples were categorized based on the adenoCa cohort, in which

BC-high and BC-low samples were determined as described in Methods. **B.** Overall survival of BC-high SqCa (red) *vs* BC-low SqCa patients (blue) **C.** Overall survival of BC-high adenoCa (red), BC-low adenoCa and SqCa. p values were determined by the log-rank test; the number of individuals in each group are indicated.

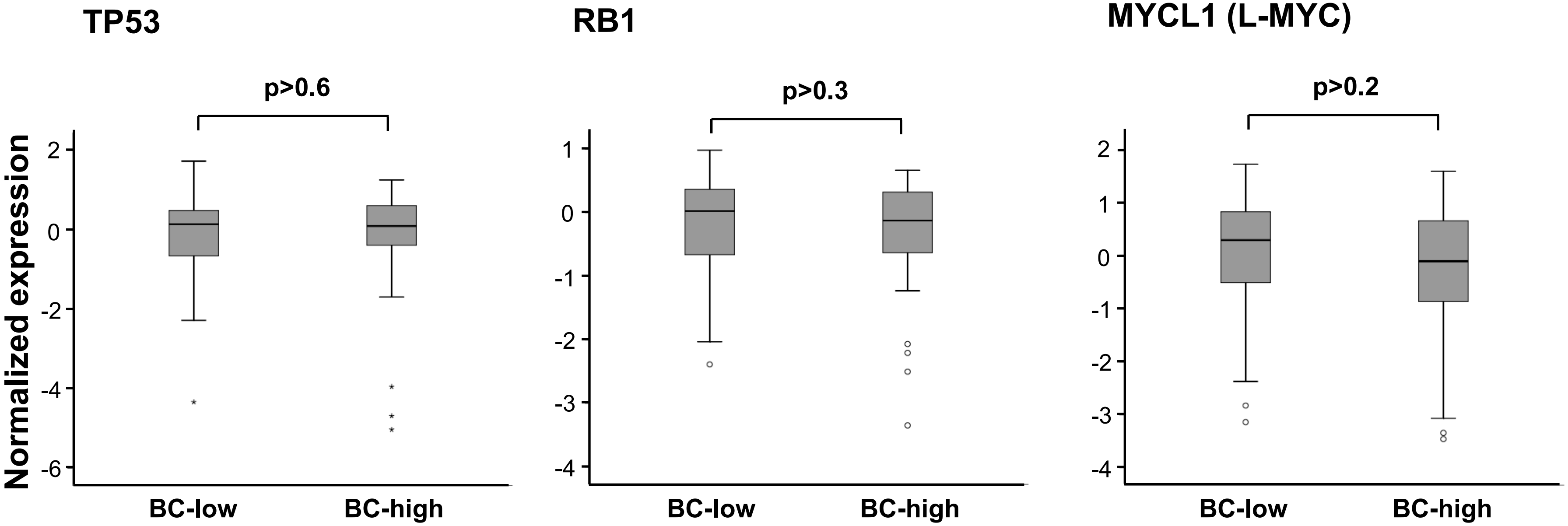
Supplemental Figure 1

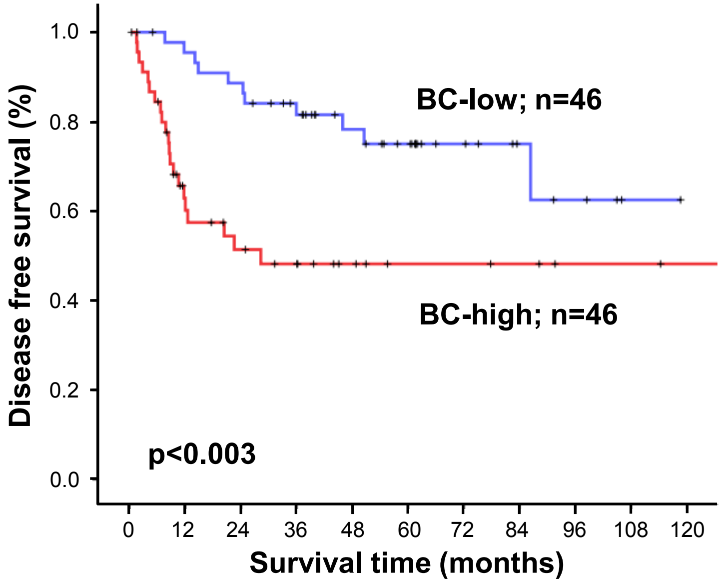



BC index
800
700
600
500
400
300
200
100
0
0
1
2
3
Normalized NKX2-1 expression
r -0.43
p<0.001

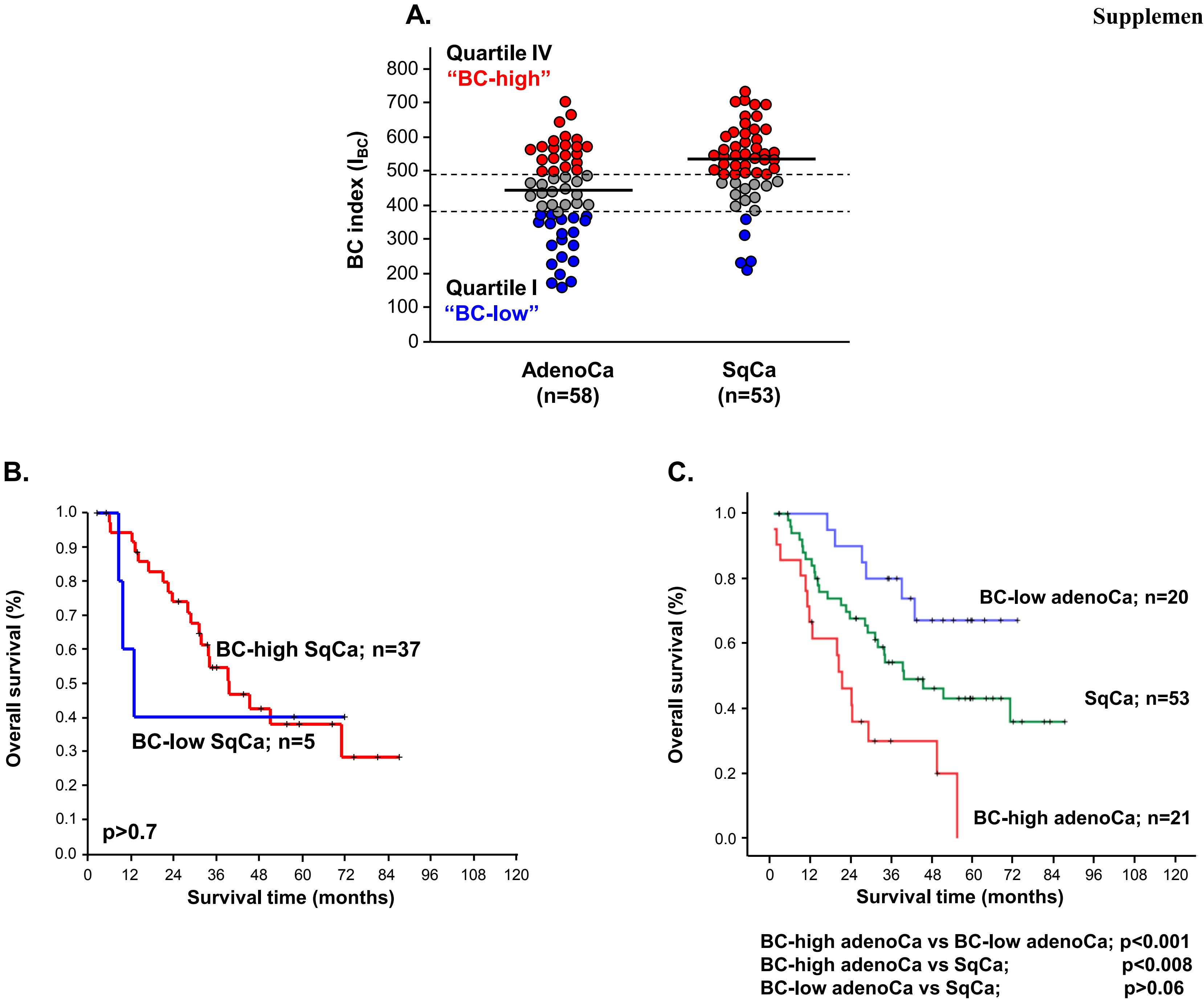

Normal airways
BC-low adenoCa
BC-high adenoCa
SqCa
TTF-1
TP63

Supplemental Figure 4




1.0
0.8
0.6
0.4
0.2
0.0
Disease free survival (%)
BC-low; n=46
BC-high; n=46
p<0.003
0
12
24
36
48
60
72
84
96
108
120
Survival time (months)

