CANCER GENOMICS & PROTEOMICS *12*: 231-242 (2015)

Gene-expression Profiling in Non-small Cell Lung Cancer with Invasion of Mediastinal Lymph Nodes for Prognosis Evaluation

MADALINA GRIGOROIU[1], REBECCA TAGETT[2], SORIN DRAGHICI[1], SIMONA DIMA[1], ANCA NASTASE[1], RALUCA FLOREA[1], ANDREI SOROP[1], VERONICA ILIE[1], NICOLAE BACALBASA[1], VALERIA TICA[1], VIKTORIA LASZLO[1], AUDREY MANSUET-LUPO[3], DIANNE DAMOTTE[3], WALTER KLEPETKO[4], IRINEL POPESCU[1] and JEAN FRANCOIS REGNARD[1]

[1]*Center of Digestive Diseases and Liver Transplantation, Fundeni Clinical Institute, Bucharest, Romania;*
[2]*Intelligent Systems and Bioinformatics Laboratory (ISBL), Department of Computer Science - Wayne State University, Detroit, MI, U.S.A.;*
[3]*Pathology Department, Hôtel-Dieu Hospital, Paris, France;*
[4]*Department of Thoracic Surgery, Vienna General Hospital, Vienna, Austria*

Abstract. *Background/Aim: The aim of the study was to determine the pathways and expression profile of the genes that might predict response to neoadjuvant chemotherapy in patients with stage IIIA non-small cell lung cancer (NSCLC). Materials and Methods: We evaluated, by microarray, the gene-expression profile of tumoral mediastinal lymph node samples surgically removed from 27 patients with stage IIIA NSCLC before neoadjuvant chemotherapy treatment. Depending on the response to the induction treatment, the patients were divided in two groups: group A: patients whose disease evolved, stabilized or who had minor response to chemotherapy, and group B: patients whose disease stabilized or had major response to chemotherapy. Results: The microarray experiments identified 1,127 genes with a modified expression in the tumoral tissue compared to normal tissue with p≤0.05 and 44 genes with p≤0.01. The identified up-regulated genes between tumoral versus normal tissue included collagen, type I, alpha 1 (COL1A1), inhibin beta A (INHBA) and thioredoxin interacting protein (TXNIP). Pathways identified with a false-discovery rate of <0.005 included: cytokine pathways, focal adhesion or extracellular matrix receptor interaction. Conclusion: Our approach identified important characteristics of NSCLC and pointed-out molecular differences between sub-groups of patients based on their response to therapy.*

The ultimate aim of reducing the high mortality rate in lung cancer is related to an increased efficiency of prevention/determination of cancer risk factor strategies and early detection of lung cancer while still in a curable stage. The major risk factors for lung cancer occurrence are smoking and high air pollution in large cities (1). Despite prevention campaigns and improved survival during the past two decades, lung cancer still represents the second leading cause of disease from cancer and the highest cancer-related death rate in Western countries. Approximately 55-60% of patients are diagnosed with disease in incurable stages, with distant metastases. Therefore, the 5-year survival rate considering all stages is only 13-16%. Non-small cell lung cancer (NSCLC) represents the most frequent type of bronchial tumor. It is classified into two major histological sub-types: adenocarcinoma and squamous cell carcinoma. Both sub-types are very different in regard to copy number, DNA methylation, genetic mutations, transcriptome, proteome and biomarkers. Defining different lung cancer entities based on clinical, pathological and molecular alterations also determines disease evolution and treatment options. In spite of significant progress in the development of targeted-therapy, the high mortality rate in lung cancer firmly emphasizes the need for prevention and efficient detection of lung cancer, as well as a better classification, enabling patients to benefit from more specific therapy.

Microarrays have been used for more than two decades in pre-clinical research to determine gene-expression profiling associated with different pathological sub-groups or diverse clinical evolution and to evaluate biological and cellular functions determined by these gene panels.

In NSCLC with invasion of mediastinal lymph nodes (N2+), the currently accepted treatment attitude is a multimodal approach, based on platinum salts chemotherapy, followed by complete surgical resection for the patients that

Correspondence to: Madalina Grigoroiu, MD, Fundeni Clinical Institute, Fundeni Ave, No. 258, 022328 Bucharest, Romania. Tel: +40 213180417, Fax: +40 213180417, E-mail: madalinalidia@yahoo.com

Key Words: Microarray, non-small cell lung cancer, gene signature.

Table I. *Exclusion causes*.

Exclusion cause	Number of patients
Lack of tumoral tissue in mediastinal lymph nodes collected by surgical biopsy (pN0)	28
Benign lesions in mediastinal lymph nodes	1
Small-cell lung carcinoma	6
Advanced invasion of mediastinal lymph nodes (bulky N2)	2
Lymphoma in mediastinal lymph nodes	1
Using frozen tumoral lymph node tissue for diagnosis	1
Lack of tumoral tissue in the frozen lymph node tissue	2
Random discovery of type N3 lymph node invasion	3
Total	44

have responded or have been stabilized by chemotherapy. After neoadjuvant chemotherapy, approximately 5% of patients will experience a full response. Most patients will experience a partial response (more than 50% tumor reduction); even so, the number of patients that are only to be stabilized or experience a minor response is still very important (approximately 40%). Five-year global survival rate of patients with NSCLC with the invasion of mediastinal lymph nodes (N2+) is 20-25%, but the sub-group analysis shows that 5-year survival rate of those who respond to chemotherapy is 42% *versus* 10% of those who do not respond to chemotherapy (2).

The objective of the present study was to determine the gene expression profile in patients with stage IIIA NSCLC (N2) and to characterize the patients who responded to neoadjuvant chemotherapeutic treatment in relation to those who did not respond or who experienced a minor response to this treatment based on the determined profiles.

We performed microarray experiments on tumoral tissue samples derived from tumoral mediastinal lymph node samples surgically removed before treatment by neoadjuvant chemotherapy from 27 patients with stage IIIA NSCLC. Depending on the response to neoadjuvant treatment according to World Health Organization criteria (3), two groups of patients were determined: group A: patients who had a partial response, whose disease stabilized or had not responded to chemotherapy; group B: patients who had a major or complete response to induction treatment.

Microarray data were also analyzed for the identification of the main pathways involved in the differentiation of the two patient groups.

Materials and Methods

Study population. All of the tissue samples and the relevant clinical data were obtained prospectively based on a written informed consent from the Departments of Thoracic Surgery at Hotel Dieu Hospital in Paris affiliated to Paris 5 Rene Descartes Medicine Faculty and at Vienna General Hospital, Austria between January 2011-December 2012. A total of 73 patients diagnosed with or suspected of having clinical stage IIIA NSCLC were originally included in the study. All patients benefited from a surgical procedure (mediastinoscopy or thoracoscopy) for the histological confirmation of the mediastinal lymph node invasion. Forty-four patients were subsequently excluded from the study following the histological analysis of the collected mediastinal lymph node tissue (Table I).

