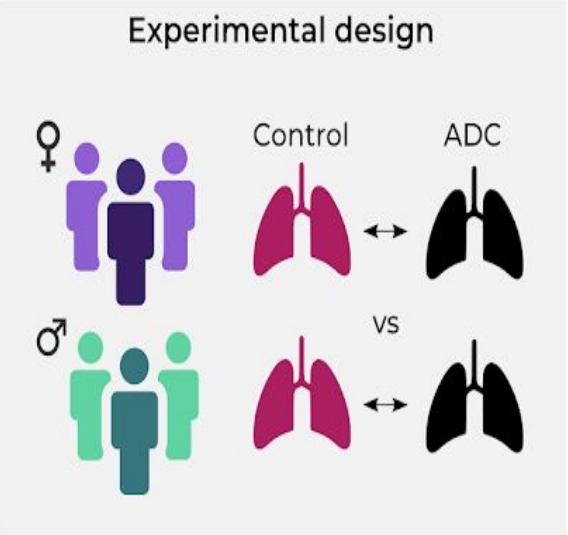
Functional signatures in non-small-cell lung cancer: a systematic review and meta-analysis of sex-based differences in transcriptomic studies

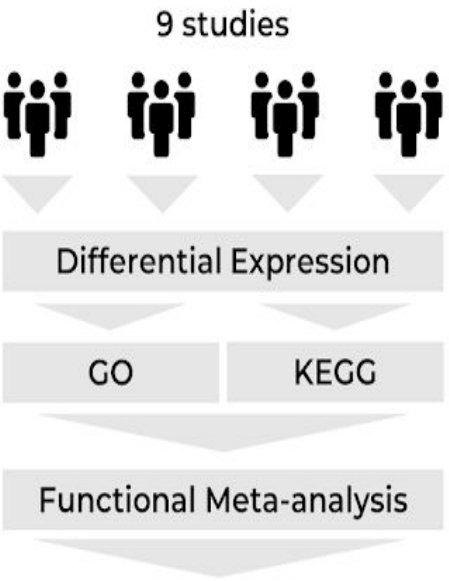
Irene Pérez-Díez [1,2], Marta R. Hidalgo [1], Pablo Malmierca-Merlo [1,3], Rosa Farràs [4], María de la Iglesia-Vayá [2], Mariano Provencio [5], Atocha Romero [5] and Francisco García-García [1,6*]

[1] Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF). Valencia, Spain.
[2] Biomedical Imaging Unit FISABIO-CIPF. Fundación para el Fomento de la Investigación Sanitario y Biomédica de la Comunidad Valenciana. Valencia, Spain.
[3] Atos Research Innovation (ARI), Spain.
[4] Department of Oncogenic Signalling, Principe Felipe Research Center (CIPF). Valencia, Spain.
[5] Hospital Universitario Puerta de Hierro-Majadahonda. Majadahonda, Spain.
[6] Spanish National Bioinformatics Institute, ELIXIR-Spain (INB, ELIXIR-ES).
* Correspondence: fgarcia@cipf.es

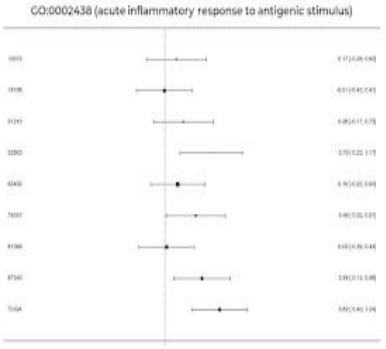
Abstract: While studies have established the existence of differences in the epidemiological and clinical patterns of lung adenocarcinoma between male and female patients, we know relatively little regarding the molecular mechanisms underlying such sex-based differences. In this study, we explore said differences through a meta-analysis of transcriptomic data. We performed a meta-analysis of the functional profiling of nine public datasets that included 1,366 samples from Gene Expression Omnibus and The Cancer Genome Atlas databases. Meta-analysis results show an enrichment of Gene Ontology terms and Kyoto Encyclopedia of Genes and Genomes pathways related to the immune response, nucleic acid metabolism, and purinergic signaling. We discovered the overrepresentation of terms associated with the immune response, particularly with acute inflammatory response, and purinergic signaling in female lung adenocarcinoma patients, which could influence reported clinical differences. Further evaluations of the identified differential biological processes and pathways could lead to the discovery of new biomarkers and therapeutic targets. Our findings also emphasize the relevance of sex-specific analyses in biomedicine, which represents a crucial aspect influencing biological variability in disease.

Keywords: NSCLC; biomarkers; functional profiling; meta-analysis, sex characteristics





GO ID	Name	LOR	LOR IC	FDR	Sex
0045410	positive regulation of interleukin-6 biosynthetic process	0.477	[0.28, 0.67]	0.008	F
0043374	CD8-positive, alpha-beta T cell differentiation	0.5	[0.25, 0.75]	0.031	F
0002863	positive regulation of inflammatory response to antigenic stimulus	0.612	[0.34, 0.88]	0.012	F
0006183	small nucleolar ribonucleoprotein complex assembly	0.504	[0.21, 0.79]	0.023	M



1. Introduction

Lung cancer is the most frequently diagnosed cancer and the leading cause of cancer-related death worldwide, representing 18.4% of all cancer deaths [1]. Exposure to tobacco, domestic biomass fuels, asbestos, and radon represent the most relevant lung cancer risk factors [1–3].

Lung cancer exhibits differences in clinical characteristics and outcomes depending on sex, with better survival observed in women [3–5]. While lung cancer incidence worldwide is higher in men, there exists an increasing trend in women that cannot be solely explained by tobacco consumption [1,2]. Furthermore, studies have reported sex-dependent differences in estrogen receptors and their impact on lung cancer [6–8]; however, conflicting results have attributed susceptibility of lung cancer in women to genetic variants, hormonal factors, molecular abnormalities, and oncogenic viruses [3,9–11].

Adenocarcinoma represents the most frequent non-small cell lung cancer (NSCLC) subtype in both sexes [12], with a higher predominance in women compared to men (41% of cases in women versus 34% in men) [3,5,11]. Interestingly, Wheatley-Price et al. demonstrated a more pronounced survival rate difference between male and female lung adenocarcinoma patients when compared to other tumor types [5].

A range of transcriptomic studies have described those biological processes contributing to the pathology of NSCLC or lung adenocarcinoma; however, technical variability among studies, inter-individual biological variability, and reduced sample size represent potential confounding factors in the evaluation of these contributions [13–19]. In addition, limited efforts have been made to explore the molecular mechanisms underlying sex-based differences in lung adenocarcinoma, with few studies considering this differential perspective [20,21]. These limitations can be partially addressed through meta-analysis, a robust methodology that combines information from related but independent studies to derive results with increased statistical power and precision [22,23].

To explore the molecular mechanisms underlying sex-based differences in lung adenocarcinoma, we performed a meta-analysis based on functional profiles of transcriptomic studies. After an exhaustive review and selection, we retrieved and analyzed nine studies from Gene Expression Omnibus (GEO) [24] and The Cancer Genome Atlas (TCGA) [25], and then combined the results in a random-effects meta-analysis. This approach allowed the identification of functional alterations caused by lung adenocarcinoma in men and women.

2. Results

We have organized our findings into three sections: the first describes the studies reviewed and selected in the systematic review; the second section reports on the results of the bioinformatic analysis of each of these selected studies as follows: i) exploratory analysis, ii) differential expression, and iii) functional characterization; and the third section presents the results of the differential functional profiling by sex.

2.1. Study Search and Selection

The systematic review identified 207 non-duplicated studies, of which 48.8% included both male and female patients (**Figure S1**). After applying inclusion and exclusion criteria (see **Figure 1**), we selected nine transcriptomic studies for further analysis (**Table 1**). The selected studies represented a population of 1,366 samples (369 controls and 997 cases), of which 44% were from men and 56% from women (**Figure 2**). 71.16% of the samples were from early-stage patients, and 91.29% were from patients older than 50 years. To reduce variability, we focused our study on early-stage samples. Further information regarding the selected studies is detailed in **Table 1**, **Figure 2**, and **Supplementary Table S1**.

