

Repeated observation of immune gene sets enrichment in women with non-small cell lung cancer

Jhajaira M. Araujo¹, Alexandra Prado¹, Nadezhda K. Cardenas², Mayer Zaharia¹, Richard Dyer³, Franco Doimi³, Leny Bravo², Luis Pinillos¹, Zaida Morante⁴, Alfredo Aguilar¹, Luis A. Mas¹, Henry L. Gomez¹, Carlos S. Vallejos¹, Christian Rolfo⁵, Joseph A. Pinto¹

¹Unidad de Investigación Básica y Traslacional, Oncosalud-AUNA, San Borja, Lima 41, Peru

²Escuela de Medicina Humana, Universidad Privada San Juan Bautista, Chorrillos, Lima 09, Peru

³Departamento de Patología, Oncosalud-AUNA, San Borja, Lima 41, Peru

⁴Departamento de Medicina Oncológica, Instituto Nacional de Enfermedades Neoplásicas, Surquillo, Lima 41, Peru

⁵Phase I – Early Clinical Trials Unit, Antwerp University Hospital, Antwerp, Edegem 2650, Belgium

Correspondence to: Christian Rolfo, **e-mail:** christian.rolfo@uza.be

Keywords: non-small cell lung cancer, gender, GSEA, CIBERSORT, immune gene sets

Received: November 19, 2015

Accepted: February 11, 2016

Published: March 06, 2016

ABSTRACT

There are different biological and clinical patterns of lung cancer between genders indicating intrinsic differences leading to increased sensitivity to cigarette smoke-induced DNA damage, mutational patterns of KRAS and better clinical outcomes in women while differences between genders at gene-expression levels was not previously reported. Here we show an enrichment of immune genes in NSCLC in women compared to men. We found in a GSEA analysis (by biological processes annotated from Gene Ontology) of six public datasets a repeated observation of immune gene sets enrichment in women. "Immune system process", "immune response", "defense response", "cellular defense response" and "regulation of immune system process" were the gene sets most over-represented while *APOBEC3G*, *APOBEC3F*, *LAT*, *CD1D* and *CCL5* represented the top-five core genes. Characterization of immune cell composition with the platform CIBERSORT showed no differences between genders; however, there were differences when tumor tissues were compared to normal tissues. Our results suggest different immune responses in NSCLC between genders that could be related with the different clinical outcome.

INTRODUCTION

Lung cancer is the most common malignancy and the leading cause of death worldwide where tobacco smoking is the main risk factor for the development of this cancer [1, 2].

It is well known that lung cancer shows different patterns of clinical characteristics and outcomes according to sex. Women have higher susceptibility to cigarette smoke-induced DNA damage and an increased risk for lung cancer, including higher levels of DNA adducts and higher frequency of the KRAS G12C mutation than men [3, 4]. In addition, female patients tend to present lung cancer at a younger age and in more advanced stages; however, women have better prognostic than men [5, 6].

Due to different susceptibility to DNA damage, according to sex, our research group previously evaluated differences between genders in gene expression and mutational status in genes involved in DNA repair without observing differences. In contrast, in a global evaluation, some gene sets were differentially enriched in women, including immune gene sets [7]. Despite the great advances in the knowledge of the genomic landscape of lung cancer, transcriptional differences between genders have not been previously explored.

With the aim of exploring transcriptional differences of Non-small cell lung cancer (NSCLC) between the genders, we performed a comprehensive analysis of differentially enriched immune gene sets annotated by Gene Ontology Biological Process (C5 BP).

RESULTS

Gene set enrichment analysis (GSEA) reveals enrichment of immune gene sets in NSCLC in women

The structure of the data for GSEA analysis is shown in Table 1. Immune gene sets enrichment was not observed in men. In contrast, we found 40 immune gene sets enriched in women where the most over-represented where “Immune Response” ([GO:0006955]; p -values < 0.00000001 ; FDR’s 0.4%–24.2%) and “Immune System Process” ([GO:0002376]; p -values < 0.00000001 ; FDR’s: 0.4%–24.3%) in 06 subsets; “Defense Response” ([GO:0006952; p -values < 0.00000001 ; FDR’s: 0.3%–13.6%), “Cellular Defense Response” ([GO:0006968]; p -values between < 0.00000001 to 0.00784929) and “Regulation of Immune System Process” ([GO:0002682]; p -values between 0.00176991 to 0.0060423) in 05 subsets (Figure 1 and Figure 2). List of all datasets with immune gene sets enrichment in ≥ 2 subsets are shown in Tables S1 and S2. We were not able to find differences in the dataset GSE7670 (without information about smoking status). It could be due to smoking is a potential confounding factor hiding differences.

Immune-related genes over-represented in several datasets

Due to the gene redundancy across immune gene sets, we identified “core genes” (genes associated with the enrichment signal) over-represented in different subsets for immune biological processes. The top-ten rank core genes included *APOBEC3G* (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G) that exerts innate immune activity against retroviruses and has shown tumor suppressive effects in human hepatocellular carcinoma and enhance cell radio resistance in lymphomas; [8, 9] *APOBEC3F* (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3F) that showed to inhibit HIV-1 DNA Integration; [10] *CCL5* (Chemokine (C-C motif) ligand 5, a well characterized chemotactic chemokine; *CD1D* (CD1d molecule) mediate the presentation of self or microbial antigens to T cells; *LAT* (Linker for activation of T cells), a major transporter for essential amino acids into activated human T cells; [11] *TRAF1* (T cell receptor associated transmembrane adaptor 1), important to transport of CTLA-4 to the cell surface; [12] *IL32* (Interleukin 32), whose expression is increased after the activation of T-cells by mitogens or the activation of NK cells by IL-2; *CRTAM* (cytotoxic and regulatory T cell molecule), upregulated in CD4 and CD8 T cells; *CFHR1* (complement factor H-related 1), whose protein product binds to *Pseudomonas aeruginosa* elongation factor Tuf together with plasminogen and *CCR2* (chemokine

[C-C motif] receptor 2) that mediates monocyte infiltration. The complete list of core genes involved in immune gene sets is shown in Table S3.

