[Frontiers in Bioscience E3, 46-50, January 1, 2011]

Bioinformatics approach for the validation of non-small cell lung cancer biomarkers

Carolina Beatriz Muller, Rafael Longhi Sampaio de Barros, Jose Claudio Fonseca Moreira, Fabio Klamt

Centro de Estudos em Estresse Oxidativo, Departamento de Bioquimica, Universidade Federal do Rio Grande do Sul (UFRGS), Rua Ramiro Barcelos 2600-anexo, Porto Alegre 90035-003, Brazil

TABLE OF CONTENTS

1. ABSTRACT

Non-small cell lung cancer (NSCLC) accounts for nearly 1 million deaths annually, worldwide. Conventional treatments offer limited benefits and patients have a survival rate of approximately 1 year. A biomarker for NSCLC could provide the potential benefits of early diagnosis, prognosis and could lead to important applications such as drug targeting. In a search for a biomarker with prognostic value, we reviewed the literature and tested potential biomarkers by performing a meta-data analysis using public databank of NSCLC biopsies containing gene expression data and clinical and pathologic information from 111 patients. We generated standard Kaplan–Meier mortality curves by clustering patients according to either biomarker expression levels or NSCLC stage grouping. Our statistical analyses show that all 60 potential biomarkers analyzed here have no prognostic value for NSCLC patient outcome.

2. INTRODUCTION

Lung cancer is the leading cause of cancer deaths worldwide, accounting for 1.3 million deaths annually (data from World Health Organization, 2008). The high mortality associated with this disease is primarily due to the fact that the majority of lung cancers are not detected until they have progressed to an advanced stage (1). Non-small cell lung cancer (NSCLC) represents nearly 85% of lung cancer cases. Compared with other major types of cancer such as colon, prostate, and breast cancers, the clinical outcome of conventional therapies for NSCLC remains poor, with a median survival of 9-11 months (2).

Therefore, there is an urgent need for more effective therapies, drugs, or treatments that could help decrease the incidence of NSCLC. Alternatively, cancer biomarker gives good guidance on many areas of cancer biology. Unlike uniformity of long-established TNM

system, the international standard protocol that allows the staging of carcinoma according to the extent of disease in the patient, cancer biomarkers are considered to be more suitable to the heterogeneous nature of cancer (3). In 2008, Sawyers discussed the three types of cancer biomarkers. These can be used for prognosis, to predict the natural course of a tumor, indicating whether the outcome for the patient is likely to be good or poor. They can also be used in prediction, helping to decide which patients are likely to respond to a given drug and at what dose it might be most effective (4). A biomarker for NSCLC could provide the potential benefits of early diagnosis, considering that the disease is usually detected in late stages when surgical removal of the tumor is no longer an option, and additionally, could lead to other important applications such as prognosis and drug targeting (5). An impressive number of molecular markers have been implicated in the prognosis of NSCLC; however, the results reported in literature are conflicting and none are in use in clinics. Thus, further investigation, newer molecular assays and the development of appropriate panel of molecular markers are still required (6).

Systematic analysis of gene expression using high-throughput screening of cDNA microarray libraries has been considered as an effective approach for identifying and validating potentials biomarkers for NSCLC (7). However, final validation should be done by testing a collection of well-defined clinical samples. Herein, we describe a bioinformatics-based approach to test and validate the prognostic value of potential NSCLC biomarkers. Our research group have been studying many aspects of tumor biology by different bioinformatics approaches (8-9). Here our approach is to correlate data collected from the literature with data on gene expression of a large and well-defined collection of NSCLC biopsies containing information on patients' clinical status and pathology to clinically evaluate the efficacy of potential biomarkers to predict patients outcome. Validation by clinical trials in large cohorts of patients is necessary before cancer-related phenotypes can be translated into the clinic as reliable biomarkers.

3. SEARCHING FOR NSCLC POTENTIAL BIOMARKER

3.1. Literature search

The list of genes presented in (Table 1) was compiled by searching the PubMed database for articles published in English between January 1985 and December 2009. Search criteria included subject heading terms for "biomarker", "prognosis", "gene expression" and "lung cancer". Genes reported in two or more articles during the period of our search, or in one article at least during the last 3 years were selected. Those articles describing pooled biomarkers into the same analysis were excluded from the list (*i.e.* combined performance for multiple genes). A total of 60 NSCLC was tested.