Supplemental Figure 6



Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
200650_s_at	lactate dehydrogenase A	57.01	Above 2x median
201666_at	TIMP metallopeptidase inhibitor 1	51.60	Above 2x median
200872_at	S100 calcium binding protein A10	44.97	Above 2x median
211945_s_at	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	36.19	Above 2x median
217466_x_at	ribosomal protein S2	32.40	Above 2x median
211936_at	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	30.89	Above 2x median
217963_s_at	nerve growth factor receptor (TNFRSF16) associated protein 1	29.77	Above 2x median
208679_s_at	actin related protein 2/3 complex, subunit 2, 34kDa	29.23	Above 2x median
209386_at	transmembrane 4 L six family member 1	28.61	Above 2x median
211597_s_at	HOP homeobox	27.42	Above 2x median
201215_at	plastin 3	27.17	Above 2x median
200634_at	profilin 1	27.14	Above 2x median
212185_x_at	metallothionein 2A	24.94	Above 2x median
201105_at	lectin, galactoside-binding, soluble, 1	24.82	Above 2x median
217294_s_at	enolase 1, (alpha)	23.20	Above 2x median
200989_at	hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	22.58	Above 2x median
200700_s_at	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2	20.86	Above 2x median
201268_at	NME1-NME2 readthrough / non-metastatic cells 2, protein (NM23B) expressed in	20.20	Above 2x median
200600_at	moesin	19.30	Above 2x median
200886_s_at	phosphoglycerate mutase 1 (brain)	19.10	Above 2x median
201818_at	lysophosphatidylcholine acyltransferase 1	18.96	Above 2x median
211980_at	collagen, type IV, alpha 1	18.70	Above 2x median
222229_x_at	ribosomal protein L26	18.66	Above 2x median
200779_at	activating transcription factor 4 (tax-responsive enhancer element B67)	18.48	Above 2x median
200625_s_at	CAP, adenylate cyclase-associated protein 1 (yeast)	17.90	Above 2x median
208700_s_at	transketolase	17.61	Above 2x median
214853_s_at	SHC (Src homology 2 domain containing) transforming protein 1	17.28	Above 2x median
201069_at	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	17.09	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 2)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
200632_s_at	N-myc downstream regulated 1	16.88	Above 2x median
203108_at	G protein-coupled receptor, family C, group 5, member A	16.47	Above 2x median
210986_s_at	tropomyosin 1 (alpha)	16.04	Above 2x median
218205_s_at	MAP kinase interacting serine/threonine kinase 2	15.99	Above 2x median
208540_x_at	S100 calcium binding protein A11	15.83	Above 2x median
209373_at	mal, T-cell differentiation protein-like	15.08	Above 2x median
220924_s_at	solute carrier family 38, member 2	14.79	Above 2x median
211964_at	collagen, type IV, alpha 2	14.78	Above 2x median
208581_x_at	metallothionein 1X	14.74	Above 2x median
212859_x_at	metallothionein 1E	14.53	Above 2x median
200931_s_at	vinculin	14.42	Above 2x median
216241_s_at	transcription elongation factor A (SII), 1	14.25	Above 2x median
209154_at	Tax1 (human T-cell leukemia virus type I) binding protein 3	13.96	Above 2x median
216640_s_at	protein disulfide isomerase family A, member 6	13.95	Above 2x median
201150_s_at	TIMP metallopeptidase inhibitor 3	13.42	Above 2x median
208892_s_at	dual specificity phosphatase 6	13.23	Above 2x median
200999_s_at	cytoskeleton-associated protein 4	13.15	Above 2x median
202442_at	adaptor-related protein complex 3, sigma 1 subunit	13.14	Above 2x median
221923_s_at	nucleophosmin (nucleolar phosphoprotein B23, numatrin)	13.07	Above 2x median
220755_s_at	chromosome 6 open reading frame 48	12.97	Above 2x median
209016_s_at	keratin 7	12.68	Above 2x median
201286_at	syndecan 1	12.62	Above 2x median
201416_at	SRY (sex determining region Y)-box 4	12.56	Above 2x median
201506_at	transforming growth factor, beta-induced, 68kDa	12.40	Above 2x median
203484_at	Sec61 gamma subunit	12.27	Above 2x median
207738_s_at	NCK-associated protein 1	12.25	Above 2x median
210512_s_at	vascular endothelial growth factor A	12.16	Above 2x median
208152_s_at	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21	12.14	Above 2x median
200788_s_at	phosphoprotein enriched in astrocytes 15	12.04	Above 2x median
201631_s_at	immediate early response 3	12.02	Above 2x median
219293_s_at	Obg-like ATPase 1	11.69	Above 2x median
208636_at	actinin, alpha 1	11.60	Above 2x median
214091_s_at	glutathione peroxidase 3 (plasma)	11.51	Above 2x median
200870_at	serine/threonine kinase receptor associated protein	11.43	Above 2x median
201577_at	non-metastatic cells 1, protein (NM23A) expressed in	11.38	Above 2x median
200790_at	ornithine decarboxylase 1	11.21	Above 2x median
200743_s_at	tripeptidyl peptidase I	10.95	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 3)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
201828_x_at	family with sequence similarity 127, member A	10.91	Above 2x median
208726_s_at	eukaryotic translation initiation factor 2, subunit 2 beta, 38kDa	10.69	Above 2x median
208693_s_at	glycyl-tRNA synthetase	10.49	Above 2x median
218084_x_at	FXYD domain containing ion transport regulator 5	10.42	Above 2x median
200606_at	desmoplakin	10.38	Above 2x median
208801_at	signal recognition particle 72kDa	10.01	Above 2x median
222231_s_at	leucine rich repeat containing 59	9.96	Above 2x median
208886_at	H1 histone family, member 0	9.95	Above 2x median
200885_at	ras homolog gene family, member C	9.91	Above 2x median
213404_s_at	Ras homolog enriched in brain	9.87	Above 2x median
211284_s_at	granulin	9.85	Above 2x median
208161_s_at	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	9.71	Above 2x median
212473_s_at	microtubule associated monoxygenase, calponin and LIM domain containing 2	9.63	Above 2x median
200771_at	laminin, gamma 1 (formerly LAMB2)	9.60	Above 2x median
201587_s_at	interleukin-1 receptor-associated kinase 1	9.54	Above 2x median
200853_at	H2A histone family, member Z	9.54	Above 2x median
201560_at	chloride intracellular channel 4	9.51	Above 2x median
201015_s_at	junction plakoglobin	9.50	Above 2x median
213603_s_at	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	9.42	Above 2x median
211456_x_at	metallothionein 1 pseudogene 2	9.37	Above 2x median
202231_at	eukaryotic translation initiation factor 3, subunit M	9.30	Above 2x median
212265_at	quaking homolog, KH domain RNA binding (mouse)	9.14	Above 2x median
201667_at	gap junction protein, alpha 1, 43kDa	9.06	Above 2x median
221739_at	chromosome 19 open reading frame 10	9.06	Above 2x median
217938_s_at	potassium channel modulatory factor 1	9.02	Above 2x median
202720_at	testis derived transcript (3 LIM domains)	9.02	Above 2x median
217975_at	WW domain binding protein 5	8.94	Above 2x median
201189_s_at	inositol 1,4,5-triphosphate receptor, type 3	8.81	Above 2x median
201088_at	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	8.75	Above 2x median
207714_s_at	serpin peptidase inhibitor, clade H (heat shock protein 47), member 1, (collagen binding protein 1)	8.63	Above 2x median
205569_at	lysosomal-associated membrane protein 3	8.59	Above 2x median
216396_s_at	etoposide induced 2.4 mRNA	8.58	Above 2x median
212995_x_at	mitotic spindle organizing protein 2B	8.54	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 4)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
217995_at	sulfide quinone reductase-like (yeast)	8.47	Above 2x median
201064_s_at	poly(A) binding protein, cytoplasmic 4 (inducible form)	8.46	Above 2x median
210406_s_at	RAB6A, member RAS oncogene family / RAB6C, member RAS oncogene family	8.37	Above 2x median
203744_at	high-mobility group box 3	8.32	Above 2x median
201300_s_at	prion protein	8.22	Above 2x median
203510_at	met proto-oncogene (hepatocyte growth factor receptor)	8.14	Above 2x median
206461_x_at	metallothionein 1H	8.04	Above 2x median
216449_x_at	heat shock protein 90kDa beta (Grp94), member 1	7.99	Above 2x median
212610_at	protein tyrosine phosphatase, non-receptor type 11	7.92	Above 2x median
218718_at	platelet derived growth factor C	7.92	Above 2x median
202525_at	protease, serine, 8	7.78	Above 2x median
221269_s_at	SH3 domain binding glutamic acid-rich protein like 3	7.78	Above 2x median
201921_at	guanine nucleotide binding protein (G protein), gamma 10	7.60	Above 2x median
201319_at	myosin, light chain 12A, regulatory, non-sarcomeric	7.58	Above 2x median
212086_x_at	lamin A/C	7.58	Above 2x median
200802_at	seryl-tRNA synthetase	7.56	Above 2x median
202655_at	mesencephalic astrocyte-derived neurotrophic factor	7.52	Above 2x median
203476_at	trophoblast glycoprotein	7.43	Above 2x median
200832_s_at	stearoyl-CoA desaturase (delta-9-desaturase)	7.22	Above 2x median
218618_s_at	fibronectin type III domain containing 3B	7.20	Above 2x median
217901_at	desmoglein 2	7.18	Above 2x median
213923_at	RAP2B, member of RAS oncogene family	7.13	Above 2x median
201272_at	aldo-keto reductase family 1, member B1 (aldose reductase)	7.10	Above 2x median
217730_at	transmembrane BAX inhibitor motif containing 1	7.06	Above 2x median
209786_at	high mobility group nucleosomal binding domain 4	6.93	Above 2x median
218368_s_at	tumor necrosis factor receptor superfamily, member 12A	6.93	Above 2x median
202381_at	ADAM metallopeptidase domain 9	6.86	Above 2x median
201540_at	four and a half LIM domains 1	6.85	Above 2x median
201655_s_at	heparan sulfate proteoglycan 2	6.82	Above 2x median
201850_at	capping protein (actin filament), gelsolin-like	6.80	Above 2x median
208727_s_at	cell division cycle 42 (GTP binding protein, 25kDa)	6.78	Above 2x median
216336_x_at	hypothetical protein LOC100505584 / metallothionein 1E	6.76	Above 2x median
204745_x_at	metallothionein 1G	6.71	Above 2x median
32811_at	myosin IC	6.70	Above 2x median
201144_s_at	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	6.64	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 5)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
201393_s_at	insulin-like growth factor 2 receptor	6.63	Above 2x median
216505_x_at	ribosomal protein S10	6.48	Above 2x median
212048_s_at	tyrosyl-tRNA synthetase	6.47	Above 2x median
213746_s_at	filamin A, alpha	6.30	Above 2x median
209181_s_at	Rab geranylgeranyltransferase, beta subunit	6.29	Above 2x median
208786_s_at	microtubule-associated protein 1 light chain 3 beta	6.20	Above 2x median
213812_s_at	calcium/calmodulin-dependent protein kinase kinase 2, beta	6.19	Above 2x median
201963_at	acyl-CoA synthetase long-chain family member 1	6.18	Above 2x median
204528_s_at	nucleosome assembly protein 1-like 1	6.14	Above 2x median
206100_at	carboxypeptidase M	6.10	Above 2x median
208789_at	polymerase I and transcript release factor	6.09	Above 2x median
200744_s_at	guanine nucleotide binding protein (G protein), beta polypeptide 1	6.09	Above 2x median
209190_s_at	diaphanous homolog 1 (Drosophila)	6.07	Above 2x median
222108_at	adhesion molecule with Ig-like domain 2	6.06	Above 2x median
219032_x_at	opsin 3	6.04	Above 2x median
218677_at	S100 calcium binding protein A14	6.03	Above 2x median
212698_s_at	septin 10	5.98	Above 2x median
201849_at	BCL2/adenovirus E1B 19kDa interacting protein 3	5.96	Above 2x median
218336_at	prefoldin subunit 2	5.94	Above 2x median
212012_at	peroxidasin homolog (Drosophila)	5.88	Above 2x median
203065_s_at	caveolin 1, caveolae protein, 22kDa	5.73	Above 2x median
217755_at	hematological and neurological expressed 1	5.73	Above 2x median
208074_s_at	adaptor-related protein complex 2, sigma 1 subunit	5.72	Above 2x median
202345_s_at	fatty acid binding protein 5 (psoriasis-associated)	5.71	Above 2x median
202670_at	mitogen-activated protein kinase kinase 1	5.71	Above 2x median
202006_at	protein tyrosine phosphatase, non-receptor type 12	5.69	Above 2x median
221676_s_at	coronin, actin binding protein, 1C	5.69	Above 2x median
221830_at	RAP2A, member of RAS oncogene family	5.68	Above 2x median
218611_at	immediate early response 5	5.67	Above 2x median
201976_s_at	myosin X	5.65	Above 2x median
201013_s_at	phosphoribosylaminoimidazole carboxylase, phosphoribosylaminoimidazole succinocarboxamide synthetase	5.65	Above 2x median
207121_s_at	mitogen-activated protein kinase 6	5.63	Above 2x median
209210_s_at	fermitin family member 2	5.62	Above 2x median
35820_at	GM2 ganglioside activator	5.62	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 6)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
208961_s_at	Kruppel-like factor 6	5.55	Above 2x median
212242_at	tubulin, alpha 4a	5.50	Above 2x median
201761_at	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	5.50	Above 2x median
217738_at	Nicotinamide phosphoribosyltransferase	5.48	Above 2x median
201324_at	epithelial membrane protein 1	5.47	Above 2x median
209270_at	laminin, beta 3	5.42	Above 2x median
212640_at	protein tyrosine phosphatase-like (proline instead of catalytic arginine), member b	5.40	Above 2x median
212024_x_at	flightless I homolog (Drosophila)	5.40	Above 2x median
203786_s_at	tumor protein D52-like 1	5.37	Above 2x median
215780_s_at	SET translocation (myeloid leukemia-associated) pseudogene / SET nuclear oncogene	5.32	Above 2x median
218728_s_at	cornichon homolog 4 (Drosophila)	5.31	Above 2x median
201579_at	FAT tumor suppressor homolog 1 (Drosophila)	5.28	Above 2x median
209080_x_at	glutaredoxin 3	5.25	Above 2x median
203535_at	S100 calcium binding protein A9	5.24	Above 2x median
221577_x_at	growth differentiation factor 15	5.24	Above 2x median
202201_at	biliverdin reductase B (flavin reductase (NADPH))	5.24	Above 2x median
209146_at	sterol-C4-methyl oxidase-like	5.23	Above 2x median
210105_s_at	FYN oncogene related to SRC, FGR, YES	5.14	Above 2x median
200889_s_at	signal sequence receptor, alpha	5.13	Above 2x median
212250_at	metadherin	5.12	Above 2x median
200733_s_at	protein tyrosine phosphatase type IVA, member 1	5.09	Above 2x median
221827_at	RanBP-type and C3HC4-type zinc finger containing 1	5.09	Above 2x median
209803_s_at	pleckstrin homology-like domain, family A, member 2	5.09	Above 2x median
37408_at	mannose receptor, C type 2	5.08	Above 2x median
202949_s_at	four and a half LIM domains 2	5.02	Above 2x median
212566_at	microtubule-associated protein 4	5.02	Above 2x median
203323_at	caveolin 2	4.97	Above 2x median
212680_x_at	protein phosphatase 1, regulatory (inhibitor) subunit 14B	4.96	Above 2x median
204135_at	filamin A interacting protein 1-like	4.93	Above 2x median
206074_s_at	high mobility group AT-hook 1	4.93	Above 2x median
205263_at	B-cell CLL/lymphoma 10	4.88	Above 2x median
221059_s_at	coactosin-like 1 (Dictyostelium)	4.86	Above 2x median
218309_at	calcium/calmodulin-dependent protein kinase II inhibitor 1	4.80	Above 2x median
201853_s_at	cell division cycle 25 homolog B (S. pombe)	4.79	Above 2x median
218237_s_at	solute carrier family 38, member 1	4.78	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 7)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
217776_at	retinol dehydrogenase 11 (all-trans/9-cis/11-cis)	4.77	Above 2x median
201110_s_at	thrombospondin 1	4.71	Above 2x median
212846_at	ribosomal RNA processing 1 homolog B (S. cerevisiae)	4.70	Above 2x median
212765_at	calmodulin regulated spectrin-associated protein 1-like 1	4.69	Above 2x median
220199_s_at	axin interactor, dorsalization associated	4.69	Above 2x median
201263_at	threonyl-tRNA synthetase	4.68	Above 2x median
203910_at	Rho GTPase activating protein 29	4.68	Above 2x median
203287_at	ladinin 1	4.66	Above 2x median
217165_x_at	metallothionein 1F	4.63	Above 2x median
212470_at	sperm associated antigen 9	4.61	Above 2x median
202693_s_at	serine/threonine kinase 17a	4.59	Above 2x median
202620_s_at	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	4.58	Above 2x median
209191_at	tubulin, beta 6	4.54	Above 2x median
216348_at	ribosomal protein S17	4.47	Above 2x median
200755_s_at	calumenin	4.42	Above 2x median
201645_at	tenascin C	4.41	Above 2x median
200808_s_at	zyxin	4.38	Above 2x median
201573_s_at	eukaryotic translation termination factor 1	4.37	Above 2x median
210058_at	mitogen-activated protein kinase 13	4.33	Above 2x median
202771_at	family with sequence similarity 38, member A	4.33	Above 2x median
202352_s_at	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	4.32	Above 2x median
212160_at	exportin, tRNA (nuclear export receptor for tRNAs)	4.31	Above 2x median
212245_at	multiple coagulation factor deficiency 2	4.31	Above 2x median
212190_at	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	4.30	Above 2x median
201595_s_at	zinc finger CCCH-type containing 15	4.25	Above 2x median
212018_s_at	ribosomal L1 domain containing 1	4.25	Above 2x median
212298_at	neuropilin 1	4.25	Above 2x median
202933_s_at	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	4.24	Above 2x median
210396_s_at	bolA homolog 2 (E. coli) / PI-3-kinase-related kinase SMG-1 pseudogene / PI-3-kinase-related kinase SMG-1 pseudogene	4.22	Above 2x median
221510_s_at	glutaminase	4.20	Above 2x median
212739_s_at	non-metastatic cells 4, protein expressed in	4.19	Above 2x median
202284_s_at	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	4.19	Above 2x median
213532_at	ADAM metallopeptidase domain 17	4.18	Above 2x median
52164_at	chromosome 11 open reading frame 24	4.16	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 8)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
221654_s_at	ubiquitin specific peptidase 3	4.16	Above 2x median
202437_s_at	cytochrome P450, family 1, subfamily B, polypeptide 1	4.14	Above 2x median
201751_at	Josephin domain containing 1	4.14	Above 2x median
221580_s_at	TATA box binding protein (TBP)-associated factor, RNA polymerase I, D, 41kDa	4.13	Above 2x median
209295_at	tumor necrosis factor receptor superfamily, member 10b	4.12	Above 2x median
209135_at	aspartate beta-hydroxylase	4.12	Above 2x median
209836_x_at	bolA homolog 2 (E. coli) / bolA homolog 2B (E. coli)	4.11	Above 2x median
209409_at	growth factor receptor-bound protein 10	4.07	Above 2x median
200615_s_at	adaptor-related protein complex 2, beta 1 subunit	4.07	Above 2x median
203414_at	monocyte to macrophage differentiation-associated	4.06	Above 2x median
205479_s_at	plasminogen activator, urokinase	4.04	Above 2x median
214435_x_at	v-ral simian leukemia viral oncogene homolog A (ras related)	4.03	Above 2x median
214214_s_at	complement component 1, q subcomponent binding protein	4.02	Above 2x median
209100_at	interferon-related developmental regulator 2	3.96	Above 2x median
201980_s_at	Ras suppressor protein 1	3.95	Above 2x median
201474_s_at	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	3.84	Above 2x median
203574_at	nuclear factor, interleukin 3 regulated	3.80	Above 2x median
205324_s_at	FtsJ homolog 1 (E. coli)	3.78	Above 2x median
212104_s_at	RNA binding protein, fox-1 homolog (C. elegans) 2	3.78	Above 2x median
205105_at	mannosidase, alpha, class 2A, member 1	3.77	Above 2x median
217988_at	cyclin B1 interacting protein 1	3.76	Above 2x median
209344_at	tropomyosin 4	3.73	Above 2x median
212072_s_at	casein kinase 2, alpha 1 polypeptide	3.71	Above 2x median
202347_s_at	ubiquitin-conjugating enzyme E2K (UBC1 homolog, yeast)	3.69	Above 2x median
211043_s_at	clathrin, light chain B	3.68	Above 2x median
202686_s_at	AXL receptor tyrosine kinase	3.67	Above 2x median
202679_at	Niemann-Pick disease, type C1	3.67	Above 2x median
212887_at	Sec23 homolog A (S. cerevisiae)	3.65	Above 2x median
205807_s_at	tuftelin 1	3.65	Above 2x median
217913_at	vacuolar protein sorting 4 homolog A (S. cerevisiae)	3.62	Above 2x median
202148_s_at	pyrroline-5-carboxylate reductase 1	3.62	Above 2x median
213671_s_at	methionyl-tRNA synthetase	3.60	Above 2x median
212658_at	lipoma HMGIC fusion partner-like 2	3.59	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 9)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
202369_s_at	translocation associated membrane protein 2	3.58	Above 2x median
218421_at	ceramide kinase	3.55	Above 2x median
202431_s_at	v-myc myelocytomatosis viral oncogene homolog (avian)	3.51	Above 2x median
210740_s_at	inositol 1,3,4-triphosphate 5/6 kinase	3.47	Above 2x median
202575_at	cellular retinoic acid binding protein 2	3.47	Above 2x median
37152_at	peroxisome proliferator-activated receptor delta	3.46	Above 2x median
218982_s_at	mitochondrial ribosomal protein S17 / zinc finger protein 713	3.46	Above 2x median
216268_s_at	jagged 1	3.46	Above 2x median
202224_at	v-crk sarcoma virus CT10 oncogene homolog (avian)	3.45	Above 2x median
217943_s_at	MAP7 domain containing 1	3.44	Above 2x median
212355_at	KH and NYN domain containing	3.44	Above 2x median
218131_s_at	GATA zinc finger domain containing 2A	3.42	Above 2x median
203725_at	growth arrest and DNA-damage-inducible, alpha	3.41	Above 2x median
208925_at	claudin domain containing 1	3.35	Above 2x median
201695_s_at	purine nucleoside phosphorylase	3.33	Above 2x median
208824_x_at	cyclin-dependent kinase 16	3.33	Above 2x median
218018_at	pyridoxal (pyridoxine, vitamin B6) kinase	3.33	Above 2x median
209365_s_at	extracellular matrix protein 1	3.29	Above 2x median
201516_at	spermidine synthase	3.27	Above 2x median
221489_s_at	sprouty homolog 4 (Drosophila)	3.18	Above 2x median
209264_s_at	tetraspanin 4	3.18	Above 2x median
204319_s_at	regulator of G-protein signaling 10	3.18	Above 2x median
209457_at	dual specificity phosphatase 5	3.18	Above 2x median
202696_at	oxidative-stress responsive 1	3.18	Above 2x median
204331_s_at	mitochondrial ribosomal protein S12	3.17	Above 2x median
209651_at	transforming growth factor beta 1 induced transcript 1	3.16	Above 2x median
204034_at	ethylmalonic encephalopathy 1	3.16	Above 2x median
202855_s_at	solute carrier family 16, member 3 (monocarboxylic acid transporter 4)	3.16	Above 2x median
221931_s_at	SEH1-like (S. cerevisiae)	3.16	Above 2x median
201481_s_at	phosphorylase, glycogen; brain	3.15	Above 2x median
203740_at	M-phase phosphoprotein 6	3.13	Above 2x median
218732_at	peptidyl-tRNA hydrolase 2	3.11	Above 2x median
212110_at	solute carrier family 39 (zinc transporter), member 14	3.11	Above 2x median
200924_s_at	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	3.11	Above 2x median
217492_s_at	phosphatase and tensin homolog / phosphatase and tensin	3.11	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 10)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
	homolog pseudogene 1		
204334_at	Kruppel-like factor 7 (ubiquitous)	3.08	Above 2x median
210916_s_at	CD44 molecule (Indian blood group)	3.07	Above 2x median
209260_at	stratifin	3.06	Above 2x median
218656_s_at	lipoma HMGIC fusion partner	3.04	Above 2x median
218092_s_at	ArfGAP with FG repeats 1	3.04	Above 2x median
202894_at	EPH receptor B4	3.04	Above 2x median
203085_s_at	transforming growth factor, beta 1	3.03	Above 2x median
214696_at	chromosome 17 open reading frame 91	3.02	Above 2x median
218826_at	solute carrier family 35, member F2	3.02	Above 2x median
209187_at	down-regulator of transcription 1, TBP-binding (negative cofactor 2)	3.01	Above 2x median
206688_s_at	cleavage and polyadenylation specific factor 4, 30kDa	2.98	Above 2x median
219648_at	melanoregulin	2.97	Above 2x median
205361_s_at	prefoldin subunit 4	2.96	Above 2x median
219121_s_at	epithelial splicing regulatory protein 1	2.94	Above 2x median
203279_at	ER degradation enhancer, mannosidase alpha-like 1	2.93	Above 2x median
201195_s_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	2.93	Above 2x median
203256_at	cadherin 3, type 1, P-cadherin (placental)	2.91	Above 2x median
203550_s_at	family with sequence similarity 189, member B	2.89	Above 2x median
221667_s_at	heat shock 22kDa protein 8	2.89	Above 2x median
219933_at	glutaredoxin 2	2.88	Above 2x median
221753_at	slingshot homolog 1 (Drosophila)	2.88	Above 2x median
203002_at	angiomotin like 2	2.86	Above 2x median
213524_s_at	G0/G1switch 2	2.85	Above 2x median
203821_at	heparin-binding EGF-like growth factor	2.81	Above 2x median
207826_s_at	inhibitor of DNA binding 3, dominant negative helix-loop-helix protein	2.80	Above 2x median
202998_s_at	lysyl oxidase-like 2	2.78	Above 2x median
221050_s_at	GTP binding protein 2	2.78	Above 2x median
203405_at	proteasome (prosome, macropain) assembly chaperone 1	2.77	Above 2x median
218145_at	tribbles homolog 3 (Drosophila)	2.75	Above 2x median
218193_s_at	golgi transport 1B	2.73	Above 2x median
204748_at	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	2.73	Above 2x median
203297_s_at	jumonji, AT rich interactive domain 2	2.73	Above 2x median
218046_s_at	mitochondrial ribosomal protein S16	2.71	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 11)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
212612_at	REST corepressor 1	2.71	Above 2x median
209369_at	annexin A3	2.70	Above 2x median
202754_at	R3H domain containing 1	2.69	Above 2x median
204616_at	ubiquitin carboxyl-terminal esterase L3 (ubiquitin thiolesterase)	2.69	Above 2x median
204214_s_at	RAB32, member RAS oncogene family	2.68	Above 2x median
217996_at	pleckstrin homology-like domain, family A, member 1	2.68	Above 2x median
202934_at	hexokinase 2	2.68	Above 2x median
201016_at	eukaryotic translation initiation factor 1A, X-linked	2.67	Above 2x median
207172_s_at	cadherin 11, type 2, OB-cadherin (osteoblast)	2.67	Above 2x median
218401_s_at	zinc finger protein 281	2.66	Above 2x median
212556_at	scribbled homolog (Drosophila)	2.66	Above 2x median
209096_at	ubiquitin-conjugating enzyme E2 variant 2	2.65	Above 2x median
210970_s_at	inhibitor of Bruton agammaglobulinemia tyrosine kinase	2.63	Above 2x median
213457_at	malignant fibrous histiocytoma amplified sequence 1	2.60	Above 2x median
210517_s_at	A kinase (PRKA) anchor protein 12	2.59	Above 2x median
204971_at	cystatin A (stefin A)	2.59	Above 2x median
218669_at	RAP2C, member of RAS oncogene family	2.57	Above 2x median
212845_at	sterile alpha motif domain containing 4A	2.57	Above 2x median
202806_at	drebrin 1	2.55	Above 2x median
202734_at	thyroid hormone receptor interactor 10	2.54	Above 2x median
203373_at	suppressor of cytokine signaling 2	2.53	Above 2x median
213734_at	WD repeat and SOCS box-containing 2	2.53	Above 2x median
212647_at	related RAS viral (r-ras) oncogene homolog	2.52	Above 2x median
205047_s_at	asparagine synthetase (glutamine-hydrolyzing)	2.52	Above 2x median
218051_s_at	5'-nucleotidase domain containing 2	2.50	Above 2x median
202557_at	heat shock protein 70kDa family, member 13	2.48	Above 2x median
219024_at	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1	2.47	Above 2x median
202628_s_at	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	2.47	Above 2x median
210869_s_at	melanoma cell adhesion molecule	2.46	Above 2x median
212590_at	related RAS viral (r-ras) oncogene homolog 2	2.46	Above 2x median
208962_s_at	fatty acid desaturase 1	2.45	Above 2x median
212607_at	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	2.45	Above 2x median
205729_at	oncostatin M receptor	2.43	Above 2x median
37462_i_at	splicing factor 3a, subunit 2, 66kDa	2.43	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 12)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
201984_s_at	epidermal growth factor receptor	2.42	Above 2x median
218995_s_at	endothelin 1	2.41	Above 2x median
213554_s_at	CDV3 homolog (mouse)	2.40	Above 2x median
202779_s_at	ubiquitin-conjugating enzyme E2S	2.40	Above 2x median
217875_s_at	prostate transmembrane protein, androgen induced 1	2.39	Above 2x median
209444_at	RAP1, GTP-GDP dissociation stimulator 1	2.38	Above 2x median
212992_at	AHNAK nucleoprotein 2	2.37	Above 2x median
203585_at	zinc finger protein 185 (LIM domain)	2.37	Above 2x median
203258_at	DR1-associated protein 1 (negative cofactor 2 alpha)	2.35	Above 2x median
222162_s_at	ADAM metallopeptidase with thrombospondin type 1 motif, 1	2.35	Above 2x median
214431_at	guanine monphosphate synthetase	2.34	Above 2x median
204194_at	BTB and CNC homology 1, basic leucine zipper transcription factor 1	2.34	Above 2x median
216088_s_at	proteasome (prosome, macropain) subunit, alpha type, 7	2.34	Above 2x median
201505_at	laminin, beta 1	2.32	Above 2x median
218238_at	GTP binding protein 4	2.32	Above 2x median
201791_s_at	7-dehydrocholesterol reductase	2.31	Above 2x median
218849_s_at	protein phosphatase 1, regulatory (inhibitor) subunit 13 like	2.31	Above 2x median
218499_at	serine/threonine protein kinase MST4	2.26	Above 2x median
203234_at	uridine phosphorylase 1	2.25	Above 2x median
221437_s_at	mitochondrial ribosomal protein S15	2.25	Above 2x median
204977_at	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	2.24	Above 2x median
212527_at	PPPDE peptidase domain containing 2	2.23	Above 2x median
204976_s_at	Alport syndrome, mental retardation, midface hypoplasia and elliptocytosis chromosomal region gene 1	2.23	Above 2x median
202912_at	adrenomedullin	2.22	Above 2x median
201389_at	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	2.21	Above 2x median
201692_at	sigma non-opioid intracellular receptor 1	2.20	Above 2x median
218526_s_at	RAN guanine nucleotide release factor	2.19	Above 2x median
202543_s_at	glia maturation factor, beta	2.19	Above 2x median
217784_at	YKT6 v-SNARE homolog (S. cerevisiae)	2.17	Above 2x median
205463_s_at	platelet-derived growth factor alpha polypeptide	2.17	Above 2x median
209324_s_at	regulator of G-protein signaling 16	2.15	Above 2x median
219155_at	phosphatidylinositol transfer protein, cytoplasmic 1	2.15	Above 2x median
213757_at	Eukaryotic translation initiation factor 5A	2.14	Above 2x median
202158_s_at	CUGBP, Elav-like family member 2	2.13	Above 2x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 13)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
209383_at	DNA-damage-inducible transcript 3	2.11	Above 2x median
201380_at	cartilage associated protein	2.11	Above 2x median
209278_s_at	tissue factor pathway inhibitor 2	2.11	Above 2x median
217437_s_at	transforming, acidic coiled-coil containing protein 1	2.10	Above 2x median
208002_s_at	acyl-CoA thioesterase 7	2.09	Above 2x median
202847_at	phosphoenolpyruvate carboxykinase 2 (mitochondrial)	2.08	Above 2x median
204285_s_at	phorbol-12-myristate-13-acetate-induced protein 1	2.07	Above 2x median
203672_x_at	thiopurine S-methyltransferase	2.06	Above 2x median
218777_at	receptor accessory protein 4	2.05	Above 2x median
218647_s_at	yrdC domain containing (E. coli)	2.05	Above 2x median
207030_s_at	cysteine and glycine-rich protein 2	2.04	Above 2x median
204647_at	homer homolog 3 (Drosophila)	2.02	Above 2x median
204268_at	S100 calcium binding protein A2	2.02	Above 2x median
221539_at	eukaryotic translation initiation factor 4E binding protein 1	2.02	Above 2x median
218796_at	fermitin family member 1	2.01	Above 2x median
203741_s_at	adenylate cyclase 7	1.99	
204255_s_at	vitamin D (1,25- dihydroxyvitamin D3) receptor	1.99	
218451_at	CUB domain containing protein 1	1.97	
203939_at	5'-nucleotidase, ecto (CD73)	1.94	
218247_s_at	mex-3 homolog C (C. elegans)	1.94	
202146_at	interferon-related developmental regulator 1	1.94	
202647_s_at	neuroblastoma RAS viral (v-ras) oncogene homolog	1.94	
218498_s_at	ERO1-like (S. cerevisiae)	1.92	
202314_at	cytochrome P450, family 51, subfamily A, polypeptide 1	1.92	
217266_at	ribosomal protein L15	1.90	
206571_s_at	mitogen-activated protein kinase kinase kinase kinase 4	1.89	
205239_at	amphiregulin	1.88	
201787_at	fibulin 1	1.87	
202131_s_at	RIO kinase 3 (yeast)	1.86	
203693_s_at	E2F transcription factor 3	1.85	
201387_s_at	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	1.84	
202267_at	laminin, gamma 2	1.84	
216620_s_at	Rho guanine nucleotide exchange factor (GEF) 10	1.83	
213062_at	N-terminal asparagine amidase	1.83	
201656_at	integrin, alpha 6	1.83	
219099_at	chromosome 12 open reading frame 5	1.81	
205016_at	transforming growth factor, alpha	1.81	

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 14)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
203695_s_at	deafness, autosomal dominant 5	1.81
218888_s_at	neuropilin (NRP) and tolloid (TLL)-like 2	1.79
207357_s_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 (GalNAc-T10)	1.79
210026_s_at	caspase recruitment domain family, member 10	1.78
204693_at	CDC42 effector protein (Rho GTPase binding) 1	1.77
219856_at	chromosome 1 open reading frame 116	1.76
31845_at	E74-like factor 4 (ets domain transcription factor)	1.76
209308_s_at	BCL2/adenovirus E1B 19kDa interacting protein 2	1.75
204435_at	nucleoporin like 1	1.74
202613_at	CTP synthase	1.71
200690_at	heat shock 70kDa protein 9 (mortalin)	1.71
203074_at	annexin A8 / annexin A8-like 1 / annexin A8-like 2	1.71
216383_at	ribosomal protein L18a /// ribosomal protein L18a pseudogene 3	1.70
219256_s_at	SH3 domain and tetratricopeptide repeats 1	1.69
209283_at	crystallin, alpha B	1.69
202728_s_at	latent transforming growth factor beta binding protein 1	1.69
218038_at	ATP5S-like	1.69
213094_at	G protein-coupled receptor 126	1.68
206884_s_at	sciellin	1.67
202196_s_at	dickkopf homolog 3 (Xenopus laevis)	1.67
204475_at	matrix metallopeptidase 1 (interstitial collagenase)	1.67
203726_s_at	laminin, alpha 3	1.65
209433_s_at	phosphoribosyl pyrophosphate amidotransferase	1.65
206714_at	arachidonate 15-lipoxygenase, type B	1.64
215084_s_at	leucine rich repeat containing 42	1.63
219045_at	ras homolog gene family, member F (in filopodia)	1.63
218828_at	phospholipid scramblase 3	1.62
213010_at	protein kinase C, delta binding protein	1.62
203592_s_at	follistatin-like 3 (secreted glycoprotein)	1.61
205349_at	guanine nucleotide binding protein (G protein), alpha 15 (Gq class)	1.61
203562_at	fasciculation and elongation protein zeta 1 (zygin I)	1.60
203348_s_at	ets variant 5	1.60
203370_s_at	PDZ and LIM domain 7 (enigma)	1.60
204030_s_at	IQ motif containing J-schwannomin interacting protein 1 read-through transcript / schwannomin interacting protein	1.59

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 15)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
	1	
222235_s_at	chondroitin sulfate N-acetylgalactosaminyltransferase 2	1.58
203105_s_at	dynamin 1-like	1.58
219038_at	MORC family CW-type zinc finger 4	1.58
204955_at	sushi-repeat-containing protein, X-linked	1.58
214097_at	ribosomal protein S21	1.56
214075_at	neuron derived neurotrophic factor	1.56
211071_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11	1.55
209453_at	solute carrier family 9 (sodium/hydrogen exchanger), member 1	1.54
218717_s_at	leprecan-like 1	1.54
221039_s_at	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	1.51
219437_s_at	ankyrin repeat domain 11	1.50
218823_s_at	potassium channel tetramerisation domain containing 9	1.47
213139_at	snail homolog 2 (Drosophila)	1.47
212702_s_at	bicaudal D homolog 2 (Drosophila)	1.46
203963_at	carbonic anhydrase XII	1.45
205289_at	bone morphogenetic protein 2	1.45
202701_at	bone morphogenetic protein 1	1.45
202422_s_at	acyl-CoA synthetase long-chain family member 4	1.45
209890_at	tetraspanin 5	1.44
205180_s_at	ADAM metallopeptidase domain 8	1.43
219496_at	ankyrin repeat domain 57	1.43
203180_at	aldehyde dehydrogenase 1 family, member A3	1.43
214168_s_at	tight junction protein 1 (zona occludens 1)	1.43
218156_s_at	TSR1, 20S rRNA accumulation, homolog (S. cerevisiae)	1.43
212290_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1	1.43
202735_at	emopamil binding protein (sterol isomerase)	1.43
209894_at	leptin receptor	1.42
204686_at	insulin receptor substrate 1	1.42
212563_at	block of proliferation 1	1.42
212094_at	paternally expressed 10	1.40
219412_at	RAB38, member RAS oncogene family	1.39
204928_s_at	solute carrier family 10 (sodium/bile acid cotransporter family), member 3	1.39

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 16)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
212657_s_at	interleukin 1 receptor antagonist	1.39
205264_at	CD3e molecule, epsilon associated protein	1.39
209053_s_at	Wolf-Hirschhorn syndrome candidate 1	1.38
213696_s_at	mediator complex subunit 8	1.38
221156_x_at	cell cycle progression 1	1.38
204257_at	fatty acid desaturase 3	1.37
203543_s_at	Kruppel-like factor 9	1.37
213848_at	dual specificity phosphatase 7	1.37
219991_at	solute carrier family 2 (facilitated glucose transporter), member 9	1.37
202218_s_at	fatty acid desaturase 2	1.36
217948_at	family with sequence similarity 127, member B	1.36
215411_s_at	TRAF3 interacting protein 2	1.36
40284_at	forkhead box A2	1.36
201819_at	scavenger receptor class B, member 1	1.35
203851_at	insulin-like growth factor binding protein 6	1.35
213793_s_at	homer homolog 1 (Drosophila)	1.35
202787_s_at	mitogen-activated protein kinase-activated protein kinase 3	1.34
219483_s_at	porcupine homolog (Drosophila)	1.34
221268_s_at	sphingosine-1-phosphate phosphatase 1	1.33
213927_at	mitogen-activated protein kinase kinase kinase 9	1.33
209301_at	carbonic anhydrase II	1.33
205404_at	hydroxysteroid (11-beta) dehydrogenase 1	1.32
206247_at	MHC class I polypeptide-related sequence B	1.32
206170_at	adrenergic, beta-2-, receptor, surface	1.32
218644_at	pleckstrin 2	1.30
209524_at	hepatoma-derived growth factor, related protein 3	1.30
216548_x_at	high-mobility group box pseudogene 1	1.30
204203_at	CCAAT/enhancer binding protein (C/EBP), gamma	1.30
210854_x_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 8	1.29
204222_s_at	GLI pathogenesis-related 1	1.29
217790_s_at	signal sequence receptor, gamma (translocon-associated protein gamma)	1.28
218817_at	signal peptidase complex subunit 3 homolog (S. cerevisiae)	1.28
39402_at	interleukin 1, beta	1.27
202712_s_at	creatine kinase, mitochondrial 1A / creatine kinase,	1.27

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
	mitochondrial 1B	
221027_s_at	phospholipase A2, group XIIA	1.25
204639_at	adenosine deaminase	1.25
219275_at	programmed cell death 5	1.25
206467_x_at	regulator of telomere elongation helicase 1 /// tumor necrosis factor receptor superfamily, member 6b, decoy	1.23
213100_at	unc-5 homolog B (C. elegans)	1.22
209441_at	Rho-related BTB domain containing 2	1.22
209714_s_at	cyclin-dependent kinase inhibitor 3	1.22
217094_s_at	itchy E3 ubiquitin protein ligase homolog (mouse)	1.22
218392_x_at	sideroflexin 1	1.21
220892_s_at	phosphoserine aminotransferase 1	1.21
204058_at	malic enzyme 1, NADP(+)-dependent, cytosolic	1.21
219901_at	FYVE, RhoGEF and PH domain containing 6	1.20
218055_s_at	WD repeat domain 41	1.19
204819_at	FYVE, RhoGEF and PH domain containing 1	1.19
219366_at	apoptosis, caspase activation inhibitor	1.19
219959_at	molybdenum cofactor sulfurase	1.18
218705_s_at	sorting nexin 24	1.17
218077_s_at	zinc finger, DHHC-type containing 3	1.17
216971_s_at	Plectin	1.17
205904_at	MHC class I polypeptide-related sequence A	1.17
210357_s_at	spermine oxidase	1.16
201397_at	phosphoglycerate dehydrogenase	1.16
219892_at	transmembrane 6 superfamily member 1	1.15
213230_at	cerebellar degeneration-related protein 2-like	1.13
203650_at	protein C receptor, endothelial	1.13
218700_s_at	RAB7, member RAS oncogene family-like 1	1.11
37966_at	parvin, beta	1.11
205234_at	solute carrier family 16, member 4 (monocarboxylic acid transporter 5)	1.11
219431_at	Rho GTPase activating protein 10	1.11
203888_at	thrombomodulin	1.11
214866_at	plasminogen activator, urokinase receptor	1.10
206857_s_at	FK506 binding protein 1B, 12.6 kDa	1.10
201181_at	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	1.09
219569_s_at	transmembrane protein 22	1.08

