Microarray analysis identifies pathways in progression of early stage lung adenocarcinoma: the importance of focal adhesion and ECM-receptor interactions

Douglas, Susan E.; Bethune, Drew C.; Xu, Zhaolin

Publisher's version / Version de l'éditeur:
https://doi.org/10.17140/PRRMOJ-1-104
Pulmonary Research and Respiratory Medicine, 1, 1, pp. 21-31, 2014-12-09

Openventio PUBLISHERS

PULMONARY RESEARCH AND RESPIRATORY MEDICINE

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

Research

*Corresponding author:
Susan E. Douglas, PhD
Principal Research Officer
Human Health Therapeutics Portfolio
National Research Council Halifax
1411 Oxford Street, Halifax, NS,
Nova Scotia, B3H 3Z1, Canada
E-mail: susan.douglas@nrc-cnrc.gc.ca

Volume 1 : Issue 1
Article Ref. #: 1000PRRMOJ1104

Article History:
Received: October 22nd, 2014
Accepted: December 7th, 2014
Published: December 9th, 2014

Citation:
Douglas SE, Bethune DC, Xu Z. Microarray analysis identifies pathways in progression of early stage lung adenocarcinoma: the importance of focal adhesion and ECM-receptor interactions. *Pulm Res Respir Med Open J.* 2014; 1(1): 21-31.

Microarray Analysis Identifies Pathways In Progression of Early Stage Lung Adenocarcinoma: The Importance of Focal Adhesion and ECM-Receptor Interactions

Susan E. Douglas[1#*], Drew C. Bethune[2#] and Zhaolin Xu[3#]

#equally contributed

[1]National Research Council Halifax, 1411 Oxford Street, Halifax, NS, B3H 3Z1, Canada

[2]Department of Surgery, Queen Elizabeth II Health Sciences Centre and Dalhousie University, 1278 Tower Rd, Halifax, NS, B3H 2Y9, Canada

[3]Department of Pathology, Queen Elizabeth II Health Sciences Centre and Dalhousie University, 5788 University Ave., Halifax, NS, B3H 1V8, Canada

ABSTRACT

Recurrence after lung cancer surgery is high, even among Non-Small Cell Lung Cancer (NSCLC) adenocarcinoma patients diagnosed early as Stage I, where there has been no spread to lymph nodes. Understanding the biological underpinnings of aggressivity and recurrence in this subset of tumours may enable the identification of patients who would benefit from adjuvant therapy. The purpose of this study was to identify differentially expressed molecular biomarkers that might underlie recurrence of Stage I tumours by comparing gene expression in later-stage tumours with those expressed in early-stage tumours. Gene expression in tissue biopsy samples from five Stage I and five Stage II/III NSCLC adenocarcinoma patients was analysed using an oligonucleotide microarray containing 17,000 probes printed in duplicate. Analyses were performed on total RNA isolated from tumour tissue of each patient using universal human RNA as a reference. Compared to normal tissues, the transcriptome of Stage I NSCLC adenocarcinomas showed enrichment in general pathways in cancer, whereas in Stage II/III more specific cancer pathways such as focal adhesion and ECM-receptor interaction pathways were enriched and components of the PPAR signalling pathway were depleted. Relative to early-stage NSCLC, Stage II/III adenocarcinomas showed up-regulation of genes of the basic transcriptional and translational machinery, particularly the "cancer testis antigen" PASD1 transcription factor. The actin cytoskeleton re-organisation and interleukin-6 pathways were also up-regulated whereas there was a generalized down-regulation of immune effectors and genes involved in immune system development. This small-scale transcriptome study provides important information about the pathways and molecules likely to be involved in the more metastatic propensity of those Stage I NSCLC adenocarcinomas that recur.

KEYWORDS: Adenocarcinoma; Biomarkers; Microarray; Non-small cell lung cancer; Recurrence; Transcriptome

ABBREVIATIONS: MMP: Matrix Metalloproteases; NSCLC: Non-Small Cell Lung Cancer; WG: Web Gestalt

Openventio PUBLISHERS

PULMONARY RESEARCH AND RESPIRATORY MEDICINE

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

INTRODUCTION

Lung cancer remains the leading cause of cancer-related death worldwide accounting for over a quarter (27%) of all cancer deaths each year.[1] Non-Small Cell Lung Cancer (NSCLC) comprises ~80% of all lung cancers, and the majority of patients are diagnosed at an advanced stage that is associated with a poor clinical outcome and low survival (<16% overall 5-year survival.[1] Even for patients classified as Stage I, the rate of post-operative recurrence is high, with a 5-year survival of only 52%. Risk factors for completely resected Stage IA tumours include poor differentiation, vascular invasion, wedge resection and minimal margins.[2] Existing methods of classification and staging such as Tumour, Node, Metastasis (TNM)[3] have great prognostic utility; however, patients with tumours of identical histology, differentiation, location and stage classifications may differ widely in their survival time or response to therapy.[4] Misclassification of Stage I tumours can occur when lymph node metastases are small and escape detection. There is a need to incorporate molecular profiling with standardized criteria from radiology and medical oncology into the classification of NSCLC.[5] Characterization of the basic underlying molecular alterations that occur during progression of early-stage NSCLC is urgently needed.

Various approaches have been attempted to aid in the classification, diagnosis and prognosis of NSCLC, including assessment of cells and nucleic acids (DNA, mRNA and microRNA) in sputum, bronchial biopsies, bronchial washing and brushing specimens, bronchial lavage fluid, blood, pleural effusions and solid tumour biopsies.[6] Several cellular tumour markers have shown potential in predicting survival in NSCLC.[7] Positive immunohistochemical staining of mTOR has been proposed as a marker of poor outcome in early stage NSCLC.[8] Apo lipoprotein E is over-expressed in lung adenocarcinomas with malignant pleural effusion and is associated with poorer survival in these patients.[9] SOX2 is over-expressed in the subset of Stage I adenocarcinomas from patients with poorer outcome and may help predict recurrence.[10]

Circulating plasma nucleic acid is elevated in lung cancer patients relative to controls and when detected at the time of diagnosis has been shown to be prognostic for poorer survival.[11] RT-PCR assays have detected KRT19 and TTF-1 mRNA[12] and TRIM28[13] transcripts in peripheral blood of NSCLC patients, indicative of circulating tumour cells. However, circulating nucleic acid markers lack organ specificity, have poor sensitivity and there is often overlap between markers originating from benign and malignant tumours.