Table 1. Studies selected after the systematic review

Study	Platform	Publication
GSE10072	Affymetrix Human Genome U133A Array	[26]
GSE19188	Affymetrix Human Genome U133 Plus 2.0 Array	[27]
GSE31210	Affymetrix Human Genome U133 Plus 2.0 Array	[28,29]
GSE32863	Illumina HumanWG-6 v3.0 Expression BeadChip	[30]
GSE63459	Illumina HumanRef-8 v3.0 Expression BeadChip	[31]
GSE75037	Illumina HumanWG-6 v3.0 Expression BeadChip	[32]
GSE81089	Illumina HiSeq 2500	[33]
GSE87340	Illumina HiSeq 2000	[34]
TCGA	Illumina HiSeq 2000	[25]

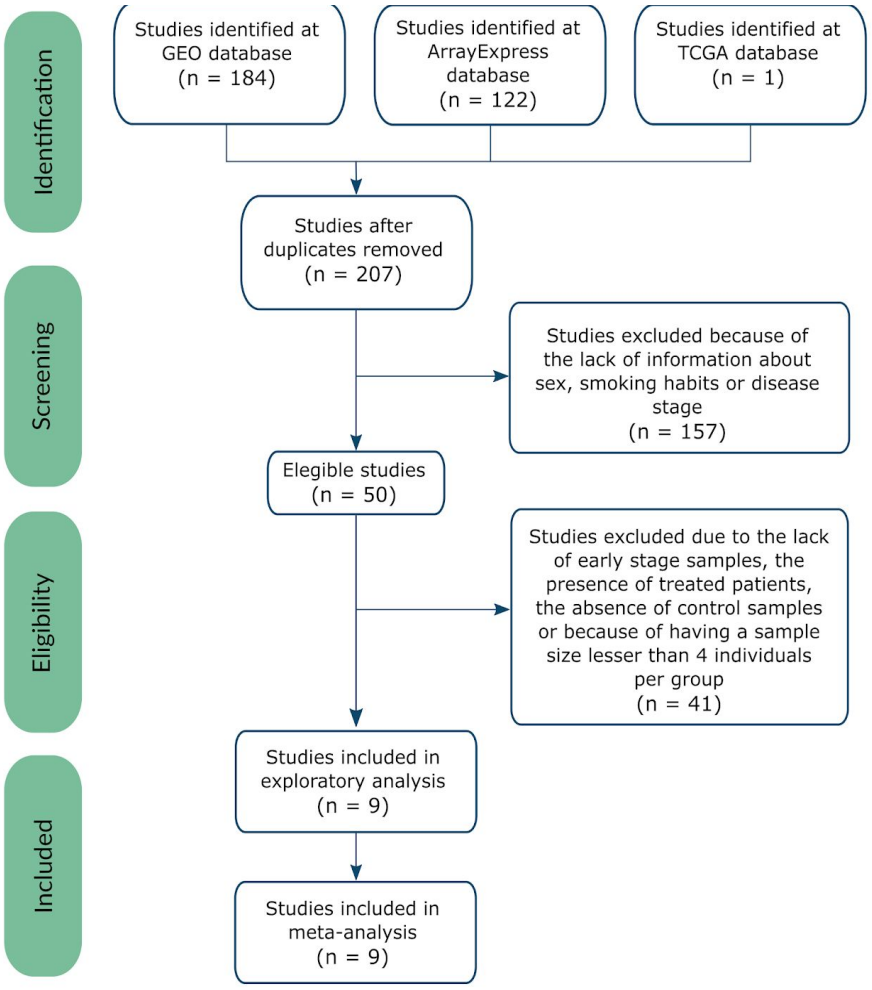


Figure 1. Flow of information through the different phases of the systematic review, following PRISMA Statement guidelines [35].

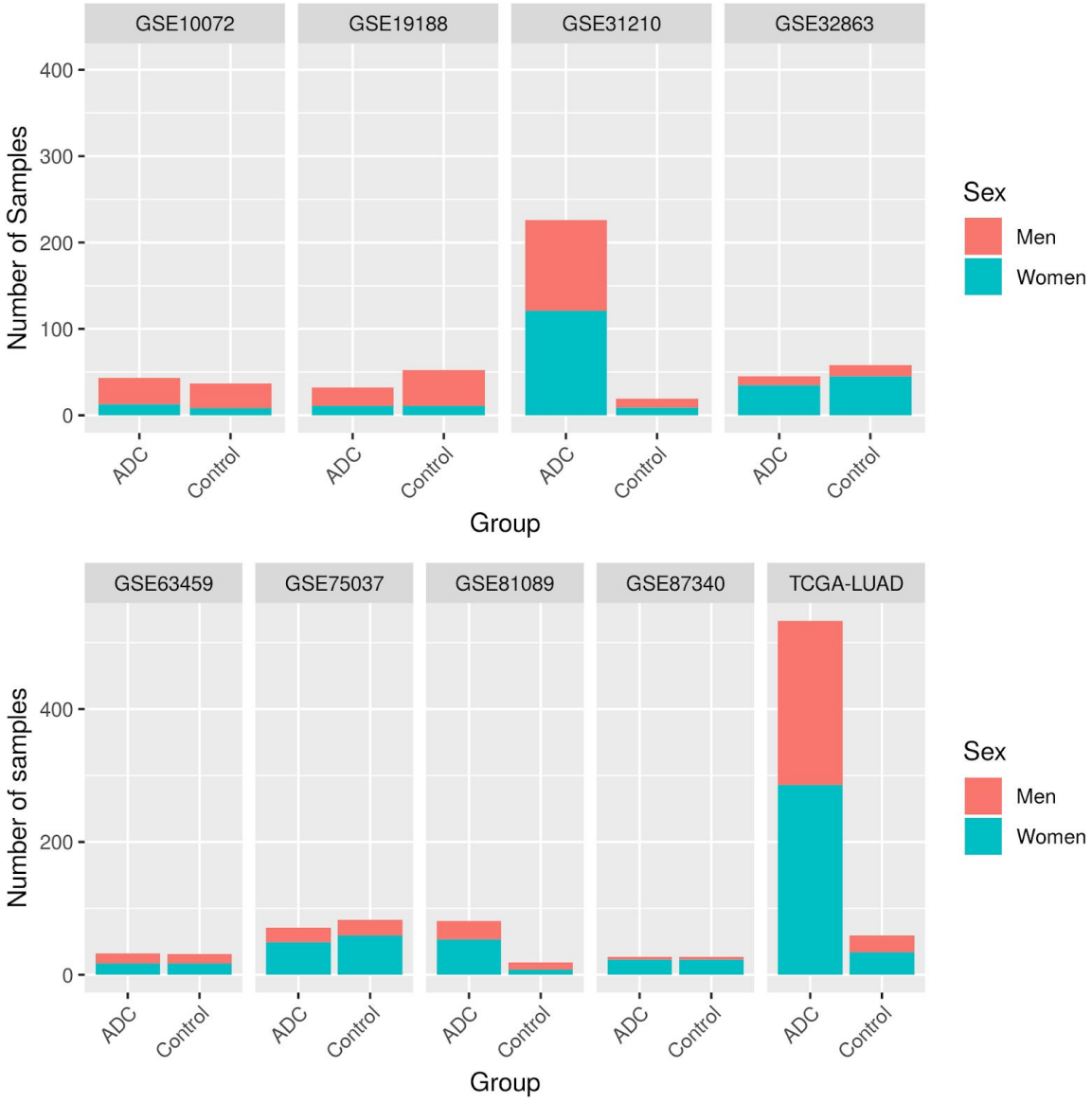


Figure 2. Number of samples per study, divided by sex and experimental group (ADC – lung adenocarcinoma samples).

2.2. Individual Analysis

Exploratory analysis found a lack of abnormal behavior except for three samples in principal component analysis (PCA) and unsupervised clustering; therefore, we excluded the GSM47570 and GSM47578 samples in study GSE19188, and the GMS773784 sample in study GSE31210 from further analysis.

The differential expression results for each study demonstrated a large number of differentially expressed genes when comparing female lung adenocarcinoma patients to female control patients, and male adenocarcinoma patients to male control patients (**Supplementary Table S2**). However, the evaluation of sex-based differences in lung adenocarcinoma patients provided a small number of significantly affected genes (see **Supplementary Table S3**), with no intersecting genes. While this was a gene set analysis based on logistic models, we were able to analyze all genes in each study, ordered by their differential expression level (regardless of significance), to discover sets of genes involved in the same biological function that share the same expression pattern.

Individual functional enrichment analysis of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways revealed a diversity of significant results among datasets, which we have summarized in **Table 2**. When analyzing intersections, UpSet plots (**Figure 3**) established that most results were exclusive to a single study or common between two to four studies.

Table 2. Summary of functional enrichment analysis results by Gene Ontology functions (BP: Biological Process, MF: Molecular Functions, CC: Cellular Component), and KEGG pathways. "Up" terms are overrepresented in female lung adenocarcinoma patients, while "Down" terms are overrepresented in male lung adenocarcinoma patients.