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 18)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
219878_s_at	Kruppel-like factor 13	1.08
210705_s_at	tripartite motif-containing 5	1.06
221803_s_at	nuclear receptor binding factor 2	1.05
220610_s_at	leucine rich repeat (in FLII) interacting protein 2	1.04
219271_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 14 (GalNAc-T14)	1.03
204005_s_at	PRKC, apoptosis, WT1, regulator	1.02
213358_at	KIAA0802	1.02
202570_s_at	discs, large (Drosophila) homolog-associated protein 4	1.02
219390_at	FK506 binding protein 14, 22 kDa	1.02
202067_s_at	low density lipoprotein receptor	1.00
209825_s_at	uridine-cytidine kinase 2	1.00
220789_s_at	transforming growth factor beta regulator 4	1.00
204244_s_at	DBF4 homolog (S. cerevisiae)	0.99
204527_at	myosin VA (heavy chain 12, myoxin)	0.99
212816_s_at	cystathionine-beta-synthase	0.98
202843_at	DnaJ (Hsp40) homolog, subfamily B, member 9	0.98
210041_s_at	phosphoglucomutase 3	0.97
221902_at	G protein-coupled receptor 153	0.97
205618_at	proline rich Gla (G-carboxyglutamic acid) 1	0.97
205579_at	histamine receptor H1	0.97
221009_s_at	angiopoietin-like 4	0.95
218775_s_at	WW and C2 domain containing 2	0.95
208084_at	integrin, beta 6	0.94
219582_at	opioid growth factor receptor-like 1	0.94
220239_at	kelch-like 7 (Drosophila)	0.94
204120_s_at	adenosine kinase	0.94
203889_at	secretogranin V (7B2 protein)	0.93
219058_x_at	tubulointerstitial nephritis antigen-like 1	0.93
209371_s_at	SH3-domain binding protein 2	0.93
220122_at	multiple C2 domains, transmembrane 1	0.93
222173_s_at	TBC1 domain family, member 2	0.93
218772_x_at	transmembrane protein 38B	0.92
204653_at	transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)	0.91
219938_s_at	proline-serine-threonine phosphatase interacting protein 2	0.91
204958_at	polo-like kinase 3	0.91
205567_at	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	0.91

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 19)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
215243_s_at	gap junction protein, beta 3, 31kDa	0.91
219654_at	protein tyrosine phosphatase-like (proline instead of catalytic arginine), member A	0.90
219361_s_at	apoptosis enhancing nuclease	0.89
213012_at	neural precursor cell expressed, developmentally down-regulated 4	0.89
205330_at	meningioma (disrupted in balanced translocation) 1	0.89
204750_s_at	desmocollin 2	0.88
203023_at	NOP16 nucleolar protein homolog (yeast)	0.88
219979_s_at	chromosome 11 open reading frame 73	0.87
210074_at	cathepsin L2	0.87
221779_at	MICAL-like 1	0.86
209946_at	vascular endothelial growth factor C	0.86
204462_s_at	solute carrier family 16, member 2 (monocarboxylic acid transporter 8)	0.86
209589_s_at	EPH receptor B2	0.85
209885_at	ras homolog gene family, member D	0.84
209758_s_at	microfibrillar associated protein 5	0.83
218484_at	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2	0.83
203597_s_at	WW domain binding protein 4 (formin binding protein 21)	0.83
214073_at	cortactin	0.81
217678_at	solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	0.81
212662_at	poliovirus receptor	0.80
218818_at	four and a half LIM domains 3	0.80
219911_s_at	solute carrier organic anion transporter family, member 4A1	0.80
221127_s_at	dickkopf homolog 3 (Xenopus laevis)	0.79
210005_at	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	0.79
206295_at	interleukin 18 (interferon-gamma-inducing factor)	0.78
206482_at	PTK6 protein tyrosine kinase 6	0.78
213338_at	transmembrane protein 158 (gene/pseudogene)	0.78
205421_at	solute carrier family 22 (extraneuronal monoamine transporter), member 3	0.77
213419_at	amyloid beta (A4) precursor protein-binding, family B, member 2	0.77

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
210017_at	mucosa associated lymphoid tissue lymphoma translocation gene 1	0.77
203946_s_at	arginase, type II	0.76
204944_at	protein tyrosine phosphatase, receptor type, G	0.76
209152_s_at	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	0.76
219181_at	lipase, endothelial	0.75
204165_at	WAS protein family, member 1	0.74
217253_at	SH3-domain binding protein 2	0.74
201564_s_at	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpuratus)	0.74
205376_at	inositol polyphosphate-4-phosphatase, type II, 105kDa	0.74
204602_at	dickkopf homolog 1 (Xenopus laevis)	0.73
206034_at	serpin peptidase inhibitor, clade B (ovalbumin), member 8	0.73
217027_x_at	karyopherin (importin) beta 1	0.73
41037_at	TEA domain family member 4	0.72
205083_at	aldehyde oxidase 1	0.71
205865_at	AT rich interactive domain 3A (BRIGHT-like)	0.70
214452_at	branched chain amino-acid transaminase 1, cytosolic	0.70
220658_s_at	aryl hydrocarbon receptor nuclear translocator-like 2	0.69
205195_at	adaptor-related protein complex 1, sigma 1 subunit	0.69
205174_s_at	glutaminyl-peptide cyclotransferase	0.69
206204_at	growth factor receptor-bound protein 14	0.69
213523_at	cyclin E1	0.68
204078_at	leprecan-like 4	0.68
206125_s_at	kallikrein-related peptidase 8	0.67
213680_at	keratin 6B	0.67
204347_at	adenylate kinase 4	0.66
213262_at	spastic ataxia of Charlevoix-Saguenay (sacsin)	0.65
208510_s_at	peroxisome proliferator-activated receptor gamma	0.64
217127_at	cystathionase (cystathionine gamma-lyase)	0.64
215812_s_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 10 (pseudogene) / solute carrier family 6 (neurotransmitter transporter, creatine), member 8	0.64
219921_s_at	dedicator of cytokinesis 5	0.63
217999_s_at	pleckstrin homology-like domain, family A, member 1	0.63
219973_at	arylsulfatase family, member J	0.62
211031_s_at	CAP-GLY domain containing linker protein 2	0.62

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 21)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
213112_s_at	sequestosome 1	0.62
221078_s_at	coiled-coil domain containing 88A	0.61
34187_at	RNA binding motif, single stranded interacting protein 2	0.61
214599_at	involucrin	0.61
213992_at	collagen, type IV, alpha 6	0.60
219631_at	low density lipoprotein receptor-related protein 12	0.60
213847_at	peripherin	0.59
220255_at	Fanconi anemia, complementation group E	0.59
215082_at	ELOVL family member 5, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	0.58
206173_x_at	GA binding protein transcription factor, beta subunit 1	0.58
215167_at	mediator complex subunit 14	0.58
206412_at	fer (fps/fes related) tyrosine kinase	0.58
204733_at	kallikrein-related peptidase 6	0.58
204920_at	carbamoyl-phosphate synthase 1, mitochondrial	0.58
205341_at	EH-domain containing 2	0.57
207020_at	heat shock transcription factor 2 binding protein	0.57
219522_at	four jointed box 1 (Drosophila)	0.57
218537_at	host cell factor C1 regulator 1 (XPO1 dependent)	0.55
200951_s_at	cyclin D2	0.55
205503_at	protein tyrosine phosphatase, non-receptor type 14	0.54
222242_s_at	kallikrein-related peptidase 5	0.54
219944_at	CAP-GLY domain containing linker protein family, member 4	0.54
209595_at	general transcription factor IIF, polypeptide 2, 30kDa	0.53
217150_s_at	neurofibromin 2 (merlin)	0.53
205394_at	CHK1 checkpoint homolog (S. pombe)	0.53
207463_x_at	protease, serine, 3	0.53
204136_at	collagen, type VII, alpha 1	0.52
203071_at	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B	0.52
214095_at	serine hydroxymethyltransferase 2 (mitochondrial)	0.51
206561_s_at	aldo-keto reductase family 1, member B10 (aldose reductase)	0.51
219225_at	piggyBac transposable element derived 5	0.50
218744_s_at	protein kinase C and casein kinase substrate in neurons 3	0.50
205402_x_at	protease, serine, 2 (trypsin 2)	0.50
210015_s_at	microtubule-associated protein 2	0.50

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 22)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
213984_at	PDS5, regulator of cohesion maintenance, homolog A (S. cerevisiae)	0.49	Below 0.5x median
208025_s_at	high mobility group AT-hook 2	0.49	Below 0.5x median
204854_at	leprecan-like 2	0.49	Below 0.5x median
209544_at	receptor-interacting serine-threonine kinase 2	0.49	Below 0.5x median
219270_at	ChaC, cation transport regulator homolog 1 (E. coli)	0.48	Below 0.5x median
206023_at	neuromedin U	0.48	Below 0.5x median
205157_s_at	keratin 17	0.48	Below 0.5x median
211905_s_at	integrin, beta 4	0.48	Below 0.5x median
216470_x_at	protease, serine, 1 (trypsin 1) / protease, serine, 2 (trypsin 2) / protease, serine, 3	0.48	Below 0.5x median
205064_at	small proline-rich protein 1B	0.46	Below 0.5x median
215918_s_at	spectrin, beta, non-erythrocytic 1	0.46	Below 0.5x median
206686_at	pyruvate dehydrogenase kinase, isozyme 1	0.44	Below 0.5x median
219702_at	placenta-specific 1	0.44	Below 0.5x median
214967_at	CDNA FLJ12055 fis, clone HEMBB1002049	0.43	Below 0.5x median
201820_at	keratin 5	0.42	Below 0.5x median
222269_at	apolipoprotein O-like	0.42	Below 0.5x median
217559_at	ribosomal protein L10-like	0.42	Below 0.5x median
209646_x_at	aldehyde dehydrogenase 1 family, member B1	0.42	Below 0.5x median
205796_at	t-complex 11 (mouse)-like 1	0.41	Below 0.5x median
203967_at	cell division cycle 6 homolog (S. cerevisiae)	0.41	Below 0.5x median
209652_s_at	placental growth factor	0.41	Below 0.5x median
209125_at	keratin 6A	0.41	Below 0.5x median
214126_at	mitochondrial carrier triple repeat 1	0.40	Below 0.5x median
204420_at	FOS-like antigen 1	0.39	Below 0.5x median
207382_at	tumor protein p63	0.39	Below 0.5x median
201249_at	solute carrier family 2 (facilitated glucose transporter), member 1	0.39	Below 0.5x median
207526_s_at	interleukin 1 receptor-like 1	0.39	Below 0.5x median
215199_at	caldesmon 1	0.38	Below 0.5x median
219998_at	galectin-related protein	0.38	Below 0.5x median
222309_at	hypothetical LOC100506935	0.38	Below 0.5x median
219987_at	Similar to Uncharacterized protein LP9056	0.38	Below 0.5x median
203736_s_at	PTPRF interacting protein, binding protein 1 (liprin beta 1)	0.37	Below 0.5x median
204455_at	dystonin	0.37	Below 0.5x median
207345_at	follistatin	0.36	Below 0.5x median
219936_s_at	G protein-coupled receptor 87	0.35	Below 0.5x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 23)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
209631_s_at	G protein-coupled receptor 37 (endothelin receptor type B-like)	0.35	Below 0.5x median
208425_s_at	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2	0.34	Below 0.5x median
205691_at	synaptogyrin 3	0.34	Below 0.5x median
214297_at	chondroitin sulfate proteoglycan 4	0.34	Below 0.5x median
205429_s_at	membrane protein, palmitoylated 6 (MAGUK p55 subfamily member 6)	0.34	Below 0.5x median
211846_s_at	poliovirus receptor-related 1 (herpesvirus entry mediator C)	0.33	Below 0.5x median
220638_s_at	Cas-Br-M (murine) ecotropic retroviral transforming sequence c	0.33	Below 0.5x median
208260_at	arginine vasopressin receptor 1B	0.33	Below 0.5x median
204952_at	LY6/PLAUR domain containing 3	0.33	Below 0.5x median
203184_at	fibrillin 2	0.32	Below 0.5x median
205899_at	cyclin A1	0.32	Below 0.5x median
205595_at	desmoglein 3	0.32	Below 0.5x median
208511_at	pituitary tumor-transforming 3 (pseudogene)	0.31	Below 0.5x median
206032_at	desmocollin 3	0.31	Below 0.5x median
202684_s_at	RNA (guanine-7-) methyltransferase	0.31	Below 0.5x median
206371_at	folate receptor 3 (gamma)	0.31	Below 0.5x median
209877_at	synuclein, gamma (breast cancer-specific protein 1)	0.31	Below 0.5x median
207717_s_at	plakophilin 2	0.30	Below 0.5x median
219555_s_at	centromere protein N	0.30	Below 0.5x median
221414_s_at	defensin, beta 126	0.30	Below 0.5x median
219926_at	popeye domain containing 3	0.30	Below 0.5x median
210113_s_at	NLR family, pyrin domain containing 1	0.29	Below 0.5x median
220058_at	chromosome 17 open reading frame 39	0.29	Below 0.5x median
217553_at	STEAP family protein MGC87042	0.28	Below 0.5x median
206832_s_at	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	0.28	Below 0.5x median
220198_s_at	eukaryotic translation initiation factor 5A2	0.28	Below 0.5x median
215058_at	DENN/MADD domain containing 5B	0.27	Below 0.5x median
206080_at	phospholipase C, eta 2	0.27	Below 0.5x median
215465_at	ATP-binding cassette, sub-family A (ABC1), member 12	0.26	Below 0.5x median
205778_at	kallikrein-related peptidase 7	0.26	Below 0.5x median
203438_at	stanniocalcin 2	0.26	Below 0.5x median
210138_at	regulator of G-protein signaling 20	0.26	Below 0.5x median
219909_at	matrix metallopeptidase 28	0.26	Below 0.5x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 24)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
214240_at	galanin prepropeptide	0.26	Below 0.5x median
204560_at	FK506 binding protein 5	0.25	Below 0.5x median
205122_at	transmembrane protein with EGF-like and two follistatin-like domains 1	0.25	Below 0.5x median
215074_at	myosin IB	0.25	Below 0.5x median
221994_at	PDZ and LIM domain 5	0.25	Below 0.5x median
204614_at	serpin peptidase inhibitor, clade B (ovalbumin), member 2	0.24	Below 0.5x median
220484_at	mucolipin 3	0.24	Below 0.5x median
209792_s_at	kallikrein-related peptidase 10	0.24	Below 0.5x median
209909_s_at	transforming growth factor, beta 2	0.24	Below 0.5x median
219247_s_at	zinc finger, DHHC-type containing 14	0.23	Below 0.5x median
206091_at	matrilin 3	0.23	Below 0.5x median
210233_at	interleukin 1 receptor accessory protein	0.23	Below 0.5x median
205827_at	cholecystokinin	0.22	Below 0.5x median
214078_at	p21 protein (Cdc42/Rac)-activated kinase 3	0.22	Below 0.5x median
220431_at	transmembrane protease, serine 11E	0.22	Below 0.5x median
210248_at	wingless-type MMTV integration site family, member 7A	0.21	Below 0.5x median
220685_at	family with sequence similarity 120C	0.21	Below 0.5x median
210286_s_at	solute carrier family 4, sodium bicarbonate cotransporter, member 7	0.21	Below 0.5x median
206172_at	interleukin 13 receptor, alpha 2	0.21	Below 0.5x median
218602_s_at	HAUS augmin-like complex, subunit 6	0.21	Below 0.5x median
208231_at	neuregulin 1	0.20	Below 0.5x median
205830_at	calmegin	0.20	Below 0.5x median
205767_at	epiregulin	0.20	Below 0.5x median
220362_at	psoriasis susceptibility 1 candidate 1	0.20	Below 0.5x median
206581_at	basonuclin 1	0.20	Below 0.5x median
204422_s_at	fibroblast growth factor 2 (basic)	0.19	Below 0.5x median
204818_at	hydroxysteroid (17-beta) dehydrogenase 2	0.19	Below 0.5x median
218964_at	AT rich interactive domain 3B (BRIGHT-like)	0.18	Below 0.5x median
220102_at	forkhead box L2	0.18	Below 0.5x median
205493_s_at	dihydropyrimidinase-like 4	0.18	Below 0.5x median
214521_at	hairy and enhancer of split 2 (Drosophila)	0.17	Below 0.5x median
220214_at	zinc finger protein 215	0.17	Below 0.5x median
210654_at	tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain	0.17	Below 0.5x median
210355_at	parathyroid hormone-like hormone	0.16	Below 0.5x median
220672_at	protein phosphatase 4, regulatory subunit 4	0.16	Below 0.5x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 25)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
222265_at	tensin 4	0.16	Below 0.5x median
212821_at	pleckstrin homology domain containing, family G (with RhoGef domain) member 3	0.15	Below 0.5x median
219550_at	roundabout, axon guidance receptor, homolog 3 (Drosophila)	0.15	Below 0.5x median
209800_at	keratin 16	0.15	Below 0.5x median
219855_at	nudix (nucleoside diphosphate linked moiety X)-type motif 11	0.14	Below 0.5x median
206307_s_at	forkhead box D1	0.14	Below 0.5x median
210244_at	cathelicidin antimicrobial peptide	0.13	Below 0.5x median
207017_at	RAB27B, member RAS oncogene family	0.13	Below 0.5x median
204726_at	cadherin 13, H-cadherin (heart)	0.13	Below 0.5x median
210587_at	inhibin, beta E	0.12	Below 0.5x median
202235_at	solute carrier family 16, member 1 (monocarboxylic acid transporter 1)	0.12	Below 0.5x median
221291_at	UL16 binding protein 2	0.12	Below 0.5x median
216564_at	similar to Argininosuccinate synthase (Citrulline--aspartate ligase)	0.12	Below 0.5x median
218990_s_at	small proline-rich protein 3	0.12	Below 0.5x median
204411_at	kinesin family member 21B	0.12	Below 0.5x median
203331_s_at	inositol polyphosphate-5-phosphatase, 145kDa	0.12	Below 0.5x median
210237_at	Artemin	0.12	Below 0.5x median
204760_s_at	nuclear receptor subfamily 1, group D, member 1 /// thyroid hormone receptor, alpha (erythroblastic leukemia viral (v-erb-a) oncogene homolog, avian)	0.11	Below 0.5x median
213865_at	discoidin, CUB and LCCL domain containing 2	0.11	Below 0.5x median
204855_at	serpin peptidase inhibitor, clade B (ovalbumin), member 5	0.10	Below 0.5x median
206400_at	lectin, galactoside-binding, soluble, 7 / lectin, galactoside-binding, soluble, 7B	0.10	Below 0.5x median
206569_at	interleukin 24	0.10	Below 0.5x median
202755_s_at	glypican 1	0.09	Below 0.5x median
219930_at	Kruppel-like factor 8	0.09	Below 0.5x median
213201_s_at	troponin T type 1 (skeletal, slow)	0.09	Below 0.5x median
206354_at	solute carrier organic anion transporter family, member 1B3	0.08	Below 0.5x median
208539_x_at	small proline-rich protein 2B	0.08	Below 0.5x median
205460_at	neuronal PAS domain protein 2	0.08	Below 0.5x median
220289_s_at	absent in melanoma 1-like	0.08	Below 0.5x median
204636_at	collagen, type XVII, alpha 1	0.07	Below 0.5x median

Supplemental Gene List I. Airway Basal Cell Signature Genes Analyzed (cont. page 26)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
206122_at	SRY (sex determining region Y)-box 15	0.07	Below 0.5x median
220620_at	cysteine-rich C-terminal 1	0.06	Below 0.5x median
221784_at	widely interspaced zinc finger motifs	0.06	Below 0.5x median
219532_at	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4	0.06	Below 0.5x median
206877_at	MAX dimerization protein 1	0.06	Below 0.5x median
206421_s_at	serpin peptidase inhibitor, clade B (ovalbumin), member 7	0.05	Below 0.5x median
221854_at	plakophilin 1 (ectodermal dysplasia/skin fragility syndrome)	0.05	Below 0.5x median
207302_at	sarcoglycan, gamma (35kDa dystrophin-associated glycoprotein)	0.05	Below 0.5x median
214595_at	potassium voltage-gated channel, subfamily G, member 1	0.05	Below 0.5x median
213796_at	small proline-rich protein 1A	0.05	Below 0.5x median
207080_s_at	peptide YY	0.05	Below 0.5x median
219554_at	Rh family, C glycoprotein	0.04	Below 0.5x median
215729_s_at	vestigial like 1 (Drosophila)	0.04	Below 0.5x median
219601_s_at	chromosome 10 open reading frame 12	0.04	Below 0.5x median
206008_at	transglutaminase 1 (K polypeptide epidermal type I, protein-glutamine-gamma-glutamyltransferase)	0.04	Below 0.5x median
222383_s_at	arachidonate lipoxygenase 3	0.04	Below 0.5x median
206969_at	keratin 34	0.03	Below 0.5x median
205138_s_at	uronyl-2-sulfotransferase	0.03	Below 0.5x median
206866_at	cadherin 4, type 1, R-cadherin (retinal)	0.03	Below 0.5x median
220664_at	small proline-rich protein 2C (pseudogene)	0.03	Below 0.5x median
206376_at	solute carrier family 6 (neutral amino acid transporter), member 15	0.02	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
208306_x_at	Major histocompatibility complex, class II, DR beta 3	90.73	Above 2x median
208894_at	major histocompatibility complex, class II, DR alpha	64.23	Above 2x median
217757_at	alpha-2-macroglobulin	45.84	Above 2x median
201137_s_at	major histocompatibility complex, class II, DP beta 1	29.14	Above 2x median
217767_at	complement component 3	28.17	Above 2x median
211991_s_at	major histocompatibility complex, class II, DP alpha 1	25.16	Above 2x median
213975_s_at	lysozyme (renal amyloidosis) ; leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 1	23.65	Above 2x median
202291_s_at	matrix Gla protein	23.48	Above 2x median
201288_at	Rho GDP dissociation inhibitor (GDI) beta	22.68	Above 2x median
203892_at	WAP four-disulfide core domain 2	22.46	Above 2x median
202833_s_at	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	21.44	Above 2x median
218211_s_at	melanophilin	20.57	Above 2x median
204124_at	solute carrier family 34 (sodium phosphate), member 2	19.17	Above 2x median
202017_at	epoxide hydrolase 1, microsomal (xenobiotic)	18.82	Above 2x median
202411_at	interferon, alpha-inducible protein 27	18.17	Above 2x median
217478_s_at	major histocompatibility complex, class II, DM alpha	17.85	Above 2x median
209118_s_at	tubulin, alpha 3	17.46	Above 2x median
200795_at	SPARC-like 1 (mast9, hevin)	16.43	Above 2x median
208998_at	uncoupling protein 2 (mitochondrial, proton carrier)	15.61	Above 2x median
204437_s_at	folate receptor 1 (adult)	15.38	Above 2x median
211959_at	insulin-like growth factor binding protein 5	11.93	Above 2x median
203416_at	CD53 antigen	11.46	Above 2x median
214433_s_at	selenium binding protein 1 ; selenium binding protein 1	11.38	Above 2x median
217414_x_at	hemoglobin, alpha 2	11.26	Above 2x median
212950_at	G protein-coupled receptor 116	10.82	Above 2x median
220532_s_at	LR8 protein	10.24	Above 2x median
209116_x_at	hemoglobin, beta ; hemoglobin, beta	9.23	Above 2x median
204070_at	retinoic acid receptor responder (tazarotene induced) 3	8.94	Above 2x median
217989_at	dehydrogenase/reductase (SDR family) member 8	8.42	Above 2x median
208791_at	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)	8.13	Above 2x median
39249_at	aquaporin 3	8.06	Above 2x median
207761_s_at	DKFZP586A0522 protein	7.63	Above 2x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 2)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
203799_at	CD302 antigen	7.25	Above 2x median
204174_at	arachidonate 5-lipoxygenase-activating protein	6.98	Above 2x median
204846_at	ceruloplasmin (ferroxidase)	6.97	Above 2x median
201641_at	bone marrow stromal cell antigen 2	6.79	Above 2x median
205081_at	cysteine-rich protein 1 (intestinal)	6.56	Above 2x median
209392_at	ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)	6.38	Above 2x median
203854_at	I factor (complement)	6.25	Above 2x median
212588_at	protein tyrosine phosphatase, receptor type, C	5.87	Above 2x median
203758_at	cathepsin O	5.70	Above 2x median
209568_s_at	ral guanine nucleotide dissociation stimulator-like 1	5.62	Above 2x median
208885_at	lymphocyte cytosolic protein 1 (L-plastin)	5.51	Above 2x median
219666_at	membrane-spanning 4-domains, subfamily A, member 6A	5.37	Above 2x median
202086_at	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) ; myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	5.37	Above 2x median
222125_s_at	hypoxia-inducible factor prolyl 4-hydroxylase	5.27	Above 2x median
209108_at	tetraspanin 6	5.26	Above 2x median
205890_s_at	ubiquitin D	5.26	Above 2x median
213059_at	cAMP responsive element binding protein 3-like 1	5.24	Above 2x median
210096_at	cytochrome P450, family 4, subfamily B, polypeptide 1	5.22	Above 2x median
203824_at	tetraspanin 8	5.13	Above 2x median
203915_at	chemokine (C-X-C motif) ligand 9	4.75	Above 2x median
212353_at	sulfatase 1	4.58	Above 2x median
218454_at	hypothetical protein FLJ22662	4.53	Above 2x median
204646_at	dihydropyrimidine dehydrogenase	4.51	Above 2x median
202357_s_at	B-factor, properdin	4.41	Above 2x median
205798_at	interleukin 7 receptor	4.30	Above 2x median
205725_at	secretoglobin, family 1A, member 1 (uteroglobin)	3.95	Above 2x median
203953_s_at	claudin 3	3.82	Above 2x median
210272_at	cytochrome P450, family 2, subfamily B, polypeptide 7 pseudogene 1	3.79	Above 2x median
211596_s_at	leucine-rich repeats and immunoglobulin-like domains 1	3.74	Above 2x median
204160_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)	3.53	Above 2x median
205313_at	transcription factor 2, hepatic; LF-B3; variant hepatic nuclear factor	3.53	Above 2x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 3)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
209355_s_at	phosphatidic acid phosphatase type 2B	3.49	Above 2x median
209616_s_at	carboxylesterase 1 (monocyte/macrophage serine esterase 1)	3.47	Above 2x median
218729_at	latexin	3.45	Above 2x median
206385_s_at	ankyrin 3, node of Ranvier (ankyrin G)	3.44	Above 2x median
214453_s_at	interferon-induced protein 44	3.43	Above 2x median
212224_at	aldehyde dehydrogenase 1 family, member A1	3.42	Above 2x median
211796_s_at	T cell receptor beta variable 21-1 ; T cell receptor beta variable 19 ; T cell receptor beta variable 5-4 ; T cell receptor beta variable 3-1 ; T cell receptor beta constant 1	3.40	Above 2x median
205831_at	CD2 antigen (p50), sheep red blood cell receptor	3.37	Above 2x median
34210_at	CD52 antigen (CAMPATH-1 antigen)	3.36	Above 2x median
218901_at	phospholipid scramblase 4	3.36	Above 2x median
210972_x_at	T cell receptor alpha locus ; T cell receptor delta variable 2 ; T cell receptor alpha variable 20 ; T cell receptor alpha joining 17 ; T cell receptor alpha constant	3.36	Above 2x median
211548_s_at	hydroxyprostaglandin dehydrogenase 15-(NAD)	3.31	Above 2x median
205776_at	flavin containing monooxygenase 5	3.29	Above 2x median
203547_at	CD4 antigen (p55)	3.24	Above 2x median
205489_at	crystallin, mu	3.23	Above 2x median
205206_at	Kallmann syndrome 1 sequence	3.20	Above 2x median
205597_at	solute carrier family 44, member 4	3.18	Above 2x median
203088_at	fibulin 5	3.13	Above 2x median
213539_at	CD3D antigen, delta polypeptide (TiT3 complex)	3.09	Above 2x median
207808_s_at	protein S (alpha)	3.04	Above 2x median
202743_at	phosphoinositide-3-kinase, regulatory subunit 3 (p55, gamma)	3.03	Above 2x median
221872_at	retinoic acid receptor responder (tazarotene induced) 1	2.97	Above 2x median
209670_at	T cell receptor alpha constant	2.97	Above 2x median
212148_at	Pre-B-cell leukemia transcription factor 1	2.96	Above 2x median
203561_at	Fc fragment of IgG, low affinity IIa, receptor (CD32)	2.95	Above 2x median
202947_s_at	glycophorin C (Gerbich blood group)	2.93	Above 2x median
211742_s_at	ecotropic viral integration site 2B	2.92	Above 2x median
221698_s_at	C-type lectin domain family 7, member A	2.90	Above 2x median
209447_at	spectrin repeat containing, nuclear envelope 1	2.89	Above 2x median
218805_at	GTPase, IMAP family member 5	2.89	Above 2x median
219014_at	placenta-specific 8	2.83	Above 2x median
218885_s_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-	2.78	Above 2x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 4)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
	acetylgalactosaminyltransferase 12 (GalNAc-T12)		
208498_s_at	amylase, alpha 1A; salivary ; amylase, alpha 1B; salivary ; amylase, alpha 1C; salivary ; amylase, alpha 2A; pancreatic ; amylase, alpha 2B; pancreatic	2.73	Above 2x median
213566_at	ribonuclease, RNase A family, k6	2.58	Above 2x median
203153_at	interferon-induced protein with tetratricopeptide repeats 1 ; interferon-induced protein with tetratricopeptide repeats 1	2.49	Above 2x median
213106_at	ATPase, aminophospholipid transporter (APLT), Class I, type 8A, member 1	2.44	Above 2x median
202018_s_at	lactotransferrin	2.42	Above 2x median
205158_at	ribonuclease, RNase A family, 4	2.41	Above 2x median
204502_at	SAM domain and HD domain 1	2.32	Above 2x median
208146_s_at	carboxypeptidase, vitellogenic-like ; carboxypeptidase, vitellogenic-like	2.29	Above 2x median
204774_at	ecotropic viral integration site 2A	2.24	Above 2x median
205668_at	lymphocyte antigen 75	2.21	Above 2x median
204103_at	chemokine (C-C motif) ligand 4	2.19	Above 2x median
204972_at	2'-5'-oligoadenylate synthetase 2, 69/71kDa	2.16	Above 2x median
218584_at	hypothetical protein FLJ21127	2.16	Above 2x median
205392_s_at	chemokine (C-C motif) ligand 14 ; chemokine (C-C motif) ligand 15	2.14	Above 2x median
213733_at	myosin IF	2.10	Above 2x median
220330_s_at	SAM domain, SH3 domain and nuclear localisation signals, 1	2.10	Above 2x median
203908_at	solute carrier family 4, sodium bicarbonate cotransporter, member 4	2.08	Above 2x median
221087_s_at	apolipoprotein L, 3	2.06	Above 2x median
221530_s_at	basic helix-loop-helix domain containing, class B, 3	2.06	Above 2x median
209539_at	Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6	2.03	Above 2x median
217109_at	mucin 4, tracheobronchial	2.02	Above 2x median
214719_at	hypothetical protein LOC283537	2.02	Above 2x median
204220_at	glia maturation factor, gamma	2.01	Above 2x median
207222_at	phospholipase A2, group X	1.99	
204576_s_at	clusterin associated protein 1	1.99	
203236_s_at	lectin, galactoside-binding, soluble, 9 (galectin 9)	1.97	
214440_at	N-acetyltransferase 1 (arylamine N-acetyltransferase)	1.94	
208920_at	sorcin	1.94	
210166_at	toll-like receptor 5	1.93	