3.2. Tumor samples and microarray data

For the clinical validation of potential NSCLC prognostic biomarkers, we used a large, homogeneous, well-defined collection of samples from lung cancer biopsies, along with respective gene expression data and relevant clinical and pathologic information -such as age, sex, cancer histological type, and NSCLC staging in a cohort follow-up period of 72 months- on 111 patients (10). Data were obtained from the GEO database (http://www.ncbi.nlm.nih.gov/projects/geo/; Series GSE3141) and from the Duke Institute for Genome Sciences & Policy website (http://data.cgt.duke.edu/oncogene.php). All gene array data was on Affymetrix U133A GeneChip, from core biopsies of patients' tumor. A *sine qua non* condition for select a given gene was the presence of two or more microarray probes in NSCLC cohort used.

3.3. Survival data analysis

We used the SPSS software (SPSS for Windows, release 14.0.0, SPSS Inc., Chicago, IL) to generate Standard Kaplan–Meier mortality curves with their significance levels, for patient clusters. Survival curves were compared using the log-rank test; patients were clustered according to biomarker expression level or NCSLC stage grouping (*i.e.* initial and advanced staging).

4. DISCUSSION

Through a systematic MEDLINE literature inspection, we selected 60 genes as potential biomarkers to be validated using a clinical databank. These biomarkers constitute a diverse group of genes involved in different cellular functions. They code proteins such as transcription factors (*TFAP2B*, *MYC*), protein kinases (*PIK3CA*, *STAT1*, *KRAS*, *STYK1*, *LCK*), protein phosphatases (*DUSP6*), receptors (*AGER*, *EGFR*, *AMFR*), and several DNA repair systems (*ALKBH5*, *ALKBH3*, *FGFR1OP*, *ERCC1*) (See table 1 for complete list of potential prognostic biomarkers). These potential biomarkers are related in the most distinct ways with lung cancer; for example, mutations in the genetic region encoding the kinase domain of the epidermal growth-factor receptor (*EGFR*) predict the sensitivity of lung tumors to erlotinib or gefitinib (11), as well different mutation in *KRAS* predict that patients with lung cancer will fail to respond to these inhibitors (12).

Some genes of DNA repair system are also considered to be potential biomarkers to NSCLC. They have been constantly described as being related with sensitivity to chemotherapeutic drugs (13-14), specially alkylating agents, which are the most largely used anti-cancer drug for NSCLC treatment (15). These drugs are mutagenic, genotoxic and have the ability to damage DNA (16). Cisplatin is an alkylating agent widely used in NSCLC treatment; however, this type of cancer can presents inherently resistance against it. Like many DNA alkylators, cisplatin acts inhibiting DNA replication, which is the critical target in cancer treatment. In this case, the resistance against cisplatin is mainly determinate by the expression levels of nucleotide excision repair genes (NER). ERCC, a component of the NER complex, is a potential marker involved in prediction of resistance to cisplatin, which has been described to be related with its mRNA expression (17-19).