Immune cell composition inferred from the transcriptional background

There were no differences in relative frequencies of immune cell composition in both genders according to the LM22 signature (Figure 3). An important difference was seen between healthy lung tissues vs NSCLC, regardless the smoking status or gender. The main component in tumors was plasma B cells and macrophages. In contrast, healthy lung tissues have a higher proportion of T cells CD8, mast cells and a lower proportion of plasma cells, compared to lung tumors. Detailed values of immune cell composition are described in the Table S3.

DISCUSSION

Although gender is a prognostic factor in some malignancies, there is a lack of biological basis explaining this phenomenon [5, 13, 14]. On the other hand, the immune system plays an important role in the efficacy of the therapy, for example, tumor infiltration by lymphocytes is associated with better response to chemotherapy and trastuzumab, and also is associated with a better prognosis in breast cancer and other cancers [15, 16, 17].

There are few reports describing the value of tumor-infiltrating lymphocytes in NSCLC where a more significant prognostic factor is the presence of tumor-induced bronchus-associated lymphoid tissue (Ti-BALT) composed by mature dendritic cell (DC)/T-cell clusters adjacent to B-cell follicles [18]. Despite of the controversial value of tumor-infiltrating lymphocytes in NSCLC, peripheral leukocytes have more prognostic relevance. A recent meta-analysis study described that a high neutrophils/lymphocyte ratio in peripheral blood is related with a poor prognosis [19]. Also, gene expression patterns of peripheral blood mononuclear cells are highly correlated with the tumor burden in NSCLC and this biological signal could disappear after tumor resection [20, 21].

A different composition of tumor infiltrating immune cells between genders might be a good explanation for our findings in the GSEA analyses and could be more easily linked to the outcomes in women patients with NSCLC; however, the similar composition of immune cells suggests differences in activation of immune pathways and interaction between immune and tumoral cells.

There are several methodologies for the analysis of various datasets to minimize the batch effect [22]. In this work we preferred to construct subsets and perform analyses in each subset individually instead of pooling data. An overall comparison in each dataset could lead to biased results due to smoking status is the main cofounder

Table 1: Characteristics and composition of subsets included in the GSEA analysis

GEO Accesion		GSE10072		GSE32863		TCGA	GSE50081	GSE47115	GSE7670
Array		Affymetrix Human Genome U133A		Illumina HumanWG-6 v3.0		IlluminaHiSeq_RNASeqV2	Affymetrix Human Genome U133 Plus 2.0	Illumina HumanHT-12 WG-DASL V4.0 R2	Affymetrix Human Genome U133A
Histology		Adenocarcinoma		Adenocarcinoma		Adenocarcinoma	NSCLC	NSCLC	Adenocarcinoma
Stage		IA-IV		IA-IV			IA-IIIB	IB-IIIB	
Samples evaluated		n = 71		116		447	81	16	52
Subsets in dataset	-	4		4		2	2	1	2
Never Smokers	Normal Tissue	31 (43.7%)	15 (48.4%)	59 (50.1%)	30 (50.8%)				
	Male		4 (26.7%)		7 (23.3%)				
	Female		11 (73.3%)		23 (76.7%)				
	Tumoral Tissue		16 (51.6%)		29 (49.2%)	69 (15.4%)	24 (29.6%)		
Current Smokers	Normal Tissue	40 (56.3%)	16 (40%)	57 (49.9%)	28 (49.1%)				
	Male		4 (25%)		6 (21.4%)				
	Female		12 (75%)		22 (78.6%)				
	Tumoral Tissue		24 (60%)		29 (50.1%)	378 (84.6%)	57 (70.4%)	16 (100%)	
Unknown	Normal Tissue								26 (50%)
	Male								5 (19.2%)
	Female								21 (80.8%)
	Tumoral Tissue								26 (50%)
Unknown	Male								5 (19.2%)
	Female								21 (80.8%)

and difficult to control. Although we decide not to use a more stringent statistics (FDR < 5%), the strength of our analysis strategy is based on the multiple validation that lead to be able to identify subtle differences in biological signals that could be masked with the batch of samples or more stringent statistics.

Our results are supported by a recent report by Qu et al. (2015), providing the first evidence of gender differences of immune regulation by elements that escape X chromosome inactivation and trigger regulatory networks and activation of genes with immune function in other autosomes in a study done in primary humans T cells [23]. Our results suggest different activities of immune gene sets regardless the immune cell composition in the tumor. Genes such as *APOBEC3G* and *APOBEC3F* seem to play an important role in the immune response in NSCLC, while over-representation of *CCL5* (a gene widely studied in other malignancies), *CD1D*, *LAT*, *TRAT1*, *IL32* and others, suggest different regulatory activities of T lymphocytes in NSCLC in women compared to men. Although “Immune Response” and “Immune System Process” where the gene sets more over-represented, which is logical because include more genes that produce redundancy across gene sets, we found differences in more specific subsets such as “Regulation of Immune System

Process” “T-Cell Activation”, “Regulation of Cytokine Biosynthetic Process” (Figure 2).

Recently, immune checkpoints inhibitors such as nivolumab and pembrolizumab are included in the National Comprehensive Cancer Network (NCCN) guidelines for the treatment of metastatic NSCLC subsequent to first line of chemotherapy [24]. A phase III study comparing nivolumab with docetaxel in patients with advanced non-squamous NSCLC showed that nivolumab not improved the overall survival in the subgroup of women patients (HR:0.78, CI: 0.58–1.04) [25]. Likewise, the results of KEYNOTE-010 trial comparing pembrolizumab versus docetaxel in patients previously treated with platinum-doublet chemotherapy showed that pembrolizumab does not increase the progression free survival in women (HR: 1.02, CI: 0.78–1.32). Conversely, both of these immune checkpoint inhibitors improved the outcome of the subgroup of men patients [26]. This contrasting results might be explained by differences in expression of immune genes.