Transcriptome analyses have proved useful in the diagnosis and prognosis of NSCLC. The transcriptomic signature differs between malignant, normal and lung metastatic tumours of non-pulmonary origin. In addition, transcriptomics can be used to distinguish between different histological subtypes and stages of lung cancer, to improve prediction of clinical outcome and response to therapy.[14] However, recent reviews of transcriptomic studies found that there was little consensus when it came to prognostic signatures,[2,14] although a meta-analysis of seven different microarray studies did reveal a set of 64 genes whose expression was associated with survival of Stage I NSCLC patients.[15] Most biomarker studies in NSCLC comprise patient cohorts including all stages and/or all histological subtypes of lung cancer;[16] there are few microarray studies only comparing patients with early stage adenocarcinoma.[17-18] One recent study compared three Stage IA and thirteen Stage IB NSCLC samples and found there was only one down-regulated gene (DSG3; desmoglein3) that discriminated between them.[19]

Biomarkers that predict metastatic potential of NSCLC at an early stage could significantly improve survival by identifying patients at high risk for recurrence and/or metastasis who may benefit from adjuvant chemotherapy.[20] Since the pattern of gene expression in higher stage tumours can be informative for predicting risk of recurrence in Stage I tumours,[21] we performed a pilot study that compared the transcriptomic profiles of tumours from patients with Stage I and Stage II/III adenocarcinoma and related these findings to the biological processes that may be involved in early recurrence/metastasis.

MATERIALS AND METHODS

Study population

Patient biopsy samples from five Stage I (with no lymph node involvement) and five Stage II/III (with lymph node involvement) patients were selected from the Queen Elizabeth II Health Sciences Center Lung Tumour Bank based on the following criteria: adenocarcinoma, age >45 years; all but one were current or past smokers (Table 1).

ID	Stage	Age	Gender	Smoking Status	Recurrence
L202	IB	76	Male	Current	no
L218	IA	65	Male	Past	yes
L252	IA	81	Male	Past	no
L272	IA	79	Female	Past	no
L278	IA	69	Male	Past	no
L229	IIA	66	Male	Past	yes
L240	IIA	74	Male	Past	no
L258	IIIA	45	Male	Never	no
L262	IIA	76	Female	Past	no
L300	IIB	54	Female	Past	no

Table 1: Patient characteristics and samples used for transcriptome analysis.

This study was approved by the Capital Health Research Ethics Board (CDHA-RS/2004-336), and all participating individuals signed informed consent.

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

RNA extraction

Frozen lung tissue (50 mg) was pulverized in a Multi Sample Bio Pulverizer (BioSpec) and homogenized in 1 mL TRIZOL® (Life Technologies) using a FastPrep®24 (MP Biomedical). Total RNA was extracted according to the manufacturer's protocol, treated with TURBO DNase (Applied Bio systems), purified with the Total RNA Purification Kit (Norgen Biotech Corp.) and quantified on a NanoDrop-1000 (Nano Drop Products).

Microarray analysis

Equal amounts (1 µg) of Universal Human RNA (Invitrogen) and lung tumor RNA samples were converted to oligo-modified cDNA using the 3DNA Array 900 Kit (Genisphere). cDNA samples were co-hybridized at 56 °C for 16 h on an oligonucleotide microarray containing 17,000 probes printed in duplicate (Atlantic Cancer Research Institute, Moncton, NB). A second hybridization was performed to bind the 3DNA Capture Regents, at 56 °C for 4 h. After washing, microarrays were scanned in an Axon GenePix 4200A scanner (Molecular Devices), gridded using GenePix software (version 6.0), and the .gpr output files were loaded into the ArrayPipe[22] server at the National Research Council Halifax. Control spots and low quality spots were flagged, the remaining spots were background-corrected using the limma normexp option with a cutoff of 50, intensity data were normalized using the limma loess sub grid option, and duplicate spot data were merged. Spots that had a normalized corrected intensity value of >100 in either channel 1 or channel 2 were analysed using Limma's empirical Bayes moderated t-test to identify spots that were significantly ($p<0.05$) differentially expressed between tumour and control reference RNA. Empirical Bayesian methods are used to provide stable results even when the number of arrays is small[23] as in this study. The Student's t-test module of MeV[24] was then used to identify spots that were significantly differentially expressed between the Stage I and Stage II/III samples. Gene enrichment analysis to determine signalling pathways and biological processes involved in early stage NSCLC progression was performed of genes that were greater than two-fold differentially expressed using Web Gestalt (http://bioinfo.vanderbilt.edu/webgestalt) and DAVID (http://david.abcc.ncifcrf.gov/).

RESULTS AND DISCUSSION

In order to minimize effects of individual genetic differences and potential confounding signals that can arise from residual cancer cells in the normal tissue surrounding a tumours,[25-26] RNAs from tumour samples were compared to control universal RNA rather than adjacent normal tissue and then differences in these expression ratios were compared between stages. After background correction, normalization and filtering, 16752 genes could be compared in the microarrays from Stage I and Stage II/III NSCLC tumours. The data are available in GEO under the accession number GSE28956.

In Stage I tumours, 647 genes were up-regulated and 711 genes were down-regulated and in Stage II/III tumours, 868 genes were up-regulated and 1051 genes were down-regulated more than two-fold relative to the control universal reference RNA ($p<0.05$; Supplemental Tables 1 and 2).

Genes differentially expressed in Stage I and Stage II/III NSCLC adenocarcinomas

Cytokeratins are valuable markers of lung cancer and have long been used to differentiate between different sub-types of this cancer; more recently they have shown promise as prognostic markers in lung cancer. They are major components of the cytoskeletal system and play a role in cell migration, invasion and metastasis. As expected, cytokeratins 7, 8, 16 and 19 (typical markers of NSCLC) were up-regulated in both stages relative to control universal reference RNA (Table 2; Supplemental Tables 1 and 2).