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 5)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
204703_at	tetratricopeptide repeat domain 10	1.90
213488_at	sushi, nidogen and EGF-like domains 1	1.88
205997_at	ADAM metallopeptidase domain 28	1.87
209200_at	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C)	1.87
205364_at	acyl-Coenzyme A oxidase 2, branched chain	1.87
205258_at	inhibin, beta B (activin AB beta polypeptide)	1.87
205590_at	RAS guanyl releasing protein 1 (calcium and DAG-regulated)	1.87
207191_s_at	immunoglobulin superfamily containing leucine-rich repeat	1.84
206932_at	cholesterol 25-hydroxylase	1.83
209606_at	pleckstrin homology, Sec7 and coiled-coil domains, binding protein	1.81
218437_s_at	leucine zipper transcription factor-like 1	1.80
204788_s_at	protoporphyrinogen oxidase	1.80
205640_at	aldehyde dehydrogenase 3 family, member B1	1.79
205114_s_at	chemokine (C-C motif) ligand 3 ; chemokine (C-C motif) ligand 3-like 1	1.79
205464_at	sodium channel, nonvoltage-gated 1, beta (Liddle syndrome)	1.78
212603_at	mitochondrial ribosomal protein S31	1.73
203523_at	lymphocyte-specific protein 1	1.73
205692_s_at	CD38 antigen (p45)	1.72
205316_at	Solute carrier family 15 (H+/peptide transporter), member 2	1.72
203151_at	microtubule-associated protein 1A	1.72
206991_s_at	chemokine (C-C motif) receptor 5	1.69
211902_x_at	T cell receptor alpha locus	1.69
203708_at	phosphodiesterase 4B, cAMP-specific (phosphodiesterase E4 dunce homolog, Drosophila)	1.69
218599_at	REC8-like 1 (yeast)	1.68
211685_s_at	neurocalcin delta	1.64
209504_s_at	pleckstrin homology domain containing, family B (evectins) member 1	1.63
206765_at	potassium inwardly-rectifying channel, subfamily J, member 2	1.63
213317_at	Chloride intracellular channel 5	1.62
221188_s_at	cell death-inducing DFFA-like effector b	1.62

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 6)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
204179_at	myoglobin	1.62
52975_at	chromosome 9 open reading frame 28	1.62
203413_at	NEL-like 2 (chicken) ; NEL-like 2 (chicken)	1.60
220173_at	chromosome 14 open reading frame 45	1.59
219243_at	GTPase, IMAP family member 4	1.58
204202_at	IQ motif containing E	1.56
203924_at	glutathione S-transferase A1	1.55
204964_s_at	sarcospan (Kras oncogene-associated gene)	1.53
218999_at	hypothetical protein FLJ11000	1.51
203474_at	IQ motif containing GTPase activating protein 2	1.51
203979_at	cytochrome P450, family 27, subfamily A, polypeptide 1	1.51
221558_s_at	lymphoid enhancer-binding factor 1	1.50
220751_s_at	chromosome 5 open reading frame 4	1.50
219602_s_at	family with sequence similarity 38, member B	1.49
216920_s_at	T cell receptor gamma constant 2 ; T cell receptor gamma variable 9 ; similar to T-cell receptor gamma chain C region PT-gamma-1/2 ; similar to T-cell receptor gamma chain V region PT-gamma-1/2 precursor ; TCR gamma alternate reading frame protein	1.48
205758_at	CD8 antigen, alpha polypeptide (p32)	1.45
215783_s_at	alkaline phosphatase, liver/bone/kidney	1.45
211003_x_at	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)	1.45
211581_x_at	leukocyte specific transcript 1	1.44
219542_at	NIMA (never in mitosis gene a)- related kinase 11	1.42
204304_s_at	prominin 1	1.42
218820_at	chromosome 14 open reading frame 132	1.41
220072_at	centrosome and spindle pole associated protein 1	1.38
219478_at	WAP four-disulfide core domain 1	1.38
219182_at	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 5	1.36
203951_at	calponin 1, basic, smooth muscle	1.34
212415_at	septin 6	1.33
203687_at	chemokine (C-X3-C motif) ligand 1	1.33
205894_at	arylsulfatase E (chondrodysplasia punctata 1)	1.33
213186_at	zinc finger DAZ interacting protein 3	1.32
220306_at	family with sequence similarity 46, member C	1.32
222304_x_at	olfactory receptor, family 7, subfamily E, member 47 pseudogene	1.31

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 7)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
38241_at	butyrophilin, subfamily 3, member A3	1.31
205141_at	angiogenin, ribonuclease, RNase A family, 5 ; ribonuclease, RNase A family, 4	1.30
214745_at	phospholipase C-like 3	1.30
213416_at	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	1.29
204041_at	monoamine oxidase B	1.27
220108_at	guanine nucleotide binding protein (G protein), alpha 14	1.25
219799_s_at	dehydrogenase/reductase (SDR family) member 9	1.25
203680_at	protein kinase, cAMP-dependent, regulatory, type II, beta	1.25
211471_s_at	RAB36, member RAS oncogene family	1.24
213784_at	RAB, member of RAS oncogene family-like 4	1.24
204783_at	Myeloid leukemia factor 1	1.22
218218_at	DIP13 beta	1.22
218789_s_at	hypothetical protein FLJ20010	1.20
213432_at	mucin 5, subtype B, tracheobronchial	1.20
220998_s_at	unc-93 homolog B1 (C. elegans) ; unc-93 homolog B1 (C. elegans)	1.20
212805_at	KIAA0367	1.20
205406_s_at	sperm autoantigenic protein 17	1.19
203716_s_at	dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2)	1.18
205328_at	claudin 10	1.18
205488_at	granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3) ; granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)	1.17
213429_at	CDNA FLJ26539 fis, clone KDN09310	1.17
207069_s_at	SMAD, mothers against DPP homolog 6 (Drosophila)	1.17
212613_at	butyrophilin, subfamily 3, member A2	1.16
203632_s_at	G protein-coupled receptor, family C, group 5, member B	1.16
205698_s_at	mitogen-activated protein kinase kinase 6	1.15
215559_at	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	1.15
220393_at	glutamate-ammonia ligase (glutamine synthetase) domain containing 1	1.14
221879_at	calmodulin-like 4	1.14
205821_at	killer cell lectin-like receptor subfamily K, member 1	1.13
205269_at	lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa)	1.12

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 8)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
219279_at	dedicator of cytokinesis 10	1.10
219543_at	MAWD binding protein	1.10
219777_at	GTPase, IMAP family member 6	1.09
209869_at	adrenergic, alpha-2A-, receptor ; adrenergic, alpha-2A-, receptor	1.08
219684_at	28kD interferon responsive protein	1.06
209737_at	membrane associated guanylate kinase, WW and PDZ domain containing 2	1.06
212980_at	Ubiquitin specific peptidase 34	1.05
205384_at	FXYD domain containing ion transport regulator 1 (phospholemman)	1.05
207651_at	G protein-coupled receptor 171	1.05
205130_at	renal tumor antigen	1.04
221241_s_at	BCL2-like 14 (apoptosis facilitator)	1.04
219734_at	SID1 transmembrane family, member 1	1.04
212686_at	protein phosphatase 1H (PP2C domain containing)	1.03
212886_at	DKFZP434C171 protein	1.02
219870_at	activating transcription factor 7 interacting protein 2	1.01
218272_at	hypothetical protein FLJ20699	1.01
221187_s_at	hypothetical protein FLJ22688	0.99
207330_at	pregnancy-zone protein	0.98
206315_at	cytokine receptor-like factor 1	0.97
206540_at	galactosidase, beta 1-like	0.97
213385_at	Chimerin (chimaerin) 2	0.97
206197_at	non-metastatic cells 5, protein expressed in (nucleoside-diphosphate kinase)	0.97
213761_at	Mdm4, transformed 3T3 cell double minute 1, p53 binding protein (mouse)	0.97
204890_s_at	lymphocyte-specific protein tyrosine kinase	0.96
220500_s_at	RAB, member of RAS oncogene family-like 2B ; RAB, member of RAS oncogene family-like 2A	0.96
213228_at	phosphodiesterase 8B	0.96
57540_at	ribokinase	0.96
207522_s_at	ATPase, Ca++ transporting, ubiquitous	0.93
218532_s_at	hypothetical protein FLJ20152	0.93
212336_at	erythrocyte membrane protein band 4.1-like 1	0.92
207490_at	tubulin, alpha 4	0.92
204591_at	cell adhesion molecule with homology to L1CAM (close	0.91

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 9)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
	homolog of L1)	
214467_at	G protein-coupled receptor 65	0.89
219396_s_at	nei endonuclease VIII-like 1 (E. coli)	0.88
213888_s_at	TRAF3 interacting protein 3	0.87
215239_x_at	zinc finger protein 273	0.87
220177_s_at	transmembrane protease, serine 3	0.86
204798_at	v-myb myeloblastosis viral oncogene homolog (avian)	0.86
219687_at	hedgehog acyltransferase	0.86
205381_at	leucine rich repeat containing 17	0.85
209460_at	4-aminobutyrate aminotransferase	0.85
206496_at	flavin containing monooxygenase 3	0.85
213839_at	KIAA0500 protein	0.85
204187_at	guanosine monophosphate reductase	0.84
221185_s_at	IQ motif containing G	0.83
214053_at	V-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	0.83
209167_at	glycoprotein M6B	0.82
219860_at	lymphocyte antigen 6 complex, locus G5C	0.81
219469_at	dynein, cytoplasmic, heavy polypeptide 2	0.80
205898_at	chemokine (C-X3-C motif) receptor 1	0.80
219584_at	phospholipase A1 member A	0.79
205906_at	forkhead box J1	0.78
210176_at	toll-like receptor 1	0.76
213661_at	regeneration associated muscle protease	0.76
207996_s_at	chromosome 18 open reading frame 1	0.75
205769_at	solute carrier family 27 (fatty acid transporter), member 2	0.75
204529_s_at	thymus high mobility group box protein TOX	0.74
206076_at	B7 gene	0.74
210129_s_at	tubulin tyrosine ligase-like family, member 3	0.73
213412_at	tight junction protein 3 (zona occludens 3)	0.72
205673_s_at	ankyrin repeat and SOCS box-containing 9	0.72
206274_s_at	ciliary rootlet coiled-coil, rootletin	0.70
205896_at	solute carrier family 22 (organic cation transporter), member 4	0.70
204802_at	Ras-related associated with diabetes	0.69
220441_at	hypothetical protein FLJ13236	0.69
213248_at	hypothetical protein LOC221362	0.68
207795_s_at	killer cell lectin-like receptor subfamily D, member 1	0.68

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 10)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa
204754_at	Hepatic leukemia factor	0.67
202652_at	amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	0.67
213996_at	yippee-like 1 (Drosophila)	0.67
216356_x_at	BAI1-associated protein 3	0.67
213601_at	slit homolog 1 (Drosophila)	0.66
205839_s_at	benzodiazapine receptor (peripheral) associated protein 1	0.66
211855_s_at	solute carrier family 25 (mitochondrial carrier, brain), member 14	0.66
206483_at	leucine rich repeat containing 6	0.65
209829_at	chromosome 6 open reading frame 32	0.64
219389_at	sushi domain containing 4	0.64
222220_s_at	translin-associated factor X interacting protein 1	0.64
212745_s_at	Bardet-Biedl syndrome 4	0.64
214076_at	Hypothetical protein MGC11335	0.64
218876_at	brain specific protein ; brain specific protein	0.63
216119_s_at	chromosome 20 open reading frame 28	0.63
219313_at	hypothetical protein DKFZp434C0328	0.63
213512_at	chromosome 14 open reading frame 79	0.62
220024_s_at	periaxin	0.61
222291_at	DKFZP564J102 protein	0.61
220005_at	purinergic receptor P2Y, G-protein coupled, 13	0.61
210044_s_at	lymphoblastic leukemia derived sequence 1	0.61
206631_at	prostaglandin E receptor 2 (subtype EP2), 53kDa	0.60
209880_s_at	selectin P ligand	0.60
211734_s_at	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide	0.60
49049_at	deltex 3 homolog (Drosophila)	0.58
210641_at	calpain 9	0.58
207624_s_at	retinitis pigmentosa GTPase regulator	0.58
205073_at	cytochrome P450, family 2, subfamily J, polypeptide 2	0.57
203485_at	reticulon 1	0.57
212695_at	cryptochrome 2 (photolyase-like)	0.57
219951_s_at	chromosome 20 open reading frame 12	0.57
219753_at	stromal antigen 3	0.56
206693_at	interleukin 7	0.56
208530_s_at	retinoic acid receptor, beta	0.56
205208_at	aldehyde dehydrogenase 1 family, member L1	0.56