Biological differences of lung tumors among genders should be deeply explored in order to improve the immunotherapeutic approaches. Our study provides evidence of biological differences of NSCLC between genders and the basis for the distinct clinical outcome.

MATERIALS AND METHODS

NSCLC datasets

With the aim of explore differences in immune gene sets in NSCLC between genders in a functional approach, we retrieved normalized gene expression data of five NSCLC datasets (GSE10072, GSE32863, GSE50081, GSE47115, and GSE7670) from the NCBI GEO website (<http://www.ncbi.nlm.nih.gov/geo/>). One RNASeq Level 3 dataset of lung adenocarcinoma was downloaded from the TCGA website (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/).

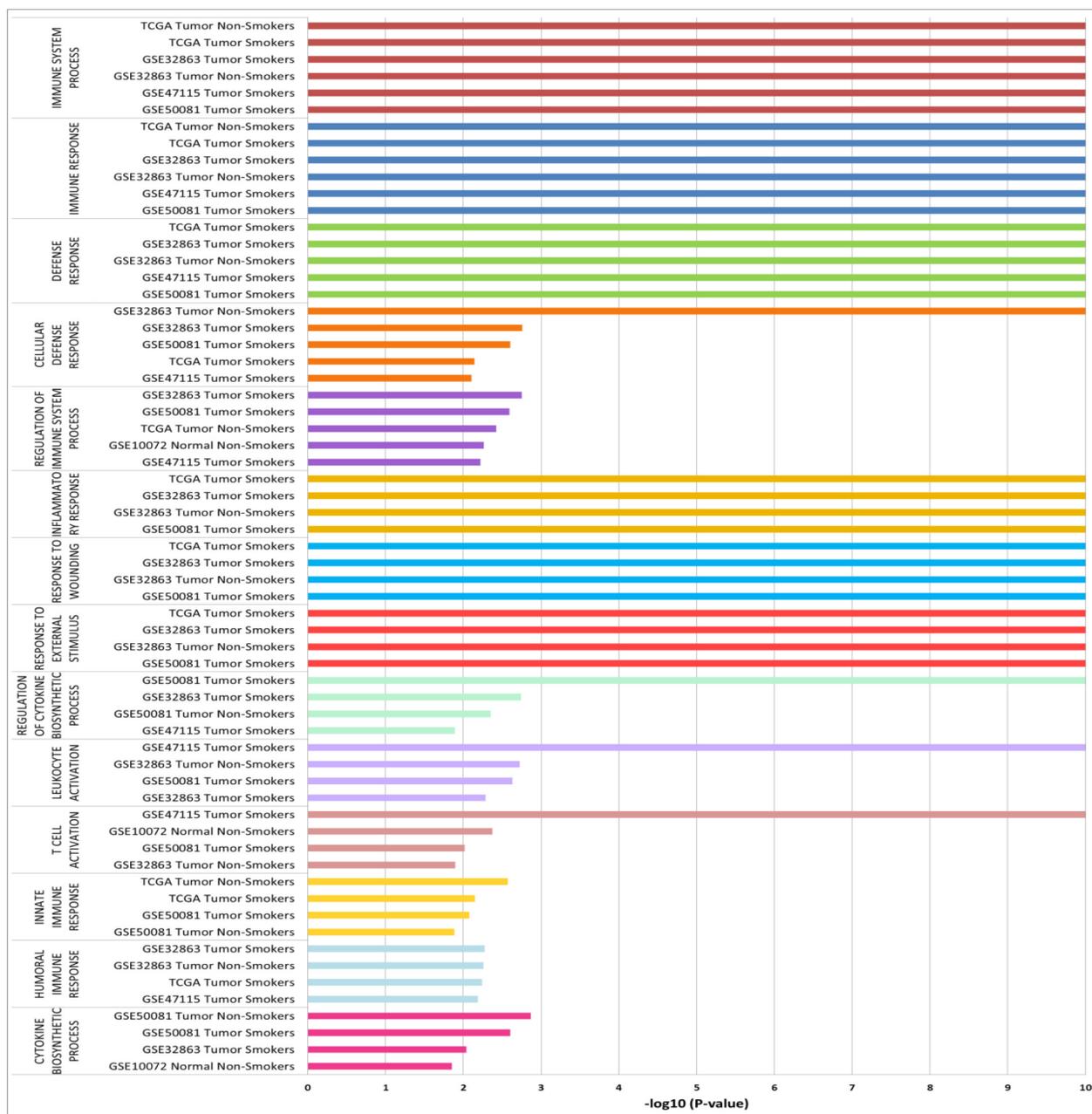
Data preprocessing

Values from datasets downloaded from the NCBI GEO were \log_2 transformed and median centered. In the TCGA dataset, expression values of “zero” were set to the overall minimum value and all data were \log_2 transformed and median centered.