Table 1. Potential NSCLC biomarkers previously described

Symbol[1]	Gene Name[1]	Gene ID[1]	*P-value*[2]	Ref.
ABCC1	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	4363	0.6239	[27]
AGER	advanced glycosylation end product-specific receptor	177	0.0894	[28]
ALKBH1	alkB. alkylation repair homolog 1	8846	0.5685	[29]
ALKBH5	alkB. alkylation repair homolog 5	54890	0.1704	[29]
ALKBH7	alkB. alkylation repair homolog 7	84266	0.6970	[29]
ALKBH8	alkB. alkylation repair homolog 8	91801	0.5752	[29]
AMFR	autocrine motility factor receptor	267	0.6829	[30,31]
BCL2L1	BCL2-like 1	598	0.0661	[22]
BIRC5	baculoviral IAP repeat-containing 5	332	0.3584	[32,33]
BRCA1	breast cancer 1, early onset	672	0.5103	[34]
CALB1	calbindin 1, 28kDa	793	0.3541	[35]
CAV1	caveolin 1, caveolae protein	857	0.0808	[36]
CBLC	Cas-Br-M (murine) ecotropic retroviral transforming sequence c	23624	0.0608	[7]
CCNB2	cyclin B2	9133	0.8282	[28]
CCND1	cyclin D1	595	0.1177	[29]
CD9	CD9 molecule	928	0.7940	[20,26]
CDK8	cyclin-dependent kinase 8	1024	0.4337	[38]
CRABP1	cellular retinoic acid binding protein 1	1381	0.5485	[27]
CTSB	cathepsin B	1508	0.6717	[39]
DNAJB4	DnaJ (Hsp40) homolog, subfamily B, member 4	11080	0.2820	[40]
DUSP6	dual specificity phosphatase 6	1848	0.2842	[41]
EGFR	epidermal growth factor receptor	1956	0.6074	[42,43]
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2.	2064	0.2379	[42,10]
ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	2065	0.5191	[44]
ERCC1	excision repair cross-complementing rodent repair deficiency.	2067	0.4269	[43,45,46]
FGFR1OP	FGFR1 oncogene partner	11116	0.1871	[47]
FOLR1	folate receptor 1 (adult)	2348	0.3262	[48]
GSTA1	glutathione S-transferase alpha 1	2938	0.6353	[35]
HMGA2	high mobility group AT-hook 2	8091	0.7014	[49]
HSP90AA1	heat shock protein 90kDa alpha (cytosolic), class A member 1	3320	0.4367	[50]
IFI44	interferon-induced protein 44	10561	0.9626	[51]
IGF2BP1	insulin-like growth factor 2 mRNA binding protein 1	10642	0.8041	[52]
IL1A	interleukin 1. alpha	3552	0.4672	[29]
ILF3	interleukin enhancer binding factor 3, 90kDa	3609	0.7057	[35]
KIF14	kinesin family member 14	9928	0.9346	[53]
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog	3845	0.4472	[54]
LARS2	leucyl-tRNA synthetase 2, mitochondrial	23395	0.8316	[27]
LCK	lymphocyte-specific protein tyrosine kinase	3932	0.5878	[44]
LST1	leukocyte specific transcript 1	7940	0.2105	[27]
MBD2	methyl-CpG binding domain protein 2	8932	0.4473	[55]
MMD	monocyte to macrophage differentiation-associated	23531	0.1662	[41]
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	4609	0.4586	[56,57]
P53AIP1	p53-regulated apoptosis-inducing protein 1	63970	0.8171	[58]
PIK3CA	phosphoinositide-3-kinase.	5290	0.1111	[59]
PLAU	plasminogen activator. urokinase	5328	0.1314	[60]
PRDX2	peroxiredoxin 2	7001	0.9166	[61]
PRSS3	protease, serine, 3	5646	0.5814	[62]
RRM1	ribonucleotide reductase M1 polypeptide	6240	0.8498	[43,45]
SLC1A7	solute carrier family 1, member 7	6512	0.4883	[35]
SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	6513	0.7739	[27]
SP100	SP100 nuclear antigen	6672	0.9490	[7]
STAT1	signal transducer and activator of transcription 1. 91kDa	6772	0.9245	[41]
STC1	stanniocalcin 1	6781	0.1508	[27]
STYK1	serine/threonine/tyrosine kinase 1	55359	0.6834	[63]
TAL2	T-cell acute lymphocytic leukemia 2	6887	0.4593	[35]
TERF2	telomeric repeat binding factor 2	7014	0.1028	[64]
TERT	telomerase reverse transcriptase	7015	0.7542	[65]
TFAP2B	transcription factor AP-2 beta	7021	0.1814	[9]
TOP2A	topoisomerase (DNA) II alpha	7153	0.1998	[66]
TYMS	thymidylate synthetase	7298	0.1742	[67,68]

[1]Gene symbols and names according to HUGO Gene Nomenclature Committee, HGNC database (http://www.genenames.org).
[2]Differences in survival were assessed with the log-rank test using SPSS version 14.0. In all statistical analyses, *P*-value less than 0.05

Despite the large literature about markers in NSCLC, our study shows that none of the genes we tested have their mRNA levels directly correlated with patient outcome. It is important to state that our analysis was based on gene expression status, not taking into account other relevant parameters, like gene mutation pattern, methylation, or cohort subgroups. As an example, *ERBB2* gene has prognostic value in patients with NSCLC when considering specific cohort gender (20). Likewise, *ERCC1* is effective in stage IIB-IIIA-IIIB of NSCLC (21) and *CD9* has prognostic value given the mutation status of *KRAS* gene (22).