Interestingly, KRT19 was expressed almost two-fold higher in Stage II/III compared to Stage I NSCLC tumours. Consistent with our results, high serum concentrations of KRT19 fragments (CYFRA 21-1) are prognostic of poor outcome in adenocarcinoma and tumours from patients with high pre-operative CYFRA 21-1 are larger and more poorly differentiated, indicative of a more aggressive nature.[27]

Cellular adhesion molecules (CEACAM1, 5, 6, 7, 8) and extracellular matrix proteins (LAMB3, LY6D, OLFM4, PSG1, SFN, VCAN, COL1A1, NAPSB, PKP) were also highly expressed in tumours relative to normal. Adhesion pathways are associated with recurrence[28] and elevated levels of extracellular matrix proteins in the serum are associated with poor prognosis.[29] A recent microarray study identified cell adhesion molecule 1 (CADM1) to be significantly associated with poor survival in NSCLC, particularly Stage I adenocarcinoma, and its prognostic value was verified by immunohistochemistry in tissue microarrays.[30] In our study CADM1 showed 3.2-fold higher expression in Stage 1 tumours than normal (Supplemental Table 1). CEACAMs are important regulators of invasion and metastasis and CEACAM6 inhibits cell-cell contact inhibition mediated by CEACAM1 in A549 lung adenocarcinoma cells,[31] inducing cellular proliferation.[32] CEACAM6 was up-regulated 37-fold in Stage I and 9-fold in Stage II/III tumours; interestingly, both CEACAM6 and surfactant proteins that it interacts with are targets of transcription factor TTF-1[33] and all are up-regulated.

The putative oncogene RHEB[34] plays a role in growth and cell cycle progression through the AKT/MTOR pathway; it was up-regulated over 14-fold in Stage II/III tumours. Induction of angiogenesis, one of the hallmarks of cancer, is a major contributor to solid tumour development. CHI3L1, which

Openventio PUBLISHERS

PULMONARY RESEARCH AND RESPIRATORY MEDICINE

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

Stage I		Stage II/III	
Gene	**Ratio**	**Gene**	**Ratio**
Carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6)	37.65	Pregnancy specific beta-1-glycoprotein 1 (PSG1)	47.04
Stratifin (SFN)	25.71	Matrix metalloprotease 1 (MMP1)	19.64
Cytokeratin 16 (KRT16)	21.75	Cytokeratin 16 (KRT16)	17.35
Prostaglandin-endoperoxide synthase 2 (PTGS2; COX2)	21.42	Cytokeratin 19 (KRT19)	16.12
Olfactomedin 4 (OLFM4)	18.18	Ras homolog enriched in brain (RHEB)	14.46
Solute carrier family 6, member 14SLC6A14	16.81	FOS-like antigen 1 (FOSL1)	13.68
Lymphocyte antigen 6 (LY6D)	15.20	Steroid 5 alpha-reductase 2-like (SRD5A3)	13.45
Seizure related 6 (SEZ6L)	14.54	Marginal Zone B And B1 Cell-Specific (MZB1)	12.83
Serine protease inhibitor, clade B, member 5 (SERPINB5)	14.46	Serine protease inhibitor, clade B, member 5 (SERPINB5)	12.45
Granzyme B (GZMB)	11.78	Galectin 4 (LGALS4)	12.29
Pregnancy specific beta-1-glycoprotein 1 (PSG1)	10.50	Stratifin (SFN)	11.78
Laminin B3 (LAMB3)	10.49	Chemokine (C-X-C) ligand 17 (CXCL17)	11.42
Cellular retinoic acid binding protein 2 (CRABP2)	10.06	Oncostatin M receptor (OSMR)	10.34
Cytokeratin 8 (KRT8)	9.85	Chitinase-like 3 (CHI3L1; YKL40)	10.17
Cytokeratin 7 (KRT7)	9.50	TNF alpha-induced protein 2 (TNFAIP2)	9.77
Ligand dependent nuclear receptor corepressor (LCOR)	9.47	Neuromedin U (NMU)	9.71
Pleckstrin homology domain interacting protein (PHIP)	9.38	Carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6)	9.44
Forkhead box P1 (FOXP1)	9.07	Versican (VCAN)	9.36
Interleukin 7 receptor (IL7R)	8.67	Collagen, type I, alpha 1 (COL1A1)	9.25
Solute carrier family 12, member 2 (SLC12A2)	8.63	Cellular retinoic acid binding protein 2 (CRABP2)	9.24
Cytokeratin 19 (KRT19)	8.61	Laminin B3 (LAMB3)	9.19
CEP57L1	8.44	Napsin B (NAPSB)	9.17
S100 Ca^{++} binding protein P (S100P)	8.43	S100 Ca^{++} binding protein P (S100P)	8.89
Glutathione peroxidase 2 (GPX2)	8.02	X (inactive)-specific transcript (XIST)	8.71
Chemokine (C-X-C) ligand 9 (CXCL9)	7.96	Plakophilin 3 (PKP)	8.57

Table 2: Top 25 up-regulated genes in Stage I and Stage II/III NSCLC adenocarcinomas relative to control RNA (identified using Limma's empirical Bayes moderated t-test at $p<0.05$).

promotes tumour angiogenesis[35] and has been shown to be prognostic for low survival in NSCLC,[36,37] was up-regulated more than 10-fold in Stage II/III tumours and 6-fold in Stage I tumours. Several galectins (LGALS), which are also implicated in tumour angiogenesis as well as progression, were up-regulated in tumours of both stages.

The S100 family are calcium-binding proteins with varied roles in cancer invasion, metastasis and recurrence. They regulate a host of intracellular processes including enzyme activities, components of cytoskeleton, motility and cell cycle. An earlier microarray study showed calcium-binding proteins S100P and S100A2, and trypsinogens TRY6 and PRSS3 to be correlated with progression to metastasis in NSCLC.[20] In our study, expression of S100P was more than 8-fold higher in both Stage I and Stage II/III tumours but S100A2 was up-regulated only in Stage II/III tumours (5.3-fold). Other S100 proteins such as S100A10, which is essential for migration of tumour-promoting macrophages into tumor sites,[38] showed increased gene expression in Stage II/III tumours. Although TRY6 and PRSS3 trypsinogens were not differentially regulated in our study, matrix metalloproteases (MMPs) were dramatically up-regulated in Stage I and Stage II/III tumours, particularly the latter. MMPs also play a crucial role in metastasis by degrading extracellular matrix thereby allowing cell migration. Proteomics and immunohistochemical staining have shown that members of the annexin family (ANX) promote cancer cell invasion and metastasis in cancer, particularly ANXA1, ANXA2, ANXA4 and ANXA5.[31] We found expression of ANXA1, ANXA2, and ANXA3 was higher in tumour tissue than normal.