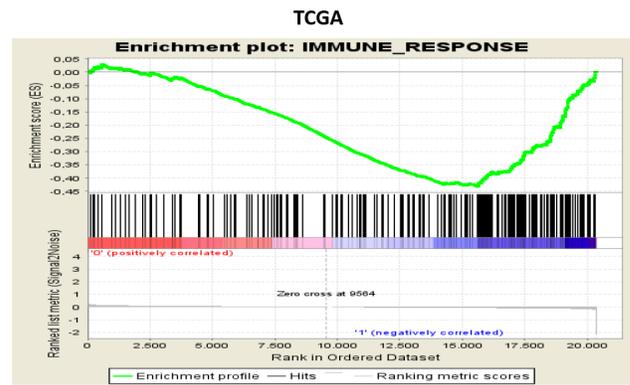
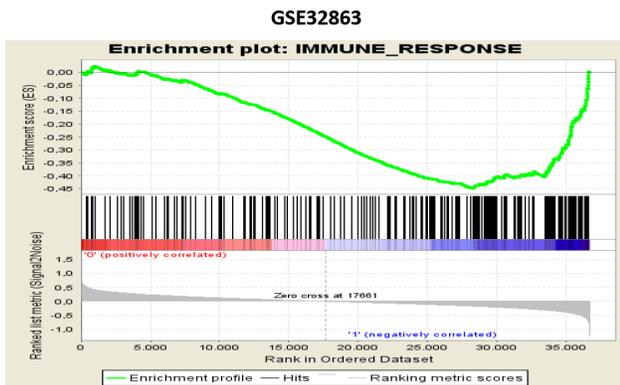
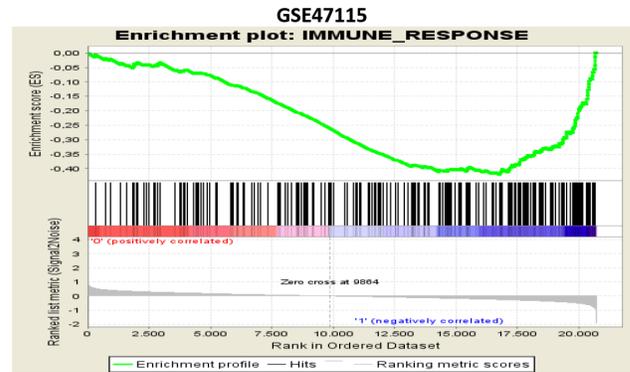
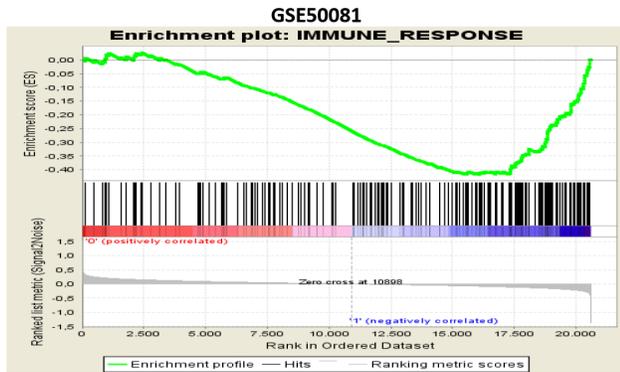
GSEA analysis

We evaluated 825 gene sets in biological processes annotated from Gene Ontology (C5 BP) in the Molecular Signature Database (MSigDB; <http://www.broad>).

A



B Tumor-Smokers



C Tumor-Non smokers

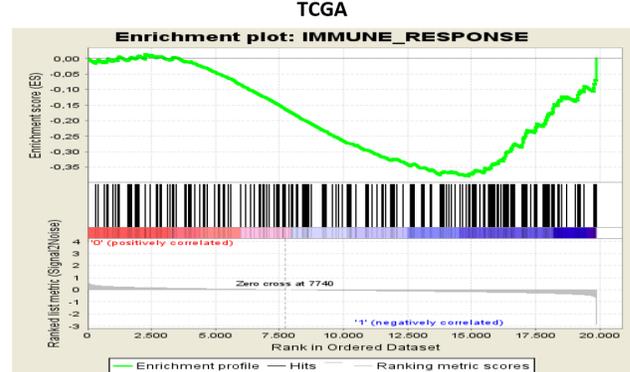
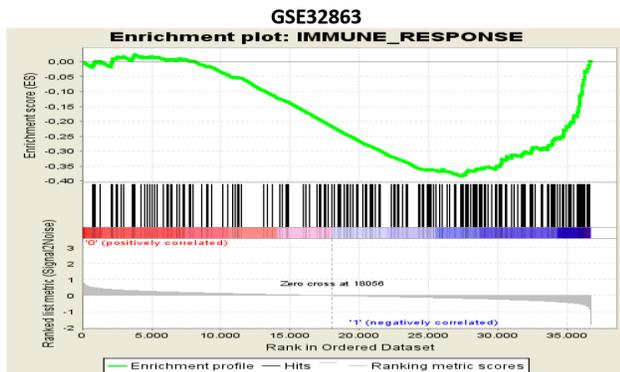
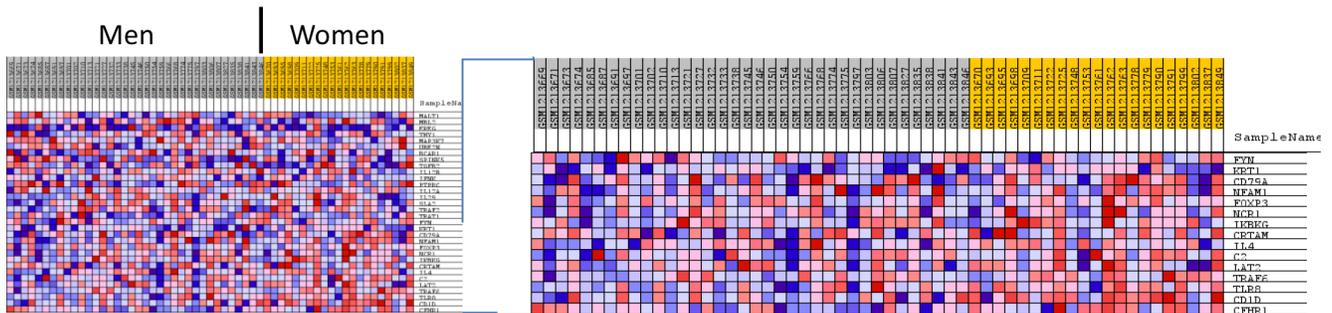
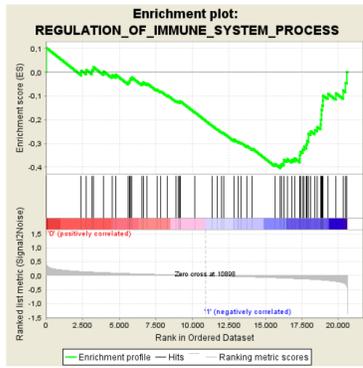
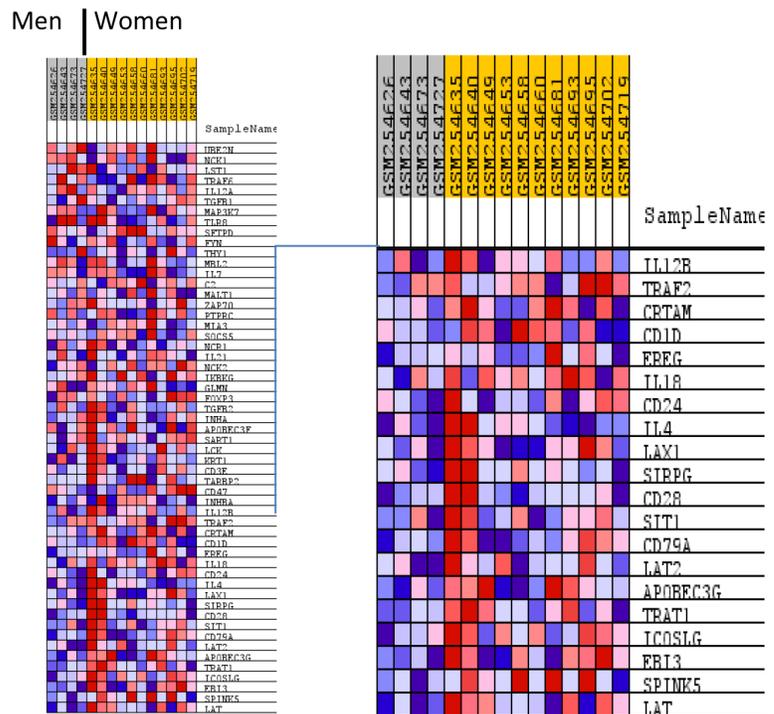
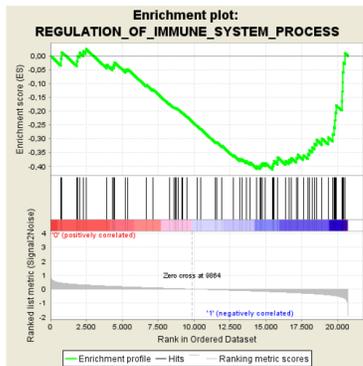