Different experimental approaches have been used to establish each of the genes listed in table 1 as potentials biomarkers for predicting patient outcome. This

approach shows that biomarker candidates should be carefully tested in clinical samples and exemplifies a rational use of public high-throughput clinical data. In theory, it could be applied to validate any possible biomarker, optimizing the use of the information available in public databanks and serving as standard tool to guide future clinical trials. Thus, we would be maximizing the use of information already generated and increasing its applicability.

The panel generated by this tool must be further analyzed. Microarrays are well described as capable of determining the expression levels of thousands of genes simultaneously (23) and the ability to define cancer subtypes, recurrence of disease and response to specific therapies using DNA microarray-based gene expression signature has been demonstrated in multiple studies (24). Bild & Col. described the activation status of several oncogenic pathway based on the statistical combination of gene expression signatures (8). Nevertheless, we believe that the gene signatures should be obtained based on biological (not statistical) combination of high-throughput screening of cDNA microarray probes. In this scenario, fluctuation of gene expression within biological networks can be evaluated by landscape analysis, which can represent different functional states of the same gene network (25)

In summary, our research describes a reliable tool able to discriminate biomarkers performance, revealing that none of the 60 genes individually tested had shown sufficient statistical power to be safely included in clinical use, when compared to TNM system, which is considered gold standard by physicians. Therefore, these approach may strengthens the development of new biomarkers, since up-to-date there is still no prognostic biomarker (based on gene expression) available for NSCLC. As stated by Dr. Goldstraw in the last World Conference on Lung Cancer (26), it is still uncertain how to integrate the predictive information from biomarkers with the anatomical extent of disease described by the TNM system, which rises the possibility that T, N, and M could be joined by a B (biological) factor. Biomarkers will probably be the next major development in NSCLC staging.

5. ACKNOWLEDGEMENTS

This project is supported by Brazilian MCT/CNPq Universal funds (479860/2006-8 & 476114/2008-0) by MCT/CNPq INCT-TM (573671/2008-7) funds. C.B.M. was supported by a training grant from FAPERGS (08511895) and F.K. received a fellowship from MCT/CNPq (303613/2008-4). The author would like to thanks Dr Márcia Triunfol for critical review of this manuscript.

6. REFERENCES

1. DC Ihde: Chemotherapy of lung cancer. *N Engl J Med* 327, 1434-1441 (1992)

2. A Spira, DS Ettinger: Multidisciplinary management of lung cancer - Reply. *N Engl J Med* 350, 2009-2010 (2004)

3. HJ Sung, JY Cho: Biomarkers for the lung cancer diagnosis and their advances in proteomics. *BMP Rep* 41, 615-625 (2008)

4. CL Sawyers: The cancer biomarker problem. *Nature* 452, 548-552 (2008)

5. CW Chi-Shing: Potentially useful biomarkers for the diagnosis, treatment and prognosis of lung cancer. *Biomed Pharmacother* 61, 515-519 (2007)

6. J Niklinski, W Niklinska, J Laudanski, E Chyczewska, L Chyczewski: Prognostic molecular markers in non-small cell lung cancer. *Lung Cancer* 34, 53-58 (2001)

7. B Kim, HJ Lee, HY Choi, Y Shin, S Nam, G Seo, D Son, J Jo, J Kim, J Lee, J Kim, K Kim, S Lee: Clinical validity of the lung cancer biomarkers identified by bioinformatics analysis of public expression data. *Cancer Res* 67, 7431-7438 (2007)

8. AH Bild, G Yao, JT Chang, Q Wang, A Potti, D Chasse, M Joshi, D Harpole, J M Lancaster, A Berchuck, J A Olson Jr, J R Marks, H K Dressman, M West, J R Nevin: Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 439, 353-357 (2006)

9. WG Deng, G Jayachandran, G Wu, K Xu, JA Roth, L Ji: Tumor-specific activation of human telomerase reverses transcriptase promoter activity by activating enhancer-binding protein-2beta in human lung cancer cells. *J Biol Chem* 282, 26460-70 (2007)

10. D Vallbohmer, J Brabender, DY Yang, K Danenberg, PM Schneider, R Metzger, AH Holscher, PV Danenberg: Sex differences in the predictive power of the molecular prognostic factor HER2/neu in patients with non-small-cell lung cancer. *Clin Lung Cancer* 7, 332-337 (2006)

11. SV Sharma, DW Bell, J Settleman, DA Haber: Epidermal growth factor receptor mutations in lung cancer. *Nature Rev. Cancer* 7, 169–181 (2007)