Many of the up-regulated genes in the Table 2 such as LY6D, GZMB, IL7R, CXCL9, CXCL17, MZB1, and LGALS4, may be derived from tumour-infiltrating immune cells. Such immune genes may be prognostic in NSCLC, and IL7R and CXCR4 were identified as key players in the tumour microenvironment by gene profiling.[39] In our study, CXCR4 showed a 2.8-fold increase in gene expression in Stage II/III tumours compared to Stage I. Interestingly, several of the genes were also included in a list of the top 20 up-regulated genes in the lung cancer

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

transcriptome derived by RNAseq.[40] These included KRT16, MMP1, Plakophilin (PKP), and Cellular Retinoic Acid Binding Protein 2 (CRABP2), the latter of which was proposed as a putative biomarker for lung adenocarcinoma. There was also considerable overlap with the down-regulated genes (data not shown). A recent RNASeq whole transcriptome study of six normal, adenocarcinoma *in situ* and invasive adenocarcinoma samples also identified CRABP2 as up-regulated in adenocarcinoma *in situ* compared to normal lung,[41] strongly implicating it in NSCLC progression.

Pathways differentially expressed in Stage I and Stage II/III NSCLC adenocarcinomas

To correlate differential gene expression with the biological pathways that are affected, gene enrichment analysis was performed. For Stage I tumours (Table 3) two major up-regulated KEGG pathways were identified: 34 genes in the cancer pathway and 13 in the more restricted small cell lung cancer pathway.

For Stage II/III tumours, two additional more specific cancer pathways were implicated: ECM-receptor interaction and focal adhesion, indicating the increased emphasis on cell migration and invasion in the later stage tumours. These pathways included many structural proteins (collagens, integrins, laminins, actinin4 ACTN4, tenascin C) as well as signalling molecules such as ERBB2, thrombopsondins 2 and 3, WNT5B, KRAS, BIRC2, cyclin D1, MAP kinase 9, protein phosphatase 1 PP1CA) and beta catenin (Supplemental Table 3). The effect on WNT, KRAS, PI3K-AKT, and MAPK signalling pathways would result in increased cell proliferation whereas increased ERBB2 signalling and anti-apoptotic protein BIRC2 would result in increased cell survival see Figure 1.

(A) 647 genes twofold up-regulated in Stage I.

KEGG analysis	#	p DAVID	p WG
Pathways in cancer	34	2.11E-06	3.29e-05
Small cell lung cancer	13	1.91E-04	9.0e-04

(B) 711 genes twofold down-regulated in Stage I.

KEGG analysis	#	p DAVID	p WG
PPAR signaling pathway	15	1.98E-06	1.49E-05
Parkinson's disease	18	6.24E-05	5.15E-05
Ascorbate and aldarate metabolism	7	7.02E-05	2.0E-04
Drug metabolism	12	9.44E-05	4.0E-04
Pentose and glucuronate interconversions	7	1.01E-04	2.0E-04
Oxidative phosphorylation	17	2.55E-04	2.0E-04
Glycolysis / Gluconeogenesis	11	3.36E-04	1.0E-03
Huntington's disease	20	5.11E-04	7.0E-04

(C) 868 genes twofold up-regulated in Stage II/III.

KEGG analysis	#	p DAVID	p WG
Pathways in cancer	43	1.07E-06	4.56e-06
ECM-receptor interaction	18	6.10E-06	5.66e-06
Small cell lung cancer	17	2.57E-05	3.23e-05
Focal adhesion	28	4.47E-05	3.23e-05

(D) 1051 genes twofold down-regulated in Stage II/III.

KEGG analysis	#	p DAVID	p WG
PPAR signaling pathway	14	1.67E-04	6.5e-03
Ascorbate and aldarate metabolism	7	2.96E-04	6.5e-03
Drug metabolism	12	8.99E-04	1.66e-02
Arginine and proline metabolism	11	9.33E-04	1.66e-02

Table 3: Enrichment analysis using Web Gestalt (WG) and DAVID of differentially regulated genes in Stage I and Stage II/III NSCLC adenocarcinomas relative to control RNA. #, number of genes; p, p value

Openventio PUBLISHERS

PULMONARY RESEARCH AND RESPIRATORY MEDICINE

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

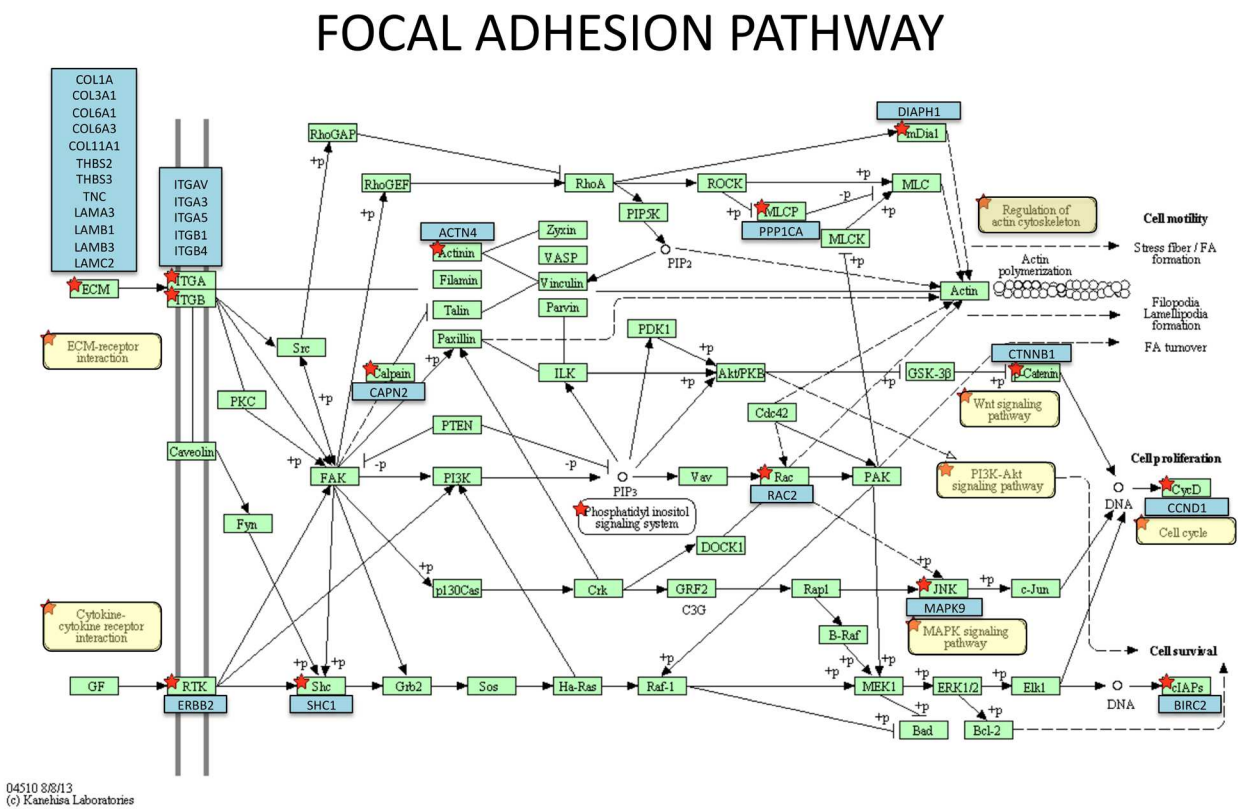

Figure 1: Focal adhesion pathway showing genes up-regulated in Stage II/III NSCLC adenocarcinomas relative to control RNA. The underlying KEGG pathway was generated by DAVID and components of the pathway containing enriched genes are indicated by red stars. The relevant genes (blue boxes) are superimposed over it.