Figure 1: Immune gene sets enrichment more overrepresented in women. $-\log_{10}$ of P -values are shown in different subsets (A). Enrichment profile generated with GSEA in the gene set “immune process” comparing NSCLC in men vs NSCLC in women shown enrichment in female in smokers (B) and in non-smokers (C).

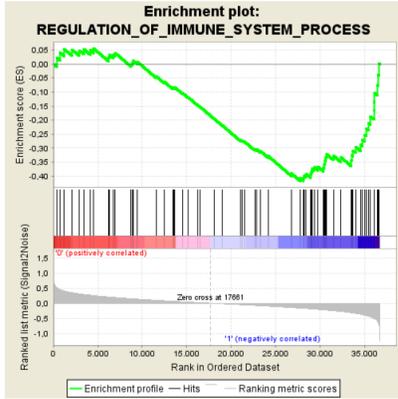
A GSE50081 (Tumor/Smokers)



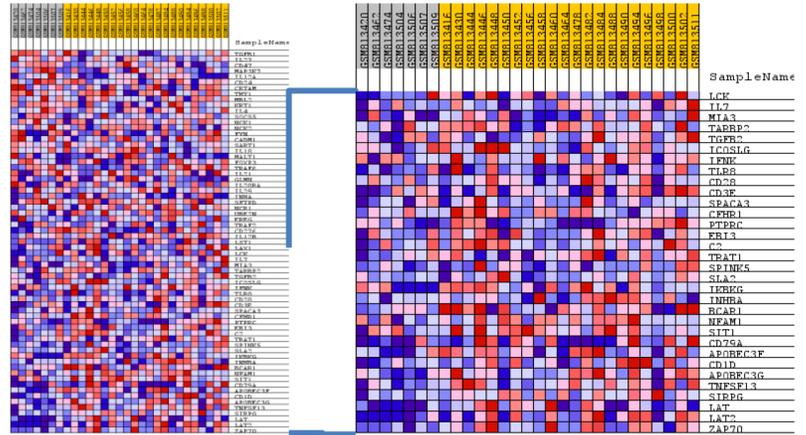
B GSE47115 (Tumor/Smokers)



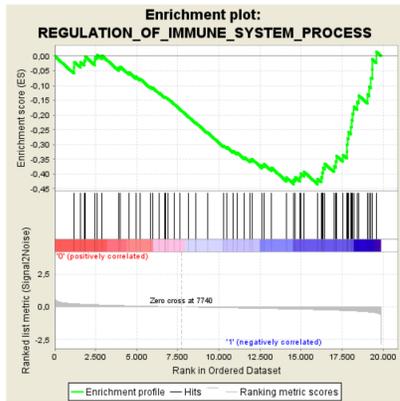
C GSE32863 (Tumor/Smokers)