All patients included in the study had clinical stage IIIA disease (invasion of mediastinal lymph nodes, N2+) with pathological confirmation. The basic features of the 29 patients are detailed in Table II. The average age of the patients was 61 years old, with male patients being predominant (n=18; 62%). Adenocarcinoma (n=22; 75.8%) was the most frequent histological type.

Neoadjuvant chemotherapy. The average number of administered chemotherapy cycles was three (2-4 interval): one (3.2%) patient received two cycles, 21 (72.6%) patients received three cycles, six (21%) patients received four cycles and one patient (3.2%) received five. All patients were treated with a combination of two drugs, the first chemotherapeutic being, invariably, cisplatin (n=23, 79%) or carboplatin (n=5, 21%), the combination of cisplatin/pemetrexed being administered in most cases (n=15; 52%). The second chemotherapeutic was most frequently etoposide (n=7, 24%), followed by gemcitabine (n=3, 10%) and vinorelbine (n=3, 10%). Twenty eight patients concluded induction treatment by chemotherapy. The assessment of treatment response was performed by CT scan/PET and showed that disease in 13 patients (46.4%) evolved under chemotherapy, in seven patients (25%) it was stabilized and eight patients (28.6%) had partial responses. No patient had complete clinical response. The global rate of clinical response to neoadjuvant chemotherapy was 15 patients (51.7%). These patients subsequently received surgical intervention.

Surgery. One patient had pleural carcinosis discovered at thoracotomy and was waived surgical resection. The other 14 (50%) patients underwent surgical resection: 11 lobectomies (78.5%), one bilobectomy (7%) and two pneumonectomies (14.5%). Systematic lymphadenectomy of mediastinal lymph nodes was performed on every patient. All patients had complete resection (R0).

Histological response. Residual N2 disease was present in seven (25%) of the patients. In seven patients (25%), a mediastinal down-staging was noted: N1 (n=1; 3.5%) or N0 (n=6; 21.5%) and they were included in group B of the patients sensitive to chemotherapy. Of these, two (7%) had a complete histological response. The other 21 patients (14

Table II. *General features of the studied population.*

Features	Number (%)
Patients	29
Gender	
Male	18 (62)
Female	11 (38)
Average age, years (range)	61 (40-75)
Histology	
Adenocarcinoma	22 (75.8)
Squamous cell carcinoma	3 (10)
Large cell neuroendocrine carcinoma	3 (10)
Large cell carcinoma	1 (3.2)
Clinical stage of the tumor	
T1a	3 (10)
T1b	8 (27.5)
T2a	10 (34.5)
T2b	2 (7)
T3	6 (20)

patients whose disease evolved under chemotherapy and seven who did not present a mediastinal down-staging by chemotherapy) were included in group A of the patients who did not respond to chemotherapy. The pathological stage was: no residual disease in two cases, stage IA in three cases, IIA in three and IIIA in six.

Total RNA extraction. Total RNA was isolated from the collected tissues, by a two-step extraction protocol: the first step was carried out using Trizol (Invitrogen, Life Technology, Carlsbad, CA, USA) extraction and the second by RNA purification (clean-up) using Absolutely RNA Microprep kit (Agilent Technologies, Santa Clara, CA, USA). RNA samples were quantitatively and qualitatively assessed by NanoDrop ND1000 (Thermo Scientific, Arlington, TX, USA) and Agilent Bioanlyzer 2100 (Santa Clara, CA, USA).

Microarray analysis. Agilent One-Color Microarray-Based Gene Expression Analysis, version 5.5/February 2007 technology was used for gene–expression profiling. Microarray slides with Whole Human Genome Oligo Microarray with SurePrint Technology 4 ×44 K (Agilent Technologies, Santa Clara, CA, USA) were used.

Quality control and normalization. Quality control was done using the Agi4×44 Preprocess R package [10]. For cases with more than one specimen for a sample, the "best" sample was used, based on quality control. The final matrix of log2 data consisted of 35 samples (columns) of 17,808 unique entry IDs (rows).

Data analysis. In order to analyze data, .tif images generated by Scan Control software were imported to Feature Extraction 9.1 software (Agilent Technologies). GE1-v5_91_0806 protocol and 026652_D_F_20110325 grid were used. Signaling pathways analysis was performed by Pathway Guide (www.advaitabio.com), using the "All Genes" option (4, 5). With this option, a cut-off value is not required to define a gene subset for input; all of the genes and their fold changes are inputs. The algorithm takes into account only gene interaction, but not over-representation of significant genes in a molecular pathway.

Table III. *Affiliation of the 24 patients to one of the two groups: group A: non-responders, group B: responders. Samples highlighted in grey were subsequently removed from analysis.*

Sample name	Gender	Group	Normal tissue available
LYN_65	F	A	No
LYN_77	F	A	No
T11B00790	F	A	Yes
T11B02702	F	A	Yes
T11B02941	F	A	Yes
T11B04855	F	A	No
T12B03590	F	A	No
T12B0 5600	F	A	No
T12B4727	F	A	No
T12B5696	F	A	No
LYN_101	M	A	No
LYN_70	M	A	No
LYN_72	M	A	No
LYN_82	M	A	No
LYN_84	M	A	No
T11B02967	M	A	No
T11B03023	M	A	No
T11B05089	M	A	Yes
T12B02200	M	A	No
T12B02876	M	A	Yes
LYN_48	F	B	No
LYN_43	M	B	No
T11B00602	M	B	Yes
T11B04901	M	B	No
T11B1622	M	B	Yes
T11B2052	M	B	Yes
T12B03843	M	B	Yes

M: Male patients, F: female patients.

Results

Profiles of responders versus non-responders. Before microarray analysis, two samples were removed due to poor quality, and three samples were excluded due to poor hybridization after microarray experiment (Figures 1 and 2).

A total of 24 patients were analyzed (Table III). A total of 1,127 genes were identified with modified expression in tumoral tissue compared to normal tissue at $p \leq 0.05$ and 44 genes at $p \leq 0.01$ (Table IV).

Out of the identified genes, those with the highest difference in gene expression identified between tumoral tissue *versus* normal tissue included collagen, type I, alpha 1 (*COL1A1*), inhibin beta A (*INHBA*) and thioredoxin interacting protein (*TXNIP*).

Differential Expression. There were two groups of patients: those who did not respond to neoadjuvant chemotherapy treatment (non-responders, group A) and those who did (responders, group B).