	Significant GO BP terms			**Significant GO MF terms**			**Significant GO CC terms**			**Significant KEGG pathways**		
	Up	Down	Total	Up	Down	Total	Up	Down	Total	Up	Down	Total
GSE10072	26	153	179	29	8	37	16	40	56	1	5	6
GSE19188	7	12	19	0	3	3	11	20	31	1	4	5
GSE31210	21	2	23	0	0	0	4	0	4	8	2	10
GSE32863	428	51	479	28	5	33	40	41	81	27	6	33
GSE63459	0	26	26	0	3	3	8	27	35	1	4	5
GSE75037	245	35	280	14	4	18	14	21	35	22	5	27
GSE81089	2	1	3	7	1	8	7	4	11	1	1	2
GSE87340	178	62	240	3	0	3	48	13	61	26	1	27
TCGA	294	228	522	28	30	58	21	70	91	30	17	47

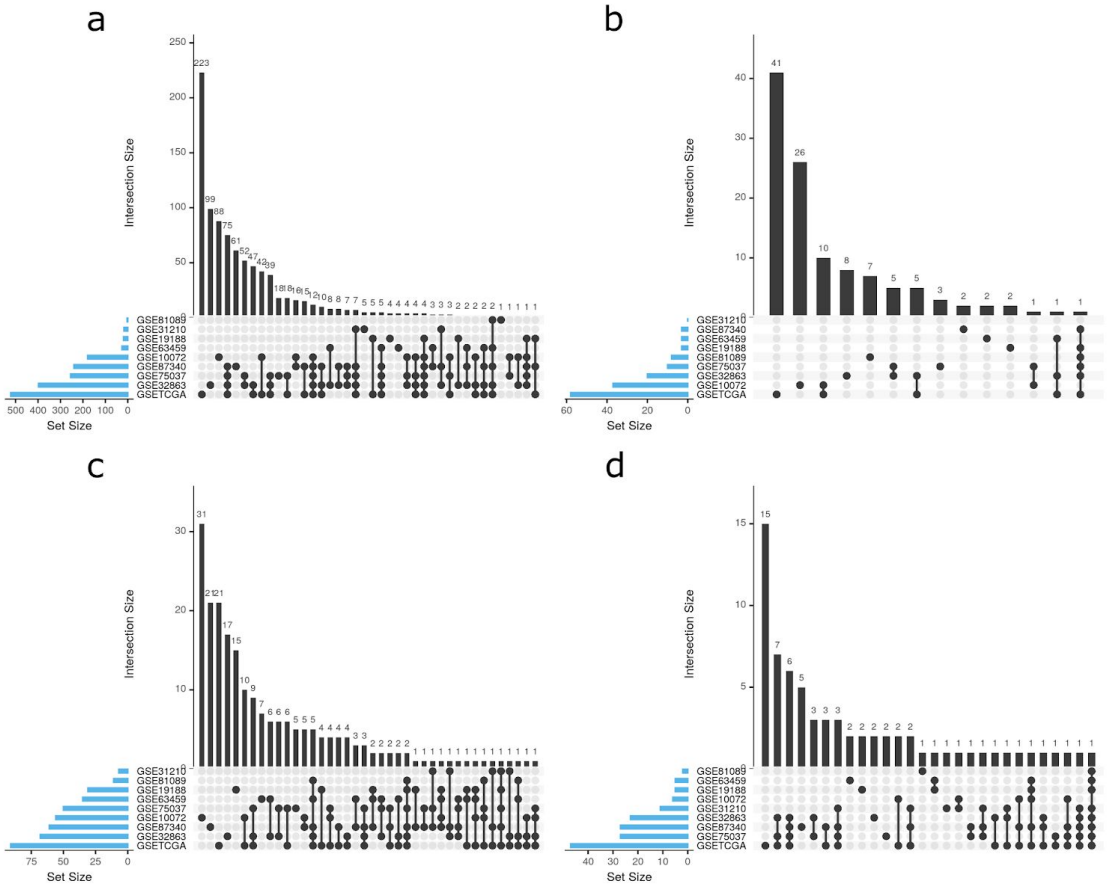


Figure 3. UpSet plots for **(a)** GO biological process, **(b)** GO molecular functions, **(c)** GO cellular components, and **(d)** KEGG pathways. UpSet plots detailing the number of common elements among GO terms in our functional enrichment analysis. Horizontal bars indicate the number of significant elements in each study. The vertical bars indicate the common elements in the sets, indicated with dots under each bar. The single points represent the number of unique elements in each group.

2.3. Meta-analysis

We performed a functional meta-analysis for each of the 8,672 GO functions and KEGG pathways, including every term found in at least two studies. Results with a false discovery rate (FDR) of < 0.05 included 106 GO biological processes (BP), 3 GO molecular functions (MF), and 20 KEGG pathways, which were associated with 21 wider functional groups. We rejected potential bias on the significant results after the inspection of funnel plots; furthermore, sensitivity analyses failed to indicate alterations in the results due to the inclusion of any study. The results for the 129 significant GO terms and KEGG pathways are further detailed in **Supplementary Table S4**, including FDR, the log odds ratio (LOR), and its 95% confidence interval (CI), and the standard error (SE) of the LOR.

2.3.1. Upregulated Functions

We discovered that 43.88% of detected functions related to immune response (**Supplementary Table S4** and **Figure 4**), which were all upregulated in female lung adenocarcinoma patients. The results provided evidence for the positive regulation of an acute inflammatory response, with CD8+ alpha-beta T cell differentiation and activation, B cell proliferation and activation, and an increase of interleukin (IL) biosynthesis, including IL-2, 6, 8, 10 and 17.

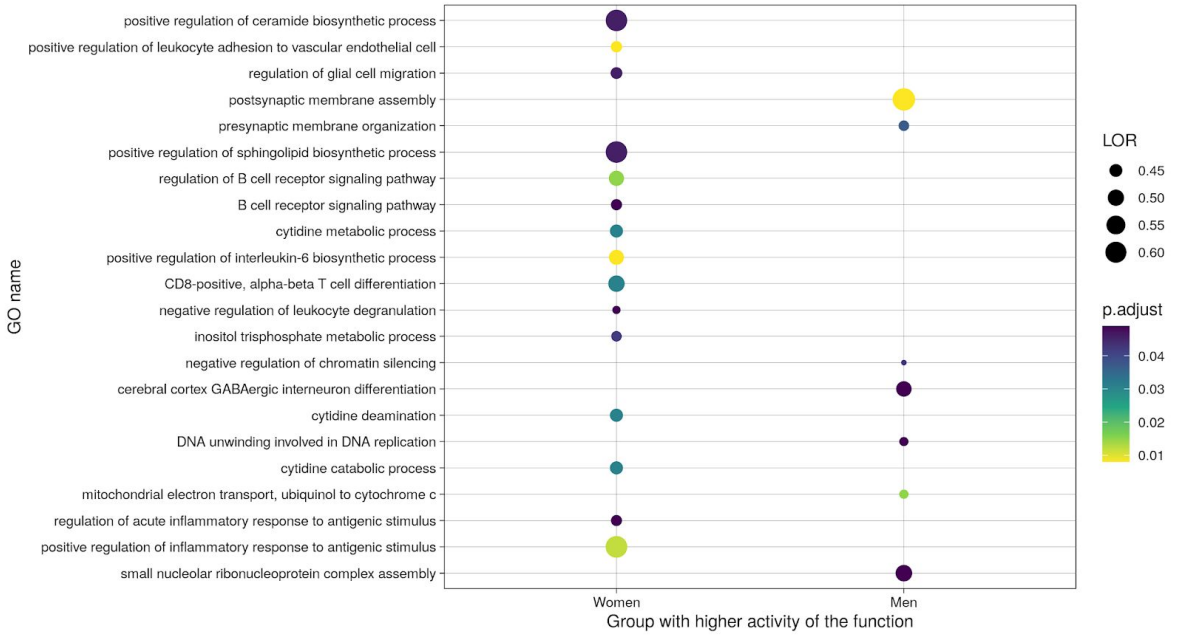


Figure 4. Summary dot plot of GO BP meta-analysis results. Only those significant terms with a LOR over 0.4 are shown.

"Metabolism - Nucleic acids metabolism and signaling" was the second most abundant functional group upregulated in female lung adenocarcinoma patients, comprising 8.63% of the altered functions. These GO terms and KEGG pathways are mainly related to purinergic signaling through G protein-coupled receptors and cytidine metabolism. Other functional groups upregulated in female lung adenocarcinoma patients include lipid metabolism, cell migration, and homeostasis.

2.3.2. Downregulated Functions

23.26% of the significant functions exhibited lower activity in female when compared to male

lung adenocarcinoma patients. Downregulated functional groups include those related to apoptosis, cell cycle progression, cell junctions, DNA repair and telomere protection, mitochondrial processes, neural development, post-translational changes, post-transcriptional changes, protein degradation, and transcription regulation (**Supplementary Table S4** and **Figure 4**).