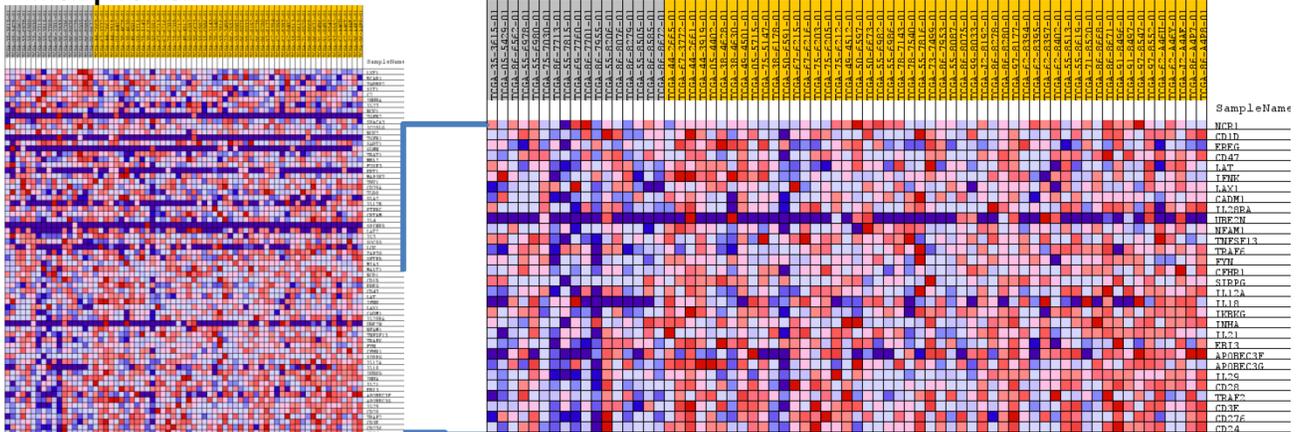
Men | Women



D TCGA (Tumor/Non-Smokers)



Men | Women



E GSE10072 (Normal/Non-Smokers)

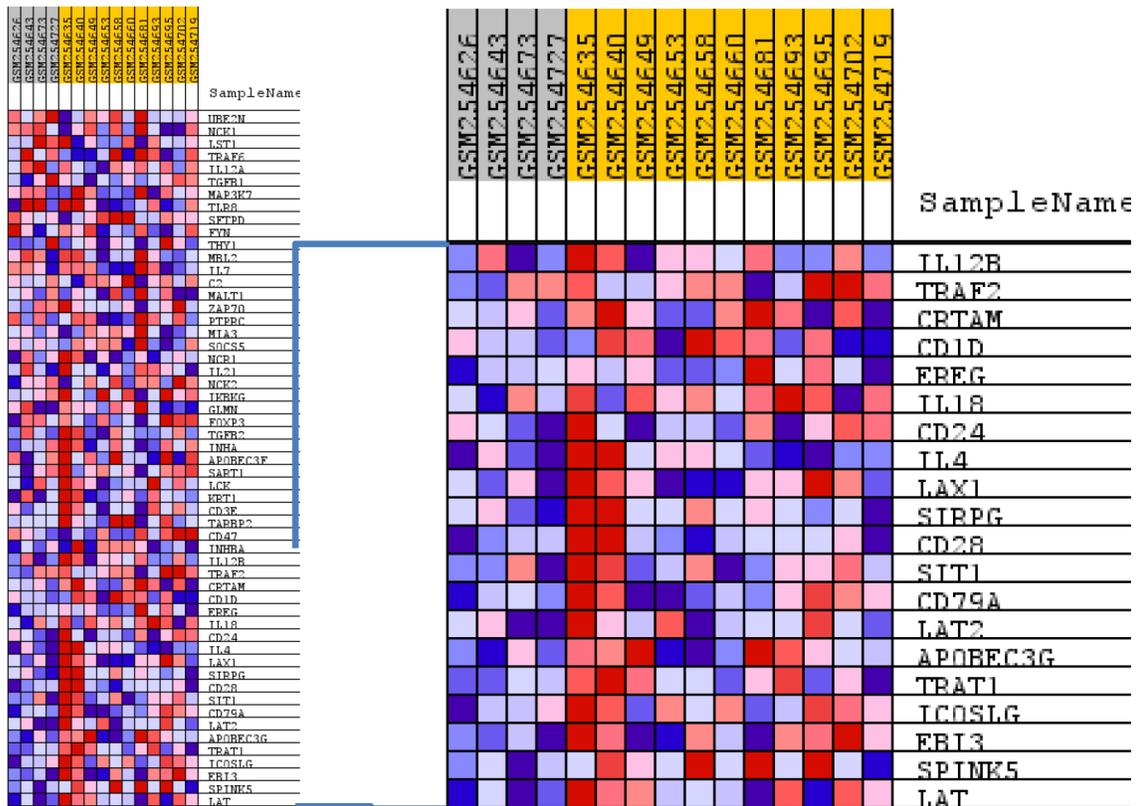
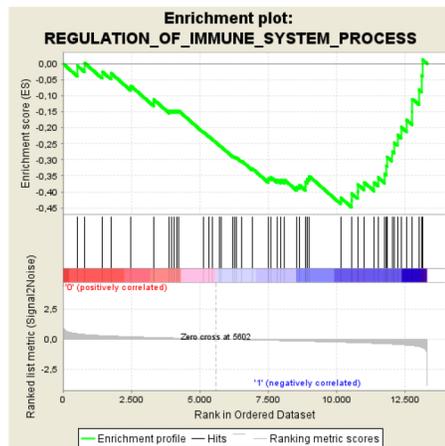


Figure 2: Enrichment profile and heatmaps generated with comparing the biological process: “REGULATION_OF IMMUNE_SYSTEM_PROCESS” between NSCLC in men vs NSCLC in women. This enrichment was observed in subsets of smokers (Figures A, B and C), non-smokers (D) and healthy tissue from non-smokers (E).

mit.edu/gsea/msigdb/msigdb_index.html). Because of the small sample size of the subsets, the GSEA was conducted with 1000 gene set permutations. The GSEA analyses were performed using the Java GSEA implementation downloaded from www.broad.mit.edu/gsea/msigdb/. Gene sets from samples of men vs women were compared. A gene set was considered enriched when it was included in the top 50 rank in at least two subsets with a p -value < 0.05 and a False Discovery Rate (FDR) $< 25\%$ and when was represented in only 1 gender.

To avoid the confounding effect of smoking and possible batch effects with the pooling of samples, we decided to divide each datasets in subsets according to smoking status (smokers vs no smokers) and type of tissue (normal vs tumoral) when these samples were available. Patients with unknown smoking status or former smokers were excluded from this analysis, except for patients from the dataset GSE7670 whose samples lack this information (Table 1).