Table IV. *List of 44 genes with FDR-adjusted p-value ≤0.01 identified by microarray analysis in tumoral tissue compared to normal tissue.*

Entrez Gene ID	Gene symbol	LogFC	Average expression
22821	*RASA3*	–1.1732	11.1369
27286	*SRPX2*	2.5458	8.8903
26508	*HEYL*	1.8766	9.1929
2191	*FAP*	4.0222	9.4378
1277	*COL1A1*	4.2312	12.9965
7434	*VIPR2*	–0.9410	6.7205
55379	*LRRC59*	1.0043	10.2502
7076	*TIMP1*	2.8775	12.5334
1741	*DLG3*	1.1230	8.5964
284217	*LAMA1*	1.3678	6.6742
6659	*SOX4*	1.6679	8.0695
54972	*TMEM132A*	2.6656	9.8623
11215	*AKAP11*	–0.8144	9.1058
7083	*TK1*	2.6764	12.2276
204	*AK2*	0.6798	9.7956
10848	*PPP1R13L*	1.8321	9.1308
1278	*COL1A2*	2.7989	13.0190
4192	*MDK*	2.3983	11.7570
57406	*ABHD6*	–1.2653	8.9651
101060503	*LOC101060503*	–1.7051	14.1635
10628	*TXNIP*	–1.7051	14.1635
10631	*POSTN*	3.5406	10.4058
113452	*TMEM54*	2.4189	10.6740
376267	*RAB15*	1.8299	10.1181
55055	*ZWILCH*	0.9057	7.6660
56262	*LRRC8A*	0.9120	11.5497
157	*ADRBK2*	–1.6111	10.0741
2050	*EPHB4*	1.1489	8.3348
5295	*PIK3R1*	–1.3390	9.2707
5754	*PTK7*	1.0388	6.9971
541471	*MIR4435-2HG*	1.1854	10.3847
51203	*NUSAP1*	1.8986	8.2627
51330	*TNFRSF12A*	2.4010	10.3348
79682	*CENPU*	1.5400	8.3846
22861	*NLRP1*	–0.9628	7.9008
80325	*ABTB1*	–0.9842	10.5884
332	*BIRC5*	2.6018	9.9363
6790	*AURKA*	1.7214	8.0565
3624	*INHBA*	3.4296	9.5855
128408	*BHLHE23*	–1.4437	11.4465
57648	*KIAA1522*	2.3820	9.5881
2150	*F2RL1*	1.8129	7.5159
84627	*ZNF469*	1.5757	9.1298
79686	*LINC00341*	–1.4099	10.1937

FC: Fold change calculated between tumoral and normal tissues. Entrez Gene ID 101060503 was discontinued and replaced by entrez gene ID 10628.

Table V. *Significant Gene Ontology terms for high variance gene panels in responders.*

Biological Processes	- Keratinocyte differentiation
	- Epithelium development
	- Epidermal cell differentiation
	- Response to lipid
Molecular Functions	Structural constituents of the cytoskeleton
Cellular Components	- Cornfield envelope
	- Proteinaceous extracellular matrix
	- Extracellular region

Table VI. *Significant Gene Ontology terms for high variance gene panels to non-responders.*

Biological Processes	- Response to hormone stimulus
	- Negative regulation of hydrolase activity
	- Negative regulation of endopeptidase activity
	- Response to lipids
Molecular Functions	Endopeptidase inhibitor activity
Cellular Components	Extracellular region

It was noticed that among the responders that there was only one female (unpaired sample). Since gender is a confounding variable, all functional analysis for treatment response was performed using male patients only. Some samples were paired to samples of non-tumoral lymph nodes from the same patient, and others were not (Table III). These type of data are referred to in the literature as "partially paired" and can be approached by linear mixed effects models (6, 7).

For the purpose of simplicity however, instead applying the partially paired approach, we first considered the paired data then we considered the data as a whole, regardless of the pairing. Due to the small number of samples that were paired, we only achieved significant results with unpaired data and proceeded in this fashion.

Responders versus non-responders. Tumoral lymph node tissue. There were no differentially expressed genes after correction for multiple comparisons; therefore we built gene sets for functional analysis using differential variance. The deregulation of cancerous gene expression often results in more highly variable expression of key genes, and recently researchers have been trying to take advantage of tumor heterogeneity in order to characterize disease (36). Gene sets were prepared for gene ontology analysis (GO) (bioinfo.vanderbilt.edu/webgestalt) (8, 9), using combinations of high- or low-variance genes in responders and non-responders. Twelve different groups were defined and 50 representative genes (highest and lowest variance) were selected as input for each group. Tables V and VI show the significant GO terms for responders and non-responders, respectively, using genes with a high variance. It was observed that there were only Molecular Process results for the low variance gene sets, and in all cases, including responders and non-responders, the "transmembrane signaling receptor activity"

Table VII. *The FDR adjusted p-value is shown for KEGG [37] pathways that are significant at FDR < 0.05, for at least one of the 5 input datasets, labeled A through E. Values below the cut-off value are underlined. The first line of the header shows the gender of the patients of the subgroup. The second line shows the lymph node type. The last line of the header shows the number of non-responders/responders considered. There was only one woman in the responder group; therefore, a female subgroup in particular was not considered.*

KEGG Pathways	A MF All 24/11	B M All 9/10	C MF Tumoral 18/7	D M Tumoral 7/6	E M Non-tumoral 2/4
Graft-*versus*-host disease	0.04	0.08	0.07	0.09	0.06
Rheumatoid polyarthritis	0.04	0.05	0.07	0.07	0.06
Allograft rejection	0.04	0.07	0.07	0.09	0.09
Autoimmune thyroiditis	0.04	0.05	0.07	0.09	0.06
Systemic lupus erythematous	0.04	0.26	0.13	0.32	0.24
Type I diabetes	0.05	0.08	0.07	0.10	0.06
Cytokine signaling pathway	0.97	0.05	0.93	0.24	0.96

M: Male patients, MF: both male and female patients.

and related groups (*e.g.* "receptor activity") were significant. Many of the GO categories were significant for some of the groups; therefore, we filtered results in order to produce the summaries in Tables V and VI. We kept GO terms with FDR adjusted *p*-values<0.01, and at least 5 genes in a term for Biological Process, at least 4 genes in a term for Molecular Function, and at least 3 genes in a term for Cellular Component.

The different minimum number of genes per category was selected because there was a preponderance of Biological Process results, and less so for Molecular Function and Cellular Component, respectively. For the sake of non-redundancy, only one category is shown in the case where there are several similar results. For example, "extracellular region" (Cellular Component) is used alone instead of citing "extracellular space", "extracellular region part", and "extracellular region". The GO enrichment analysis showed that genes for transmembrane signaling do not vary from patient to patient, independent of response status. The highly variable responder genes involve keratin and cornification. The cornified envelope in epidermal skin cells is made of keratins, and is surrounded by lipids. It is associated with barrier malfunctions and results in a signal for apoptosis (10). The highly variable non-responder genes involve enzymatic activity and hormone sensitivity.