2.4. Metafun-NSCLC Web Tool

The Metafun-NSCLC web tool (https://bioinfo.cipf.es/metafun-nscls) contains information related to the nine studies and 1,329 samples involved in this study. For each study, this resource includes fold-changes of genes and LOR of functions and pathways that can be explored by users to identify profiles of interest.

We carried out a total of 8,672 meta-analyses. For each of the 129 significant functions and pathways, Metafun-NSCLC depicts the global activation level for all studies and the specific contribution of each study, using statistical indicators (LOR, CI, and p-value) and graphical representations by function as forest and funnel plots. This open resource hopes to contribute to data sharing between researchers, the elaboration of new studies, and the discovery of new findings.

3. Discussion

Cases of NSCLC, including adenocarcinoma, exhibit differences in incidence, prevalence, and severity in female and male patients [1,3,5,36]. Elucidating the molecular basis for this sex-based differential impact will have clinical relevance, as this conformation can guide/improve both diagnosis and treatment.

Biomedical research generally underrepresents female patients, with sex-based differences rarely considered [37,38]. Our systematic review of transcriptomic studies revealed that only 48.8% of NSCLC-related datasets considered both sexes, a figure similar (49%) to that reported by Woitowich et al. [38]. Sex-based differences have an impact on disease biomarkers, drug response, and treatment [37], and, therefore, sex must represent a critical component of experimental design. Added to this problem, we faced a lack of standardization among studies and detailed clinical information (mutations, smoking status, stages) when searching for suitable datasets. The consideration of Findable, Accessible, Interoperable, and Reusable (FAIR) data principles [39], a requisite for quality science, would ensure that generated data can be of further use throughout the scientific community.

To the best of our knowledge, only two studies have attempted to address the functional alterations caused by NSCLC in both male and female patients - Shi et al. [21], taking only into account women and Araujo et al. [20]. Shi et al. [21] integrated the samples of two datasets for a differential expression analysis followed by a functional enrichment analysis, whereas Araujo et al. [20] independently processed six datasets and jointly analyzed their results.

In this study, we addressed sex-based differences in NSCLC patients through meta-analysis to address previous limitations and improve on those approaches employed by others. This robust methodology integrates groups of data and provides results with higher statistical power and precision [22,23] and reveals findings that cannot be obtained through the intersection or addition of results in individual studies.

Our results demonstrate an enrichment of immune response-related terms in female lung adenocarcinoma patients, which agrees with the findings of Araujo et al. [20]. The analyzed functions suggest the positive regulation of CD8+ alpha-beta T cell activation and differentiation in female lung adenocarcinoma patients, which play an essential role in antitumor immunity [39,40]. Furthermore, Ye et al. [40] discovered a more abundant population of effector memory CD8+ T cells in female lung adenocarcinoma patients, which agrees with our results. A previous study described CD8+ lymphocyte levels as a prognostic biomarker in NSCLC [41], and specifically in lung adenocarcinoma [42], with a correlation between higher levels of CD8+ lymphocytes with higher survival rates and lower disease recurrence. Elevated levels of active CD8+ T cells in female lung adenocarcinoma patients could form part of the molecular mechanism underlying higher survival rates when compared to male lung adenocarcinoma patients. Activation of the Notch signaling pathway decreases CD8+ T lymphocyte activity in lung adenocarcinoma [43]; therefore, the downregulation of the Notch signaling pathway discovered in female lung adenocarcinoma

patients could explain higher CD8+ T activity when compared to male lung adenocarcinoma patients

Concerning the immune response, we also detected differences that supported the increased production of IL-2, which is known to stimulate T cell proliferation and the production of effector T cells, thereby amplifying the lymphocytic response [41]. Higher levels of IL-2 could also relate to increased activity of CD8+ T cells in female lung adenocarcinoma patients. Increased levels of IL-10 are also supported in female lung adenocarcinoma patients and, although IL-10 has anti-inflammatory and anti-immune activities [42,43], studies have suggested a dual role in cancer. In advanced lung adenocarcinoma, high expression of IL-10 receptor 1 correlates with worse prognosis [42], while IL-10 expression by T-regulatory cells inhibits apoptosis through Programmed death-ligand 1 inhibition [43]. Nevertheless, IL-10 correlates with better prognosis when expressed by CD8+ T cells in early stage NSCLC [44] and it seems to activate the antitumor control of CD8+ T cells [45]. IL-2 and IL-10 could have an increased activity in early stage female patients

, alongside with a higher population of active CD8+ T cells than males, conferring women a survival advantage.

We also detected the positive regulation of IL-6 biosynthesis in female lung adenocarcinoma patients, with increased IL-6 levels correlating with worse prognosis in NSCLC patients in previous studies [46,47]. Network analysis in non-smoking female lung adenocarcinoma patients described IL-6 as one of the pathology's central nodes [21], and these findings agree with our results, which provide evidence of the critical role of IL-6 in tumor progression in female lung adenocarcinoma patients. IL-8 and IL-17 also exhibit increased production and biosynthesis in female lung adenocarcinoma patients, with said interleukins known to influence tumor growth and metastasis and correlate with worse prognosis [48–50].

Although altered immune responses can positively and negatively influence tumor progression, our findings, and others, have detected GO terms that point to an elevated acute immune response in female compared to male lung adenocarcinoma patients. Of note, sex-based immunological differences in lung adenocarcinoma might have an impact on immunotherapy response. Different studies have addressed the role of sex in immunotherapy [40,51–53], establishing improved survival for female NSCLC patients. The discovered molecular pathways differentially activated between male and female lung adenocarcinoma patients may underlie phenotypic differences regarding immunotherapy response.

We also detected an upregulation of purinergic signaling and nucleic acid metabolism in female lung adenocarcinoma patients, a finding not described in previous studies. In NSLC has been described An antitumor effect of the P2X4 receptor has been described in NSLC [54], which also exhibits sexual dimorphism in murine brain microglia [55]. Other P2 and A2 receptors play a role in NSCLC [54], but evidence of sex-based differences in receptor expression in human NSCLC patients has yet to be reported. Thus, the exploration of the differential role of purinergic signaling between male and female lung adenocarcinoma patients may represent an interesting proposition.

Our study has characterized functional differences between sexes in lung adenocarcinoma, shedding light on the functional basis behind this pathology in men and women. While our meta-analysis findings confirmed the conclusions of other studies, we also report previously undescribed alterations in biological processes that may broaden this field of study. Further knowledge regarding how those factors related to the functional mechanisms described above differentially impact male and female lung adenocarcinoma patients may improve our understanding of the disease and improve treatment and diagnosis through biomarker identification.

4. Materials and Methods

Bioinformatics and statistical analysis employed R software v.3.5.3 [56]. **Supplementary Table S5** details R packages and versions.

4.1. Study Search and Selection

Publicly available datasets were collected from GEO [24], ArrayExpress [57], and TCGA [25]. A systematic search of studies published in the period 2004-2018 was conducted in 2019 following the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [35]. Several keywords were employed in the search, including lung adenocarcinoma (ADC), non-small-cell lung carcinoma (NSCLC), Homo sapiens, and excluding cell lines.

Eleven variables were considered for each study, including the clinical characteristics of the patients (e.g., sex and smoking habit) and experimental design (e.g., sample size and sample extraction source). The final inclusion criteria were:

Sex, disease stage, and smoking habit variables registered
RNA extracted directly from human lung biopsies
Both normal and lung adenocarcinoma samples available
Patients had not undergone treatment before biopsy
Sample size of > 3 for case and control groups in both sexes

Finally, normalized gene expression data of six array NSCLC datasets (GSE10072, GSE19188, GSE31210, GSE32863, GSE63459, and GSE75037) and counts matrix of three RNA-seq NSCLC datasets (GSE81089, GSE87340, and TCGA-LUAD) were retrieved.

4.2. Individual Transcriptomics Analysis

Individual transcriptomics analysis consisted of three steps: pre-processing, differential expression analysis, and functional enrichment analysis (**Figure 5a**).

Data pre-processing included the standardization of the nomenclature of the clinical variables included in each study, normalization of RNA-seq counts matrix, and exploratory analysis. RNA-seq counts were pre-processed with the *edgeR* [58] R package using the trimmed mean of m-values (TMM) method [59]. Annotation from probe set to Entrez identifiers from the National Center for Biotechnology Information [60] database and gene symbol was carried out with the *biomaRt* [61] R package. When dealing with duplicated probe-to-Entrez mappings, the median of their expression values was calculated. The exploratory analysis included unsupervised clustering and PCA, to detect patterns of expression between samples and genes, and the presence of batch effect in each study (**Figure 5b**).