In addition, up-regulation of ACTN4, PP1CA and DIAPH1 gene expression would positively impact regulation of the actin cytoskeleton and cell motility.

Both Stage I and Stage II/III tumours showed enrichment in down-regulated genes involved in PPAR signalling and various metabolic pathways (Table 3). PPAR γ has been implicated as a tumour suppressor in NSCLC, and xenograft models of lung cancer show that it inhibits lung tumour cell proliferation and growth through a variety of metabolic activities.[42] Down-regulation of the anti-tumour PPAR signalling may enhance the ability of NSCLC tumours to grow and metastasize. PPAR ligands are under development as potential therapeutic agents for lung cancer.[42]

Genes differentially expressed between Stage I and Stage II/III NSCLC adenocarcinomas

Between Stage I and II/III tumours, there were 48 significantly differentially regulated genes (26 up-regulated and 22 down-regulated; p<0.01) (Figure 2; Supplemental Table 4).

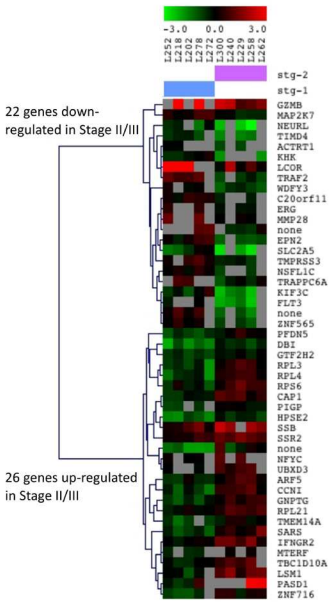

Figure 2: Heat map of differentially regulated genes (p<0.01) between Stage I (blue bar) and Stage II/III (purple bar) adenocarcinomas. Individual tumor samples are represented in the columns and genes in the rows. Red is up-regulated, green is down-regulated, and grey is missing data.

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

PASD1 exhibited the greatest up-regulation between groups (9.4-fold) whereas the remainder were between 1.4 and 2.9-fold up-regulated. PASD1 is an immunogenic "cancer testis antigen" that is thought to function as a transcription factor and is associated with cancer of the small intestine, colon, lung, head and neck as well as hematopoietic malignancies.[43] It shows promise as a target for immunotherapy against various hematopoietic cancers. PASD1 levels were significantly higher in H1299, a commonly used NSCLC cell line, than other cell lines[44] and our results confirm that targeting it could also be a promising therapeutic approach in NSCLC. On the other hand, LCOR showed the greatest down-regulation (5.9-fold) between groups with the remaining 21 genes showing between 1.6 and 4-fold down-regulation. LCOR is a ligand-dependent co-repressor of various nuclear hormone receptors and has recently been shown to regulate transcription of CDKN1A, which encodes the cell cycle regulator p21, and the cell adhesion molecule, E-cadherin;[45] reduced expression would promote transcription of cancer-related genes. Although this effect of LCOR on CDKN1A and E-cadherin has only been reported in cell lines and not yet for lung cancer, our studies suggest that this regulatory network is important in progression of NSCLC.

At a less stringent cut-off ($p<0.05$), there were 373 significantly differentially regulated genes between Stage I and II/III tumours (178 up-regulated and 195 down-regulated; Supplemental Table 5) of which the top 25 are shown in Table 4.

Up-regulated		Down-regulated	
Gene	**Ratio**	**Gene**	**Ratio**
Surfactant, pulmonary-associated protein A2 (SFTPA2)	9.94	Ubiquilin-like (UBQLNL)	0.006
PAS domain containing 1 (PASD1)	9.41	Zinc finger protein 367 (ZNF36)	0.15
Nuclear receptor co-activator 7 (NCOA7)	4.93	Ligand dependent nuclear receptor corepressor (LCOR)	0.17
Surfactant, pulmonary-associated protein D (SFTPD)	3.92	Keratin associated protein 21-1 (KRTAP21-1)	0.17
Matrix Gla protein (MGP)	3.37	Tescalcin (TESC)	0.18
Member RAS oncogene family (RAB2A)	3.33	Pleckstrin homology domain interacting protein (PHIP)	0.19
PHD finger protein 11 (PHF11), variant 2	3.19	Tumor necrosis factor receptor superfamily, member 6b, decoy (TNFRSF6B)	0.20
Src-like-adaptor (SLA), variant 3	3.06	Dihydrolipoamide branched chain transacylase E2 (DBT)	0.20
Synovial sarcoma, X breakpoint 1 (SSX1)	3.05	Family with sequence similarity 101, member B (FAM101B)	0.20
Tubulin, alpha 1a (TUBA1A)	3.04	Solute carrier family 52, riboflavin transporter, member 1 (GPR172B)	0.20
Ras-related associated with diabetes (RRAD)	3.03	Solute carrier family 30 (zinc transporter), member 7 (SLC30A7)	0.24
Keratin associated protein 19-2 (KRTAP19-2)	2.98	ATP/GTP binding protein-like 5 (AGBL5)	0.24
Heat shock protein 90kDa alpha (cytosolic), class A member 1 (HSP90AA1), variant 2	2.96	Solute carrier family 2, member 4 regulator (SLC2A4RG)	0.24
Archaemetzincins-2 (AMZ2), variant 1	2.94	Neuralized homolog (NEURL)	0.24
Ribosomal protein S4, X-linked (RPS4X)	2.85	Uridine-cytidine kinase 2 (UCK)	0.24
Pinin, desmosome associated protein (PNN)	2.81	Multiple EGF-like-domains 11 (MEGF11)	0.25
Leptin receptor overlapping transcript (LEPROT)	2.81	Growth differentiation factor 5 (GDF5)	0.25
Leukotriene A4 hydrolase (LTA4H)	2.77	Granzyme B (GZMB)	0.25
Calpain 2, (m/II) large subunit (CAPN2)	2.74	Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1 (ACE)	0.025
Peptidylglycine alpha-amidating monooxygenase (PAM), variant 1	2.73	Forkhead box C2 (FOXC2)	0.26
Niemann-Pick disease, type C2 (NPC2)	2.69	HNF1 homeobox B (HNF1B)	0.26
Dynein, light chain, LC8-type 1 (DYNLL1), variant 3	2.68	CDC42 effector protein (Rho GTPase binding) 5 (CDC42EP5; CEP5)	0.26
Histone cluster 1, H2bk (HIST1H2BK)	2.66	Hippocalcin (HPCA)	0.28
LSM1 homolog, U6 small nuclear RNA associated (LSM1)	2.66	Signal transducer and activator of transcription 5B (STAT5B)	0.28
Inositol 1,4,5-trisphosphate 3-kinase C (ITPKC)	2.65	GC-rich sequence DNA-binding factor 1 (GCFC1)	0.28