Molecular pathway analysis. We included tumoral and non-tumoral lymph nodes and tried to determine some molecular pathways that would define the difference in response to chemotherapy. Should there be a genetic signal for the response to chemotherapy, it may be a characteristic of the patient, and present in tumoral and non-tumoral lymph nodes. Data were again divided into subgroups (Table VII). The results suggest, not surprisingly, that only larger sample size allow good significance. In Table VII, significant pathways are presented only for the groups with at least 10 samples per category. There are 7 pathways that were significant (FDR corrected *p*-value < 0.05) in at least one of the subgroups, and the only subgroups with significant pathways are the largest subgroups. All these pathways are related to the immune system; one is a signaling pathway and the rest are autoimmune related or immune response to non-self tissue. The Chemokine signaling pathway is also related to autoimmune diseases (11); chemokines are involved in the process of lymphocyte recruitment. It remains unclear whether the response signal to chemotherapy is present both in invaded and non-invaded lymph nodes. Nevertheless, considering columns D and E of Table VII, we can see that the FDR adjusted *p*-values are well correlated with each other, except the Chemokine signaling pathway, which is more involved in invaded lymph nodes.

Pathways in Table VII are reminiscent of what one might expect when comparing the transcriptome of a male patient with that of a female patient, since the great difference in the gender-related transcriptome resides in the immune system processes (12, 13). Nonetheless, the pathways' significance of male patient group in subgroup B (invaded and non-invaded lymph nodes) also suggests an immune function similar to autoimmune behavior. As negative control, the Pathway Guide was run when comparing men to women. Gender-specific significant pathways are presented in Table VIII. Many of them are pathways for autoimmune diseases, known to be more frequent in women than in men (12).

Table VIII shows there are 19 significant pathways attributed to gender differences. Seven of these are also significant for patients who responded to neoadjuvant chemotherapy compared to those who did not, although some are probably significant only because in the category

Table VIII. *A total of 19 significant pathways, based on gender differences were identified.*

Rank	Name	pPERT FDR
1	Cytokine-cytokine receptor interaction	0.025
2	*Systemic lupus erythematosus	0.025
3	*Type I diabetes mellitus	0.025
4	**Rheumatoid arthritis	0.025
5	*Graft-*versus*-host disease	0.025
6	MAPK signaling pathway	0.025
7	NFκB signaling pathway	0.025
8	Neuroactive ligand-receptor interaction	0.025
9	**Autoimmune thyroid disease	0.025
10	Antigen processing and presentation	0.027
11	*Allograft rejection	0.027
12	**Cytokine signaling pathway	0.031
13	Pathways in cancer	0.034
14	Toll-like receptor signaling pathway	0.037
15	Proteoglycans in cancer	0.039
16	Small cells lung cancer	0.043
17	Viral myocarditis	0.043
18	Focal adhesion	0.047
19	Hepatitis B	0.047

pPERT FDR: False discovery rate of perturbation probability, the probability is calculated based on the amount of perturbation measured in each pathway. *Pathways that are also significant for responders/non-responders (Table VII). **Pathways that are also significant only for male responders/non-responders.

of patients who responded to neoadjuvant therapy there was only one female patient.

However, taking into account only male patients, there are three important pathways for chemotherapy response; these are also some of the most important pathways attributed to gender differences: rheumatoid arthritis, autoimmune thyroid disease and Chemokine signaling. It is interesting to visualize these three pathways, comparing male/female results to responder/non-responder results. Pathway Guide allows allows visualization of fold changes, total perturbation (14) and total perturbation accumulation (15) in KEGG (37) pathways. In cases of response to chemotherapy, no genes that differentially expressed. Thus, we will only consider results regarding the perturbation accumulation only (not taking into account the fold changes). Total perturbation accumulation summarizes the effect of the upstream genes, for each gene and deducts the fold change, while the total perturbation includes the fold change of each gene. The total perturbation accumulation in the Rheumatoid arthritis pathway is presented for male responders/non-responders (Figure 3A) and female/male patients (Figure 3B). The effect on the pathway is the same for both categories. Non-responders and females present down-regulation of the signal for T-cells (cytotoxic T-lymphocyte-associated protein 4, CTLA4 and cluster of differentiation 28, CD28) and up-regulation of angiogenesis (Fms-related tyrosine kinase, 1FLT1), when compared to responders and males, respectively. CD28 is important for cytokine production, as well as for the activation and survival of T-cells. FLT1 is involved in angiogenesis and macrophage function.

In Figure 4, total perturbation accumulation in the Autoimmune thyroid disease pathway is similarly presented. The effect on the pathway has both similarities and differences for male non-responders /responders (top) and female/male (bottom).

As for Rheumatoid arthritis pathway, non-responders and females downregulate signaling to T-cells from CD28. However, B-cell receptor signaling, as it is mediated by CD40, was down-regulated in non-responders compared to responders, but up-regulated in women versus men. CD40 plays part in immune and inflammatory responses. FAS surface receptor activity of cell death was down-regulated in non-responders compared to responders, but up-regulated in women compared to men.

This observation is in accordance with the fact that, for non- responders, there is not sufficient apoptosis for control of the disease, but women (who develop most autoimmune diseases) overexhibit apoptosis in the target tissues.

In Figure 5, total perturbation accumulation in the Chemokine signaling pathway for male non-responders/ responders and female/male patients is presented. Most effects on this pathway are opposite. (responders versus non-responders and male versus female patients): compared to responders, the non-responders have a general repression of the entire pathway and women have a general stimulation of the whole pathway when compared to men. Some inflammatory chemokines initiate immune response. Chemokine signaling repression in non-responders and its activation in women coroborates the hypothesis that in non-responders the immune response is under-active whereas in women it is over-active.

Discussion

After diagnosing lung cancer, the clinical and pathological parameters, such as the histological profile of the tumor, the stage and the localization of metastases determine the disease prognosis and the choice of treatment (16, 17). The current guidelines of clinical practice discriminate small cell lung cancer from NSCLC, as well as the most important sub-types of NSCLC: squamous and non-squamous cell carcinoma, resulted from the standard pathological analysis.

Microarray analysis of resected tumoral material allows the realization of a molecular profile. Therefore, based on transcription profiles, pulmonary adenocarcinomas were stratified into molecular subgroups that propose different cellular and prognostic characteristics (18-23). In relation to

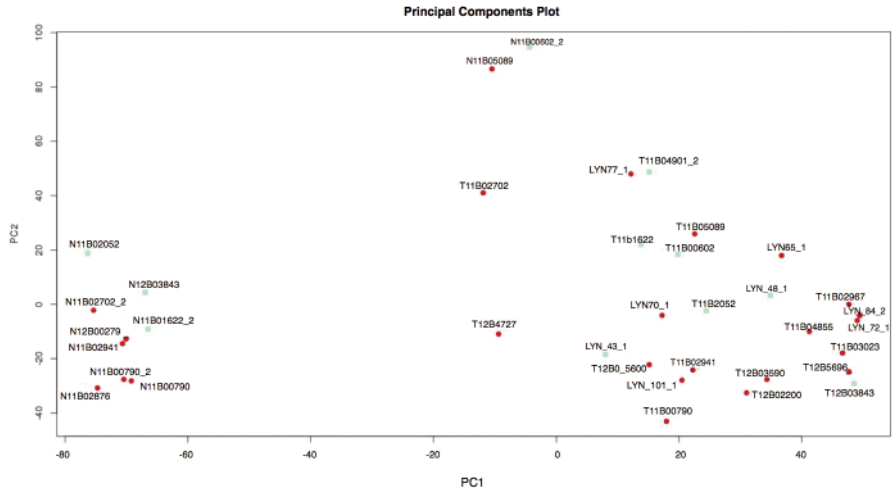

Figure 1. *Principal component analysis for quality control. Quality control was carried out using Agi4X44Preprocess R package. Samples Lyn_82 (group A), T11B04901 (group B) and Lyn_84 (group A) cluster with each other and not with tumor or normal, which are the two main groups in the figure, seen separated on the positive (tumor) and negative (normal) x-axis. These flagged for poor quality and removed from the analysis (not shown in figure). Blue spots: responders, red spots: non-responders.*