Differential expression analyses were performed using the *limma* [62] R package. To detect differentially expressed genes in male and female lung adenocarcinoma patients, the following contrast was applied:

(ADC.W - Control.W) - (ADC.M - Control.M)

where ADC.W, Control.W, ADC.M and Control.M correspond to lung adenocarcinoma affected women, control women, lung adenocarcinoma affected men and control men, respectively. Paired samples design was implemented, and tobacco consumption was included as a batch effect to reduce its impact on data. P-values were calculated and corrected for FDR [63]. This comparison allows the detection of genes and functions altered by the disease and that have higher or lower activity in women when compared to men. Significant functions and genes were considered when FDR < 0.05.

Functional enrichment analyses were performed using the Gene Set Enrichment Analysis (GSEA) implemented in the *mdgsa* [64] R package. P-values were, again, corrected for FDR. For functional annotation, two functional databases were used: the KEGG Pathway database [65] and GO [66]. GO terms were analyzed and propagated independently for each GO ontology: BPs, MFs, and cellular components (CC). Those annotations excessively specific or generic were filtered out, keeping functions with blocks of annotations between 10 and 500. Intersections within groups were analyzed with UpSet plots [67] (**Figure 5c**).

4.3. Functional Meta-analysis

Functional GSEA results were integrated into a functional meta-analysis [68] implemented with *mdgsa* and *metafor* [69] R packages. Meta-analysis was applied under the DerSimonian & Laird random-effects model [70], taking into account individual study heterogeneity. This model considers the variability of individual studies by increasing the weights of studies with less

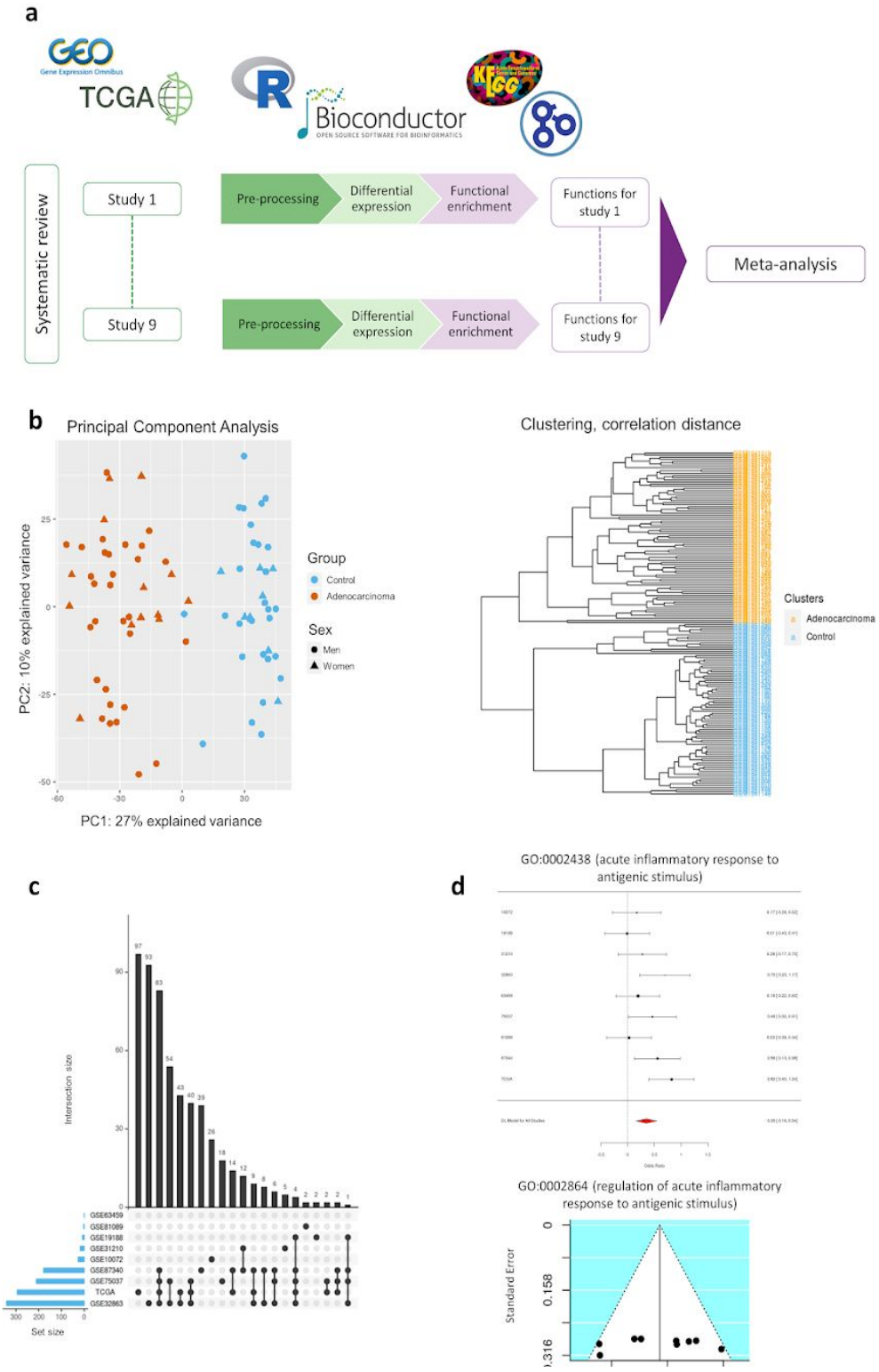


Figure 5. Workflow and analysis design. **(a)** Summary of the analysis design followed in this work, **(b)** example of exploratory analysis performed at the pre-processing stage, **(c)** example of UpSet plot as an intersection analysis for functional enrichment analysis results, and **(d)** examples of forest and funnel plots to assess meta-analysis results.

variability when computing meta-analysis results. Thus, the most robust functions between studies are highlighted.

A total of 6,467 GO BP terms, 785 GO CC terms, 1207 GO MF terms, and 213 KEGG pathways were evaluated. P-values, FDR corrected p-values, LOR, and 95% CIs of the LOR were calculated for each evaluated function. Functions and pathways with FDR < 0.05 were considered significant, and both funnel and forest plots were computed for each (**Figure 5d**). These representations were checked to assess for possible biased results, where LOR represents the effect size of a function, and the SE of the LOR serves as a study precision measure [71]. Sensitivity analysis (leave-one-out cross-validation [69]) was conducted for each significant function to verify possible alterations in the results due to the inclusion of any study.

4.4. Metafun-NSCLC Web Tool

All data and results generated in the different steps of the meta-analysis are available in the Metafun-NSCLC web tool (https://bioinfo.cipf.es/metafun-nsclc), which is freely accessible to any user and allows the confirmation of the results described in this manuscript and the exploration of other results of interest.

The front-end was developed using the Angular Framework. All graphics used in this web resource have been implemented with Plot.ly except for the exploratory analysis cluster plot, which was generated with ggplot2.

This easy-to-use resource is divided into five sections: 1) Summary of analysis results in each of the phases. Then, for each of the studies, the detailed results of the 2) exploratory analysis, 3) differential expression, and 4) functional profiling. The user can interact with the web tool through graphics and tables and also search for specific information for a gene or function. Finally, section 5) provides several indicators for the significant functions identified in the meta-analysis that inform whether they are more active in men or women.

5. Conclusions

We identified immune responses and purinergic signaling as the two main biological processes altered between female and male lung adenocarcinoma patients by a meta-analysis of transcriptomic datasets. Immune responses and purinergic signaling exhibit increased activity in female lung adenocarcinoma patients, whereas other processes (such as DNA repair) are more active in male lung adenocarcinoma. Although further biological experiments are required to verify and fully explore these findings, our results provide new clues to understand the molecular mechanisms of sex-based differences in lung adenocarcinoma patients and opens new perspectives to identify new biomarkers and therapeutic targets.

Data Availability*:* The data used for the analyses described in this work is publicly available at GEO (https://www.ncbi.nlm.nih.gov/geo/) and the Genomic Data Commons Data Portal (TCGA, https://portal.gdc.cancer.gov/). The accession numbers of the GEO datasets downloaded are: GSE10072, GSE19188, GSE31210, GSE32863, GSE63459, GSE75037, GSE81089 and GSE87340. Only the samples associated with the project TCGA-LUAD were downloaded from the TCGA.

Computer Code and Software: The code developed for the analyses described in this work is publicly available at GitLab (https://gitlab.com/ubb-cipf/metafunr). All software and versions used are detailed in **Supplementary Table S5**.

Supplementary Materials. Figure S1: Distribution of sex information among the reviewed studies, **Table S1**: Distribution of clinicopathological characteristics of each study population, **Table S2**: Summary of differential expression analysis results, **Table S3**: Significant genes differentially expressed between male and female lung adenocarcinoma patients, **Table S4**: All significant GO terms and KEGG pathways in the functional meta-analysis, **Table S5**: Software and versions used in this work.