Analysis of immune cells composition from gene expression data

To investigate if signal of enrichment of immune gene sets in women is related to a different background of immune cells in the samples, gene expression analysis with the online analytical platform CIBERSORT (<https://cibersort.stanford.edu/>) was done. CIBERSORT quantify relative levels of the abundances of distinct cell types in a mixed cell population [27].

We evaluated our subsets profiled with Affymetrix platforms (GSE10072 and GSE50081) with the LM22 gene signature that is able to identify 22 immune cell types (LM22 signature is only validated for Affymetrix microarrays data). For this analysis, labels of affymetrix probes were replaced with gene names. Due to more than one affymetrix probe could represent a gene, genes were collapsed to the highest value. After this step, the data were quintile normalized. Analyses were done with 100 permutations with default statistical parameters. The results were filtered by a maximum p -value of 0.05.

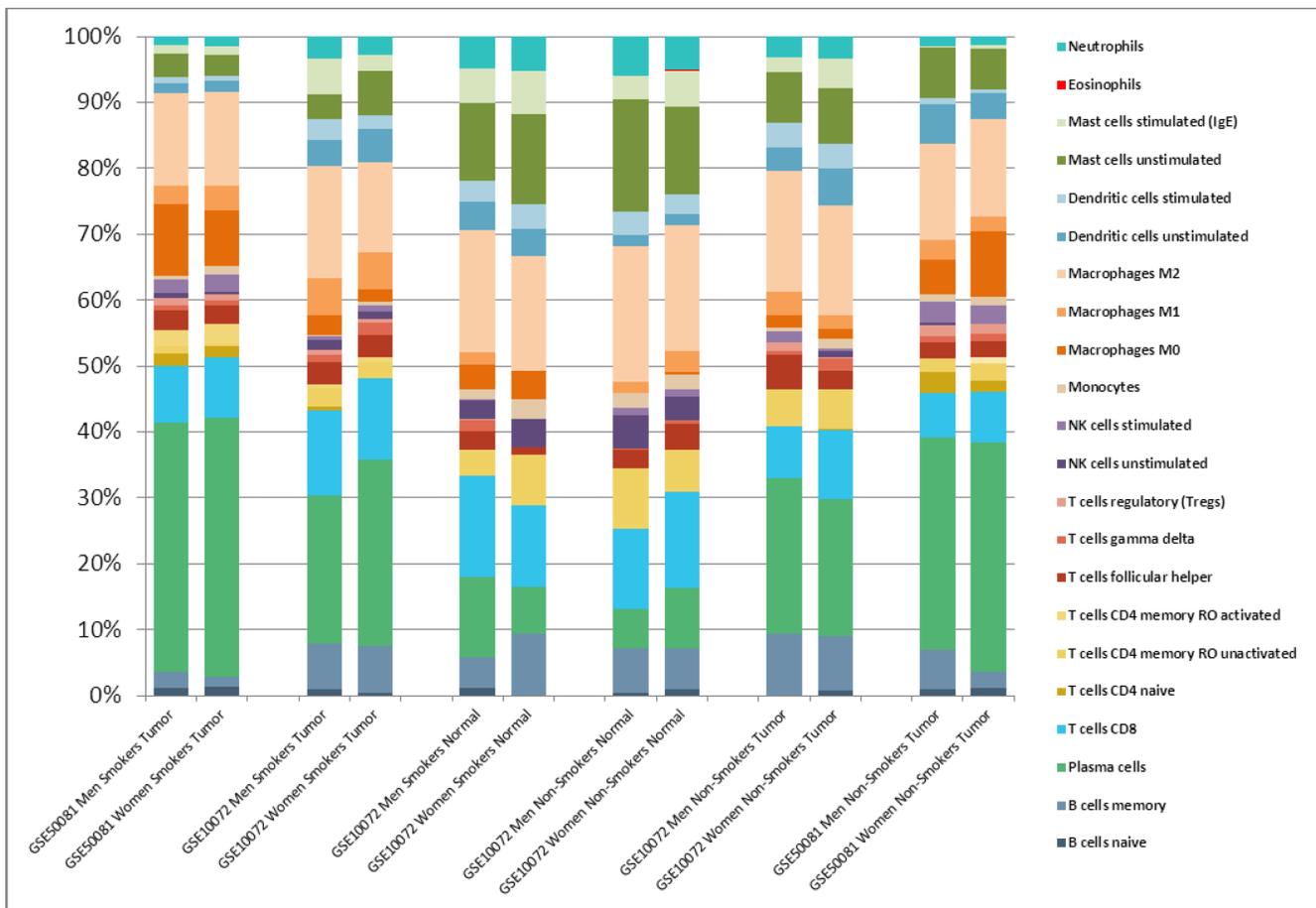


Figure 3: Relative leukocyte fractions evaluated by CIBERSORT in Affymetrix datasets to infer relative RNA fractions from 22 leukocyte subsets (LM22 signature) in each sample. Shown are the average fractions in each dataset.

Authors' contributions

J.A.P., J.M.A., H.L.G. and C.V. were responsible for the design, execution of the study, data interpretation and final manuscript writing; J.M.A., A.P., N.K.C., M.Z., R.D., F.D. and L.B. assisted with the collation of study materials, data analysis and data interpretation; J.M.A., A.P., N.K.C. and L.B., assembly and verification of the data; L.P., M.Z., Z.M., A.A., H.L.G. AND C.R., contributed to the interpretation of the results and participated to the manuscript writing. J.A.P., C.V. and C.R., supervised the study. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS AND FUNDING

This work was funded by a Research Grant of AUNA.

CONFLICTS OF INTEREST

The authors declare they have not potential conflicts of interests with this research.