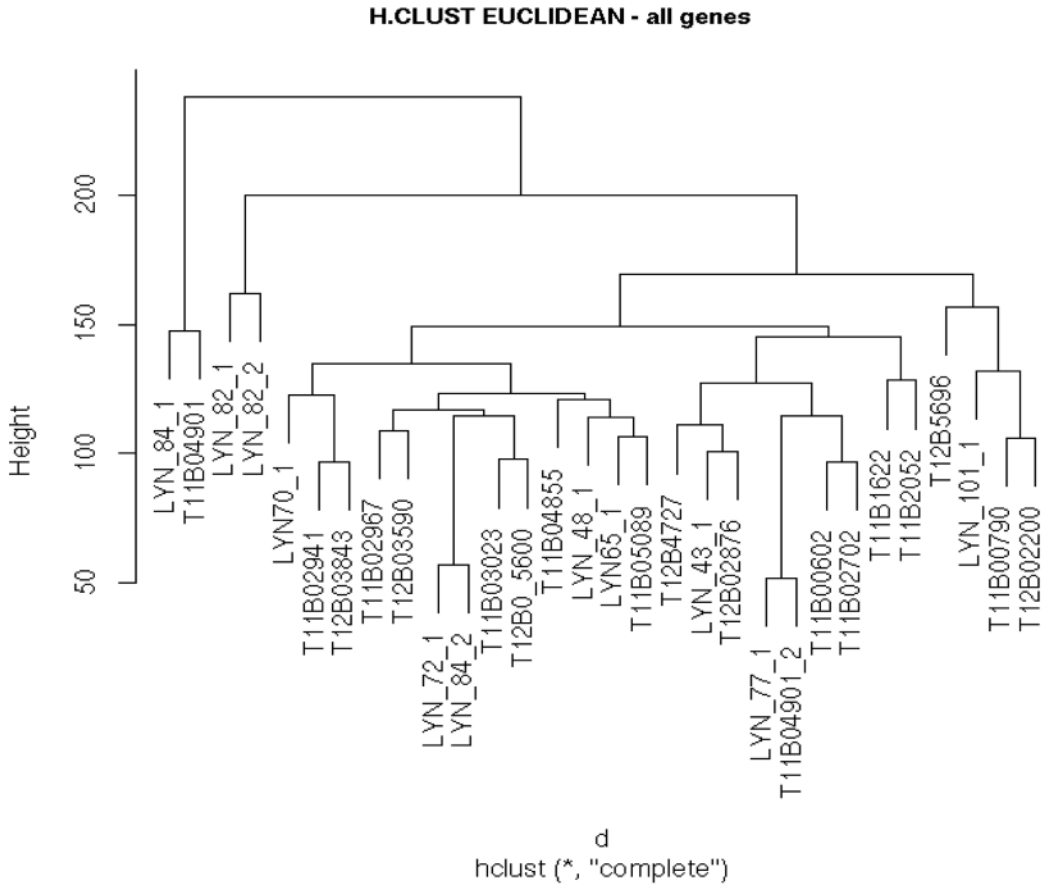

Figure 2. *Clustering dendogram (Euclidean distance) for quality control. Samples Lyn_82 (group A), T11B04901 (group B) and Lyn_84 (group A) (shown with red arrows) were of poor quality. Hybridization was repeated for all three samples (shown as Lyn_82_2, T11B04901_2 and Lyn_84_2) and quality control re-performed but the quality was still poor (also according to relative log expression for quality control-data not shown); samples Lyn_84 and T11B04901 clustered separately from the other samples.*

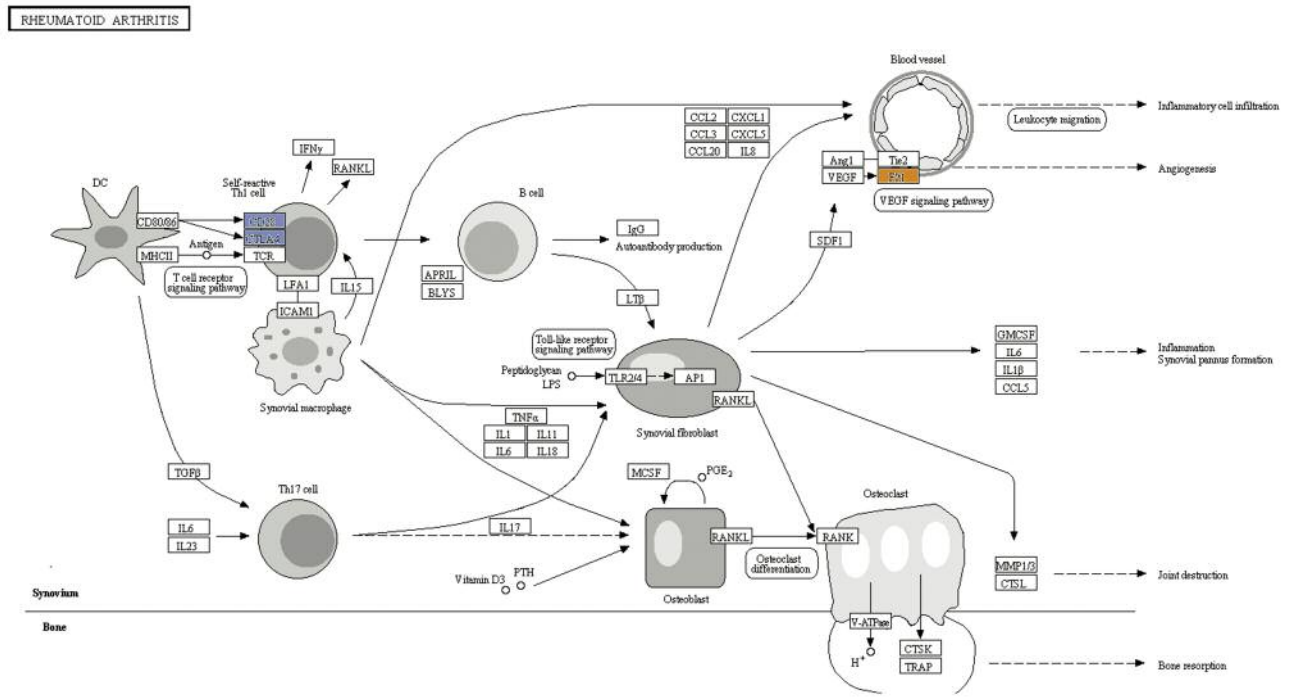

Figure 3. *Total perturbation accumulation in the Rheumatoid arthritis pathway for male responders/non-responders (A) and female/male patients (B). Cytotoxic T-lymphocyte-associated protein 4 (CTLA4) and cluster of differentiation 28 (CD28) are shown in blue, Fms-related tyrosine kinase (FLT1) in orange. This pathway can be analyzed in an interactive manner and in more detail at www.genome.jp/kegg-bin/show_pathway?hsa05323.*