Author Contributions: Conceptualization, F.G.G.; methodology, M.R.H., and F.G.G.; software, P.M.M., I.P.D.,

M.R.H., and F.G.G.; validation, P.M.M., A.R., M.B., R.F., M.R.H., and F.G.G.; formal analysis, I.P.D., M.R.H., and F.G.G; investigation, I.P.D., M.R.H., and F.G.G.; data curation, I.P.D.; writing—original draft preparation, I.P.D., M.R.H., and F.G.G.; writing—review and editing, I.P.D., M.R.H., A.R., M.B., R.F., and F.G.G.; visualization, I.P.D., M.R.H., and F.G.G.; supervision, M.R.H., and F.G.G.; funding acquisition, F.G.G.; project administration, F.G.G. All authors have read and agreed to the published version of the manuscript.

Funding: ISCIII PT17/0009/0015 FEDER.

Acknowledgments: The authors thank the Principe Felipe Research Center (CIPF) for providing access to the cluster, which is co-funded by European Regional Development Funds (FEDER) in Valencian Community 2014-2020. The results published here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **2018**, *68*, 394–424, doi:10.3322/caac.21492.
2. Jemal, A.; Miller, K.D.; Ma, J.; Siegel, R.L.; Fedewa, S.A.; Islami, F.; Devesa, S.S.; Thun, M.J. Higher Lung Cancer Incidence in Young Women Than Young Men in the United States. *N. Engl. J. Med.* **2018**, *378*, 1999–2009, doi:10.1056/NEJMoa1715907.
3. Barta, J.A.; Powell, C.A.; Wisnivesky, J.P. Global Epidemiology of Lung Cancer. *Ann. Glob. Health* **2019**, *85*, 8, doi:10.5334/aogh.2419.
4. Barquín, M.; Calvo, V.; García-García, F.; Nuñez, B.; Sánchez-Herrero, E.; Serna-Blasco, R.; Auglytė, M.; Carcereny, E.; Rodriguez-Abreu, D.; López Castro, R.; et al. Sex is a strong prognostic factor in stage IV non-small-cell lung cancer patients and should be considered in survival rate estimation. *Cancer Epidemiol.* **2020**, *67*, 101737, doi:10.1016/j.canep.2020.101737.
5. Wheatley-Price, P.; Blackhall, F.; Lee, S.-M.; Ma, C.; Ashcroft, L.; Jitlal, M.; Qian, W.; Hackshaw, A.; Rudd, R.; Booton, R.; et al. The influence of sex and histology on outcomes in non-small-cell lung cancer: a pooled analysis of five randomized trials. *Ann. Oncol.* **2010**, *21*, 2023–2028, doi:10.1093/annonc/mdq067.
6. Fasco, M.J.; Hurteau, G.J.; Spivack, S.D. Gender-dependent expression of alpha and beta estrogen receptors in human nontumor and tumor lung tissue. *Mol. Cell. Endocrinol.* **2002**, *188*, 125–140, doi:10.1016/S0303-7207(01)00750-X.
7. Stabile, L.P.; Davis, A.L.G.; Gubish, C.T.; Hopkins, T.M.; Luketich, J.D.; Christie, N.; Finkelstein, S.; Siegfried, J.M. Human Non-Small Cell Lung Tumors and Cells Derived from Normal Lung Express Both Estrogen Receptor α and β and Show Biological Responses to Estrogen. *Cancer Res.* **2002**, *62*, 2141–2150.
8. Kawai, H.; Ishii, A.; Washiya, K.; Konno, T.; Kon, H.; Yamaya, C.; Ono, I.; Minamiya, Y.; Ogawa, J. Estrogen Receptor α and β are Prognostic Factors in Non–Small Cell Lung Cancer. *Clin. Cancer Res.* **2005**, *11*, 5084–5089, doi:10.1158/1078-0432.CCR-05-0200.
9. Kligerman, S.; White, C. Epidemiology of Lung Cancer in Women: Risk Factors, Survival, and Screening. *Am. J. Roentgenol.* **2011**, *196*, 287–295, doi:10.2214/AJR.10.5412.
10. Bain, C.; Feskanich, D.; Speizer, F.E.; Thun, M.; Hertzmark, E.; Rosner, B.A.; Colditz, G.A. Lung Cancer Rates in Men and Women with Comparable Histories of Smoking. *JNCI J. Natl. Cancer Inst.* **2004**, *96*, 826–834, doi:10.1093/jnci/djh143.
11. North, C.M.; Christiani, D.C. Women and Lung Cancer: What is New? *Semin. Thorac. Cardiovasc. Surg.* **2013**, *25*, 87–94, doi:10.1053/j.semtcvs.2013.05.002.
12. Travis, W.D.; Brambilla, E.; Noguchi, M.; Nicholson, A.G.; Geisinger, K.R.; Yatabe, Y.; Beer, D.G.; Powell, C.A.; Riely, G.J.; Van Schil, P.E.; et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International Multidisciplinary Classification of Lung Adenocarcinoma. *J. Thorac. Oncol.* **2011**, *6*, 244–285, doi:10.1097/JTO.0b013e318206a221.
13. Choi, J.K.; Yu, U.; Kim, S.; Yoo, O.J. Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* **2003**, *19*, i84–i90, doi:10.1093/bioinformatics/btg1010.
14. Hong, F.; Breitling, R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics* **2008**, *24*, 374–382, doi:10.1093/bioinformatics/btm620.
15. Ramasamy, A.; Mondry, A.; Holmes, C.C.; Altman, D.G. Key Issues in Conducting a Meta-Analysis of Gene Expression Microarray Datasets. *PLOS Med.* **2008**, *5*, e184, doi:10.1371/journal.pmed.0050184.
16. Campain, A.; Yang, Y.H. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics* **2010**, *11*, 408, doi:10.1186/1471-2105-11-408.
17. Rau, A.; Marot, G.; Jaffrézic, F. Differential meta-analysis of RNA-seq data from multiple studies. *BMC Bioinformatics* **2014**, *15*, 91, doi:10.1186/1471-2105-15-91.
18. Poli-Neto, O.B.; Meola, J.; Rosa-e-Silva, J.C.; Tiezzi, D. Transcriptome meta-analysis reveals differences of immune profile between eutopic endometrium from stage I-II and III-IV endometriosis independently of hormonal milieu. *Sci. Rep.* **2020**, *10*, 313, doi:10.1038/s41598-019-57207-y.
19. Sass, P.A.; Dąbrowski, M.; Charzyńska, A.; Sachadyn, P. Transcriptomic responses to wounding: meta-analysis of gene expression microarray data. *BMC Genomics* **2017**, *18*, 850, doi:10.1186/s12864-017-4202-8.
20. Araujo, J.M.; Prado, A.; Cardenas, N.K.; Zaharia, M.; Dyer, R.; Doimi, F.; Bravo, L.; Pinillos, L.; Morante, Z.; Aguilar, A.; et al. Repeated observation of immune gene sets enrichment in women with non-small cell lung cancer. *Oncotarget* **2016**, *7*, 20282–20292, doi:10.18632/oncotarget.7943.
21. Shi, K.; Li, N.; Yang, M.; Li, W. Identification of Key Genes and Pathways in Female Lung Cancer Patients Who Never Smoked by a Bioinformatics Analysis. *J. Cancer* **2019**, *10*, 51–60, doi:10.7150/jca.26908.
22. Normand, S.-L.T. Meta-analysis: formulating, evaluating, combining, and reporting. *Stat. Med.* **1999**, *18*, 321–359, doi:10.1002/(SICI)1097-0258(19990215)18:3<321::AID-SIM28>3.0.CO;2-P.
23. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*; Higgins, J., Green, S., Eds.; The Cochrane Collaboration;
24. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **2002**, *30*, 207–210, doi:10.1093/nar/30.1.207.
25. Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.M.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **2013**, *45*,