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 11)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
207704_s_at	growth arrest-specific 7	0.56	
208396_s_at	phosphodiesterase 1A, calmodulin-dependent	0.55	
219370_at	reprimo, TP53 dependant G2 arrest mediator candidate	0.54	
220622_at	leucine rich repeat containing 31	0.53	
221928_at	acetyl-Coenzyme A carboxylase beta	0.53	
216663_s_at	zinc finger, MYND-type containing 10	0.52	
210033_s_at	sperm associated antigen 6	0.52	
212935_at	MCF.2 cell line derived transforming sequence-like	0.52	
208518_s_at	period homolog 2 (Drosophila)	0.51	
204378_at	breast carcinoma amplified sequence 1	0.51	
219670_at	chromosome 1 open reading frame 165	0.51	
219820_at	solute carrier family 6, member 16	0.51	
204792_s_at	WD and tetratricopeptide repeats 2	0.51	
204364_s_at	chromosome 2 open reading frame 23	0.51	
216807_at	KIAA1751	0.50	
214032_at	zeta-chain (TCR) associated protein kinase 70kDa	0.50	
209074_s_at	TU3A protein	0.50	
205043_at	cystic fibrosis transmembrane conductance regulator, ATP-binding cassette (sub-family C, member 7)	0.50	
41577_at	protein phosphatase 1, regulatory (inhibitor) subunit 16B	0.49	Below 0.5x median
219703_at	meiosis-specific nuclear structural 1	0.47	Below 0.5x median
221122_at	HRAS-like suppressor 2	0.47	Below 0.5x median
215695_s_at	glycogenin 2	0.47	Below 0.5x median
205969_at	arylacetamide deacetylase (esterase)	0.47	Below 0.5x median
206004_at	transglutaminase 3 (E polypeptide, protein-glutamine-gamma-glutamyltransferase)	0.45	Below 0.5x median
201283_s_at	OGT(O-Glc-NAc transferase)-interacting protein 106 KDa	0.45	Below 0.5x median
214222_at	dynein, axonemal, heavy polypeptide 7	0.45	Below 0.5x median
210505_at	alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide	0.44	Below 0.5x median
220197_at	ATPase, H+ transporting, lysosomal V0 subunit a isoform 4	0.44	Below 0.5x median
215363_x_at	folate hydrolase (prostate-specific membrane antigen) 1	0.44	Below 0.5x median
221107_at	cholinergic receptor, nicotinic, alpha polypeptide 9	0.43	Below 0.5x median
220003_at	leucine rich repeat containing 36	0.43	Below 0.5x median
206336_at	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	0.43	Below 0.5x median
219824_at	solute carrier family 13 (sodium/sulfate symporters),	0.43	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 12)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
	member 4		
220581_at	chromosome 6 open reading frame 97	0.43	Below 0.5x median
213438_at	neurofascin	0.43	Below 0.5x median
209828_s_at	interleukin 16 (lymphocyte chemoattractant factor)	0.43	Below 0.5x median
213679_at	hypothetical protein FLJ13946	0.42	Below 0.5x median
219695_at	sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingomyelinase II)	0.42	Below 0.5x median
206078_at	kalirin, RhoGEF kinase	0.41	Below 0.5x median
205979_at	secretoglobin, family 2A, member 1	0.40	Below 0.5x median
210323_at	tektin 2 (testicular)	0.40	Below 0.5x median
219455_at	hypothetical protein FLJ21062	0.40	Below 0.5x median
215266_at	dynein, axonemal, heavy polypeptide 3	0.40	Below 0.5x median
205662_at	B9 protein	0.40	Below 0.5x median
219115_s_at	interleukin 20 receptor, alpha	0.39	Below 0.5x median
220107_s_at	chromosome 14 open reading frame 140	0.39	Below 0.5x median
205243_at	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 3	0.39	Below 0.5x median
206938_at	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	0.39	Below 0.5x median
219768_at	V-set domain containing T cell activation inhibitor 1	0.39	Below 0.5x median
209789_at	coronin, actin binding protein, 2B	0.38	Below 0.5x median
208109_s_at	chromosome 15 open reading frame 5	0.38	Below 0.5x median
220059_at	BCR downstream signaling 1	0.38	Below 0.5x median
205848_at	growth arrest-specific 2	0.38	Below 0.5x median
220908_at	hypothetical protein FLJ32855	0.37	Below 0.5x median
217016_x_at	hypothetical LOC389177	0.37	Below 0.5x median
220958_at	unc-51-like kinase 4 (C. elegans)	0.37	Below 0.5x median
214551_s_at	CD7 antigen (p41)	0.37	Below 0.5x median
214261_s_at	alcohol dehydrogenase 6 (class V)	0.37	Below 0.5x median
217245_at	early lymphoid activation protein	0.36	Below 0.5x median
207902_at	interleukin 5 receptor, alpha	0.36	Below 0.5x median
220390_at	ATP/GTP binding protein-like 2	0.36	Below 0.5x median
220636_at	dynein, axonemal, intermediate polypeptide 2	0.36	Below 0.5x median
221173_at	Usher syndrome 1C (autosomal recessive, severe)	0.36	Below 0.5x median
219136_s_at	hypothetical protein FLJ12681	0.35	Below 0.5x median
210409_at	chromosome 6 open reading frame 124	0.35	Below 0.5x median
219743_at	hairy/enhancer-of-split related with YRPW motif 2	0.34	Below 0.5x median
220416_at	ATPase, Class I, type 8B, member 4	0.34	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 13)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
219882_at	tubulin tyrosine ligase-like family, member 7	0.34	Below 0.5x median
205844_at	vanin 1	0.34	Below 0.5x median
206424_at	cytochrome P450, family 26, subfamily A, polypeptide 1	0.33	Below 0.5x median
207636_at	serpin peptidase inhibitor, clade I (pancpin), member 2	0.33	Below 0.5x median
210693_at	signal peptide peptidase-like 2B	0.33	Below 0.5x median
220312_at	family with sequence similarity 83, member E	0.32	Below 0.5x median
209368_at	epoxide hydrolase 2, cytoplasmic	0.32	Below 0.5x median
213959_s_at	KIAA1005 protein	0.32	Below 0.5x median
207981_s_at	estrogen-related receptor gamma	0.31	Below 0.5x median
210742_at	CDC14 cell division cycle 14 homolog A (S. cerevisiae)	0.31	Below 0.5x median
220331_at	cytochrome P450, family 46, subfamily A, polypeptide 1	0.31	Below 0.5x median
47553_at	deafness, autosomal recessive 31	0.31	Below 0.5x median
206773_at	lymphocyte antigen 6 complex, locus H	0.30	Below 0.5x median
204796_at	echinoderm microtubule associated protein like 1	0.30	Below 0.5x median
219871_at	hypothetical protein FLJ13197	0.30	Below 0.5x median
221161_at	achaete-scute complex (Drosophila) homolog-like 3	0.30	Below 0.5x median
205623_at	aldehyde dehydrogenase 3 family, memberA1	0.30	Below 0.5x median
220196_at	mucin 16	0.30	Below 0.5x median
207772_s_at	HMT1 hnRNP methyltransferase-like 4 (S. cerevisiae)	0.30	Below 0.5x median
214598_at	claudin 8	0.30	Below 0.5x median
205922_at	vanin 2	0.29	Below 0.5x median
206556_at	clusterin-like 1 (retinal)	0.29	Below 0.5x median
219857_at	chromosome 10 open reading frame 81	0.29	Below 0.5x median
37547_at	parathyroid hormone-responsive B1	0.28	Below 0.5x median
205744_at	double C2-like domains, alpha	0.28	Below 0.5x median
206005_s_at	chromosome 6 open reading frame 84	0.28	Below 0.5x median
209693_at	astrotactin 2	0.28	Below 0.5x median
219587_at	tetratricopeptide repeat domain 12	0.28	Below 0.5x median
208597_at	ciliary neurotrophic factor	0.28	Below 0.5x median
205921_s_at	solute carrier family 6 (neurotransmitter transporter, taurine), member 6	0.27	Below 0.5x median
205795_at	neurexin 3	0.27	Below 0.5x median
206457_s_at	deiodinase, iodothyronine, type I	0.27	Below 0.5x median
206262_at	alcohol dehydrogenase 1A (class I), alpha polypeptide ; alcohol dehydrogenase IB (class I), beta polypeptide ; alcohol dehydrogenase 1C (class I), gamma polypeptide	0.27	Below 0.5x median
219564_at	potassium inwardly-rectifying channel, subfamily J, member 16	0.26	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed
(cont. page 14)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
205715_at	bone marrow stromal cell antigen 1	0.26	Below 0.5x median
207244_x_at	cytochrome P450, family 2, subfamily A, polypeptide 6	0.25	Below 0.5x median
215904_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 4	0.25	Below 0.5x median
215043_s_at	SMA3 ; SMA5	0.25	Below 0.5x median
219663_s_at	hole gene	0.25	Below 0.5x median
220308_at	coiled-coil domain containing 19	0.25	Below 0.5x median
215143_at	Hypothetical protein FLJ36166	0.25	Below 0.5x median
214204_at	PARK2 co-regulated	0.25	Below 0.5x median
222205_x_at	Chromosome 20 open reading frame 91	0.24	Below 0.5x median
205473_at	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B, isoform 1 (Renal tubular acidosis with deafness)	0.24	Below 0.5x median
214188_at	Hexamethylene bis-acetamide inducible 1	0.24	Below 0.5x median
206291_at	neurotensin	0.24	Below 0.5x median
206318_at	serine peptidase inhibitor-like, with Kunitz and WAP domains 1 (eppin)	0.24	Below 0.5x median
220321_s_at	hypothetical protein FLJ13646	0.24	Below 0.5x median
219766_at	hypothetical protein MGC4093	0.23	Below 0.5x median
206541_at	kallikrein B, plasma (Fletcher factor) 1	0.23	Below 0.5x median
205031_at	ephrin-B3	0.23	Below 0.5x median
220389_at	hypothetical protein FLJ23514	0.23	Below 0.5x median
216103_at	acyl-CoA thioesterase 11	0.23	Below 0.5x median
204557_s_at	DAZ interacting protein 1	0.22	Below 0.5x median
210082_at	ATP-binding cassette, sub-family A (ABC1), member 4	0.22	Below 0.5x median
220156_at	EF-hand calcium binding domain 1	0.22	Below 0.5x median
207900_at	chemokine (C-C motif) ligand 17	0.22	Below 0.5x median
219965_s_at	PDZ domain containing, X chromosome	0.21	Below 0.5x median
213589_s_at	hypothetical protein LOC146712 ; UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1	0.21	Below 0.5x median
214082_at	carbonic anhydrase VB, mitochondrial	0.21	Below 0.5x median
220168_at	cancer susceptibility candidate 1	0.21	Below 0.5x median
207021_at	zona pellucida binding protein	0.21	Below 0.5x median
222134_at	D-aspartate oxidase	0.20	Below 0.5x median
206243_at	TIMP metallopeptidase inhibitor 4	0.20	Below 0.5x median
220351_at	chemokine (C-C motif) receptor-like 1	0.20	Below 0.5x median
206465_at	acyl-CoA synthetase bubblegum family member 1	0.20	Below 0.5x median
206721_at	chromosome 1 open reading frame 114	0.20	Below 0.5x median
214156_at	myosin VIIA and Rab interacting protein	0.20	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed (cont. page 15)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
210390_s_at	chemokine (C-C motif) ligand 15	0.19	Below 0.5x median
206480_at	leukotriene C4 synthase	0.19	Below 0.5x median
220125_at	dynein, axonemal, intermediate polypeptide 1	0.19	Below 0.5x median
222068_s_at	leucine rich repeat containing 50	0.19	Below 0.5x median
205675_at	microsomal triglyceride transfer protein (large polypeptide, 88kDa)	0.19	Below 0.5x median
210073_at	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	0.18	Below 0.5x median
213456_at	sclerostin domain containing 1	0.18	Below 0.5x median
211303_x_at	prostate-specific membrane antigen-like	0.18	Below 0.5x median
206017_at	KIAA0319	0.18	Below 0.5x median
220302_at	male germ cell-associated kinase	0.18	Below 0.5x median
220218_at	chromosome 9 open reading frame 68	0.18	Below 0.5x median
214147_at	chromosome 1 open reading frame 175	0.17	Below 0.5x median
219511_s_at	synuclein, alpha interacting protein (synphilin)	0.17	Below 0.5x median
213217_at	adenylate cyclase 2 (brain)	0.17	Below 0.5x median
214217_at	Glutamate receptor, metabotropic 5	0.17	Below 0.5x median
210262_at	cysteine-rich secretory protein 2	0.17	Below 0.5x median
205413_at	metallophosphoesterase domain containing 2	0.17	Below 0.5x median
220539_at	chromosome 10 open reading frame 92	0.16	Below 0.5x median
220812_s_at	HERV-H LTR-associating 2	0.16	Below 0.5x median
221908_at	Hypothetical protein FLJ14627	0.16	Below 0.5x median
207958_at	UDP glucuronosyltransferase 2 family, polypeptide A1	0.16	Below 0.5x median
203995_at	chromosome 21 open reading frame 2	0.16	Below 0.5x median
207430_s_at	microseminoprotein, beta-	0.16	Below 0.5x median
220523_at	EF-hand domain (C-terminal) containing 2	0.16	Below 0.5x median
220269_at	hypothetical protein FLJ23049	0.16	Below 0.5x median
213611_at	aquaporin 5	0.16	Below 0.5x median
214608_s_at	eyes absent homolog 1 (Drosophila)	0.16	Below 0.5x median
210328_at	glycine N-methyltransferase	0.16	Below 0.5x median
215174_at	flavin containing monooxygenase 6	0.15	Below 0.5x median
205948_at	protein tyrosine phosphatase, receptor type, T	0.15	Below 0.5x median
214385_s_at	mucin 5, subtypes A and C, tracheobronchial/gastric	0.15	Below 0.5x median
220055_at	zinc finger protein 287	0.14	Below 0.5x median
207583_at	ATP-binding cassette, sub-family D (ALD), member 2	0.14	Below 0.5x median
219578_s_at	cytoplasmic polyadenylation element binding protein 1	0.14	Below 0.5x median
220991_s_at	ring finger protein 32	0.14	Below 0.5x median
206755_at	cytochrome P450, family 2, subfamily B, polypeptide 6	0.14	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed
(cont. page 16)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
207550_at	myeloproliferative leukemia virus oncogene	0.13	Below 0.5x median
220344_at	chromosome 11 open reading frame 16	0.13	Below 0.5x median
216135_at	Hypothetical protein MGC35048	0.13	Below 0.5x median
214162_at	hypothetical protein LOC284244	0.13	Below 0.5x median
221180_at	Yeast Sps1/Ste20-related kinase 4 (S. cerevisiae)	0.12	Below 0.5x median
205056_s_at	G protein-coupled receptor 162	0.12	Below 0.5x median
220280_s_at	ankyrin repeat and MYND domain containing 1	0.12	Below 0.5x median
205363_at	butyrobetaine (gamma), 2-oxoglutarate dioxygenase (gamma-butyrobetaine hydroxylase) 1	0.12	Below 0.5x median
210445_at	fatty acid binding protein 6, ileal (gastrotropin)	0.11	Below 0.5x median
219106_s_at	kelch repeat and BTB (POZ) domain containing 10	0.11	Below 0.5x median
219614_s_at	solute carrier family 6 (proline IMINO transporter), member 20	0.11	Below 0.5x median
222102_at	glutathione S-transferase A3	0.10	Below 0.5x median
220528_at	vanin 3	0.10	Below 0.5x median
214818_at	hypothetical protein LOC284001	0.10	Below 0.5x median
205044_at	gamma-aminobutyric acid (GABA) A receptor, pi	0.10	Below 0.5x median
207913_at	cytochrome P450, family 2, subfamily F, polypeptide 1	0.10	Below 0.5x median
207797_s_at	LRP2 binding protein	0.10	Below 0.5x median
207149_at	cadherin 12, type 2 (N-cadherin 2)	0.10	Below 0.5x median
221863_at	KIAA1193	0.09	Below 0.5x median
220614_s_at	chromosome 6 open reading frame 103	0.09	Below 0.5x median
205337_at	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	0.09	Below 0.5x median
220361_at	hypothetical protein FLJ12476	0.09	Below 0.5x median
205813_s_at	methionine adenosyltransferase I, alpha	0.09	Below 0.5x median
220652_at	kinesin family member 24	0.09	Below 0.5x median
211468_s_at	RecQ protein-like 5	0.09	Below 0.5x median
221182_at	chromosome 1 open reading frame 129	0.08	Below 0.5x median
214420_s_at	Cytochrome P450, family 2, subfamily C, polypeptide 9	0.08	Below 0.5x median
207344_at	A kinase (PRKA) anchor protein 3	0.08	Below 0.5x median
222089_s_at	hypothetical protein AF447587	0.08	Below 0.5x median
220766_at	B-cell translocation gene 4	0.08	Below 0.5x median
208327_at	cytochrome P450, family 2, subfamily A, polypeptide 13	0.07	Below 0.5x median
206810_at	zinc finger protein 44 (KOX 7)	0.07	Below 0.5x median
220426_at	hypothetical protein MGC5356	0.07	Below 0.5x median
214652_at	dopamine receptor D1	0.07	Below 0.5x median
206526_at	RIB43A domain with coiled-coils 2	0.06	Below 0.5x median

Supplemental Gene List II. Airway Non-basal Cell Signature Genes Analyzed
(cont. page 17)

Probe ID	Gene name	Median of normalized expression levels in lung adenoCa	
207896_s_at	deleted in lung and esophageal cancer 1	0.06	Below 0.5x median
220756_s_at	G protein-coupled receptor 172B	0.06	Below 0.5x median
206149_at	hepatocellular carcinoma antigen gene 520	0.06	Below 0.5x median
207177_at	prostaglandin F receptor (FP)	0.06	Below 0.5x median
221946_at	chromosome 9 open reading frame 116	0.06	Below 0.5x median
213845_at	GRIK2=GluR6 kainate receptor {3' untranslated region, TAA trinucleotide repeats} [human, caudate/putamen, mRNA Partial, 540 nt]	0.06	Below 0.5x median
209720_s_at	serpin peptidase inhibitor, clade B (ovalbumin), member 3	0.06	Below 0.5x median
216209_at	hypothetical gene supported by AK057632; AL137270; BC057846	0.06	Below 0.5x median
207328_at	arachidonate 15-lipoxygenase	0.05	Below 0.5x median
220064_at	tetratricopeptide repeat domain 21B	0.05	Below 0.5x median
207766_at	cyclin-dependent kinase-like 1 (CDC2-related kinase)	0.05	Below 0.5x median
210402_at	potassium inwardly-rectifying channel, subfamily J, member 1	0.05	Below 0.5x median
214961_at	KIAA0774	0.04	Below 0.5x median
215432_at	acyl-CoA synthetase medium-chain family member 1	0.04	Below 0.5x median
203030_s_at	protein tyrosine phosphatase, receptor type, N polypeptide 2	0.04	Below 0.5x median
206815_at	sperm associated antigen 8	0.04	Below 0.5x median
220105_at	rhabdoid tumor deletion region gene 1	0.03	Below 0.5x median
206380_s_at	properdin P factor, complement	0.03	Below 0.5x median
207261_at	cyclic nucleotide gated channel alpha 3	0.03	Below 0.5x median
205478_at	protein phosphatase 1, regulatory (inhibitor) subunit 1A	0.03	Below 0.5x median
206522_at	maltase-glucoamylase (alpha-glucosidase)	0.03	Below 0.5x median

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
37892_at	collagen, type XI, alpha 1	12.60	$2.37x10^{-9}$
209125_at	keratin 6A ; keratin 6C ; keratin 6E	8.53	$2.94x10^{-8}$
204855_at	serpin peptidase inhibitor, clade B (ovalbumin), member 5	7.94	$3.20x10^{-6}$
204475_at	matrix metallopeptidase 1 (interstitial collagenase)	6.44	$1.70x10^{-5}$
205627_at	cytidine deaminase	6.13	$9.09x10^{-8}$
206023_at	neuromedin U	6.03	$3.58x10^{-5}$
218469_at	gremlin 1, cysteine knot superfamily, homolog (Xenopus laevis)	5.46	$1.47x10^{-11}$
209942_x_at	melanoma antigen family A, 3	5.15	$4.14x10^{-3}$
213338_at	Ras-induced senescence 1	5.08	$3.25x10^{-10}$
205476_at	chemokine (C-C motif) ligand 20	4.51	$1.08x10^{-3}$
205403_at	interleukin 1 receptor, type II	4.51	$5.20x10^{-6}$
206561_s_at	aldo-keto reductase family 1, member B10 (aldose reductase)	4.50	$9.77x10^{-4}$
219555_s_at	uncharacterized bone marrow protein BM039	4.47	$5.09x10^{-7}$
214612_x_at	melanoma antigen family A, 6	4.41	$6.97x10^{-3}$
204420_at	FOS-like antigen 1	4.34	$6.57x10^{-9}$
204641_at	NIMA (never in mitosis gene a)-related kinase 2	4.25	$2.84x10^{-9}$
213711_at	keratin, hair, basic, 1	4.22	$2.56x10^{-3}$
201890_at	ribonucleotide reductase M2 polypeptide	4.21	$9.25x10^{-12}$
209825_s_at	uridine-cytidine kinase 2	4.15	$6.69x10^{-9}$
211506_s_at	interleukin 8	4.15	$3.66x10^{-7}$
202855_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 3	4.15	$2.03x10^{-13}$
219918_s_at	asp (abnormal spindle)-like, microcephaly associated (Drosophila)	4.14	$1.10x10^{-7}$
213201_s_at	troponin T type 1 (skeletal, slow)	4.12	$1.34x10^{-4}$
201710_at	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	4.08	$1.32x10^{-9}$
221520_s_at	cell division cycle associated 8	3.99	$1.44x10^{-7}$
201564_s_at	fascin homolog 1, actin-bundling protein (Strongylocentrotus purpuratus)	3.92	$3.98x10^{-9}$
209587_at	paired-like homeodomain transcription factor 1	3.88	$2.78x10^{-5}$
206155_at	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	3.79	$3.15x10^{-4}$
220431_at	transmembrane protease, serine 11E	3.76	$1.78x10^{-4}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 2)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
204162_at	kinetochore associated 2	3.74	$7.34x10^{-7}$
205874_at	inositol 1,4,5-trisphosphate 3-kinase A	3.73	$4.12x10^{-5}$
213790_at	ADAM metallopeptidase domain 12 (meltrin alpha)	3.72	$2.87x10^{-8}$
218542_at	chromosome 10 open reading frame 3	3.68	$9.68x10^{-10}$
205490_x_at	gap junction protein, beta 3, 31kDa (connexin 31)	3.67	$6.94x10^{-6}$
211571_s_at	chondroitin sulfate proteoglycan 2 (versican)	3.66	$4.21x10^{-8}$
214603_at	melanoma antigen family A, 2 ; melanoma antigen family A, 2B	3.65	$7.05x10^{-3}$
202580_x_at	forkhead box M1	3.62	$3.71x10^{-8}$
204126_s_at	CDC45 cell division cycle 45-like (S. cerevisiae)	3.62	$4.57x10^{-5}$
212992_at	chromosome 14 open reading frame 78	3.61	$1.26x10^{-10}$
212353_at	sulfatase 1	3.59	$3.11x10^{-9}$
210052_s_at	TPX2, microtubule-associated, homolog (Xenopus laevis)	3.59	$3.17x10^{-10}$
202870_s_at	CDC20 cell division cycle 20 homolog (S. cerevisiae)	3.57	$1.38x10^{-9}$
205959_at	matrix metallopeptidase 13 (collagenase 3)	3.57	$1.65x10^{-3}$
202779_s_at	ubiquitin-conjugating enzyme E2S	3.56	$2.52x10^{-12}$
204351_at	S100 calcium binding protein P	3.55	$7.27x10^{-3}$
221521_s_at	DNA replication complex GINS protein PSF2	3.55	$5.28x10^{-6}$
203819_s_at	IGF-II mRNA-binding protein 3	3.54	$2.13x10^{-3}$
204920_at	carbamoyl-phosphate synthetase 1, mitochondrial	3.54	$7.60x10^{-4}$
205064_at	small proline-rich protein 1B (cornifin)	3.50	$4.65x10^{-4}$
218726_at	Holliday junction recognition protein	3.49	$4.36x10^{-8}$
204825_at	maternal embryonic leucine zipper kinase	3.40	$1.27x10^{-8}$
211668_s_at	plasminogen activator, urokinase	3.40	$2.37x10^{-9}$
212949_at	barren homolog (Drosophila)	3.40	$5.64x10^{-8}$
204952_at	LY6/PLAUR domain containing 3	3.36	$5.09x10^{-5}$
202311_s_at	collagen, type I, alpha 1	3.35	$3.33x10^{-7}$
203876_s_at	matrix metallopeptidase 11 (stromelysin 3)	3.34	$1.41x10^{-5}$
202363_at	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican)	3.33	$1.10x10^{-7}$
221009_s_at	angiopoietin-like 4	3.33	$1.66x10^{-7}$
205046_at	centromere protein E, 312kDa	3.32	$6.09x10^{-8}$
219491_at	leucine rich repeat and fibronectin type III domain containing 4	3.31	$4.06x10^{-6}$
206504_at	cytochrome P450, family 24, subfamily A, polypeptide 1	3.30	$8.91x10^{-4}$
203535_at	S100 calcium binding protein A9 (calgranulin B)	3.30	$4.96x10^{-5}$
214370_at	S100 calcium binding protein A8 (calgranulin A)	3.27	$1.25x10^{-3}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 3)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	p value
220651_s_at	MCM10 minichromosome maintenance deficient 10 (S. cerevisiae)	3.27	5.60×10^{-7}
221730_at	collagen, type V, alpha 2	3.27	2.03×10^{-9}
203362_s_at	MAD2 mitotic arrest deficient-like 1 (yeast)	3.25	4.16×10^{-10}
201037_at	phosphofructokinase, platelet	3.24	9.25×10^{-12}
209800_at	keratin 16 (focal non-epidermolytic palmoplantar keratoderma)	3.22	6.10×10^{-4}
202620_s_at	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	3.21	5.60×10^{-11}
205328_at	claudin 10	3.20	2.77×10^{-4}
204286_s_at	phorbol-12-myristate-13-acetate-induced protein 1	3.20	2.14×10^{-6}
206025_s_at	tumor necrosis factor, alpha-induced protein 6	3.18	3.23×10^{-11}
210643_at	tumor necrosis factor (ligand) superfamily, member 11	3.18	1.70×10^{-4}
219271_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 14 (GalNAc-T14)	3.17	7.01×10^{-4}
210511_s_at	inhibin, beta A (activin A, activin AB alpha polypeptide)	3.14	1.14×10^{-5}
205542_at	six transmembrane epithelial antigen of the prostate 1	3.09	9.68×10^{-10}
213008_at	Fanconi anemia, complementation group I	3.02	5.74×10^{-6}
202828_s_at	matrix metallopeptidase 14 (membrane-inserted)	3.00	7.07×10^{-6}
202997_s_at	lysyl oxidase-like 2	3.00	1.20×10^{-5}
201387_s_at	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	2.99	9.72×10^{-6}
217901_at	Desmoglein 2	2.99	3.36×10^{-10}
218663_at	chromosome condensation protein G	2.99	1.23×10^{-7}
218741_at	chromosome 22 open reading frame 18	2.97	1.21×10^{-6}
200606_at	desmoplakin	2.95	9.91×10^{-9}
206549_at	insulin-like 4 (placenta)	2.95	1.62×10^{-3}
206439_at	dermatan sulfate proteoglycan 3	2.95	8.52×10^{-4}
204014_at	dual specificity phosphatase 4	2.94	1.80×10^{-3}
210074_at	cathepsin L2	2.94	3.90×10^{-8}
205916_at	S100 calcium binding protein A7 (psoriasin 1)	2.91	6.12×10^{-3}
204281_at	TEA domain family member 4	2.90	2.85×10^{-8}
219148_at	PDZ binding kinase	2.88	3.71×10^{-7}
205302_at	insulin-like growth factor binding protein 1	2.88	5.77×10^{-5}
205767_at	epiregulin	2.86	6.08×10^{-3}
202404_s_at	collagen, type I, alpha 2	2.84	1.72×10^{-8}
219825_at	cytochrome P450, family 26, subfamily B, polypeptide 1	2.83	7.44×10^{-6}
209596_at	matrix-remodelling associated 5	2.82	2.28×10^{-7}
218355_at	kinesin family member 4A	2.82	7.75×10^{-8}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 4)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	p value
213523_at	cyclin E1	2.80	3.04×10^{-6}
218843_at	fibronectin type III domain containing 4	2.80	2.27×10^{-7}
203213_at	Cell division cycle 2, G1 to S and G2 to M	2.80	2.63×10^{-8}
221436_s_at	cell division cycle associated 3	2.80	5.86×10^{-6}
218990_s_at	small proline-rich protein 3	2.79	1.23×10^{-3}
203418_at	cyclin A2	2.79	5.98×10^{-8}
209680_s_at	kinesin family member C1	2.78	2.46×10^{-5}
214710_s_at	cyclin B1	2.78	2.37×10^{-9}
201116_s_at	carboxypeptidase E	2.76	6.65×10^{-5}
202627_s_at	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	2.75	1.63×10^{-6}
205513_at	transcobalamin I (vitamin B12 binding protein, R binder family)	2.75	6.52×10^{-3}
204580_at	matrix metallopeptidase 12 (macrophage elastase)	2.74	4.59×10^{-4}
207717_s_at	plakophilin 2	2.74	6.69×10^{-5}
220658_s_at	aryl hydrocarbon receptor nuclear translocator-like 2	2.73	1.29×10^{-8}
204602_at	dickkopf homolog 1 (Xenopus laevis)	2.72	8.76×10^{-4}
211844_s_at	neuropilin 2	2.71	5.84×10^{-5}
203560_at	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	2.71	3.82×10^{-6}
203967_at	CDC6 cell division cycle 6 homolog (S. cerevisiae)	2.71	3.66×10^{-7}
203925_at	glutamate-cysteine ligase, modifier subunit	2.69	1.39×10^{-6}
218880_at	FOS-like antigen 2	2.68	1.94×10^{-12}
204962_s_at	centromere protein A, 17kDa	2.68	1.54×10^{-6}
214974_x_at	chemokine (C-X-C motif) ligand 5	2.68	3.61×10^{-3}
206645_s_at	nuclear receptor subfamily 0, group B, member 1	2.68	3.27×10^{-3}
203976_s_at	chromatin assembly factor 1, subunit A (p150)	2.67	1.63×10^{-5}
212489_at	collagen, type V, alpha 1	2.67	1.72×10^{-5}
209921_at	solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	2.66	9.70×10^{-4}
205650_s_at	fibrinogen alpha chain	2.66	8.28×10^{-3}
202705_at	cyclin B2	2.65	8.00×10^{-9}
202094_at	baculoviral IAP repeat-containing 5 (survivin)	2.64	2.82×10^{-5}
202504_at	tripartite motif-containing 29	2.64	1.86×10^{-3}
221909_at	ring finger protein, transmembrane 2	2.64	6.66×10^{-3}
204078_at	synaptonemal complex protein SC65	2.64	3.47×10^{-5}
204146_at	RAD51 associated protein 1	2.63	3.93×10^{-7}
201292_at	topoisomerase (DNA) II alpha 170kDa	2.63	1.48×10^{-8}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 5)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	p value
212110_at	solute carrier family 39 (zinc transporter), member 14	2.63	2.33×10^{-8}
202267_at	laminin, gamma 2	2.62	6.05×10^{-5}
204933_s_at	tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)	2.61	1.87×10^{-3}
219787_s_at	epithelial cell transforming sequence 2 oncogene	2.61	6.99×10^{-8}
219936_s_at	G protein-coupled receptor 87	2.61	4.31×10^{-3}
208103_s_at	acidic (leucine-rich) nuclear phosphoprotein 32 family, member E	2.60	1.45×10^{-7}
204709_s_at	kinesin family member 23	2.60	4.89×10^{-5}
211756_at	parathyroid hormone-like hormone	2.60	1.12×10^{-3}
210683_at	neurturin	2.59	1.93×10^{-4}
201266_at	thioredoxin reductase 1	2.59	1.54×10^{-7}
202219_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 8	2.59	1.23×10^{-3}
203764_at	discs, large homolog 7 (Drosophila)	2.58	4.47×10^{-9}
214469_at	histone 1, H2ae	2.55	5.85×10^{-3}
203691_at	peptidase inhibitor 3, skin-derived (SKALP)	2.55	7.14×10^{-3}
206074_s_at	high mobility group AT-hook 1	2.55	1.75×10^{-10}
204822_at	TTK protein kinase	2.55	5.73×10^{-8}
214240_at	galanin	2.55	4.17×10^{-5}
214639_s_at	homeo box A1	2.54	3.21×10^{-3}
208607_s_at	serum amyloid A1 ; serum amyloid A2	2.53	4.61×10^{-3}
222037_at	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)	2.53	5.08×10^{-6}
204766_s_at	nudix (nucleoside diphosphate linked moiety X)-type motif 1	2.52	1.67×10^{-4}
218009_s_at	protein regulator of cytokinesis 1	2.51	2.73×10^{-8}
205924_at	RAB3B, member RAS oncogene family	2.51	1.78×10^{-4}
202912_at	adrenomedullin	2.50	4.11×10^{-6}
218944_at	pyrroline-5-carboxylate reductase-like	2.50	2.04×10^{-4}
201761_at	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	2.50	2.50×10^{-9}
207165_at	hyaluronan-mediated motility receptor (RHAMM)	2.50	5.84×10^{-7}
204430_s_at	solute carrier family 2 (facilitated glucose/fructose transporter), member 5	2.49	2.22×10^{-4}
203699_s_at	deiodinase, iodothyronine, type II	2.49	1.33×10^{-3}
221539_at	eukaryotic translation initiation factor 4E binding protein 1	2.49	3.91×10^{-7}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 6)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
219650_at	excision repair cross-complementing rodent repair deficiency, complementation group 6-like	2.48	$4.91x10^{-5}$
204879_at	podoplanin	2.48	$1.87x10^{-3}$
210145_at	phospholipase A2, group IVA (cytosolic, calcium-dependent)	2.47	$5.02x10^{-4}$
206067_s_at	Wilms tumor 1	2.47	$6.08x10^{-3}$
202589_at	thymidylate synthetase	2.47	$2.00x10^{-8}$
213909_at	leucine rich repeat containing 15	2.47	$6.91x10^{-5}$
214463_x_at	histone 1, H4k ; histone 1, H4j	2.46	$8.43x10^{-5}$
202954_at	ubiquitin-conjugating enzyme E2C	2.46	$3.90x10^{-8}$
204385_at	kynureninase (L-kynurenine hydrolase)	2.46	$8.37x10^{-5}$
205499_at	sushi-repeat-containing protein, X-linked 2	2.46	$1.72x10^{-5}$
211150_s_at	dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydrogenase complex)	2.46	$6.15x10^{-5}$
209372_x_at	tubulin, beta 2 ; tubulin, beta polypeptide paralog	2.46	$2.57x10^{-4}$
203083_at	thrombospondin 2	2.45	$3.79x10^{-6}$
201629_s_at	acid phosphatase 1, soluble	2.45	$8.64x10^{-7}$
208079_s_at	serine/threonine kinase 6	2.44	$6.09x10^{-8}$
213640_s_at	lysyl oxidase	2.44	$9.41x10^{-5}$
214290_s_at	histone 2, H2aa	2.44	$3.19x10^{-5}$
213943_at	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	2.43	$2.93x10^{-6}$
205568_at	aquaporin 9	2.43	$4.08x10^{-4}$
204344_s_at	Sec23 homolog A (S. cerevisiae)	2.43	$2.26x10^{-5}$
215076_s_at	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)	2.42	$1.35x10^{-7}$
217640_x_at	chromosome 18 open reading frame 24	2.41	$1.58x10^{-3}$
219249_s_at	FK506 binding protein 10, 65 kDa	2.41	$4.54x10^{-4}$
206391_at	retinoic acid receptor responder (tazarotene induced) 1	2.41	$3.72x10^{-3}$
219032_x_at	opsin 3 (encephalopsin, panopsin)	2.40	$3.71x10^{-8}$
202765_s_at	fibrillin 1 (Marfan syndrome)	2.39	$3.05x10^{-3}$
203440_at	cadherin 2, type 1, N-cadherin (neuronal)	2.39	$4.98x10^{-3}$
212396_s_at	KIAA0090	2.39	$1.46x10^{-3}$
204558_at	RAD54-like (S. cerevisiae)	2.39	$1.37x10^{-4}$
205240_at	G-protein signalling modulator 2 (AGS3-like, C. elegans)	2.39	$2.55x10^{-4}$
220789_s_at	transforming growth factor beta regulator 4	2.39	$9.70x10^{-6}$
212942_s_at	KIAA1199	2.38	$4.55x10^{-5}$
219232_s_at	egl nine homolog 3 (C. elegans)	2.38	$2.09x10^{-6}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 7)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	**p value**
204026_s_at	ZW10 interactor	2.38	3.07×10^{-7}
220442_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 4 (GalNAc-T4)	2.37	3.19×10^{-6}
209955_s_at	fibroblast activation protein, alpha	2.37	6.50×10^{-8}
217818_s_at	actin related protein 2/3 complex, subunit 4, 20kDa	2.37	8.24×10^{-4}
203477_at	collagen, type XV, alpha 1	2.37	5.63×10^{-5}
202240_at	polo-like kinase 1 (Drosophila)	2.37	9.87×10^{-6}
219926_at	popeye domain containing 3	2.36	2.62×10^{-3}
209875_s_at	secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	2.35	7.48×10^{-4}
218585_s_at	denticleless homolog (Drosophila)	2.35	3.63×10^{-7}
211148_s_at	angiopoietin 2	2.35	4.16×10^{-4}
205071_x_at	X-ray repair complementing defective repair in Chinese hamster cells 4	2.34	8.24×10^{-4}
202949_s_at	four and a half LIM domains 2	2.33	2.33×10^{-6}
209270_at	laminin, beta 3	2.33	1.92×10^{-4}
220254_at	low density lipoprotein-related protein 12	2.32	2.43×10^{-3}
209714_s_at	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	2.32	1.91×10^{-9}
213139_at	snail homolog 2 (Drosophila)	2.32	1.21×10^{-5}
202503_s_at	KIAA0101	2.32	9.83×10^{-7}
40472_at	PLSC domain containing protein	2.32	2.14×10^{-4}
201489_at	peptidylprolyl isomerase F (cyclophilin F)	2.31	4.42×10^{-8}
219410_at	transmembrane protein 45A	2.31	4.96×10^{-5}
204714_s_at	coagulation factor V (proaccelerin, labile factor)	2.31	1.07×10^{-3}
213680_at	keratin 6B	2.31	3.01×10^{-3}
213765_at	microfibrillar associated protein 5	2.30	1.81×10^{-3}
205266_at	leukemia inhibitory factor (cholinergic differentiation factor)	2.30	3.68×10^{-5}
204170_s_at	CDC28 protein kinase regulatory subunit 2	2.30	7.33×10^{-8}
219366_at	apoptosis, caspase activation inhibitor	2.30	1.10×10^{-8}
203755_at	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	2.29	1.68×10^{-8}
217428_s_at	collagen, type X, alpha 1(Schmid metaphyseal chondrodysplasia)	2.29	4.33×10^{-3}
205229_s_at	coagulation factor C homolog, cochlin (Limulus polyphemus)	2.29	2.98×10^{-3}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 8)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	p value
201195_s_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	2.28	$1.06x10^{-3}$
204256_at	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	2.28	$2.11x10^{-6}$
222206_s_at	nicalin homolog (zebrafish)	2.28	$1.40x10^{-4}$
216913_s_at	KIAA0690	2.27	$3.39x10^{-4}$
204033_at	thyroid hormone receptor interactor 13	2.27	$3.82x10^{-6}$
204748_at	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	2.27	$9.15x10^{-3}$
217755_at	hematological and neurological expressed 1	2.26	$9.28x10^{-8}$
210135_s_at	short stature homeobox 2	2.26	$1.36x10^{-3}$
201930_at	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)	2.26	$5.44x10^{-9}$
202499_s_at	solute carrier family 2 (facilitated glucose transporter), member 3	2.26	$3.77x10^{-6}$
201897_s_at	CDC28 protein kinase regulatory subunit 1B	2.25	$6.15x10^{-9}$
204444_at	kinesin family member 11	2.25	$8.95x10^{-8}$
217312_s_at	collagen, type VII, alpha 1 (epidermolysis bullosa, dystrophic, dominant and recessive)	2.25	$1.07x10^{-3}$
205207_at	interleukin 6 (interferon, beta 2)	2.25	$7.02x10^{-6}$
202923_s_at	glutamate-cysteine ligase, catalytic subunit	2.25	$4.35x10^{-5}$
204114_at	nidogen 2 (osteonidogen)	2.25	$1.95x10^{-5}$
205535_s_at	BH-protocadherin (brain-heart)	2.25	$4.25x10^{-3}$
218975_at	collagen, type V, alpha 3	2.24	$3.08x10^{-4}$
211564_s_at	PDZ and LIM domain 4	2.24	$8.35x10^{-4}$
218245_at	leucine rich repeat containing 54	2.24	$6.07x10^{-5}$
205394_at	CHK1 checkpoint homolog (S. pombe)	2.24	$3.66x10^{-7}$
204457_s_at	growth arrest-specific 1	2.24	$4.54x10^{-3}$
212022_s_at	antigen identified by monoclonal antibody Ki-67	2.24	$3.26x10^{-6}$
211071_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11	2.23	$9.84x10^{-6}$
219634_at	carbohydrate (chondroitin 4) sulfotransferase 11	2.23	$1.52x10^{-6}$
221781_s_at	DnaJ (Hsp40) homolog, subfamily C, member 10	2.23	$3.71x10^{-8}$
218653_at	solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 15	2.23	$7.41x10^{-6}$
205460_at	neuronal PAS domain protein 2	2.22	$2.49x10^{-3}$
219201_s_at	twisted gastrulation homolog 1 (Drosophila)	2.22	$3.28x10^{-3}$
222077_s_at	Rac GTPase activating protein 1	2.22	$2.40x10^{-7}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 9)