some transcription similarities of pathologically defined carcinoma (broncho alveolar carcinoma, squamous cell carcinoma and large cell carcinoma), different patterns of gene expression have been associated with different adenocarcinoma subtypes (bronchioid, squamoid and magnoid, respectively) (21). Recently, an architectural classification of invasive pulmonary adenocarcinoma described five predominant patterns that were proven to have prognostic value for survival, regardless of disease stage (24, 25). The presence of different histological patterns in one tumor prevents evaluation of the predominant architecture. Moreover, different adenocarcinoma pathological patterns do not necessarily correspond to this molecular subtype.

Similar subtype analyses were performed for squamous cell pulmonary adenocarcinoma for prognosis estimation (26-28). Such adenocarcinomas were classified into four different gene-expression subtypes (primitive, classical, secretory and basal) generated by microarray analyses and were reproducible by independent microarray analyses and by RNA sequencing of approximately 600 samples (28, 29). The primitive subtype was associated with the poorest prognosis. Moreover, different patterns of gene expression of different molecular subtypes were correlated with the activation status of biological and cellular processes and oncogene pathways.

Further studies based on quantitative real time PCR and microarray have identified miRNA patterns specific to lung cancer subtypes. For instance, histological classification of adenocarcinoma and squamous cell pulmonary adenocarcinoma was realized with a 34-miRNA panel on paraffin-embedded samples for 205 male smokers (30).

In the present study, gene expression was analyzed, as well as signaling pathways of patients with NSCLC and invasion of mediastinal lymph nodes (stage IIIA, N2), depending on the histological response to neoadjuvant chemotherapy treatment. A statistically significant correlation between gene expression levels and treatment response was not established, whether using clustering, principal components or differential gene expression.

Therefore, we looked for genes that vary more in non-responders than in responders, and many genes where found that distinguish the histological status of mediastinal lymph nodes. In responders, the most variable genes concern keratinization; in non-responders, most relate to hormonal stimulation and peptidase activity. The gene-expression variance pattern of male patients can be mistaken for chemotherapy response in this study.

Molecular pathway analysis using Impact Analysis (14,15) and studying all gene alteration has led to some convincing results regarding the mechanisms of response to therapy. This

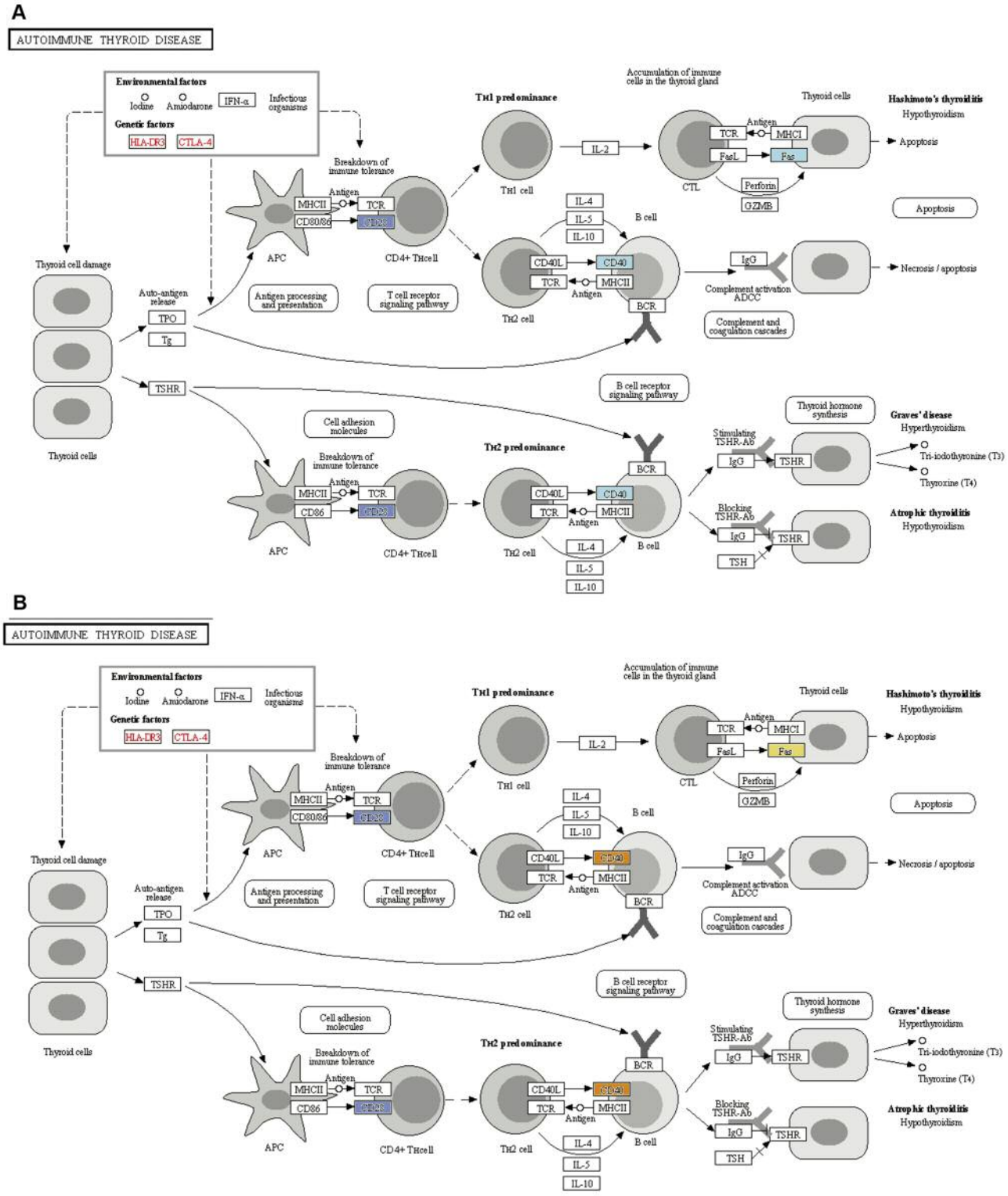

Figure 4. *Total perturbation accumulation in the Autoimmune thyroid disease pathway for male non-responders/responders (A) and female/male patients (B). Cluster of differentiation 28 (CD28) is shown in blue. CD40 and FAS are activated in women compared to men (orange), but down-regulated (blue) in non-responders compared to responders. This pathway can be analyzed in an interactive manner and in more detail at www.genome.jp/kegg-bin/show_pathway?hsa05320.*