Table 4: Top 25 differentially regulated genes in Stage II/III compared to Stage I NSCLC adenocarcinomas (identified using MeV at $p<0.05$).

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

Published gene expression studies have indicated that aggressive lung adenocarcinomas have higher levels of cell proliferation-related genes including cyclins, transcription factor TTF-1, surfactant SFTPB, and those involved in immunological function.[21] As mentioned above, TTF-1 regulates the transcription of CEACAM6 and surfactant-associated proteins. Both SFTPA2 and SFTPD were in the top five genes up-regulated in Stage II/III compared to Stage I tumours (Table 4) and cyclin I was up-regulated 2.56-fold in Stage II/III compared to Stage I tumours. MGP is a mesenchymal gene encoding an extracellular matrix protein and RT-PCR, ELISA and immunohistochemistry have shown that is over-expressed in recurrent gliomas and associated with poor outcome.[46] It has not thus far been studied in lung cancer, and may provide a novel biomarker for lung recurrence. Immunohistochemistry, qRT-PCR, Western blot analysis and RNA interference have shown that LSM1 functions as an oncogene in the progression of lung cancer[47] and members of the Ras and Src oncogene families (RAB2A, RRAD and SLA) were also up-regulated. Breast tumours shown to over-express RRAD by immunohistochemistry and Western blotting are of higher grade, size and nodal involvement and have poor prognosis.[48] SLA mediates cell migration and invasion through integrin signalling by Src and FAK tyrosine kinases and *in vitro* overexpression studies have shown that it is also a negative regulator of T and B cell-mediated responses,[49] which are crucial for anti-tumour immunity. Microarray, reverse phase protein array, ELISA and immunohistochemistry analyses have shown that patients with lung adenocarcinomas that express genes associated with an active immune response, such as RANTES, MIP-1-beta and STAT5 have better outcomes.[50-51] Consistent with this is the 2.7-fold lower STAT5B expression in Stage II/III than Stage I tumours in our study.

Pathways differentially expressed between Stage I and Stage II/III NSCLC adenocarcinomas

Later-stage tumours showed enrichment in pathways mostly involved in transcription, translation, mitochondrial electron transport chain, and actin cytoskeleton organization (Supplemental Table 6A). This reflects the overall higher metabolism and growth rate of more advanced cancers. Interestingly, there was also enrichment in genes in the IL-6 signaling pathway; these included phospholipase C gamma subunit (PLCG1), FYN oncogene, ras-related small GTP-binding protein (RAC1), heat shock protein 90kDA alpha A1 (HSP90AA1), and protein phosphatase2 regulatory subunit B, gamma (PPP2R2C). FYN is a tyrosine kinase that has been implicated in the control of cell growth. HSP90AA1 is a chaperone for tyrosine kinases EGFR, MET and ALK,[52] all of which are oncogenic drivers of lung cancer. PLCG1 is responsible for intracellular transduction of receptor-mediated tyrosine kinase activators. RAC1 is involved in control of cell growth/cytoskeletal reorganization, and PPP2R2C is also involved in tumour signal transduction pathways. IL-6 is present in the tumour microenvironment and induces tumour proliferation, angiogenesis and resistance to chemotherapy; it has recently been shown by tissue microarray analysis to be a novel prognostic biomarker in NSCLC.[53] CD88, which is a high-affinity receptor for C5a that is widely expressed on lung epithelial cells, is up-regulated in Stage II/III (3.6-fold) compared to Stage I (2.3-fold) tumours. Over-expression of this receptor, assessed by tissue microarray analysis, correlates with down-regulation of E-cadherin expression and lymph node metastasis in NSCLC patients.[54]

Stage II/III *vs.* Stage I tumours did not show statistically significant enrichment in any pathways and there was only borderline enrichment in some GO biological processes (Supplemental Table 6B).

CONCLUSIONS

By comparing the transcriptomes of early and later-stage NSCLC adenocarcinomas, we have uncovered information on pathways that are involved in recurrence in early-stage lung cancer. Many of the hallmarks of cancer[55] such as proliferative signalling, angiogenesis, invasion and metastasis were evident in the sets of genes we identified. A significant number of the differentially regulated genes participated in closely related processes, helping to validate our results. Transcript data reflected the overall higher growth rate of more advanced cancers and the involvement of pathways involved in actin cytoskeleton remodelling and cell migration. While the small number of samples analysed in this study makes it difficult to make strong conclusions, we have nonetheless found concordant results with other reports that used a variety of complementary molecular techniques and uncovered possible pathways underlying recurrence. This study provides interesting novel leads to be followed up on in larger prospective investigations.