12. W Pao, TY Wang, GJ Riely, VA Miller, Q Pan, M Ladanyi, MF Zakowski, RT Heelan, MG Kris, HE Varmus: *KRAS* mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2, e17 (2005)

13. MR Muller, J Thomale, MF Rajewsky, S Seeber: Drug Resistance and DNA repair in Leukaemia. *Cytotechnology* 27, 175-185 (1998)

14. E Felip, R Rosell: Testing for excision repair cross-complementing 1 in patients with non-small-cell lung cancer for chemotherapy response. *Expert Rev Mol Diagn* 7, 261-268 (2007)

15. G Damia, M D'Incalci: Mechanisms of resistance to alkylating agents. *Cytotechnology* 27, 165-173 (1998)
16. F Drabløs, E Feyzi, PA Aas, CB Vaagbø, B Kavli, MS Bratlie, J Pena-Diaz, M Otterlei, G Slupphaug, HE Krokan: Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)* 3, 1389-1407 (2004)

17. C Schettino, MA Bareschino, P Maione, A Rossi, F Ciardiello, C Gridelli: The potential role of pharmacogenomic and genomic in the adjuvant treatment of early stage non small cell lung cancer. *Curr Genomics* 9, 252-262 (2008)

18. R Rosell, RV Lord, M Taron, N Reguart: DNA repair and cisplatin resistance in non-small-cell lung cancer. *Lung Cancer* 38, 217-227 (2002)

19. R Rosell, M Taron, V Alberola, B Massuti, E Felip: Genetic testing for chemotherapy in non-small cell lung cancer. *Lung Cancer* 41, 97-102 (2003)

20. M Miyake, M Adachi, C Huang, M Higashiyama, K Kodama, T Taki: A novel molecular staging protocol for non-small cell lung cancer. *Oncogene* 18, 2397-2404 (1999)

21. C McNeil: New NSCLC staging raises treatment issues. *J Natl Cancer Inst* 99, 1748-1749 (2007)

22. B Karczmarek-Borowska, A Filip, J Wojcierowski, A Smolen, E Korobowicz, I Korszen-Pilecka, M Zdunek: Estimation of prognostic value of Bcl-xL gene expression in non-small cell lung cancer. *Lung Cancer* 51, 61-69 (2006)

23. M Campioni, V Ambrogi, E Pompeo, G Citro, M Castelli, EP Spugnini, A Gatti, P Cardelli, L Lorenzon, A Baldi, TC Mineo: Identification of genes down-regulated during lung cancer progression: a cDNA array study. *J Exp Clin Cancer Res* 27:38 (2008)

24. S Ramaswamy, T R Golub: DNA microarrays in clinical oncology. J. *Clin. Oncol.* 20, 1932-1941 (2002)

25. MA Castro, JL Filho, RJ Damolin, M Sinigaglia, JC Moreira, JC Mombach, RM de Almeida: ViaComplex: software for landscape analysis of gene expression networks in genomic context. *Bioinformatics* 25, 1468-1469 (2009)

26. M Adachi, T Taki, T Konishi, CI Huang, M Higashiyama, M Miyake: Novel staging protocol for non-small-cell lung cancers according to MRP-1/CD9 and KAI1/CD82 gene expression. *J Clin Oncol* 16, 1397-1406 (1998)

27. Y Lu, W Lemon, PY Liu, Y Yi, C Morrison, P Yang, Z Sun, J Szoke, WL Gerald, M Watson, R Govindan, M You: A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med* 3, e467 (2006)

28. D Stav, I Bar, J Sandbank: Usefulness of CDK5RAP3, CCNB2, and RAGE genes for the diagnosis of lung adenocarcinoma. *Int J Biol Markers* 22, 108-113 (2007)

29. F Oshita, A Sekiyama, H Saito, K Yamada, K Noda, Y Miyagi: Genome-wide cDNA microarray screening of genes related to the benefits of paclitaxel and irinotecan chemotherapy in patients with advanced non-small cell lung cancer. *J Exp Ther Oncol* 6, 49-53 (2006)

30. I Takanami, K Takeuchi: Autocrine motility factor-receptor gene expression in lung cancer. *Jpn J Thorac Cardiovasc Surg* 51, 368-373 (2003)

31. I Takanami, K Takeuchi, H Watanabe, T Yanagawa, K Takagishi, A Raz: Significance of autocrine motility factor receptor gene expression as a prognostic factor in non-small-cell lung cancer. *Int J Cancer* 95, 384-387 (2001)