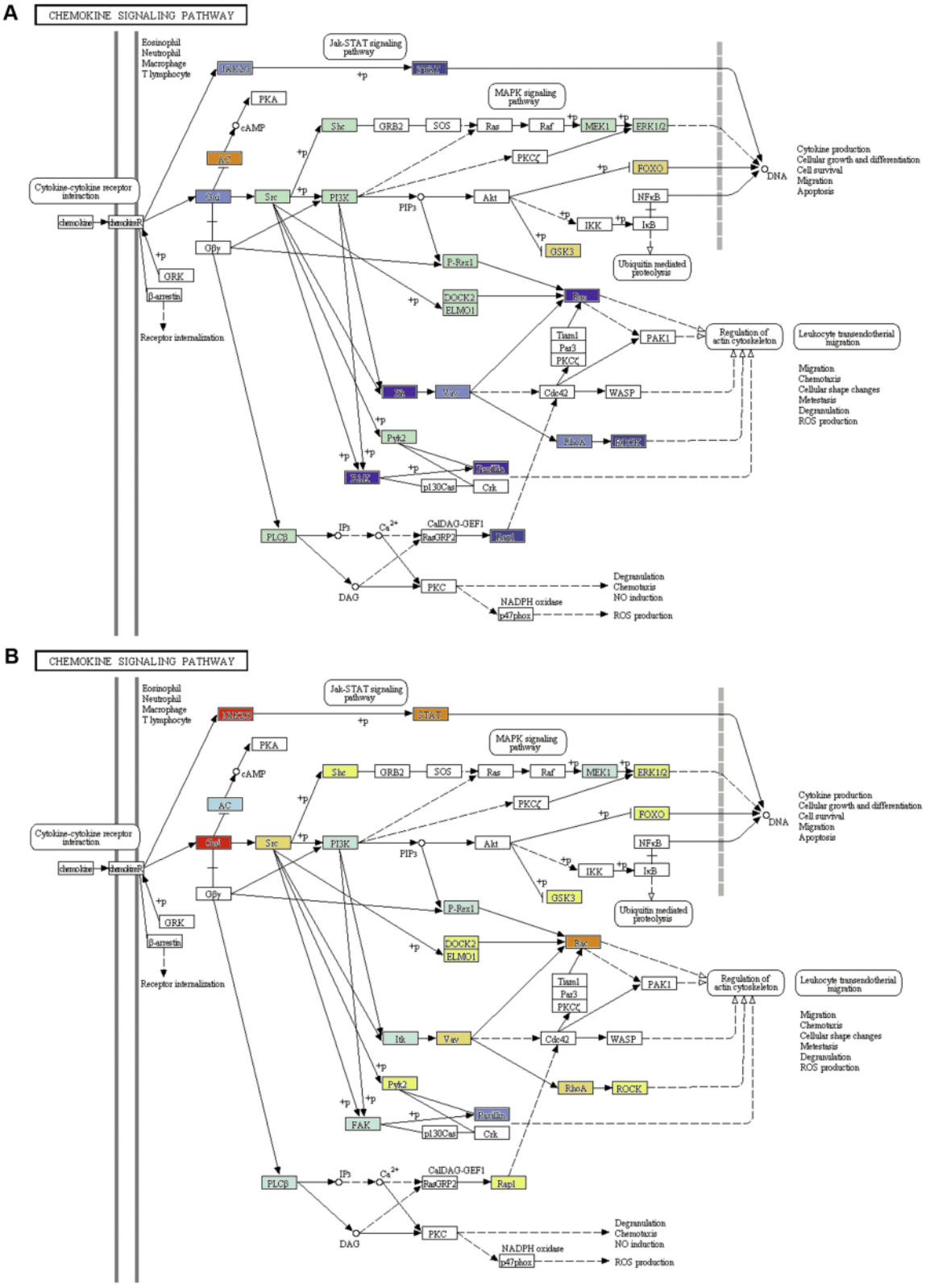

Figure 5. *Total perturbation accumulation in the Chemokine-signaling pathway for male non-responders/responders (A) and female/male patients (B). The red lines represent coherent cascades of perturbation. This pathway can be analyzed in an interactive manner and in more detail at www.genome.jp/kegg-bin/show_pathway?hsa04062.*

analysis suggests that the difference between responders and non-responders lies in the regulation of immune mechanisms. All pathways that are significant in chemotherapy response overlap with those that are also significant for gender difference. However, considering only male patients, three regulatory pathways of immune mechanisms are significant: Rheumatoid arthritis, Autoimmune thyroid disease and Chemokin signaling.

Total perturbation accumulation in the Rheumatoid arthritis pathway was basically the same for responders and non-responders, as well as for female compared to male patients. Total perturbation accumulation in the Autoimmune thyroid disease pathway was similar concerning T-cell activity, but opposite between responders and non-responders regarding the immune and inflammatory response and apoptosis. Total perturbation accumulation for Chemokine signaling pathway was completely inhibited in non-responders compared to responders, but it was completely activated in women compared to men.

The observations arising from the molecular pathways analysis suggest that sex hormones and chemotherapy response may be connected. The significance of the GO terms for hormonal stimulus among genes with high variance from non-responders supports this hypothesis. Considering that total perturbation acumulation in The Rheumatoid arthritis pathway is similar to chemotherapy response, it is interesting to notice that estrogen and the chemokine CCL13 (C-C motif ligand 13 cytokine) are involved in the severity of rheumatoid arthritis owed to apoptosis dysregulation (31). A review by Nilsson (32) discusses the importance of estrogens (both estrogen receptor α and β) in controlling the progress of inflammation, especially by the modification of leukocyte behavior and by downregulating chemokines and adhesion molecules. Hohla *et al*. found that female gender is associated with a favorable response to chemotherapy in pancreatic cancer (33). Rodriguez-Lara *et al*. have shown that pulmonary adenocarcinoma overexpresses estrogen receptor β, as well as chemokine (C-X-C motif) ligand 12 (CXCL12) and CXC chemokine ligand 4 (CXCL4) compared to normal lung tissue and that this overexpression is higher in pre-menopausal women, who present high risk, than in post-menopausal women, or in men (34). The study of Szymanowska-Narloch *et al*. showed the potential use of hormonal therapy in selected NSCLC patients (35).

The research for the determination of specific gene-expression signatures, genetic mutations and other genomic alterations is promising and will allow biomarker identification and targeted therapy for different types of tumoral architecture.

Conflicts of Interest

The Authors declare that there exist no conflicts of interest with regard to this study.