ACKNOWLEDGEMENT

We gratefully acknowledge the technical assistance of Jeffrey Gallant, Evelyn Teh and Jason Williams and bioinformatics analysis by Susanne Penny, National Research Council (NRC). We thank John Fris, database manager of the Lung Tumor Bank at Capital District Health Authority, for collating clinical annotations. Funding was provided by NRC and Dalhousie University Departments of Pathology and Surgery.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012; 62: 10-29. doi: 10.3322/caac.20138

2. Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? J. Natl. *Cancer Inst.* 2010; 102: 464-474. doi: 10.1093/jnci/djq025

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

3. Tsim S, O'Dowd CA, Milroy R, Davidson S. Staging of non-small cell lung cancer (NSCLC): A review. Respir. *Med*. 2010; 104: 1767-1774. doi: 10.1016/j.rmed.2010.08.005

4. Chen HY, Yu SL, Li KC, Yang PC. Biomarkers and transcriptome profiling of lung cancer. *Respirology*. 2012; 17: 620-626. doi: 10.1111/j.1440-1843.2012.02154.x

5. Travis WD, Brambilla E, Noguchi M, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. 2011; 6: 244-285. doi: 10.1097/JTO.0b013e318206a221

6. Holdenrieder S, Stieber P. Circulating apoptotic markers in the management of non-small cell lung cancer. *Cancer Biomarkers*. 2009; 6: 197-210. doi: 10.3233/CBM-2009-0130

7. Daigo Y, Nakamura Y. From cancer genomics to thoracic oncology: discovery of new biomarkers and therapeutic targets for lung and esophageal carcinoma. *Gen Thorac Cardiovasc Surg*. 2008; 56: 43-53. doi: 10.1007/s11748-007-0211-x

8. Dhillon T, Mauri FA, Bellezza G, et al. Overexpression of the mammalian target of rapamycin: a novel biomarker for poor survival in resected early stage non-small cell lung cancer. J. *Thorac. Oncol*. 2010; 5: 314-319. doi: 10.1097/JTO.0b013e3181ce6604

9. Su WP, Chen YT, Lai WW, Lin CC, Yan JJ, Su WC. Apolipoprotein E expression promotes lung adenocarcinoma proliferation and migration and as a potential survival marker in lung cancer. Lung Cancer. 2011; 71: 28-33. doi: 10.1016/j.lungcan.2010.04.009

10. Sholl LM, Barletta JA, Yeap BY, Chirieac LR, Hornick JL. Sox2 protein expression is an independent poor prognostic indicator in stage I lung adenocarcinoma. *Am J Surg Pathol*. 2010; 34: 1193-1198. doi: 10.1097/PAS.0b013e3181e5e024

11. van der Drift MA, Hol BEA, Klaassen CHW, et al. Circulating DNA is a non-invasive prognostic factor for survival in non-small cell lung cancer. *Lung Cancer*. 2010; 68: 283-287. doi: 10.1016/j.lungcan.2009.06.021

12. Yoon SO, Kim YT, Jung KC, Jeon YK, Kim BH, Kim CW. TTF-1 mRNA-positive circulating tumor cells in the peripheral blood predict poor prognosis in surgically resected non-small cell lung cancer patients. *Lung Cancer*. 2010; 71: 209-216. doi: 10.1016/j.lungcan.2010.04.017

13. Liu L, Zhao E, Li C, et al. TRIM28, a new molecular marker predicting metastasis and survival in early-stage non-small cell lung cancer. *Cancer Epidemiol*. 2013; 37: 71-78. doi: 10.1016/j.canep.2012.08.005

14. Singhal S, Miller D, Ramalingam S, Sun SY. Gene expression profiling of non-small cell lung cancer. *Lung Cancer*. 2008; 60: 313-324. doi: 10.1016/j.lungcan.2008.03.007

15. Lu Y, Lemon W, Liu PY, Yang P, You M. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med*. 2006; 3: e467. doi: 10.1371/journal.pone.0030880

16. Ohgino K, Soejima K, Yasuda H, et al. Expression of fibroblast growth factor 9 is associated with poor prognosis in patients with resected non-small cell lung cancer. *Lung Cancer*. 2014; 83: 90-96. doi: 10.1016/j.lungcan.2013.10.016

17. Meyerson M, Carbone D. Genomic and proteomic profiling of lung cancers: lung cancer classification in the age of targeted therapy. J. Clin. *Oncol*. 2005; 23: 3219-3226. doi: 10.1200/JCO.2005.15.511

18. Zhu CQ, Ding K, Strumpf D, et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J. Clin. Oncol*. 2010; 28: 4417-4424. doi: 10.1200/JCO.2009.26.4325

19. Sanchez-Palencia A, Gomez-Morales M, Gomez-Capilla JA, et al. Gene expression profiling reveals novel biomarkers in nonsmall cell lung cancer. *Int. J. Cancer*. 2011; 129: 355-364.

20. Diederichs S, Bulk E, Steffen B, et al. S100 family members and trypsinogens are predictors of distant metastasis and survival in early-stage non-small cell lung cancer. *Cancer Res*. 2004; 64: 5564-5569.

21. Shedden K, Taylor JM, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multisite, blinded validation study. *Nat. Med*. 2008; 14: 822-827. doi: 10.1038/nm.1790

22. Hokamp K, Roche FM, Acab M, et al. ArrayPipe: a flexible processing pipeline for microarray data. *Nucleic Acids Research*. 2004; 32: W457-W459. doi: 10.1093/nar/gkh446

23. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*. 2004; 3(3). doi: 10.2202/1544-6115.1027

24. Saeed AI, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis. *BioTechniques*. 2003; 34: 374-378.

25. Kadara H, Lacroix L, Behrens C, et al. Identification of gene signatures and molecular markers for human lung cancer prognosis using an in vitro lung carcinogenesis system.

Cancer Prev. Res. 2009; 2: 702-711. doi: 10.1158/1940-6207.CAPR-09-0084

26. Mwenifumbo JC, Marra MA. Cancer genome-sequencing study design. *Nat Rev Genet*. 2013; 14: 321-332. doi: 10.1038/nrg3445

27. Park SY, Lee JG, Kim J, et al. Preoperative serum CYFRA 21-1 level as a prognostic factor in surgically treated adenocarcinoma of lung. *Lung Cancer.* 2013; 79: 156-160.

28. Nanjundan M, Byers LA, Carey MS, et al. Proteomic profiling identifies pathways dysregulated in non-small cell lung cancer and an inverse association of AMPK and adhesion pathways with recurrence. *J Thorac Oncol.* 2010; 5: 1894-1904. doi: 10.1097/JTO.0b013e3181f2a266

29. Gogali A, Charalabopoulos K, Zampira I, et al. Soluble adhesion molecules E-cadherin, intercellular adhesion molecule-1, and E-selectin as lung cancer biomarkers. *Chest.* 2010; 138: 1173-1179. doi: 10.1378/chest.10-0157

30. Botling J, Edlund K, Lohr M, et al. Biomarker Discovery in Non-Small Cell Lung Cancer: Integrating Gene Expression Profiling, Meta-analysis, and Tissue Microarray Validation. *Clin Cancer Res.* 2013; 19: 194-204. doi: 10.1158/1078-0432.CCR-12-1139

31. Deng S, Wang J, Hou L, et al. Annexin A1, A2, A4 and A5 play important roles in breast cancer, pancreatic cancer and laryngeal carcinoma, alone and/or synergistically. *Oncol Lett.* 2013; 5: 107-112.