32. B Karczmarek-Borowska, A Filip, J Wojcierowski, Smolen A, I Pilecka, A Jablonka: Survivin antiapoptotic gene expression as a prognostic factor in non-small cell lung cancer: *in situ* hybridization study. *Folia Histochem Cytobiol* 43, 237-242 (2005)

33. M Monzo, R Rosell, E Felip, J Astudillo, JJ Sanchez, J Maestre, C Martin, A Font, A Barnadas, A Abad: A novel anti-apoptosis gene: Re-expression of survivin messenger RNA as a prognosis marker in non-small-cell lung cancers. *J Clin Oncol* 17, 2100-2104 (1999)

34. R Rosell, M Skrzypski, E Jassem, M Taron, R Bartolucci, JJ Sanchez, P Mendez, I Chaib, L Perez-Roca, A Szymanowska, W Rzyman, F Puma, G Kobierska-Gulida, R Farabi, J Jassem: BRCA1: a novel prognostic factor in resected non-small-cell lung cancer. *PLoS ONE* 2, e1129 (2007)

35. NL Guo, YW Wan, K Tosun, H Lin, Z Msiska, DC Flynn, SC Remick, V Vallyathan, A Dowlati, X Shi, V Castranova, DG Beer, Y Qian: Confirmation of Gene Expression-Based Prediction of Survival in Non-Small Cell Lung Cancer. *Clin Cancer Res* 14, 8213-8220 (2008)

36. CC Ho, SH Kuo, PH Huang, HY Huang, CH Yang, PC Yang: Caveolin-1 expression is significantly associated with drug resistance and poor prognosis in advanced non-small cell lung cancer patients treated with gemcitabine-based chemotherapy. *Lung Cancer* 59, 105-110 (2008)

37. R Li, SJ An, ZH Chen, GC Zhang, JQ Zhu, Q Nie, Z Xie, AL Guo, TS Mok, YL Wu: Expression of cyclin D1 splice variants is differentially associated with outcome in non-small cell lung cancer patients. *Human Pathology* 39, 1792-1801 (2008)

38. F Oshita, A Sekiyama, H Ito, Y Kameda, Y Miyagi: Genome-wide cDNA microarray screening of genes related to survival in patients after curative resection of non-small cell lung cancer. *Oncol Rep* 16, 817-821 (2006)

39. DJ Raz, MR Ray, JY Kim, B He, M Taron, M Skrzypski, M Segal, DR Gandara, R Rosell, DM Jablons: A

Multigene Assay Is Prognostic of Survival in Patients with Early-Stage Lung Adenocarcinoma. *Clin Cancer Res* 14, 5565-5570 (2008)

40. MF Tsai, CC Wang, GC Chang, CY Chen, HY Chen, CL Cheng, YP Yang, CY Wu, FY Shih, CC Liu, HP Lin, YS Jou, SC Lin, CW Lin, WJ Chen, WK Chan, JJ Chen, PC Yang: A new tumor suppressor DnaJ-like heat shock protein, HLJ1, and survival of patients with non-small-cell lung carcinoma. *J Natl Cancer Inst* 98, 825-838 (2006)

41. HY Chen, SL Yu, CH Chen, GC Chang, CY Chen, A Yuan, CL Cheng, CH Wang, HJ Terng, SF Kao, WK Chan, HN Li, CC Liu, S Singh, WJ Chen, JJ Chen, PC Yang: A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 356, 11-20 (2007)

42. T Shibata, S Hanada, A Kokubu, Y Matsuno, H Asamura, T Ohta, M Sakamoto, S Hirohashi: Gene expression profiling of epidermal growth factor receptor/KRAS pathway activation in lung adenocarcinoma. *Cancer Sci* 98, 985-991 (2007)

43. P Ceppi, M Volante, S Novello, I Rapa, KD Danenberg, PV Danenberg, A Cambieri, G Selvaggi, S Saviozzi, R Calogero, M Papotti, GV Scagliotti ERCC1 and RRM1 gene expressions but not EGFR are predictive of shorter survival in advanced non-small-cell lung cancer treated with cisplatin and gemcitabine. *Ann Oncol* 17, 1818-1825 (2006)