1113–1120. doi:10.1038/ng.2764.
26. Landi, M.T.; Dracheva, T.; Rotunno, M.; Figueroa, J.D.; Liu, H.; Dasgupta, A.; Mann, F.E.; Fukuoka, J.; Hames, M.; Bergen, A.W.; et al. Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. *PLOS ONE* **2008**, *3*, e1651. doi:10.1371/journal.pone.0001651.
27. Hou, J.; Aerts, J.; den Hamer, B.; van IJcken, W.; den Bakker, M.; Riegman, P.; van der Leest, C.; van der Spek, P.; Foekens, J.A.; Hoogsteden, H.C.; et al. Gene Expression-Based Classification of Non-Small Cell Lung Carcinomas and Survival Prediction. *PLOS ONE* **2010**, *5*, e10312. doi:10.1371/journal.pone.0010312.
28. Okayama, H.; Kohno, T.; Ishii, Y.; Shimada, Y.; Shiraishi, K.; Iwakawa, R.; Furuta, K.; Tsuta, K.; Shibata, T.; Yamamoto, S.; et al. Identification of Genes Upregulated in ALK-Positive and EGFR/KRAS/ALK-Negative Lung Adenocarcinomas. *Cancer Res.* **2012**, *72*, 100. doi:10.1158/0008-5472.CAN-11-1403.
29. Yamauchi, M.; Yamaguchi, R.; Nakata, A.; Kohno, T.; Nagasaki, M.; Shimamura, T.; Imoto, S.; Saito, A.; Ueno, K.; Hatanaka, Y.; et al. Epidermal Growth Factor Receptor Tyrosine Kinase Defines Critical Prognostic Genes of Stage I Lung Adenocarcinoma. *PLOS ONE* **2012**, *7*, e43923. doi:10.1371/journal.pone.0043923.
30. Selamat, S.A.; Chung, B.S.; Girard, L.; Zhang, W.; Zhang, Y.; Campan, M.; Siegmund, K.D.; Koss, M.N.; Hagen, J.A.; Lam, W.L.; et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. *Genome Res.* **2012**, *22*, 1197–1211. doi:10.1101/gr.132662.111.
31. Robles, A.I.; Arai, E.; Mathé, E.A.; Okayama, H.; Schetter, A.J.; Brown, D.; Petersen, D.; Bowman, E.D.; Noro, R.; Welsh, J.A.; et al. An Integrated Prognostic Classifier for Stage I Lung Adenocarcinoma Based on mRNA, microRNA, and DNA Methylation Biomarkers. *J. Thorac. Oncol.* **2015**, *10*, 1037–1048. doi:10.1097/JTO.0000000000000560.
32. Girard, L.; Rodriguez-Canales, J.; Behrens, C.; Thompson, D.M.; Botros, I.W.; Tang, H.; Xie, Y.; Rekhtman, N.; Travis, W.D.; Wistuba, I.I.; et al. An Expression Signature as an Aid to the Histologic Classification of Non–Small Cell Lung Cancer. *Clin. Cancer Res.* **2016**, *22*, 4880. doi:10.1158/1078-0432.CCR-15-2900.
33. Mezheyeuski, A.; Bergsland, C.H.; Backman, M.; Djureinovic, D.; Sjöblom, T.; Bruun, J.; Micke, P. Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.* **2018**, *244*, 421–431. doi:10.1002/path.5026.
34. Sun, Z.; Wang, L.; Eckloff, B.W.; Deng, B.; Wang, Y.; Wampfler, J.A.; Jang, J.; Wieben, E.D.; Jen, J.; You, M.; et al. Conserved recurrent gene mutations correlate with pathway deregulation and clinical outcomes of lung adenocarcinoma in never-smokers. *BMC Med. Genomics* **2014**, *7*, 486. doi:10.1186/1755-8794-7-32.
35. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, T.P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Med.* **2009**, *6*, e1000097. doi:10.1371/journal.pmed.1000097.
36. Sun, S.; Schiller, J.H.; Gazdar, A.F. Lung cancer in never smokers — a different disease. *Nat. Rev. Cancer* **2007**, *7*, 778–790. doi:10.1038/nrc2190.
37. Kim, A.M.; Tingen, C.M.; Woodruff, T.K. Sex bias in trials and treatment must end. *Nature* **2010**, *465*, 688–689. doi:10.1038/465688a.
38. Woitowich, N.C.; Beery, A.; Woodruff, T. A 10-year follow-up study of sex inclusion in the biological sciences. *eLife* **2020**, *9*, e56344. doi:10.7554/eLife.56344.
39. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, Ij.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. doi:10.1038/sdata.2016.18.
40. Ye, Y.; Jing, Y.; Li, L.; Mills, G.B.; Diao, L.; Liu, H.; Han, L. Sex-associated molecular differences for cancer immunotherapy. *Nat. Commun.* **2020**, *11*, 1779. doi:10.1038/s41467-020-15679-x.
41. Abbas, A.K.; Trotta, E.; R. Simeonov, D.; Marson, A.; Bluestone, J.A. Revisiting IL-2: Biology and therapeutic prospects. *Sci. Immunol.* **2018**, *3*, eaat1482. doi:10.1126/sciimmunol.aat1482.
42. Gao, Y.; Lu, J.; Zeng, C.; Yang, J.; Huang, B.; Zhang, N.; Li, L.; Fu, X. IL-10 suppresses IFN-γ-mediated signaling in lung adenocarcinoma. *Clin. Exp. Med.* **2020**, *20*, 449–459. doi:10.1007/s10238-020-00626-3.
43. Vahl, J.M.; Friedrich, J.; Mittler, S.; Trump, S.; Heim, L.; Kachler, K.; Balabko, L.; Fuhrich, N.; Geppert, C.-I.; Trufa, D.I.; et al. Interleukin-10-regulated tumour tolerance in non-small cell lung cancer. *Br. J. Cancer* **2017**, *117*, 1644–1655. doi:10.1038/bjc.2017.336.
44. Miotto, D.; Cascio, N.L.; Stendardo, M.; Querzoli, P.; Pedriali, M.; De Rosa, E.; Fabbri, L.M.; Mapp, C.E.; Boschetto, P. CD8+ T cells expressing IL-10 are associated with a favourable prognosis in lung cancer. *Lung Cancer* **2010**, *69*, 355–360. doi:10.1016/j.lungcan.2009.12.012.
45. Emmerich, J.; Mumm, J.B.; Chan, I.H.; LaFace, D.; Truong, H.; McClanahan, T.; Gorman, D.M.; Oft, M. IL-10 Directly Activates and Expands Tumor-Resident CD8^{+} T Cells without De Novo Infiltration from Secondary Lymphoid Organs. *Cancer Res.* **2012**, *72*, 3570. doi:10.1158/0008-5472.CAN-12-0721.
46. Shang, G.-S.; Liu, L.; Qin, Y.-W. IL-6 and TNF-α promote metastasis of lung cancer by inducing epithelial-mesenchymal transition. *Oncol. Lett.* **2017**, *13*, 4657–4660. doi:10.3892/ol.2017.6048.
47. Silva, E.M.; Mariano, V.S.; Pastrez, P.R.A.; Pinto, M.C.; Castro, A.G.; Syrjanen, K.J.; Longatto-Filho, A. High systemic IL-6 is associated with worse prognosis in patients with non-small cell lung cancer. *PLOS*