32. Singer BB, Scheffrahn I, Kammerer R, Suttorp N, Ergun S, Slevogt H. Deregulation of the CEACAM expression pattern causes undifferentiated cell growth in human lung adenocarcinoma cells. *PLoS One.* 2010; 5: e8747. doi: 10.1371/journal.pone.0008747

33. Chapin C, Bailey NA, Gonzales LW, Lee JW, Gonzalez RF, Ballard PL. Distribution and surfactant association of carcinoembryonic cell adhesion molecule 6 in human lung. *Am J Physiol Lung Cell Mol Physiol.* 2012; 302: L216-225. doi: 10.1152/ajplung.00055.2011

34. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep.* 2013; 3: 2650. doi: 10.1038/srep02650

35. Shao R, Hamel K, Petersen L, et al. YKL-40, a secreted glycoprotein, promotes tumor angiogenesis. *Oncogene.* 2009; 28: 4456-4468. doi: 10.1038/onc.2009.292

36. Choi IK, Kim YH, Kim JS, Seo JH. High serum YKL-40 is a poor prognostic marker in patients with advanced non-small cell lung cancer. Acta Oncol. 2010; 49: 861-864. doi: 10.3109/02841861003631503

37. Thom I, Andritzky B, Schuch G, et al. Elevated pretreatment serum concentration of YKL-40-An independent prognostic biomarker for poor survival in patients with metastatic nonsmall cell lung cancer. *Cancer.* 2010; 116: 4114-4121. doi: 10.1002/cncr.25196

38. Phipps KD, Surette AP, O'Connell PA, Waisman DM. Plasminogen receptor S100A10 is essential for the migration of tumor-promoting macrophages into tumor sites. *Cancer Res.* 2011; 71: 6676-6683. doi: 10.1158/0008-5472.CAN-11-1748

39. Suzuki K, Kachala SS, Kadota K, et al. Prognostic immune markers in non-small cell lung cancer. *Clin Cancer Res*. 2011; 17: 5247-5256. doi: 10.1158/1078-0432.CCR-10-2805

40. Han SS, Kim WJ, Hong Y, et al. RNA sequencing identifies novel markers of non-small cell lung cancer. *Lung Cancer.* 2014; 84: 229-235. doi: 10.1016/j.lungcan.2014.03.018

41. Morton ML, Bai X, Merry CR, et al. Identification of mRNAs and lincRNAs associated with lung cancer progression using next-generation RNA sequencing from laser micro-dissected archival FFPE tissue specimens. *Lung Cancer*. 2014; 85: 31-39. doi: 10.1016/j.lungcan.2014.03.020

42. Han SW, Roman J. Anticancer actions of PPARgamma ligands: Current state and future perspectives in human lung cancer. *World J Biol Chem.* 2010; 1: 31-40.

43. Khan G, Denniss F, Mills KI, Pulford K, Guinn B. PASD1: A promising target for the immunotherapy of haematological malignancies. *J. Genet. Syndr. Gene Ther.* 2013; 4. doi: 10.4172/2157-7412.1000186

44. Guinn BA, Bland EA, Lodi U, et al. Humoral detection of leukaemia-associated antigens in presentation acute myeloid leukaemia. *Biochem Biophys Res Commun.* 2005; 335: 1293-1304. doi: 10.1016/j.bbrc.2005.08.024

45. Calderon MR, Verway M, An BS, et al. Ligand-dependent corepressor (LCoR) recruitment by Kruppel-like factor 6 (KLF6) regulates expression of the cyclin-dependent kinase inhibitor CDKN1A gene. *J Biol Chem*. 2012; 287: 8662-8674. doi: 10.1074/jbc.M111.311605

46.Mertsch S, Schurgers LJ, Weber K, Paulus W, Senner V. Matrix gla protein (MGP): an overexpressed and migration-promoting mesenchymal component in glioblastoma. *BMC Cancer.* 2009; 9: 302. doi: 10.1186/1471-2407-9-302

ISSN 2377-1658 Open Journal http://dx.doi.org/10.17140/PRRMOJ-1-104

47. Watson PM, Miller SW, Fraig M, Cole DJ, Watson DK, Boylan AM. CaSm (LSm-1) overexpression in lung cancer and mesothelioma is required for transformed phenotypes. *Am J Respir Cell Mol Biol*. 2008; 38: 671-678. doi: 10.1165/rcmb.2007-0205OC

48. Tseng YH, Vicent D, Zhu J, et al. Regulation of growth and tumorigenicity of breast cancer cells by the low molecular weight GTPase Rad and nm23. *Cancer Res*. 2001; 61: 2071-2079.

49. Holland SJ, Liao XC, Mendenhall MK, et al. Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel inhibitor of antigen receptor signaling. *J Exp Med*. 2001; 194: 1263-1276. doi: 10.1084/jem.194.9.1263

50. Moran CJ, Arenberg DA, Huang CC, et al. RANTES expression is a predictor of survival in stage I lung adenocarcinoma. *Clin Cancer Res*. 2002; 8: 3803-3812.

51. He Y, Zhou Z, Hofstetter WL, et al. Aberrant expression of proteins involved in signal transduction and DNA repair pathways in lung cancer and their association with clinical parameters. *PLoS One*. 2012; 7: e31087. doi: 10.1371/journal.pone.0031087

52. Koizumi H, Yamada T, Takeuchi S, et al. Hsp90 inhibition overcomes HGF-triggering resistance to EGFR-TKIs in EGFR-mutant lung cancer by decreasing client protein expression and angiogenesis. *J Thorac Oncol*. 2012; 7: 1078-1085. doi: 10.1097/JTO.0b013e3182519a2c

53. Pei BX, Sun BS, Zhang ZF, Wang AL, Ren P. Interstitial tumor-associated macrophages combined with tumor-derived colony-stimulating factor-1 and interleukin-6, a novel prognostic biomarker in non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2014; 148: 1208-1216. doi: 10.1016/j.jtcvs.2014.05.003

54. Gu J, Ding JY, Lu CL, et al. Overexpression of CD88 predicts poor prognosis in non-small-cell lung cancer. *Lung Cancer* . 2013; 81: 259-265. doi: 10.1016/j.lungcan.2013.04.020

55. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144: 646-674. doi: 10.1016/j.cell.2011.02.013