44. H Tang, E Goldberg: Homo Sapiens Lactate Dehydrogenase c (Ldhc) Gene Expression in Cancer Cells is Regulated by Transcription Factor Sp1, CREB and CpG Island Methylation. *J Androl* 30, 157-167 (2008)

45. Z Zheng, T Chen, X Li, E Haura, A Sharma, G Bepler: DNA synthesis and repair genes RRM1 and ERCC1 in lung cancer. *N Engl J Med* 356, 800-808 (2007)

46. L Tian, M Suzuki, T Nakajima, R Kubo, Y Sekine, K Shibuya, K Hiroshima, Y Nakatani, T Fujisawa, I Yoshino: Clinical significance of aberrant methylation of prostaglandin E receptor 2 (PTGER2) in nonsmall cell lung cancer: association with prognosis, PTGER2 expression, and epidermal growth factor receptor mutation. *Cancer* 113, 1396-1403 (2008)

47. Y Mano, K Takahashi, N Ishikawa, A Takano, W Yasui, K Inai, H Nishimura, E Tsuchiya, Y Nakamura, Y Daigo Fibroblast growth factor receptor 1 oncogene partner as a novel prognostic biomarker and therapeutic target for lung cancer. *Cancer Sci* 98, 1902-1913 (2007)

48. S Iwakiri, M Sonobe, S Nagai, T Hirata, H Wada, R Miyahara: Expression status of folate receptor alpha is significantly correlated with prognosis in non-small-cell lung cancers. *Ann Surg Oncol* 15, 889-899 (2008)

49. B Meyer, S Loeschke, A Schultze, T Weigel, M Sandkamp, T Goldmann, E Vollmer, J Bullerdiek: HMGA2 overexpression in non-small cell lung cancer. *Mol Carcinog* 46, 503-511 (2007)

50. MI Gallegos Ruiz, K Floor, P Roepman, JA Rodriguez, GA Meijer, WJ Mooi, E Jassem, J Niklinski, T Muley, N van Zandwijk, EF Smit, K Beebe, L Neckers, B Ylstra, G Giaccone: Integration of gene dosage and gene expression in non-small cell lung cancer, identification of HSP90 as potential target. *PLoS ONE* 3, e0001722 (2008)

51. ES Lee, DS Son, SH Kim, J Lee, J Jo, J Han, H Kim, HJ Lee, HY Choi, Y Jung, M Park, YS Lim, K Kim, YM Shim, BC Kim, K Lee, N Huh, C Ko, K Park, JW Lee, YS Choi, J Kim: Prediction of Recurrence-Free Survival in Postoperative Non-Small Cell Lung Cancer Patients by Using an Integrated Model of Clinical Information and Gene Expression. *Clin Cancer Res* 14, 7397-7404 (2008)

52. T Kato, S Hayama, T Yamabuki, I Nshikawa, M Miyamoto, T Ito, E Tsuchiya, S Kondo, Y Nakamura, Y Daigo: Increased expression of insulin-like growth factor-II messenger RNA-binding protein 1 is associated with tumor progression in patients with lung cancer. *Clin Cancer Res* 13, 434-442 (2007)

53. TW Corson, CQ Zhu, SK Lau, FA Shepherd, MS Tsao, BL Gallie: KIF14 messenger RNA expression is independently prognostic for outcome in lung cancer. *Clin Cancer Res* 13, 3229-3234 (2007)

54. T Shibata, S Hanada, A Kokubu, Y Matsuno, H Asamura, T Ohta, M Sakamoto, S Hirohashi: Gene expression profiling of epidermal growth factor receptor/KRAS pathway activation in lung adenocarcinoma. *Cancer Sci* 98, 985-991 (2007)

55. J Xing, DJ Stewart, J Gu, C Lu, MR Spitz, X Wu: Expression of methylation-related genes is associated with overall survival in patients with non-small cell lung cancer. *Br J Cancer* 98, 1716-1722 (2008)

56. Z Sun, P Yang, MC Aubry, F Kosari, C Endo, J Molina, G Vasmatzis: Can gene expression profiling predict survival for patients with squamous cell carcinoma of the lung? *Mol Cancer* 3, 35 (2004)

57. S Tomida, K Koshikawa, Y Yatabe, T Harano, N Ogura, T Mitsudomi, M Some, K Yanagisawa, T Takahashi, H Osada, T Takahashi: Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. *Oncogene* 23, 5360-5370 (2004)