Probe ID	Gene name	BC-high adenoCa *vs* BC-low adenoCa	
		Fold-change	p value
212242_at	tubulin, alpha 1 (testis specific)	2.22	$2.79x10^{-10}$
211088_s_at	polo-like kinase 4 (Drosophila)	2.22	$1.58x10^{-3}$
214702_at	fibronectin 1	2.22	$2.87x10^{-4}$
202458_at	protease, serine, 23	2.21	$1.12x10^{-5}$
221777_at	chromosome 12 open reading frame 52	2.21	$3.88x10^{-4}$
204338_s_at	regulator of G-protein signalling 4	2.20	$9.99x10^{-3}$
206364_at	kinesin family member 14	2.20	$1.03x10^{-6}$
208850_s_at	Thy-1 cell surface antigen	2.20	$8.71x10^{-6}$
220093_at	anthrax toxin receptor 1	2.20	$7.35x10^{-4}$
222039_at	kinesin family member 18B	2.20	$8.84x10^{-7}$
214845_s_at	calumenin	2.20	$9.09x10^{-8}$
215223_s_at	superoxide dismutase 2, mitochondrial	2.20	$3.05x10^{-5}$
202154_x_at	tubulin, beta 3	2.19	$4.03x10^{-13}$
204401_at	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4	2.19	$2.13x10^{-4}$
209406_at	BCL2-associated athanogene 2	2.19	$5.12x10^{-6}$
204597_x_at	stanniocalcin 1	2.19	$1.31x10^{-5}$
206595_at	cystatin E/M	2.19	$5.96x10^{-3}$
218498_s_at	ERO1-like (S. cerevisiae)	2.18	$5.24x10^{-8}$
219874_at	solute carrier family 12 (potassium/chloride transporters), member 8	2.18	$1.40x10^{-7}$
216228_s_at	WD repeat and HMG-box DNA binding protein 1	2.18	$2.68x10^{-4}$
205774_at	coagulation factor XII (Hageman factor)	2.17	$1.22x10^{-4}$
204941_s_at	aldehyde dehydrogenase 3 family, member B2	2.16	$9.75x10^{-3}$
209877_at	synuclein, gamma (breast cancer-specific protein 1)	2.16	$3.50x10^{-4}$
202338_at	thymidine kinase 1, soluble	2.15	$1.77x10^{-8}$
205691_at	synaptogyrin 3	2.15	$2.00x10^{-3}$
206079_at	choroideremia-like (Rab escort protein 2)	2.15	$8.00x10^{-5}$
218644_at	pleckstrin 2	2.15	$9.70x10^{-6}$
203936_s_at	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	2.15	$8.34x10^{-4}$
210095_s_at	insulin-like growth factor binding protein 3	2.15	$4.17x10^{-5}$
218883_s_at	MLF1 interacting protein	2.14	$1.45x10^{-6}$
219270_at	ChaC, cation transport regulator homolog 1 (E. coli)	2.14	$6.61x10^{-5}$
214866_at	plasminogen activator, urokinase receptor	2.13	$6.57x10^{-9}$
221591_s_at	family with sequence similarity 64, member A	2.13	$5.84x10^{-6}$
200853_at	H2A histone family, member Z	2.12	$1.22x10^{-8}$
209464_at	aurora kinase B	2.12	$1.45x10^{-5}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 10)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
209850_s_at	CDC42 effector protein (Rho GTPase binding) 2	2.12	1.83×10^{-5}
220198_s_at	eukaryotic translation initiation factor 5A2	2.12	1.37×10^{-4}
213520_at	RecQ protein-like 4	2.11	2.25×10^{-3}
219311_at	chromosome 18 open reading frame 9	2.11	1.32×10^{-5}
210397_at	defensin, beta 1	2.11	4.58×10^{-3}
218980_at	formin homology 2 domain containing 3	2.11	9.55×10^{-3}
221598_s_at	cofactor required for Sp1 transcriptional activation, subunit 8, 34kDa	2.11	8.23×10^{-4}
209832_s_at	DNA replication factor	2.11	1.33×10^{-4}
218046_s_at	mitochondrial ribosomal protein S16	2.11	2.26×10^{-6}
201506_at	transforming growth factor, beta-induced, 68kDa	2.11	1.72×10^{-5}
204427_s_at	transmembrane emp24 domain trafficking protein 2	2.10	6.58×10^{-5}
209408_at	kinesin family member 2C	2.10	1.41×10^{-8}
203270_at	deoxythymidylate kinase (thymidylate kinase)	2.10	8.13×10^{-7}
201105_at	lectin, galactoside-binding, soluble, 1 (galectin 1)	2.10	1.23×10^{-8}
214981_at	Periostin, osteoblast specific factor	2.10	7.99×10^{-3}
205085_at	origin recognition complex, subunit 1-like (yeast)	2.10	1.08×10^{-3}
206103_at	ras-related C3 botulinum toxin substrate 3 (rho family, small GTP binding protein Rac3)	2.09	7.21×10^{-4}
209156_s_at	collagen, type VI, alpha 2	2.09	1.36×10^{-5}
207828_s_at	centromere protein F, 350/400ka (mitosin)	2.09	5.20×10^{-6}
219493_at	SHC SH2-domain binding protein 1	2.09	1.21×10^{-6}
210138_at	regulator of G-protein signalling 20	2.09	5.40×10^{-3}
217996_at	pleckstrin homology-like domain, family A, member 1	2.09	1.85×10^{-6}
221410_x_at	protocadherin beta 3	2.09	3.19×10^{-3}
220124_at	giant axonal neuropathy (gigaxonin)	2.08	7.57×10^{-5}
201940_at	carboxypeptidase D	2.08	1.54×10^{-5}
218039_at	nucleolar and spindle associated protein 1	2.08	5.65×10^{-8}
221845_s_at	ClpB caseinolytic peptidase B homolog (E. coli)	2.08	4.07×10^{-3}
221291_at	UL16 binding protein 2	2.07	1.89×10^{-3}
202990_at	phosphorylase, glycogen; liver (Hers disease, glycogen storage disease type VI)	2.07	1.64×10^{-7}
206389_s_at	phosphodiesterase 3A, cGMP-inhibited	2.07	6.55×10^{-3}
217744_s_at	PERP, TP53 apoptosis effector	2.07	1.12×10^{-6}
218782_s_at	ATPase family, AAA domain containing 2	2.07	2.06×10^{-5}
205066_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 1	2.06	1.21×10^{-3}
208696_at	chaperonin containing TCP1, subunit 5 (epsilon)	2.06	1.48×10^{-8}
205436_s_at	H2A histone family, member X	2.06	9.68×10^{-10}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 11)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
210115_at	ribosomal protein L39-like	2.06	$4.61x10^{-3}$
201504_s_at	translin	2.06	$2.95x10^{-4}$
205780_at	BCL2-interacting killer (apoptosis-inducing)	2.06	$3.10x10^{-3}$
37145_at	granulysin	2.06	$4.26x10^{-3}$
209803_s_at	pleckstrin homology-like domain, family A, member 2	2.06	$2.77x10^{-5}$
205552_s_at	2',5'-oligoadenylate synthetase 1, 40/46kDa	2.06	$1.77x10^{-4}$
221261_x_at	melanoma antigen family D, 4 ; melanoma antigen family D, 4	2.05	$9.60x10^{-3}$
201268_at	non-metastatic cells 2, protein (NM23B) expressed in	2.05	$4.03x10^{-13}$
214581_x_at	tumor necrosis factor receptor superfamily, member 21	2.05	$9.32x10^{-5}$
205796_at	t-complex 11 (mouse) like 1	2.04	$1.79x10^{-5}$
210517_s_at	A kinase (PRKA) anchor protein (gravin) 12	2.04	$3.68x10^{-4}$
202483_s_at	RAN binding protein 1	2.04	$3.61x10^{-8}$
204470_at	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	2.04	$2.90x10^{-3}$
204750_s_at	desmocollin 2	2.04	$1.21x10^{-4}$
209122_at	adipose differentiation-related protein	2.04	$7.91x10^{-6}$
201755_at	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	2.04	$2.57x10^{-3}$
212481_s_at	tropomyosin 4	2.03	$9.68x10^{-10}$
209260_at	stratifin	2.03	$4.64x10^{-4}$
208511_at	pituitary tumor-transforming 3	2.03	$4.67x10^{-4}$
219532_at	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4	2.03	$7.93x10^{-3}$
201263_at	threonyl-tRNA synthetase	2.02	$3.06x10^{-9}$
204341_at	tripartite motif-containing 16 ; tripartite motif containing 16-like	2.02	$1.13x10^{-5}$
220532_s_at	LR8 protein	2.02	$2.20x10^{-4}$
203414_at	monocyte to macrophage differentiation-associated	2.02	$6.09x10^{-8}$
205574_x_at	bone morphogenetic protein 1	2.02	$5.64x10^{-7}$
209567_at	RRS1 ribosome biogenesis regulator homolog (S. cerevisiae)	2.02	$1.68x10^{-5}$
219257_s_at	sphingosine kinase 1	2.02	$2.17x10^{-9}$
202431_s_at	v-myc myelocytomatosis viral oncogene homolog (avian)	2.02	$3.67x10^{-4}$
217738_at	pre-B-cell colony enhancing factor 1	2.02	$4.80x10^{-6}$
208381_s_at	sphingosine-1-phosphate lyase 1	2.02	$2.34x10^{-3}$
209642_at	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	2.02	$2.34x10^{-6}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 12)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
201577_at	non-metastatic cells 1, protein (NM23A) expressed in	2.01	$1.43x10^{-8}$
211762_s_at	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	2.01	$1.68x10^{-8}$
207463_x_at	protease, serine, 3 (mesotrypsin)	2.01	$9.83x10^{-3}$
AFFX-HUMGAPDH/M33197_5_at	glyceraldehyde-3-phosphate dehydrogenase	2.01	$1.52x10^{-11}$
204817_at	extra spindle poles like 1 (S. cerevisiae)	2.01	$4.66x10^{-4}$
202235_at	solute carrier family 16 (monocarboxylic acid transporters), member 1	2.01	$4.09x10^{-3}$
202755_s_at	glypican 1	2.01	$4.38x10^{-3}$
213599_at	Opa interacting protein 5	2.00	$1.54x10^{-6}$
217919_s_at	mitochondrial ribosomal protein L42	2.00	$3.03x10^{-8}$
217785_s_at	SNARE protein Ykt6	2.00	$3.26x10^{-6}$
203799_at	CD302 antigen	0.50	$7.83x10^{-4}$
205671_s_at	major histocompatibility complex, class II, DO beta	0.50	$5.06x10^{-4}$
209807_s_at	nuclear factor I/X (CCAAT-binding transcription factor)	0.50	$1.12x10^{-3}$
214053_at	V-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	0.50	$8.72x10^{-8}$
216278_at	KIAA0256 gene product	0.50	$4.17x10^{-3}$
209829_at	chromosome 6 open reading frame 32	0.50	$9.54x10^{-4}$
211466_at	nuclear factor I/B	0.50	$2.12x10^{-3}$
213856_at	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	0.50	$5.45x10^{-3}$
204266_s_at	choline kinase alpha	0.50	$9.58x10^{-7}$
210051_at	Rap guanine nucleotide exchange factor (GEF) 3	0.50	$3.20x10^{-3}$
213472_at	heterogeneous nuclear ribonucleoprotein H1 (H)	0.50	$1.02x10^{-3}$
215314_at	Ankyrin 3, node of Ranvier (ankyrin G)	0.50	$5.64x10^{-6}$
210108_at	calcium channel, voltage-dependent, L type, alpha 1D subunit	0.50	$9.51x10^{-4}$
203365_s_at	matrix metallopeptidase 15 (membrane-inserted)	0.50	$6.05x10^{-3}$
203994_s_at	chromosome 21 open reading frame 2	0.50	$1.22x10^{-3}$
212510_at	glycerol-3-phosphate dehydrogenase 1-like	0.50	$2.93x10^{-6}$
202222_s_at	desmin	0.49	$6.62x10^{-5}$
205259_at	nuclear receptor subfamily 3, group C, member 2	0.49	$1.35x10^{-4}$
212774_at	zinc finger protein 238	0.49	$5.77x10^{-7}$
208135_at	HNF1 homeobox B	0.49	$5.91x10^{-4}$
220221_at	vacuolar protein sorting 13D (yeast)	0.49	$1.91x10^{-4}$
212841_s_at	PTPRF interacting protein, binding protein 2 (liprin beta 2)	0.49	$3.47x10^{-5}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 13)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
204039_at	CCAAT/enhancer binding protein (C/EBP), alpha	0.49	$2.83\text{x}10^{-5}$
215536_at	major histocompatibility complex, class II, DQ beta 2	0.49	$7.87\text{x}10^{-3}$
222150_s_at	pigeon homolog (Drosophila)	0.49	$2.44\text{x}10^{-8}$
214250_at	nuclear mitotic apparatus protein 1	0.49	$3.50\text{x}10^{-3}$
204130_at	hydroxysteroid (11-beta) dehydrogenase 2	0.49	$8.33\text{x}10^{-3}$
204802_at	Ras-related associated with diabetes	0.49	$1.56\text{x}10^{-3}$
206714_at	arachidonate 15-lipoxygenase, second type	0.49	$2.01\text{x}10^{-3}$
217422_s_at	CD22 antigen ; myelin associated glycoprotein	0.49	$3.30\text{x}10^{-3}$
219127_at	proline rich 15-like	0.49	$2.23\text{x}10^{-5}$
204497_at	adenylate cyclase 9	0.48	$6.13\text{x}10^{-6}$
201540_at	four and a half LIM domains 1	0.48	$1.66\text{x}10^{-3}$
205316_at	Solute carrier family 15 (H+/peptide transporter), member 2	0.48	$2.99\text{x}10^{-5}$
205619_s_at	mesenchyme homeo box 1	0.48	$9.91\text{x}10^{-5}$
214657_s_at	Trophoblast-derived noncoding RNA	0.48	$4.02\text{x}10^{-4}$
214035_x_at	nuclear pore complex interacting protein-like 3	0.48	$1.94\text{x}10^{-7}$
219389_at	sushi domain containing 4	0.48	$4.56\text{x}10^{-3}$
220168_at	cancer susceptibility candidate 1	0.48	$4.63\text{x}10^{-5}$
219136_s_at	lipase maturation factor 1	0.48	$7.69\text{x}10^{-4}$
213212_x_at	golgin A6 family-like 5 (pseudogene)	0.48	$4.06\text{x}10^{-6}$
220677_s_at	ADAM metallopeptidase with thrombospondin type 1 motif, 8	0.48	$3.63\text{x}10^{-3}$
214964_at	KIAA1856 protein	0.48	$1.86\text{x}10^{-4}$
209840_s_at	leucine rich repeat neuronal 3	0.48	$4.87\text{x}10^{-4}$
205102_at	transmembrane protease, serine 2	0.48	$7.63\text{x}10^{-5}$
220512_at	deleted in liver cancer 1	0.48	$5.63\text{x}10^{-4}$
206290_s_at	regulator of G-protein signalling 7	0.48	$6.20\text{x}10^{-3}$
208206_s_at	RAS guanyl releasing protein 2 (calcium and DAG-regulated)	0.48	$7.07\text{x}10^{-4}$
219095_at	phospholipase A2, group IVB (cytosolic)	0.48	$1.16\text{x}10^{-5}$
220369_at	KIAA2010	0.48	$9.82\text{x}10^{-4}$
213364_s_at	sorting nexin 1	0.48	$1.20\text{x}10^{-5}$
215784_at	CD1E antigen, e polypeptide	0.48	$7.58\text{x}10^{-3}$
204081_at	neurogranin (protein kinase C substrate, RC3)	0.47	$5.48\text{x}10^{-4}$
205906_at	forkhead box J1	0.47	$1.99\text{x}10^{-4}$
221602_s_at	Fas apoptotic inhibitory molecule 3	0.47	$3.68\text{x}10^{-3}$
213121_at	small nuclear ribonucleoprotein 70kDa polypeptide (RNP antigen)	0.47	$4.85\text{x}10^{-5}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 14)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
214249_at	tripartite motif-containing 2	0.47	$3.58x10^{-3}$
215339_at	natural killer-tumor recognition sequence	0.47	$5.04x10^{-3}$
215578_at	Gephyrin	0.47	$4.42x10^{-5}$
221828_s_at	chromosome 9 open reading frame 28	0.47	$7.68x10^{-4}$
222071_s_at	solute carrier organic anion transporter family, member 4C1	0.47	$1.16x10^{-3}$
205531_s_at	glutaminase 2 (liver, mitochondrial)	0.47	$2.67x10^{-3}$
210258_at	regulator of G-protein signalling 13	0.47	$2.64x10^{-4}$
210528_at	major histocompatibility complex, class I-related	0.47	$7.82x10^{-5}$
219197_s_at	signal peptide, CUB domain, EGF-like 2	0.47	$2.93x10^{-4}$
221427_s_at	cyclin L2	0.47	$8.87x10^{-4}$
210944_s_at	calpain 3, (p94)	0.47	$6.18x10^{-8}$
39249_at	aquaporin 3	0.47	$7.01x10^{-3}$
207073_at	cyclin-dependent kinase-like 2 (CDC2-related kinase)	0.47	$1.34x10^{-3}$
220150_s_at	chromosome 6 open reading frame 60	0.47	$5.30x10^{-8}$
205101_at	class II, major histocompatibility complex, transactivator	0.47	$3.60x10^{-6}$
212599_at	autism susceptibility candidate 2	0.47	$3.34x10^{-4}$
215474_at	Chromosome 8 open reading frame 36	0.47	$2.29x10^{-3}$
205343_at	sulfotransferase family, cytosolic, 1C, member 1	0.47	$1.94x10^{-4}$
210328_at	glycine N-methyltransferase	0.46	$5.24x10^{-4}$
217929_s_at	KIAA0319-like	0.46	$5.38x10^{-3}$
220220_at	leucine rich repeat containing 37, member A4 (pseudogene)	0.46	$8.42x10^{-5}$
220296_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 (GalNAc-T10)	0.46	$5.63x10^{-5}$
205464_at	sodium channel, nonvoltage-gated 1, beta (Liddle syndrome)	0.46	$9.32x10^{-5}$
205978_at	klotho	0.46	$6.61x10^{-5}$
207896_s_at	deleted in lung and esophageal cancer 1	0.46	$4.44x10^{-4}$
215800_at	dual oxidase 1	0.46	$2.48x10^{-7}$
204201_s_at	protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase)	0.46	$3.62x10^{-5}$
220187_at	STEAP family member 4	0.46	$2.45x10^{-3}$
215785_s_at	cytoplasmic FMR1 interacting protein 2	0.46	$1.77x10^{-5}$
215807_s_at	plexin B1	0.45	$3.02x10^{-3}$
206007_at	proteoglycan 4	0.45	$1.10x10^{-3}$
215783_s_at	alkaline phosphatase, liver/bone/kidney	0.45	$5.96x10^{-4}$
221874_at	KIAA1324	0.45	$1.31x10^{-3}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 15)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
215147_at	CUG triplet repeat, RNA binding protein 2	0.45	5.43x10^{-4}
206330_s_at	SHC (Src homology 2 domain containing) transforming protein 3	0.45	1.32x10^{-4}
205414_s_at	KIAA0672 gene product	0.45	2.59x10^{-4}
220456_at	chromosome 20 open reading frame 38	0.45	4.25x10^{-4}
215782_at	Ras-like GTPase-like	0.45	3.82x10^{-5}
208451_s_at	complement component 4A ; complement component 4B ; complement component 4B, telomeric	0.44	3.55x10^{-5}
210239_at	iroquois homeobox protein 5	0.44	6.24x10^{-3}
220704_at	zinc finger protein, subfamily 1A, 1 (Ikaros)	0.44	3.35x10^{-4}
214479_at	GDNF family receptor alpha 3	0.44	8.89x10^{-3}
221577_x_at	growth differentiation factor 15	0.44	4.81x10^{-4}
207355_at	solute carrier family 1 (glutamate transporter), member 7	0.44	4.76x10^{-3}
204719_at	ATP-binding cassette, sub-family A (ABC1), member 8	0.44	9.32x10^{-4}
210323_at	tektin 2 (testicular)	0.44	4.40x10^{-5}
210084_x_at	tryptase alpha/beta 1	0.44	6.43x10^{-3}
203892_at	WAP four-disulfide core domain 2	0.44	1.87x10^{-3}
205839_s_at	benzodiazapine receptor (peripheral) associated protein 1	0.43	1.88x10^{-8}
215575_at	phosphodiesterase 4D interacting protein (myomegalin)	0.43	3.73x10^{-4}
220061_at	acyl-CoA synthetase medium-chain family member 5	0.43	5.08x10^{-5}
209368_at	epoxide hydrolase 2, cytoplasmic	0.43	5.18x10^{-4}
213106_at	ATPase, aminophospholipid transporter (APLT), Class I, type 8A, member 1	0.43	2.46x10^{-6}
204124_at	solute carrier family 34 (sodium phosphate), member 2	0.43	1.27x10^{-3}
219230_at	transmembrane protein 100	0.43	6.47x10^{-4}
206019_at	RNA binding motif protein 19	0.43	8.43x10^{-4}
218087_s_at	sorbin and SH3 domain containing 1	0.42	7.51x10^{-4}
220055_at	zinc finger protein 287	0.42	2.42x10^{-4}
214602_at	collagen, type IV, alpha 4	0.42	3.32x10^{-3}
219857_at	chromosome 10 open reading frame 81	0.42	5.92x10^{-3}
214188_at	Hexamethylene bis-acetamide inducible 1	0.42	2.52x10^{-4}
221272_s_at	chromosome 1 open reading frame 21	0.42	4.26x10^{-5}
222073_at	collagen, type IV, alpha 3 (Goodpasture antigen)	0.42	1.40x10^{-5}
206255_at	B lymphoid tyrosine kinase	0.42	1.91x10^{-3}
212448_at	neural precursor cell expressed, developmentally down-regulated 4-like	0.42	6.51x10^{-6}
205500_at	complement component 5	0.42	5.83x10^{-4}
205776_at	flavin containing monooxygenase 5	0.42	3.98x10^{-4}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 16)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
207837_at	RNA binding protein with multiple splicing	0.41	2.66×10^{-5}
205569_at	lysosomal-associated membrane protein 3	0.41	4.24×10^{-4}
208434_at	myelodysplasia syndrome 1	0.41	3.71×10^{-4}
213802_at	Protease, serine, 12 (neurotrypsin, motopsin)	0.41	2.91×10^{-3}
203911_at	RAP1, GTPase activating protein 1	0.41	6.49×10^{-6}
204428_s_at	lecithin-cholesterol acyltransferase	0.41	8.18×10^{-4}
209793_at	glutamate receptor, ionotropic, AMPA 1	0.41	2.85×10^{-5}
204865_at	carbonic anhydrase III, muscle specific	0.41	2.76×10^{-5}
205200_at	C-type lectin domain family 3, member B	0.41	3.50×10^{-4}
205883_at	zinc finger and BTB domain containing 16	0.41	1.15×10^{-5}
216134_at	FERM domain containing 4B	0.41	3.37×10^{-5}
215921_at	polycystic kidney disease 1 (autosomal dominant) pseudogene 1	0.41	2.16×10^{-4}
214890_s_at	DKFZP564J102 protein	0.40	1.53×10^{-3}
220858_at	sorbin and SH3 domain containing 2	0.40	8.72×10^{-5}
220889_s_at	carbonic anhydrase X	0.40	3.47×10^{-3}
40284_at	forkhead box A2	0.40	3.11×10^{-3}
209763_at	chordin-like 1	0.40	1.04×10^{-5}
210113_s_at	NACHT, leucine rich repeat and PYD (pyrin domain) containing 1	0.40	6.28×10^{-5}
210571_s_at	cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acetylneuraminate monooxygenase)	0.40	1.16×10^{-5}
207569_at	v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	0.40	8.83×10^{-6}
220624_s_at	E74-like factor 5 (ets domain transcription factor)	0.40	7.35×10^{-3}
203548_s_at	lipoprotein lipase	0.40	5.67×10^{-3}
215355_at	POU domain, class 2, transcription factor 3	0.39	2.79×10^{-3}
219525_at	solute carrier family 47, member 1	0.39	7.03×10^{-3}
206010_at	hyaluronan binding protein 2	0.39	5.74×10^{-5}
206069_s_at	acyl-Coenzyme A dehydrogenase, long chain	0.39	6.26×10^{-4}
213493_at	sushi, nidogen and EGF-like domains 1	0.39	1.00×10^{-4}
217418_x_at	membrane-spanning 4-domains, subfamily A, member 1	0.39	1.07×10^{-5}
201497_x_at	myosin, heavy polypeptide 11, smooth muscle	0.39	8.38×10^{-3}
215184_at	death-associated protein kinase 2	0.39	1.55×10^{-6}
205969_at	arylacetamide deacetylase (esterase)	0.39	6.26×10^{-3}
201785_at	ribonuclease, RNase A family, 1 (pancreatic)	0.38	3.94×10^{-5}
212695_at	cryptochrome 2 (photolyase-like)	0.38	1.47×10^{-5}
220630_s_at	chitinase, acidic	0.38	4.52×10^{-5}