Acknowledgements

The research presented in this article was supported by Funding Contract no. 211 with OI – ANCS signed on 20.07.2010, within POS CE Operational Programme, Priority Axis 2 – CDI competitiveness – Major Field of Intervention D.2.1. "Research-development in partnerships between research-development universities/institutes and enterprises for result attainment applicable in economy", Operation 2.1.2.: "High level scientific CD projects with specialists from abroad", Project title Gene Profile Of Non-Small Cell Primitive Bronchopulmonary Cancers And Invasion Of Mediastinal Lymph Nodes, Project number/Code SMIS 692/12650, undertaken between 20.07.2010 - 30.06.2014.

References

1 Alberg AJ, Brock MV, Ford JG, Samet JM and Spivack SD: Epidemiology of lung cancer: Diagnosis and management of lung cancer. Third Edition: American College of Chest Physicians evidence-based clinical practice guidelines. Chest *143*: e1S, e29S, 2013.

2 Stefani A, Alifano M, Bobbio A, Grigoroiu M, Jouni R, Magdeleinat P and Regnard JF: Which patients should be operated on after induction chemotherapy for N2 non-small cell lung cancer? Analysis of a 7-year experience in 175 patients. J. Thorac Cardiovasc Surg *140(2)*: 356-363, 2010.

3 Miller AB, Hoogstraten B, Staquet M and Winkler A: Reporting results of cancer treatment. Cancer *47*: 207-214, 1981.

4 Voichita C, Donato M and Draghici S: Incorporating Gene Significance in the Impact Analysis of Signalling Pathways. In Machine Learning and Applications (ICMLA), 11th International Conference *1*: 126-131, 2012.

5 Voichita C and Draghici S: ROntoTools: R Onto-Tools suite. R package version 1.1.2.

6 Johan L, Jayoun K, Sangcheol K, Donghyeon Y, Kyunga K and ByungSoo K: Detection of differentially expressed gene sets in a partially paired microarray data set. Stat Appl Genet Molec Biol 11 (3), 2012.

7 Yi S, Park T: Integrated analysis of the heterogeneous microarray data. BMC Bioinformatics *12(5)*: S3, 2011.

8 Jing W, Dexter D, Zhiao S and Bing Z: Web-based gene set analysis toolkit (webgestalt): update 2013. Nucleic Acids Res *41(W1)*: W77-W83, 2013.

9 Zhang B, Kirov S and Snoddy J: Webgestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res *33(2)*: W741-W748, 2005.

10 Candi E, Schmidt R and Melino G: The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol *6(4)*: 328-340, 2005.

11 Godessart N and Kunkel SL: Chemokines in autoimmune disease. Curr Opin Immunol *13(6)*: 670-675, 2001.

12 Oertelt-Prigione S: The influence of sex and gender on the immune response. Autoimmun Rev *11(6)*: A479-A485, 2012.

13 Jansen R, Batista S and Brooks AI: Sex differences in the human peripheral blood transcriptome. BMC Genomics *15(1)*: 33, 2014.

14 Draghici S, Khatri P and Tarca AL: A systems biology approach for pathway level analysis. Genome Res *17(10)*: 1537-1545, 2007.

15 Tarca AL, Draghici S and Khatri P: A novel signaling pathway impact analysis. Bioinformatics *25(1)*: 75-82, 2009.

16 Ettinger DS, Akerley W and Borghaei H: Non-small cell lung cancer. J Natl Compr Cancer Netw *10*: 1236-1271, 2012.
17 Kalemkerian GP, Akerley W and Bogner P: Small cell lung cancer. J Natl Compr Cancer Netw *11*: 78-98, 2013.
18 Bhattacharjee A, Richards WG and Staunton J: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci *98*: 13790-13795, 2001.
19 Beer DG, Kardia SL and Huang CC: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med *8*: 816-824, 2002.
20 Garber ME, Troyanskaya OG and Schluens K: Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci *98*: 13784-13789, 2001.
21 Hayes DN, Monti S and Parmigiani G: Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. J Clin Oncol *24*: 5079-5090, 2006.
22 Park YY, Park ES and Kim SB: Development and validation of a prognostic gene-expression signature for lung adenocarcinoma. PLoS One 7, 2012.
23 Takeuchi T, Tomida S and Yatabe Y: Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologicbehaviors. J Clin Oncol *24*: 1679-1688, 2006.
24 Warth A, Muley T and Meister M: The novel histologic International Association for the Study of Lung Cancer/ American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. J Clin Oncol *30*: 1438-1446, 2012.
25 Travis WD, Brambilla E and Riely GJ: New pathologic classification of lung cancer: Relevance for clinical practice and clinical trials. J Clin Oncol *31*: 992-1001, 2013.
26 Inamura K, Fujiwara T and Hoshida Y: Two subclasses of lung squamous cell carcinoma with different gene expression profiles and prognosis identified by hierarchical clustering and non-negative matrix factorization. Oncogene *24*: 7105-7113, 2005.
27 Raponi M, Zhang Y and Yu J: Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. Cancer Res *66*: 7466-7472, 2006.
28 Wilkerson MD, Yin X and Hoadley KA: Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. Clin Cancer Res *16*: 4864-4875, 2010.
29 The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*: 519-525, 2012.
30 Landi MT, Zhao Y and Rotunno M: MicroRNA expression differentiates histology and predicts survival of lung cancer. Clin Cancer Res *16*: 430-441, 2010.
31 Yamaguchi A, Nozawa K and Fujishiro M: Estrogen inhibits apoptosis and promotes cc motif chemokine ligand 13 expression on synovial fibroblasts in rheumatoid arthritis. Immunopharmacol Immunotoxicol *34(5)*: 852-857, 2012.
32 Nilsson BO: Modulation of the inflammatory response by estrogens with focus on the endothelium and its interactions with leukocytes. Inflammation Research *56(7)*: 269-273, 2007.
33 Hohla F, Hopfinger G and Romeder F: Female gender may predict response to folfirinox in patients with unresectable pancreatic cancer: A single institution retrospective review. International journal of oncology *44(1)*: 319-326, 2014.
34 Rodriguez-Lara V, Peña-Mirabal E and Baez-Saldana R: Estrogen receptor beta and cxcr4/cxcl12 expression: Differences by sex and hormonal status in lung adenocarcinoma. Arch Med Res, 2014.
35 Szymanowska-Narloch A, Jassem E and Skrzypski M: Molecular profiles of non-small cell lung cancers in cigarette smoking and never-smoking patients. Adv Med Sci *1*: 11, 2014.
36 Bravo HC, Pihur V, McCall M, Irizarry R and Leek J: Gene expression anti-profiles as a basis for accurate universal cancer signatures. BMC Bioinformatics, 13(1):272, 2012.
37 Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S and Tokimatsu T : KEGG for linking genomes to life and the environment. Nucleic Acids Res, 36 (suppl 1):D480-D484, 2008.

Received June 29, 2015
Revised August 10, 2015
Accepted August 18, 2015