58. SI Yamashita, Y Masuda, N Yoshida, H Matsuzaki, T Kurizaki, Y Haga, S Ikei, M Miyawaki, Y Kawano, M Chujyo, K Kawahara: p53AIP1 expression can be a prognostic marker in non-small cell lung cancer. *Clin Oncol (R Coll Radiol)* 20, 148-151 (2008)

59. B Angulo, A Suarez-Gauthier, F Lopez-Rios, PP Medina, E Conde, M Tang, G Soler, A Lopez-Encuentra, JC Cigudosa, M Sanchez-Cespedes: Expression signatures in lung cancer reveal a profile for EGFR-mutant tumours and identify selective PIK3CA overexpression by gene amplification. *J Pathol* 214, 347-356 (2008)

60. HS Hofmann, B Bartling, A Simm, R Murray, N Aziz, G Hansen, RE Silber, S Burdach: Identification and classification of differentially expressed genes in non-small cell lung cancer by expression profiling on a global human 59.620-element oligonucleotide array. *Oncol Rep* 16, 587-595 (2006)

61. JH Kim, PN Bogner, SH Baek, N Ramnath, P Liang, HR Kim, C Andrews, YM Park: Up-regulation of peroxiredoxin 1 in lung cancer and its implication as a prognostic and therapeutic target. *Clin Cancer Res* 14, 2326-2333 (2008)

62. S Diederichs, E Bulk, B Steffen, P Ji, L Tickenbrock, K Lang, KS Zanker, R Metzger, PM Schneider, V Gerke, M Thomas, WE Berdel, H Serve, C Muller-Tidow: S100 family members and trypsinogens are predictors of distant metastasis and survival in early-stage non-small cell lung cancer. *Cancer Res* 64, 5564-5569 (2004)

63. T Amachika, D Kobayashi, R Moriai, N Tsuji, N Watanabe: Diagnostic relevance of overexpressed mRNA of novel oncogene with kinase-domain (NOK) in lung cancers. *Lung Cancer* 56, 337-340 (2007)

64. X Lin, J Gu, C Lu, MR Spitz, X Wu: Expression of telomere-associated genes as prognostic markers for overall survival in patients with non-small cell lung cancer. *Clin Cancer Res* 12, 5720-5725 (2006)

65. N Miura, H Nakamura, R Sato, T Tsukamoto, T Harada, S Takahashi, Y Adachi, K Shomori, A Sano, Y Kishimoto, H Ito, J Hasegawa, G Shiota: Clinical usefulness of serum telomerase reverse transcriptase (hTERT) mRNA and epidermal growth factor receptor (EGFR) mRNA as a novel tumor marker for lung cancer. *Cancer Sci* 97, 1366-1373 (2006)

66. P Ceppi, M Longo, M Volante, S Novello, S Cappia, E Bacillo, G Selvaggi, S Saviozzi, R Calogero, M Papotti, GV Scagliotti: Excision repair cross complementing-1 and topoisomerase IIalpha gene expression in small-cell lung cancer patients treated with platinum and etoposide: a retrospective study. *J Thorac Oncol* 3, 583-589 (2008)

67. Y Shintani, M Ohta, H Hirabayashi, H Tanaka, K Iuchi, K Nakagawa, H Maeda, T Kido, S Miyoshi, H Matsuda: Thymidylate synthase and dihydropyrimidine dehydrogenase mRNA levels in tumor tissues and the efficacy of 5-fluorouracil in patients with non-small-cell lung cancer. *Lung Cancer* 45, 189-196 (2004)

68. Y Shintani, M Ohta, H Hirabayashi, H Tanaka, K Iuchi, K Nakagawa, H Maeda, T Kido, S Miyoshi, H Matsuda: New prognostic indicator for non-small-cell lung cancer, quantitation of thymidylate synthase by real-time reverse transcription polymerase chain reaction. *Int J Cancer* 104, 790-795 (2003)

Key Words: Bioinformatic approach, Non-Small Cell Lung Cancer, prognostic biomarkers

Send correspondence to: Fabio Klamt, Centro de Estudos em Estresse Oxidativo, Departamento de Bioquimica, Universidade Federal do Rio Grande do Sul (UFRGS), Rua Ramiro Barcelos 2600-anexo, Porto Alegre 90035-003, Brazil, Tel: 555133085577, Fax: 55513308 535, E-mail: 00025267@ufrgs.br

http://www.bioscience.org/current/volE3.htm