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 17)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
208791_at	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)	0.38	$1.56x10^{-3}$
216468_s_at	zinc finger protein 682	0.38	$1.78x10^{-4}$
201525_at	apolipoprotein D	0.38	$2.09x10^{-4}$
203638_s_at	fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome)	0.38	$1.03x10^{-3}$
202295_s_at	cathepsin H	0.37	$3.27x10^{-8}$
208609_s_at	tenascin XA pseudogene ; tenascin XB	0.37	$3.61x10^{-7}$
220269_at	zinc finger, B-box domain containing	0.37	$1.66x10^{-3}$
215563_s_at	macrophage stimulating, pseudogene 9	0.37	$5.89x10^{-6}$
220156_at	EF-hand calcium binding domain 1	0.37	$2.05x10^{-4}$
202992_at	complement component 7	0.37	$2.48x10^{-3}$
203924_at	glutathione S-transferase A1	0.36	$3.27x10^{-3}$
214774_x_at	trinucleotide repeat containing 9	0.36	$2.30x10^{-3}$
204487_s_at	potassium voltage-gated channel, KQT-like subfamily, member 1	0.36	$2.36x10^{-3}$
214203_s_at	proline dehydrogenase (oxidase) 1	0.36	$3.84x10^{-4}$
37004_at	surfactant, pulmonary-associated protein B	0.36	$2.08x10^{-3}$
220622_at	leucine rich repeat containing 31	0.36	$2.35x10^{-3}$
205591_at	olfactomedin 1	0.35	$1.82x10^{-3}$
209460_at	4-aminobutyrate aminotransferase	0.35	$1.14x10^{-7}$
215350_at	spectrin repeat containing, nuclear envelope 1	0.35	$1.24x10^{-5}$
208606_s_at	wingless-type MMTV integration site family, member 4	0.34	$1.31x10^{-4}$
217133_x_at	cytochrome P450, family 2, subfamily B, polypeptide 6	0.34	$1.26x10^{-6}$
213317_at	Chloride intracellular channel 5	0.34	$2.41x10^{-4}$
216373_at	transmembrane anterior posterior transformation 1	0.34	$2.53x10^{-5}$
210119_at	potassium inwardly-rectifying channel, subfamily J, member 15	0.34	$5.34x10^{-5}$
205654_at	complement component 4 binding protein, alpha	0.34	$1.52x10^{-3}$
208250_s_at	deleted in malignant brain tumors 1	0.34	$2.64x10^{-4}$
204437_s_at	folate receptor 1 (adult)	0.34	$8.16x10^{-5}$
206651_s_at	carboxypeptidase B2 (plasma, carboxypeptidase U)	0.34	$7.41x10^{-4}$
211024_s_at	thyroid transcription factor 1 ; thyroid transcription factor 1	0.33	$3.30x10^{-4}$
204424_s_at	LIM domain only 3 (rhombotin-like 2)	0.32	$2.09x10^{-5}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 18)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
216000_at	KIAA0484 protein	0.32	$2.09\mathrm{x}10^{-6}$
218922_s_at	LAG1 longevity assurance homolog 4 (S. cerevisiae)	0.32	$3.99\mathrm{x}10^{-5}$
215559_at	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	0.32	$1.13\mathrm{x}10^{-5}$
205216_s_at	apolipoprotein H (beta-2-glycoprotein I)	0.32	$4.66\mathrm{x}10^{-6}$
206742_at	c-fos induced growth factor (vascular endothelial growth factor D)	0.32	$3.76\mathrm{x}10^{-6}$
204343_at	ATP-binding cassette, sub-family A (ABC1), member 3	0.32	$7.82\mathrm{x}10^{-5}$
218835_at	surfactant, pulmonary-associated protein A2	0.32	$3.10\mathrm{x}10^{-3}$
209614_at	alcohol dehydrogenase 1A (class I), alpha polypeptide ; alcohol dehydrogenase IB (class I), beta polypeptide ; alcohol dehydrogenase 1C (class I), gamma polypeptide	0.32	$4.05\mathrm{x}10^{-6}$
210081_at	advanced glycosylation end product-specific receptor	0.31	$9.00\mathrm{x}10^{-5}$
212951_at	G protein-coupled receptor 116	0.31	$5.63\mathrm{x}10^{-7}$
208335_s_at	Duffy blood group	0.31	$1.86\mathrm{x}10^{-3}$
205496_at	KIAA0408	0.31	$6.32\mathrm{x}10^{-6}$
204379_s_at	fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism)	0.31	$4.39\mathrm{x}10^{-4}$
205489_at	crystallin, mu	0.31	$2.48\mathrm{x}10^{-7}$
215554_at	glycosylphosphatidylinositol specific phospholipase D1	0.31	$2.76\mathrm{x}10^{-5}$
204213_at	polymeric immunoglobulin receptor	0.31	$6.39\mathrm{x}10^{-7}$
213451_x_at	tenascin XB	0.30	$2.04\mathrm{x}10^{-6}$
221646_s_at	zinc finger, DHHC-type containing 11	0.30	$7.68\mathrm{x}10^{-9}$
205544_s_at	complement component (3d/Epstein Barr virus) receptor 2	0.30	$4.88\mathrm{x}10^{-4}$
205987_at	CD1C antigen, c polypeptide	0.29	$1.62\mathrm{x}10^{-4}$
210121_at	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2	0.29	$3.84\mathrm{x}10^{-4}$
209904_at	troponin C type 1 (slow)	0.29	$1.12\mathrm{x}10^{-3}$
205043_at	cystic fibrosis transmembrane conductance regulator, ATP-binding cassette (sub-family C, member 7)	0.29	$4.57\mathrm{x}10^{-4}$
214433_s_at	selenium binding protein 1 ; selenium binding protein 1	0.28	$1.02\mathrm{x}10^{-9}$
205614_x_at	macrophage stimulating 1 (hepatocyte growth factor-like)	0.27	$1.91\mathrm{x}10^{-8}$
204754_at	Hepatic leukemia factor	0.27	$3.95\mathrm{x}10^{-6}$
210382_at	secretin receptor	0.26	$9.21\mathrm{x}10^{-7}$
210067_at	aquaporin 4	0.26	$3.20\mathrm{x}10^{-3}$
205576_at	serpin peptidase inhibitor, clade D (heparin cofactor), member 1	0.25	$2.85\mathrm{x}10^{-4}$
220003_at	leucine rich repeat containing 36	0.25	$2.33\mathrm{x}10^{-6}$

Supplemental Gene List III. Transcriptomal Differences between Basal Cell (BC)-high Adenocarcinoma (adenoCa) *vs* BC-low AdenoCa (cont. page 19)

		BC-high adenoCa *vs* BC-low adenoCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
205700_at	hydroxysteroid (17-beta) dehydrogenase 6	0.25	7.40×10^{-6}
219476_at	chromosome 1 open reading frame 116	0.25	4.04×10^{-5}
206311_s_at	phospholipase A2, group IB (pancreas)	0.24	2.21×10^{-6}
221132_at	claudin 18	0.24	2.72×10^{-3}
208498_s_at	amylase, alpha 1A; salivary ; amylase, alpha 1B; salivary ; amylase, alpha 1C; salivary ; amylase, alpha 2A; pancreatic ; amylase, alpha 2B; pancreatic	0.23	5.94×10^{-7}
38691_s_at	surfactant, pulmonary-associated protein C	0.23	9.97×10^{-5}
205185_at	serine peptidase inhibitor, Kazal type 5	0.22	4.68×10^{-5}
205979_at	secretoglobin, family 2A, member 1	0.22	2.15×10^{-3}
209613_s_at	alcohol dehydrogenase IB (class I), beta polypeptide	0.21	3.81×10^{-5}
205261_at	progastricsin (pepsinogen C)	0.19	1.97×10^{-5}
205725_at	secretoglobin, family 1A, member 1 (uteroglobin)	0.18	3.28×10^{-4}
210550_s_at	Ras protein-specific guanine nucleotide-releasing factor 1	0.18	1.05×10^{-5}
210272_at	cytochrome P450, family 2, subfamily B, polypeptide 7 pseudogene 1	0.18	1.44×10^{-7}
204811_s_at	calcium channel, voltage-dependent, alpha 2/delta subunit 2	0.17	9.25×10^{-12}
214199_at	surfactant, pulmonary-associated protein D	0.17	1.25×10^{-5}
218876_at	brain specific protein	0.16	6.58×10^{-7}
204712_at	WNT inhibitory factor 1	0.14	1.37×10^{-5}
210096_at	cytochrome P450, family 4, subfamily B, polypeptide 1	0.10	2.81×10^{-9}

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival

Probe ID	Gene name	OS<2yr *vs* OS>5yr in adenoCa		Airway BC signature
		Fold-change	p value	
37892_at	collagen, type XI, alpha 1	5.62	$3.33x10^{-2}$	
219491_at	leucine rich repeat and fibronectin type III domain containing 4	3.73	$3.59x10^{-2}$	
218468_s_at	gremlin 1, cysteine knot superfamily, homolog (Xenopus laevis)	3.68	$3.69x10^{-2}$	
205403_at	interleukin 1 receptor, type II	3.64	$7.12x10^{-3}$	
219918_s_at	asp (abnormal spindle)-like, microcephaly associated (Drosophila)	3.14	$4.06x10^{-2}$	
209825_s_at	uridine-cytidine kinase 2	3.09	$2.01x10^{-2}$	O
209875_s_at	secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	2.95	$1.67x10^{-2}$	
201890_at	ribonucleotide reductase M2 polypeptide	2.94	$4.26x10^{-2}$	
218542_at	chromosome 10 open reading frame 3	2.94	$4.06x10^{-2}$	
209900_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 1	2.93	$3.00x10^{-2}$	
212396_s_at	KIAA0090	2.82	$3.69x10^{-2}$	
204641_at	NIMA (never in mitosis gene a)-related kinase 2	2.76	$3.09x10^{-2}$	
212353_at	sulfatase 1	2.74	$2.29x10^{-2}$	
213338_at	Ras-induced senescence 1	2.68	$3.15x10^{-2}$	O
202855_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 3	2.67	$2.86x10^{-2}$	O
205568_at	aquaporin 9	2.65	$3.80x10^{-2}$	
205542_at	six transmembrane epithelial antigen of the prostate 1	2.55	$3.15x10^{-2}$	
217553_at	similar to Six transmembrane epithelial antigen of prostate	2.55	$2.86x10^{-2}$	O
202363_at	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican)	2.54	$2.90x10^{-2}$	
203362_s_at	MAD2 mitotic arrest deficient-like 1 (yeast)	2.53	$4.21x10^{-2}$	
202580_x_at	forkhead box M1	2.52	$5.93x10^{-3}$	
201037_at	phosphofructokinase, platelet	2.51	$2.57x10^{-2}$	
210809_s_at	periostin, osteoblast specific factor	2.49	$4.88x10^{-2}$	
213008_at	hypothetical protein FLJ10719	2.48	$2.14x10^{-2}$	
203213_at	Cell division cycle 2, G1 to S and G2 to M	2.46	$4.08x10^{-2}$	

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival (cont. page 2)

		OS<2yr *vs* OS>5yr in adenoCa		
Probe ID	**Gene name**	**Fold-change**	**p value**	**Airway BC signature**
215812_s_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 8 ; similar to sodium- and chloride-dependent creatine transporter	2.42	4.04×10^{-2}	O
212992_at	chromosome 14 open reading frame 78	2.40	2.57×10^{-2}	O
204825_at	maternal embryonic leucine zipper kinase	2.38	4.21×10^{-2}	
210052_s_at	TPX2, microtubule-associated, homolog (Xenopus laevis)	2.38	2.39×10^{-2}	
218663_at	chromosome condensation protein G	2.36	3.04×10^{-2}	
202828_s_at	matrix metallopeptidase 14 (membrane-inserted)	2.35	3.00×10^{-2}	
201291_s_at	topoisomerase (DNA) II alpha 170kDa	2.33	2.32×10^{-2}	
217901_at	Desmoglein 2	2.33	3.82×10^{-3}	O
202094_at	baculoviral IAP repeat-containing 5 (survivin)	2.32	3.15×10^{-2}	
218843_at	fibronectin type III domain containing 4	2.31	2.65×10^{-2}	
201506_at	transforming growth factor, beta-induced, 68kDa	2.27	1.51×10^{-2}	O
202949_s_at	four and a half LIM domains 2	2.24	4.22×10^{-2}	O
204750_s_at	desmocollin 2	2.23	2.90×10^{-2}	O
205393_s_at	CHK1 checkpoint homolog (S. pombe)	2.22	4.06×10^{-2}	O
210132_at	ephrin-A3	2.22	2.14×10^{-2}	
206026_s_at	tumor necrosis factor, alpha-induced protein 6	2.21	2.29×10^{-2}	
218726_at	hypothetical protein DKFZp762E1312	2.21	4.26×10^{-2}	
218355_at	kinesin family member 4A	2.19	3.15×10^{-2}	
203806_s_at	Fanconi anemia, complementation group A ; Fanconi anemia, complementation group A	2.17	3.15×10^{-2}	
218644_at	pleckstrin 2	2.17	2.29×10^{-2}	O
204444_at	kinesin family member 11	2.16	3.15×10^{-2}	
204620_s_at	chondroitin sulfate proteoglycan 2 (versican)	2.16	3.98×10^{-2}	
215723_s_at	phospholipase D1, phophatidylcholine-specific	2.15	3.80×10^{-2}	
204146_at	RAD51 associated protein 1	2.14	3.15×10^{-2}	
218585_s_at	denticleless homolog (Drosophila)	2.14	4.02×10^{-2}	
202998_s_at	lysyl oxidase-like 2	2.11	2.14×10^{-2}	O
203418_at	cyclin A2	2.08	4.49×10^{-2}	
212242_at	tubulin, alpha 1 (testis specific)	2.08	2.29×10^{-2}	O
220658_s_at	aryl hydrocarbon receptor nuclear translocator-like 2	2.08	4.21×10^{-2}	O

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival (cont. page 3)

		OS<2yr *vs* OS>5yr in adenoCa		
Probe ID	**Gene name**	**Fold-change**	**p value**	**Airway BC signature**
213943_at	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	2.07	2.29×10^{-2}	
218009_s_at	protein regulator of cytokinesis 1	2.07	2.29×10^{-2}	
220002_at	hypothetical protein FLJ10157	2.06	4.08×10^{-2}	
214702_at	fibronectin 1	2.02	2.94×10^{-2}	O
215076_s_at	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)	2.02	4.02×10^{-2}	
205066_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 1	2.00	2.29×10^{-2}	
209955_s_at	fibroblast activation protein, alpha	2.00	4.10×10^{-2}	
210074_at	cathepsin L2	2.00	3.15×10^{-2}	O
219978_s_at	nucleolar and spindle associated protein 1	2.00	4.72×10^{-2}	
204237_at	GULP, engulfment adaptor PTB domain containing 1	1.98	3.51×10^{-2}	
204962_s_at	centromere protein A, 17kDa	1.98	4.94×10^{-2}	
219888_at	sperm associated antigen 4	1.97	3.15×10^{-2}	
210845_s_at	plasminogen activator, urokinase receptor	1.96	3.15×10^{-2}	
222077_s_at	Rac GTPase activating protein 1	1.93	2.57×10^{-2}	
217744_s_at	PERP, TP53 apoptosis effector	1.93	2.14×10^{-2}	
219032_x_at	opsin 3 (encephalopsin, panopsin)	1.93	3.33×10^{-2}	O
206364_at	kinesin family member 14	1.92	2.14×10^{-2}	
204531_s_at	breast cancer 1, early onset	1.92	4.88×10^{-2}	
204342_at	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 24	1.90	3.00×10^{-2}	
205122_at	transmembrane protein with EGF-like and two follistatin-like domains 1	1.89	3.15×10^{-2}	O
208079_s_at	serine/threonine kinase 6	1.88	3.15×10^{-2}	
202887_s_at	DNA-damage-inducible transcript 4	1.86	4.08×10^{-2}	
205574_x_at	bone morphogenetic protein 1	1.81	4.26×10^{-2}	
201669_s_at	myristoylated alanine-rich protein kinase C substrate	1.81	4.29×10^{-2}	
213125_at	olfactomedin-like 2B	1.81	4.92×10^{-2}	
219366_at	apoptosis, caspase activation inhibitor	1.80	3.15×10^{-2}	O
204298_s_at	lysyl oxidase	1.79	2.39×10^{-2}	
45633_at	hypothetical protein FLJ13912	1.79	3.00×10^{-2}	
209642_at	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	1.78	3.00×10^{-2}	

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival (cont. page 4)