ONE **2017**, *12*, e0181125, doi:10.1371/journal.pone.0181125.
48. Lewis, A.M.; Varghese, S.; Xu, H.; Alexander, H.R. Interleukin-1 and cancer progression: the emerging role of interleukin-1 receptor antagonist as a novel therapeutic agent in cancer treatment. *J. Transl. Med.* **2006**, *4*, 48, doi:10.1186/1479-5876-4-48.
49. Gottschlich, A.; Endres, S.; Kobold, S. Can we use interleukin-1β blockade for lung cancer treatment? *Transl. Lung Cancer Res.* **2018**, *7*, S160–S164.
50. Wu, F.; Xu, J.; Huang, Q.; Han, J.; Duan, L.; Fan, J.; Lv, Z.; Guo, M.; Hu, G.; Chen, L.; et al. The Role of Interleukin-17 in Lung Cancer. *Mediators Inflamm.* **2016**, *2016*, 8494079, doi:10.1155/2016/8494079.
51. Conforti, F.; Pala, L.; Bagnardi, V.; De Pas, T.; Martinetti, M.; Viale, G.; Gelber, R.D.; Goldhirsch, A. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. *Lancet Oncol.* **2018**, *19*, 737–746, doi:10.1016/S1470-2045(18)30261-4.
52. Wang, C.; Qiao, W.; Jiang, Y.; Zhu, M.; Shao, J.; Ren, P.; Liu, D.; Li, W. Effect of sex on the efficacy of patients receiving immune checkpoint inhibitors in advanced non-small cell lung cancer. *Cancer Med.* **2019**, *8*, 4023–4031, doi:10.1002/cam4.2280.
53. Conforti, F.; Pala, L.; Bagnardi, V.; Viale, G.; De Pas, T.; Pagan, E.; Pennacchioli, E.; Cocorocchio, E.; Ferrucci, P.F.; De Marinis, F.; et al. Sex-Based Heterogeneity in Response to Lung Cancer Immunotherapy: A Systematic Review and Meta-Analysis. *JNCI J. Natl. Cancer Inst.* **2019**, *111*, 772–781, doi:10.1093/jnci/djz094.
54. Vijayan, D.; Smyth, M.J.; Teng, M.W.L. Purinergic Receptors: Novel Targets for Cancer Immunotherapy. In *Oncoimmunology: A Practical Guide for Cancer Immunotherapy*; Zitvogel, L., Kroemer, G., Eds.; Springer International Publishing: Cham, 2018; pp. 115–141 ISBN 978-3-319-62431-0.
55. Crain, J.M.; Nikodemova, M.; Watters, J.J. Expression of P2 nucleotide receptors varies with age and sex in murine brain microglia. *J. Neuroinflammation* **2009**, *6*, 24, doi:10.1186/1742-2094-6-24.
56. R Core Team *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014;
57. Athar, A.; Füllgrabe, A.; George, N.; Iqbal, H.; Huerta, L.; Ali, A.; Snow, C.; Fonseca, N.A.; Petryszak, R.; Papatheodorou, I.; et al. ArrayExpress update - from bulk to single-cell expression data. *Nucleic Acids Res.* **2019**, *47*, D711–D715, doi:10.1093/nar/gky964.
58. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140, doi:10.1093/bioinformatics/btp616.
59. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **2010**, *11*, R25, doi:10.1186/gb-2010-11-3-r25.
60. Maglott, D.; Ostell, J.; Pruitt, K.D.; Tatusova, T. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.* **2005**, *33*, D54–D58, doi:10.1093/nar/gki031.
61. Durinck, S.; Moreau, Y.; Kasprzyk, A.; Davis, S.; De Moor, B.; Brazma, A.; Huber, W. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **2005**, *21*, 3439–3440, doi:10.1093/bioinformatics/bti525.
62. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47–e47, doi:10.1093/nar/gkv007.
63. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300.
64. Montaner, D.; Dopazo, J. Multidimensional Gene Set Analysis of Genomic Data. *PLOS ONE* **2010**, *5*, e10348, doi:10.1371/journal.pone.0010348.
65. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30, doi:10.1093/nar/28.1.27.
66. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29, doi:10.1038/75556.
67. Lex, A.; Gehlenborg, N.; Strobelt, H.; Vuillemot, R.; Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 1983–1992, doi:10.1109/TVCG.2014.2346248.
68. García-García, F. Methods of functional enrichment analysis in genomic studies, Universitat de València: València, 2016.
69. Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. *J. Stat. Softw. Vol 1 Issue 3 2010* **2010**.
70. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **1986**, *7*, 177–188, doi:10.1016/0197-2456(86)90046-2.
71. Sterne, J.A.C.; Egger, M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *J. Clin. Epidemiol.* **2001**, *54*, 1046–1055, doi:10.1016/S0895-4356(01)00377-8.

Supplementary Figure S1. Information regarding sex distribution among reviewed studies.

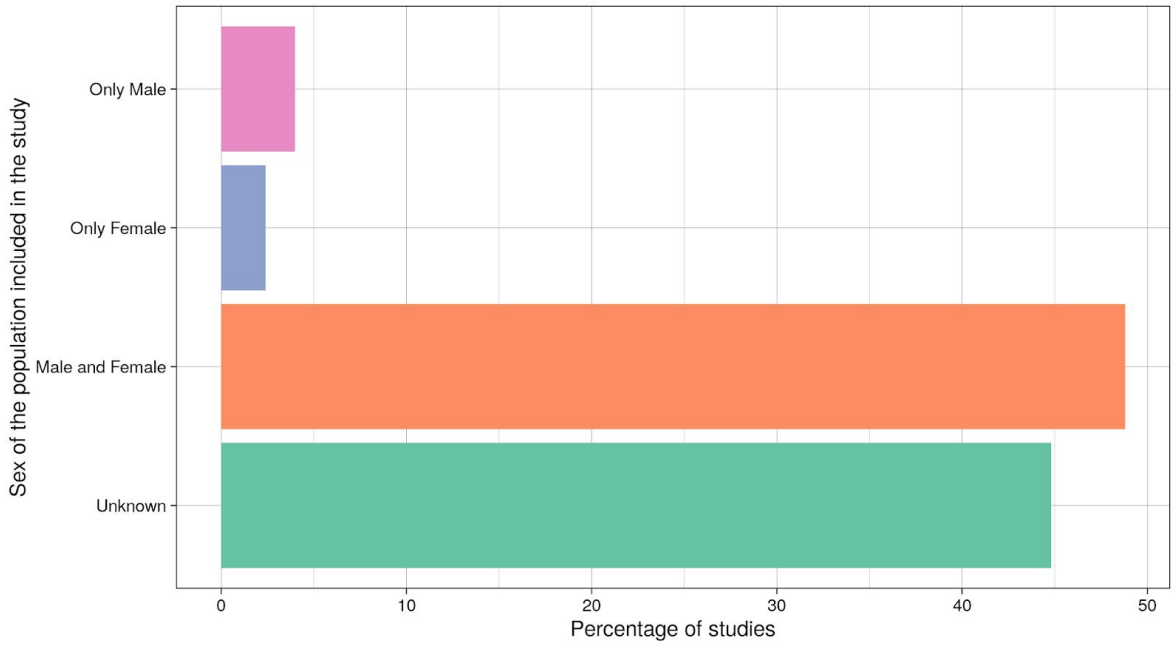


Table S2. Summary of differential expression analysis results. Two exploratory differential expression analyses were performed (ADC Women - Control Women, ADC.W -Control.W; ADC Men - Control Men, ADC.M - Control.M), together with the contrast of interest: (ADC.W - Control.W) - (ADC.M - Control.M). When performing the contrast of interest, "Up" terms are overrepresented in female lung adenocarcinoma patients, while "Down" terms are overrepresented in male lung adenocarcinoma patients.

Study		(ADC.W-ControlW)-(ADC.M-Control.M)	ADC.W-ControlW	ADC.M-Control.M
GSE10072	Up	0	1199	3182
	Down	0	1296	2688
GSE19188	Up	0	2348	5828
	Down	0	2111	3830
GSE31210	Up	0	3310	3560
	Down	0	2370	2543
GSE32863	Up	6	5243	2458
	Down	1	4507	2172
GSE63459	Up	0	2409	1561
	Down	0	2064	1611
GSE75037	Up	1	5779	4115
	Down	1	5117	3680
GSE81089	Up	3	2654	3564
	Down	0	3383	4262
GSE87340	Up	1	4887	1875
	Down	3	4958	1841
TCGA	Up	1	5861	5397
	Down	0	5569	5269

Table S3. Genes differentially expressed between male and female lung adenocarcinoma patients

ENTREZ ID	Gene Name	Up / Down	logFC	adj.pval	Study
9086	Eukaryotic translation initiation factor 1A Y-linked	Up	0.606 1.514	0.014 7.38*10-7	GSE32863 GSE75037
146330	F-box and leucine rich repeat protein 16	Down	1.247 2.157	0.028 0.003	GSE32863 GSE75037
3394	interferon regulatory factor 8	Up	1.177	0.028	GSE32863
80301	pleckstrin homology domain containing O2	Up	0.986	0.028	GSE32863
3689	integrin subunit beta 2	Up	1.393	0.043	GSE32863
11309	solute carrier organic anion transporter family member 2B1	Up	1.123	0.049	GSE32863
83706	fermitin family member 3	Up	0.866	0.049	GSE32863
252948	testis-specific transcript, Y-linked 16	Up	1.655 2.983	1.98*10-7 3.3*10-14	GSE81089 TCGA
107987337	ZFY antisense RNA 1	Up	1.609	0.001	GSE81089
6736	sex determining region Y	Up	1.519	0.007	GSE81089
694	BTG anti-proliferation factor 1	Up	0.859	0.007	GSE87340
64582	G protein-coupled receptor 135	Down	2.213	0.007	GSE87340
22979	EFR3 homolog B	Down	1.799	0.03	GSE87340
54753	zinc finger protein 853	Down	1.581	0.04	GSE87340

Table S5. Software and versions used in this study.

Software / R package	Version
R	3.5.3
AnnotationDbi	1.44.0
Biobase	2.42.0
biomaRt	2.38.0
edgeR	3.24.3
GEOQuery	2.50.5
ggdendro	0.1-20
ggpubr	0.2
hgu133plus2.db	3.2.3
illuminaHumanv3.db	1.26.0
KEGG.db	3.2.3
limma	3.38.3
mdgsa	1.14.0
metafor	2.1-0
methods	3.5.3
org.Hs.eg.db	3.7.0
reshape	0.8.8
stats	3.5.3
SummarizedExperiment	1.12.0
TCGAbiolinks	2.10.5
tidyverse	1.2.1
UpSetR	1.3.3
utils	3.5.3

The **Metafun-NSCLC** web tool is freely available at: https://bioinfo.cipf.es/metafun-nsclc

20 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).