REFERENCES

1. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst.* 1999; 91:1194–210.
2. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr>, accessed on 10/July/2015.
3. Ryberg D, Hewan A, Phillips DH, Haugen A. Different susceptibility to smoking-induced DNA damage among male and female lung cancer patients. *Cancer Res.* 1994; 54:5801–3.
4. Dogan S, Shen R, Ang DC, Johnson ML, D'Angelo SP, Paik PK, Brzostowski EB, Riely GJ, Kris MG, Zakowski MF, Ladanyi M. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin Cancer Res.* 2012; 18:6169–77.
5. Radzikowska E, Głaz P, Roszkowski K. Lung cancer in women: age, smoking, histology, performance status, stage, initial treatment and survival. Population-based study of 20,561 cases. *Ann Oncol.* 2002; 13:1087–93.
6. Patel JD. Lung cancer in women. *J Clin Oncol.* 2005; 23:32
7. Pinto J, Prado A, Cárdenas N, Valdiviezo P, Neciosup S, Aguilar A, Sarria G, Zaharia M, Flores C, Mas L. Increased susceptibility to lung cancer related to smoking in women is not explained by the expression of DNA repair genes. *J Thorac Oncol.* 2014; 9:P2.40.
8. Nowarski R, Wilner OI, Cheshin O, Shahar OD, Kenig E, Baraz L, Britan-Rosich E, Nagler A, Harris RS, Goldberg M, Willner I, Kotler M. APOBEC3G enhances lymphoma cell radioresistance by promoting cytidine deaminase-dependent DNA repair. *Blood.* 2012; 120:366–75.
9. Chang LC, Kuo TY, Liu CW, Chen YS, Lin HH, Wu PF. APOBEC3G exerts tumor suppressive effects in human hepatocellular carcinoma. *Anticancer Drugs.* 2014; 25:456–61.
10. Sato K, Takeuchi JS, Misawa N, Izumi T, Kobayashi T, Kimura Y, Iwami S, Takaori-Kondo A, Hu WS, Aihara K, Ito M, An DS8 Pathak VK, Koyanagi Y. APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. *PLoS Pathog.* 2014; 10:e1004453.
11. Hayashi K, Jutabha P, Endou H, Sagara H, Anzai N. LAT1 is a critical transporter of essential amino acids for immune reactions in activated human T cells. *J Immunol.* 2013; 191:4080–5.
12. Jarvis LB, Goodall JC, Gaston JS. Human leukocyte antigen class I-restricted immunosuppression by human CD8+ regulatory T cells requires CTLA-4-mediated interaction with dendritic cells. *Hum Immunol.* 2008; 69:687–95.
13. Elsaleh H, Joseph D, Grieco F, Zeps N, Spry N, Iacopetta B. Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer. *Lancet.* 2000; 355:1745–50.
14. Micheli A, Ciampichini R, Oberaigner W, Ciccolallo L, de Vries E, Izarzugaza I, Zamboni P, Gatta G, De Angelis R; EUROCARE Working Group. The advantage of women in cancer survival: an analysis of EUROCARE-4 data. *Eur J Cancer.* 2009; 45:1017–27.
15. Loi S, Michiels S, Salgado R, Sirtaine N, Jose V, Fumagalli D, Kellokumpu-Lehtinen PL, Bono P, Kataja V, Desmedt C, Piccart MJ, Loibl S, Denkert C, et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. *Ann Oncol.* 2014; 25:1544–50.
16. Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, Rouas G, Francis P, Crown JP, Hitre E, de Azambuja E, Quinaux E, Di Leo A, et al. Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. *J Clin Oncol.* 2013; 31:860–7.
17. Salgado R, Denkert C, Campbell C, Salgado R, Denkert C, Campbell C, Savas P, Nucifero P, Aura C, de Azambuja E, Eidtmann H, Ellis CE, Baselga J, et al. Tumor-Infiltrating Lymphocytes and Associations With Pathological Complete Response and Event-Free Survival in HER2-Positive Early-Stage Breast Cancer Treated With Lapatinib and Trastuzumab A Secondary Analysis of the NeoALTTO Trial. *JAMA Oncol.* 2015; 1:448–54.

18. Dieu-Nosjean MC, Antoine M, Danel C, Heudes D, Poulot V, Rabbe N, Laurans L, Tartour E, de Chaisemartin L, Lebecque S, Fridman WH, Cadranel J. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J Clin Oncol*. 2008; 26:4410–7.
19. Cheng H, Long F, Jaiswar M, Yang L, Wang C, Zhou Z. Prognostic role of the neutrophil-to-lymphocyte ratio in pancreatic cancer: a meta-analysis. *Sci Rep*. 2015 Jul. 31; 5:11026.
20. Kossenkov AV, Dawany N, Evans TL, Kucharczuk JC, Albelda SM, Showe LC, Showe MK, Vachani A. Peripheral immune cell gene expression predicts survival of patients with non-small cell lung cancer. *PLoS One*. 2012; 7:e34392.
21. Kossenkov AV, Vachani A, Chang C, Nichols C, Billouin S, Horng W, Rom WN, Albelda SM, Showe MK, Showe LC. Resection of non-small cell lung cancers reverses tumor-induced gene expression changes in the peripheral immune system. *Clin Cancer Res*. 2011; 17:5867–77.
22. Chen C1, Grennan K, Badner J, Zhang D, Gershon E, Jin L, Liu C. Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. *PLoS One*. 2011; 28;6:e17238.
23. Qu K, Zaba LC, Giresi PG, Li R, Longmire M, Kim YH, Greenleaf WJ, Chang HY. Individuality and Variation of Personal Regulomes in Primary Human T Cells. *Cell Systems*. 2015; 1:51–61.
24. National Comprehensive Cancer Network. National Comprehensive Cancer Network (Version 4.2016). http://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed January. 26, 2016.
25. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, Chow LQ, Vokes EE, Felip E, Holgado E, Barlesi F, Kohlhäufel M, Arrieta O, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *N Engl J Med*. 2015; 373:1627–39.
26. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, Molina J, Kim JH, Arvis CD, Ahn MJ, Majem M, Fidler MJ, de Castro G Jr, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2015. pii: S0140-673601281–7.
27. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12:453–7.