		OS<2yr *vs* OS>5yr in adenoCa		
Probe ID	**Gene name**	**Fold-change**	**p value**	**Airway BC signature**
203755_at	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	1.78	$4.26x10^{-2}$	
203414_at	monocyte to macrophage differentiation-associated	1.77	$4.02x10^{-2}$	O
201930_at	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)	1.77	$2.29x10^{-2}$	
222118_at	uncharacterized bone marrow protein BM039	1.77	$3.46x10^{-2}$	
201250_s_at	solute carrier family 2 (facilitated glucose transporter), member 1	1.76	$3.15x10^{-2}$	
204244_s_at	activator of S phase kinase	1.76	$2.69x10^{-2}$	O
201202_at	proliferating cell nuclear antigen	1.75	$2.01x10^{-2}$	
203358_s_at	enhancer of zeste homolog 2 (Drosophila)	1.75	$3.15x10^{-2}$	
200737_at	phosphoglycerate kinase 1	1.74	$1.67x10^{-2}$	
202087_s_at	cathepsin L	1.73	$4.06x10^{-2}$	
218349_s_at	Zwilch, kinetochore associated, homolog (Drosophila)	1.73	$3.96x10^{-2}$	
221565_s_at	family with sequence similarity 26, member B	1.73	$4.06x10^{-2}$	
219257_s_at	sphingosine kinase 1	1.72	$2.29x10^{-2}$	
220893_at	uncharacterized gastric protein ZA52P	1.72	$2.14x10^{-2}$	
203145_at	sperm associated antigen 5	1.71	$3.00x10^{-2}$	
218854_at	squamous cell carcinoma antigen recognized by T cells 2	1.70	$6.25x10^{-3}$	
219306_at	kinesin family member 15	1.69	$4.02x10^{-2}$	
215084_s_at	leucine rich repeat containing 42	1.69	$2.39x10^{-2}$	O
218193_s_at	golgi transport 1 homolog B (S. cerevisiae)	1.68	$2.39x10^{-2}$	O
205024_s_at	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)	1.67	$3.15x10^{-2}$	
AFFX-HUMGAPDH/M33197_5_at	glyceraldehyde-3-phosphate dehydrogenase	1.65	$4.80x10^{-2}$	
214266_s_at	PDZ and LIM domain 7 (enigma)	1.65	$3.04x10^{-2}$	
203594_at	RNA terminal phosphate cyclase domain 1	1.64	$4.50x10^{-2}$	
220926_s_at	chromosome 1 open reading frame 22	1.64	$1.51x10^{-2}$	
202338_at	thymidine kinase 1, soluble	1.64	$2.39x10^{-2}$	
203137_at	Wilms tumor 1 associated protein	1.64	$4.49x10^{-2}$	
213004_at	angiopoietin-like 2	1.64	$4.02x10^{-2}$	
200897_s_at	palladin	1.60	$4.92x10^{-2}$	
201136_at	proteolipid protein 2 (colonic epithelium-enriched)	1.60	$3.57x10^{-2}$	

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival (cont. page 5)

		OS<2yr *vs* OS>5yr in adenoCa		
Probe ID	**Gene name**	**Fold-change**	**p value**	**Airway BC signature**
200756_x_at	calumenin	1.59	$2.65\mathrm{x}10^{-2}$	
200776_s_at	basic leucine zipper and W2 domains 1 ; similar to basic leucine zipper and W2 domains 1	1.58	$3.15\mathrm{x}10^{-2}$	
209709_s_at	hyaluronan-mediated motility receptor (RHAMM)	1.58	$4.08\mathrm{x}10^{-2}$	
206796_at	WNT1 inducible signaling pathway protein 1	1.58	$3.15\mathrm{x}10^{-2}$	
219874_at	solute carrier family 12 (potassium/chloride transporters), member 8	1.58	$2.14\mathrm{x}10^{-2}$	
41037_at	TEA domain family member 4	1.57	$4.43\mathrm{x}10^{-2}$	O
208854_s_at	serine/threonine kinase 24 (STE20 homolog, yeast)	1.54	$2.46\mathrm{x}10^{-2}$	
209091_s_at	SH3-domain GRB2-like endophilin B1	1.53	$3.15\mathrm{x}10^{-2}$	
208636_at	Actinin, alpha 1	1.52	$3.15\mathrm{x}10^{-2}$	O
202536_at	chromatin modifying protein 2B	1.52	$4.02\mathrm{x}10^{-2}$	
201238_s_at	capping protein (actin filament) muscle Z-line, alpha 2	1.50	$3.70\mathrm{x}10^{-2}$	
201486_at	reticulocalbin 2, EF-hand calcium binding domain	1.49	$1.51\mathrm{x}10^{-2}$	
208810_at	DnaJ (Hsp40) homolog, subfamily B, member 6	1.49	$2.39\mathrm{x}10^{-2}$	
210639_s_at	ATG5 autophagy related 5 homolog (S. cerevisiae)	1.48	$2.39\mathrm{x}10^{-2}$	
212021_s_at	antigen identified by monoclonal antibody Ki-67	1.45	$4.26\mathrm{x}10^{-2}$	
221891_x_at	heat shock 70kDa protein 8	1.45	$2.32\mathrm{x}10^{-2}$	
202897_at	protein tyrosine phosphatase, non-receptor type substrate 1	1.44	$2.24\mathrm{x}10^{-2}$	
201461_s_at	mitogen-activated protein kinase-activated protein kinase 2	1.43	$3.62\mathrm{x}10^{-2}$	
211945_s_at	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	1.40	$4.59\mathrm{x}10^{-2}$	O
201020_at	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	1.38	$1.48\mathrm{x}10^{-2}$	
205202_at	protein-L-isoaspartate (D-aspartate) O-methyltransferase	1.37	$2.29\mathrm{x}10^{-2}$	
200059_s_at	ras homolog gene family, member A ; ras homolog gene family, member A	1.37	$3.21\mathrm{x}10^{-2}$	
202142_at	COP9 constitutive photomorphogenic homolog subunit 8 (Arabidopsis)	1.34	$4.88\mathrm{x}10^{-2}$	
214629_x_at	reticulon 4	1.33	$4.22\mathrm{x}10^{-2}$	
218803_at	checkpoint with forkhead and ring finger domains	1.30	$3.46\mathrm{x}10^{-2}$	
219540_at	zinc finger protein 267	1.30	$5.17\mathrm{x}10^{-3}$	

Supplemental Gene List IV. Transcriptomal Differences between Lung Adenocarcinoma (adenoCa) Patients with Poor Survival *vs* Those with Good Survival (cont. page 6)

		OS<2yr *vs* OS>5yr in adenoCa		
Probe ID	**Gene name**	**Fold-change**	**p value**	**Airway BC signature**
201179_s_at	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	1.28	$3.15x10^{-2}$	
202529_at	phosphoribosyl pyrophosphate synthetase-associated protein 1	1.27	$3.33x10^{-2}$	

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma

		BC-high adenoCa *vs* SqCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
209278_s_at	tissue factor pathway inhibitor 2	6.29	$1.92x10^{-3}$
209016_s_at	keratin 7	6.19	$1.26x10^{-4}$
203108_at	G protein-coupled receptor, family C, group 5, member A	3.61	$4.25x10^{-4}$
205239_at	amphiregulin (schwannoma-derived growth factor)	3.52	$1.25x10^{-3}$
222108_at	adhesion molecule with Ig-like domain 2	3.52	$4.23x10^{-5}$
208161_s_at	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	3.26	$3.46x10^{-5}$
219032_x_at	opsin 3 (encephalopsin, panopsin)	2.93	$7.61x10^{-4}$
203939_at	5'-nucleotidase, ecto (CD73)	2.92	$1.25x10^{-3}$
209803_s_at	pleckstrin homology-like domain, family A, member 2	2.89	$2.13x10^{-5}$
219045_at	ras homolog gene family, member F (in filopodia)	2.59	$7.15x10^{-5}$
209135_at	aspartate beta-hydroxylase	2.56	$8.17x10^{-4}$
212298_at	neuropilin 1	2.56	$2.96x10^{-4}$
205767_at	epiregulin	2.37	$9.77x10^{-3}$
208892_s_at	dual specificity phosphatase 6	2.33	$7.92x10^{-4}$
205174_s_at	homeodomain-only protein	2.32	$2.05x10^{-2}$
211597_s_at	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	2.32	$4.62x10^{-2}$
206884_s_at	sciellin	2.27	$2.05x10^{-2}$
203910_at	Rho GTPase activating protein 29	2.26	$3.09x10^{-3}$
201474_s_at	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	2.21	$1.49x10^{-4}$
203889_at	carboxypeptidase M	2.2	$3.48x10^{-2}$
206100_at	secretory granule, neuroendocrine protein 1 (7B2 protein)	2.2	$8.59x10^{-3}$
201818_at	hypothetical protein FLJ12443	2.18	$5.44x10^{-3}$
202525_at	protease, serine, 8 (prostasin)	2.14	$4.40x10^{-4}$
212473_s_at	microtubule associated monoxygenase, calponin and LIM domain containing 2	2.09	$1.46x10^{-3}$
213524_s_at	G0/G1switch 2	2.08	$9.77x10^{-3}$
209457_at	dual specificity phosphatase 5	2.07	$1.53x10^{-3}$
217875_s_at	transmembrane, prostate androgen induced RNA	2.05	$1.95x10^{-3}$
218826_at	solute carrier family 35, member F2	2.01	$1.40x10^{-6}$
202437_s_at	cytochrome P450, family 1, subfamily B, polypeptide 1	1.98	$4.03x10^{-2}$
205234_at	solute carrier family 16 (monocarboxylic acid transporters), member 4	1.98	$6.46x10^{-3}$
202949_s_at	regulator of G-protein signalling 10	1.96	$1.46x10^{-3}$
204319_s_at	four and a half LIM domains 2	1.96	$3.03x10^{-3}$

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 2)

Probe ID	Gene name	BC-high adenoCa *vs* SqCa Fold-change	p value
203510_at	met proto-oncogene (hepatocyte growth factor receptor)	1.94	8.63×10^{-3}
218611_at	immediate early response 5	1.87	1.46×10^{-3}
218309_at	calcium/calmodulin-dependent protein kinase II inhibitor 1	1.84	2.51×10^{-2}
201666_at	TIMP metallopeptidase inhibitor 1	1.82	1.79×10^{-4}
218084_x_at	FXYD domain containing ion transport regulator 5	1.79	1.51×10^{-3}
203741_s_at	adenylate cyclase 7	1.77	5.60×10^{-3}
218368_s_at	tumor necrosis factor receptor superfamily, member 12A	1.76	1.19×10^{-2}
205180_s_at	ADAM metallopeptidase domain 8	1.75	8.59×10^{-3}
214091_s_at	glutathione peroxidase 3 (plasma)	1.73	3.71×10^{-2}
221510_s_at	glutaminase	1.72	4.62×10^{-2}
212859_x_at	metallothionein 1E (functional)	1.69	2.13×10^{-2}
201506_at	caspase recruitment domain family, member 10	1.68	3.76×10^{-2}
210026_s_at	transforming growth factor, beta-induced, 68kDa	1.68	9.77×10^{-3}
204255_s_at	cell cycle progression 1	1.67	2.21×10^{-3}
214866_at	plasminogen activator, urokinase receptor	1.67	2.17×10^{-3}
221156_x_at	vitamin D (1,25- dihydroxyvitamin D3) receptor	1.67	2.13×10^{-5}
218718_at	platelet derived growth factor C	1.66	3.05×10^{-2}
221059_s_at	coactosin-like 1 (Dictyostelium)	1.64	7.41×10^{-3}
209295_at	tumor necrosis factor receptor superfamily, member 10b	1.63	3.24×10^{-4}
205376_at	inositol polyphosphate-4-phosphatase, type II, 105kDa	1.6	9.77×10^{-3}
201850_at	capping protein (actin filament), gelsolin-like	1.59	2.14×10^{-2}
209264_s_at	tetraspanin 4	1.59	2.95×10^{-2}
200700_s_at	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2	1.58	4.47×10^{-3}
203234_at	uridine phosphorylase 1	1.56	1.80×10^{-2}
204135_at	downregulated in ovarian cancer 1	1.56	4.82×10^{-2}
203414_at	monocyte to macrophage differentiation-associated	1.55	8.17×10^{-3}
212647_at	sulfide quinone reductase-like (yeast)	1.54	3.31×10^{-3}
217995_at	zinc finger protein 281	1.54	1.48×10^{-2}
218401_s_at	related RAS viral (r-ras) oncogene homolog	1.54	5.22×10^{-4}
211945_s_at	chromosome 1 open reading frame 116	1.53	6.25×10^{-5}
213010_at	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	1.53	6.70×10^{-3}
219856_at	protein kinase C, delta binding protein	1.53	3.17×10^{-2}
200733_s_at	acyl-CoA synthetase long-chain family member 4	1.51	4.62×10^{-2}
202422_s_at	protein tyrosine phosphatase type IVA, member 1	1.51	2.34×10^{-2}
200872_at	S100 calcium binding protein A10 (annexin II ligand, calpactin I, light polypeptide (p11))	1.5	2.72×10^{-3}

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 3)

		BC-high adenoCa *vs* SqCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
207357_s_at	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 (GalNAc-T10)	1.5	2.25×10^{-3}
212242_at	tubulin, alpha 1 (testis specific)	1.49	1.49×10^{-2}
202686_s_at	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	1.48	3.13×10^{-2}
213603_s_at	AXL receptor tyrosine kinase	1.48	2.29×10^{-2}
209589_s_at	melanoma cell adhesion molecule	1.47	1.80×10^{-2}
210869_s_at	EPH receptor B2	1.47	2.84×10^{-2}
213230_at	paraneoplastic antigen	1.47	2.87×10^{-2}
203650_at	protein C receptor, endothelial (EPCR)	1.46	3.84×10^{-2}
212845_at	sterile alpha motif domain containing 4	1.44	3.17×10^{-2}
212658_at	lipoma HMGIC fusion partner-like 2	1.42	4.19×10^{-3}
202006_at	pyridoxal (pyridoxine, vitamin B6) kinase	1.41	2.48×10^{-2}
203002_at	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B	1.41	5.44×10^{-3}
203071_at	angiomotin like 2	1.41	1.59×10^{-2}
218018_at	protein tyrosine phosphatase, non-receptor type 12	1.41	9.69×10^{-3}
219256_s_at	SH3 domain and tetratricopeptide repeats 1	1.4	4.62×10^{-2}
200600_at	follistatin-like 3 (secreted glycoprotein)	1.39	2.56×10^{-2}
203592_s_at	moesin	1.39	9.44×10^{-3}
218817_at	signal peptidase complex subunit 3 homolog (S. cerevisiae)	1.39	2.29×10^{-2}
211980_at	collagen, type IV, alpha 1	1.38	2.47×10^{-2}
202720_at	tripartite motif-containing 5	1.37	6.72×10^{-3}
210705_s_at	testis derived transcript (3 LIM domains)	1.37	8.63×10^{-3}
202655_at	multiple coagulation factor deficiency 2	1.36	5.32×10^{-3}
210970_s_at	GTP binding protein 2	1.36	4.46×10^{-3}
212245_at	arginine-rich, mutated in early stage tumors	1.36	1.73×10^{-3}
221050_s_at	inhibitor of Bruton agammaglobulinemia tyrosine kinase	1.36	2.72×10^{-3}
201980_s_at	Ras suppressor protein 1	1.35	1.73×10^{-2}
212739_s_at	non-metastatic cells 4, protein expressed in	1.35	4.82×10^{-2}
52164_at	chromosome 11 open reading frame 24	1.34	3.17×10^{-2}
200788_s_at	multiple C2-domains with two transmembrane regions 1	1.32	2.94×10^{-2}
220122_at	phosphoprotein enriched in astrocytes 15	1.32	2.05×10^{-2}
204120_s_at	proline rich Gla (G-carboxyglutamic acid) 1	1.3	2.29×10^{-2}
205618_at	adenosine kinase	1.3	4.01×10^{-2}
221739_at	chromosome 19 open reading frame 10	1.3	1.09×10^{-2}
200625_s_at	CAP, adenylate cyclase-associated protein 1 (yeast)	1.28	4.14×10^{-2}
208801_at	signal recognition particle 72kDa	1.27	4.16×10^{-2}
202787_s_at	BH3-only member B protein	1.24	1.56×10^{-2}

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 4)

Probe ID	Gene name	BC-high adenoCa *vs* SqCa Fold-change	p value
218775_s_at	mitogen-activated protein kinase-activated protein kinase 3	1.24	3.84×10^{-2}
209100_at	interferon-related developmental regulator 2	1.23	2.29×10^{-2}
201319_at	myosin regulatory light chain MRCL3	1.22	4.55×10^{-2}
209190_s_at	flightless I homolog (Drosophila)	1.19	3.50×10^{-2}
212024_x_at	diaphanous homolog 1 (Drosophila)	1.19	1.32×10^{-2}
34187_at	RNA binding motif, single stranded interacting protein 2	1.15	1.44×10^{-2}
209544_at	receptor-interacting serine-threonine kinase 2	1.14	1.09×10^{-2}
220198_s_at	eukaryotic translation initiation factor 5A2	0.89	2.29×10^{-2}
218964_at	AT rich interactive domain 3B (BRIGHT- like)	0.88	2.92×10^{-2}
221414_s_at	defensin, beta 126 ; defensin, beta 126	0.86	9.34×10^{-3}
205138_s_at	pleckstrin homology domain containing, family G (with RhoGef domain) member 3	0.85	1.46×10^{-3}
212821_at	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4	0.85	2.96×10^{-3}
219532_at	uronyl-2-sulfotransferase	0.85	4.64×10^{-2}
203597_s_at	WW domain binding protein 4 (formin binding protein 21)	0.83	2.82×10^{-2}
220484_at	mucolipin 3	0.82	4.29×10^{-3}
205493_s_at	dihydropyrimidinase-like 4	0.81	7.97×10^{-3}
211846_s_at	poliovirus receptor-related 1 (herpesvirus entry mediator C; nectin)	0.79	4.62×10^{-2}
214095_at	serine hydroxymethyltransferase 2 (mitochondrial)	0.78	2.72×10^{-3}
215411_s_at	TRAF3 interacting protein 2	0.78	3.20×10^{-2}
214595_at	chromosome 8 open reading frame 20	0.77	5.57×10^{-3}
218421_at	potassium voltage-gated channel, subfamily G, member 1	0.77	2.29×10^{-2}
218777_at	ceramide kinase	0.77	9.77×10^{-3}
206376_at	ankyrin repeat domain 11	0.76	5.57×10^{-3}
219437_s_at	solute carrier family 6, member 15	0.76	2.26×10^{-2}
220255_at	Fanconi anemia, complementation group E	0.76	2.12×10^{-3}
200924_s_at	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	0.75	3.71×10^{-2}
203297_s_at	Jumonji, AT rich interactive domain 2	0.75	8.63×10^{-3}
203438_at	stanniocalcin 2	0.74	2.70×10^{-2}
219038_at	MORC family CW-type zinc finger 4	0.74	2.95×10^{-2}
200951_s_at	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1	0.71	3.84×10^{-2}
206832_s_at	mitogen-activated protein kinase 6	0.71	8.52×10^{-3}
207121_s_at	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	0.71	3.45×10^{-2}
210237_at	artemin	0.71	7.62×10^{-3}

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 5)

		BC-high adenoCa *vs* SqCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
219024_at	cyclin D2	0.71	$3.13\text{x}10^{-2}$
202067_s_at	placental growth factor, vascular endothelial growth factor-related protein	0.7	$3.21\text{x}10^{-2}$
209652_s_at	MICAL-like 1	0.7	$4.69\text{x}10^{-2}$
219855_at	low density lipoprotein receptor (familial hypercholesterolemia)	0.7	$2.13\text{x}10^{-5}$
220892_s_at	nudix (nucleoside diphosphate linked moiety X)-type motif 11	0.7	$1.59\text{x}10^{-2}$
221779_at	phosphoserine aminotransferase 1	0.7	$3.66\text{x}10^{-2}$
204203_at	CCAAT/enhancer binding protein (C/EBP), gamma	0.69	$3.30\text{x}10^{-2}$
214431_at	guanine monphosphate synthetase	0.68	$3.17\text{x}10^{-2}$
202755_s_at	schwannomin interacting protein 1	0.67	$8.63\text{x}10^{-3}$
204030_s_at	glypican 1	0.67	$2.68\text{x}10^{-2}$
215780_s_at	SET translocation (myeloid leukemia-associated)	0.67	$7.91\text{x}10^{-3}$
206581_at	basonuclin 1	0.66	$2.08\text{x}10^{-4}$
201249_at	fibulin 1	0.65	$2.21\text{x}10^{-3}$
201787_at	solute carrier family 2 (facilitated glucose transporter), member 1	0.65	$7.89\text{x}10^{-3}$
202934_at	hexokinase 2	0.63	$3.19\text{x}10^{-2}$
203562_at	fasciculation and elongation protein zeta 1 (zygin I)	0.63	$2.93\text{x}10^{-2}$
220664_at	small proline-rich protein 2C	0.62	$3.17\text{x}10^{-2}$
206307_s_at	forkhead box D1	0.61	$2.29\text{x}10^{-2}$
201656_at	integrin, alpha 6	0.59	$4.69\text{x}10^{-2}$
204165_at	WAS protein family, member 1	0.58	$1.16\text{x}10^{-2}$
206122_at	SRY (sex determining region Y)-box 15	0.58	$5.68\text{x}10^{-4}$
213358_at	KIAA0802	0.58	$5.11\text{x}10^{-3}$
201397_at	meningioma (disrupted in balanced translocation) 1	0.57	$2.87\text{x}10^{-4}$
205330_at	phosphoglycerate dehydrogenase	0.57	$4.39\text{x}10^{-3}$
202779_s_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11	0.54	$7.89\text{x}10^{-3}$
211071_s_at	ubiquitin-conjugating enzyme E2S	0.54	$4.62\text{x}10^{-2}$
212190_at	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	0.53	$2.94\text{x}10^{-2}$
201667_at	hepatoma-derived growth factor, related protein 3	0.49	$2.06\text{x}10^{-2}$
204136_at	collagen, type VII, alpha 1 (epidermolysis bullosa, dystrophic, dominant and recessive)	0.49	$2.02\text{x}10^{-4}$
207382_at	gap junction protein, alpha 1, 43kDa (connexin 43)	0.49	$4.30\text{x}10^{-5}$
209283_at	tumor protein p73-like	0.49	$4.19\text{x}10^{-3}$
209524_at	crystallin, alpha B	0.49	$6.26\text{x}10^{-3}$

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 6)

		BC-high adenoCa *vs* SqCa	
Probe ID	**Gene name**	**Fold-change**	**p value**
201387_s_at	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	0.48	$3.68x10^{-2}$
203786_s_at	tumor protein D52-like 1	0.47	$1.19x10^{-2}$
212702_s_at	bicaudal D homolog 2 (Drosophila)	0.47	$5.55x10^{-5}$
213992_at	collagen, type IV, alpha 6	0.46	$5.68x10^{-4}$
207345_at	follistatin	0.45	$2.28x10^{-6}$
202912_at	adrenomedullin	0.44	$3.03x10^{-3}$
213139_at	snail homolog 2 (Drosophila)	0.44	$1.56x10^{-2}$
202712_s_at	creatine kinase, mitochondrial 1B ; creatine kinase, mitochondrial 1A	0.42	$2.02x10^{-4}$
204614_at	serpin peptidase inhibitor, clade B (ovalbumin), member 2	0.41	$6.51x10^{-3}$
209800_at	Rhesus blood group, C glycoprotein	0.41	$4.62x10^{-2}$
219554_at	keratin 16 (focal non-epidermolytic palmoplantar keratoderma)	0.41	$3.70x10^{-4}$
216268_s_at	jagged 1 (Alagille syndrome)	0.4	$3.09x10^{-3}$
210355_at	parathyroid hormone-like hormone	0.39	$1.37x10^{-4}$
203184_at	fibrillin 2 (congenital contractural arachnodactyly)	0.38	$1.02x10^{-3}$
207030_s_at	cysteine and glycine-rich protein 2	0.38	$2.08x10^{-4}$
202345_s_at	annexin A8	0.36	$1.43x10^{-2}$
203074_at	fatty acid binding protein 5 (psoriasis-associated)	0.36	$1.06x10^{-2}$
215812_s_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 8	0.35	$1.07x10^{-5}$
204268_at	serpin peptidase inhibitor, clade B (ovalbumin), member 5	0.32	$3.87x10^{-2}$
204855_at	S100 calcium binding protein A2	0.32	$1.90x10^{-2}$
208539_x_at	small proline-rich protein 2B	0.3	$2.06x10^{-3}$
210854_x_at	NADH:ubiquinone oxidoreductase MLRQ subunit homolog	0.28	$5.49x10^{-6}$
218484_at	solute carrier family 6 (neurotransmitter transporter, creatine), member 8	0.28	$1.55x10^{-4}$
205595_at	desmoglein 3 (pemphigus vulgaris antigen)	0.25	$1.37x10^{-4}$
206561_s_at	aldo-keto reductase family 1, member B10 (aldose reductase)	0.23	$3.17x10^{-2}$
204455_at	cystatin A (stefin A)	0.19	$9.46x10^{-4}$
204971_at	dystonin	0.19	$7.10x10^{-4}$
205064_at	small proline-rich protein 3	0.18	$2.12x10^{-3}$
218990_s_at	small proline-rich protein 1B (cornifin)	0.18	$5.55x10^{-5}$
205157_s_at	keratin 17	0.17	$3.96x10^{-3}$
206032_at	keratin 6B	0.13	$1.33x10^{-5}$
213680_at	desmocollin 3	0.13	$1.46x10^{-3}$

Supplemental Gene List V. Differences in Airway BC-signature Gene Expression between BC-high Adenocarcinoma (adenoCa) *vs* Squamous Cell Carcinoma (cont. page 7)

Probe ID	**Gene name**	**BC-high adenoCa *vs* SqCa** **Fold-change**	 **p value**
213796_at	plakophilin 1 (ectodermal dysplasia/skin fragility syndrome)	0.12	$1.10x10^{-4}$
221854_at	small proline-rich protein 1A	0.12	$8.01x10^{-6}$
209125_at	keratin 6A ; keratin 6C ; keratin 6E	0.1	$4.25x10^{-4}$
201820_at	keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne types)	0.06	$6.08x10^{-5}$