A novel role for the tumor suppressor gene *ITF2* in lung tumorigenesis through the action of Wnt signaling pathway.

Olga Pernía[1,2]*, Ana Sastre-Perona[3]*, Carlos Rodriguez-Antolín[1,2], Alvaro García-Guede[1,2], María Palomares-Bralo[4,5], Rocío Rosas[1,2], Darío Sanchez-Cabrero[2], Carmen Rodriguez[1,2], Rubén Martín-Arenas[4,5], Verónica Pulido[1,2], Javier de Castro[2], Pilar Santisteban[3,6], Olga Vera[1,2]+, Inmaculada Ibáñez de Cáceres[1,2]+

[1]Laboratorio de Epigenétic, INGEMM, Hospital La PAZ. Madrid.
[2]Terapias Experimentales y Biomarcadores en Cáncer. N33. IdiPAZ. Madrid.
[3]Instituto de Investigaciones Biomedicas CSIC/UAM. Madrid.
[4]Ciber de Enfermedades Raras (CIBERER). Instituto de Salud Carlos III (ISCIII), Madrid.
[5]Laboratorio de Genómica Estructural y Funcional, INGEMM. IdiPAZ. Madrid
[6] Ciber de Cáncer (CIBERONC) Instituto de Salud Carlos III (ISCIII), Madrid.

*** Both authors contributed equally to this work.**
+ Corresponding authors

Corresponding authors:
Inmaculada Ibañez de Caceres E-mail: inma.ibanezca@salud.madrid.org
Olga Vera Puente E-mail: olga.verapuente@moffitt.org
Cancer Epigenetics Laboratory, INGEMM
Biomarkers and Experimental Therapeutics in Cancer, IdiPAZ
Paseo de la Castellana 261, 28046 Madrid, Spain
Phone +34-91-2071010-248
Fax +34-91-2071010

One-sentence summary: Our study identifies Wnt-signaling pathway as a molecular mechanism behind the development of platinum-resistance through ITF2 action; as well as to present two potential therapeutic targets for further studies, ITF2 and HOXD9.

Abbreviations used in the manuscript: HULP: Hospital Universitario La Paz; TCF4: Transcription Factor 4; CDDP: cisplatin; qRT-PCR: quantitative real-time PCR; qMSP: quantitative methylation-specific PCR; NSCLC: non-small cell lung cancer; ATT: adjacent tumor tissues;

Total number of Figures and Tables: 6 Figures, 1 Tables; 5 Supplementary Figures and 5 Supplementary Tables.

World count: 5230

Declaration of interest:

All the authors have read the journal's authorship statement and declare no potential conflicts of interest.

SUMMARY

Despite often leading to a platinum resistance, platinum-based chemotherapy continues to be the standard treatment for non-small cell lung cancer (NSCLC) and ovarian cancer. In this study, we used array-CGH and qRT-PCR methodologies to identify and validate cytogenetic alterations that arise after cisplatin treatment in four paired cisplatin-sensitive/resistant cell lines. We found the presence of a common deletion of the gene *ITF2* in cisplatin-resistant cell lines. Our translational approach included the expression analysis in a total of 35 lung and ovarian primary tumors, showing a frequent downregulation of *ITF2*. Whole transcriptome sequencing (RNA-seq) and Wnt-pathway analysis in a subgroup of NSCLC patients identified four genes with a potential role in the development of this malignancy. Further functional assays overexpressing *ITF2* in NSCLC cisplatin-resistant cells suggest the regulation of *HOXD9* by the Wnt pathway and its implication on NSCLC progression.

Keywords: TCF4, *ITF2*, Wnt pathway, cisplatin-resistance, Lung cancer.

Abreviations: ATT, Adjacent Tumor Tissues; CGH, Comparative genomic hybridization; NSCLC, Non-Small Cell Lung Cancer; NLM, Normal lung mean; TCF4, transcription factor 4; HOXD9, RIOX1, Ribosomal Oxygenase 1.

INTRODUCTION

Non-small cell lung cancer (NSCLC) and ovarian cancer are two of the most deadly cancers plaguing our society: the former accounting for more than 80% of primary lung-cancer cases and the latter boasting the highest mortality of the gynecological malignancies worldwide (1). Although platinum-based chemotherapy still plays an important role in the treatment of these two tumor types, the disease progresses to a platinum-resistant state in a high percentage of the diagnosed cases (2,3). Cisplatin (CDDP) is a platinum compound that induces apoptosis in cancer cells (4,5) but leads to cytogenetic alterations, such as deletions or amplifications, which contribute to the development of CDDP-resistance (6,7). Cisplatin treatment also causes changes in DNA copy numbers, altering the expression of genes involved in the drug efflux (ABC transporters family) or the antiapoptotic pathway (BCL2) (8). Moreover cisplatin treatment induce complex aberrations, including deletions of genes involved in tumor progression, metastasis and drug resistance (9) as a consequence of its binding to the N7 position of the guanines and crosslinking DNA in cancer cells (4,5). In this study, by performing a high resolution million feature array-Comparative Genomic Hybridization (aCGH) with four NSCLC and ovarian cancer sensitive/resistant paired cell lines, previously reported by our group (10) we explored the chromosomic deletions that differ in the resistant subtypes. We were able to find a common deletion that includes the Transcription Factor 4, *TCF4* (hereafter called *ITF2*). *ITF2* belongs to the Class 1 of bHLH transcription factors involved in the neurogenesis (11). Chromosomic alterations and mutations in this gene are associated with intellectual disability (12-14).

Little is known about how *ITF2* regulates different cellular pathways as epithelial-mesenchymal transition (11), besides its role in different tumors, such as being regulated by *CXCR4* in breast cancer (15) or β-catenin in colon and ovarian cancers (16,17). Most importantly, *ITF2* is a downstream target gene of the Wnt/β-catenin/TCF pathway acting as a tumor suppressor, negatively regulating its activity (16,17).

Therefore, further study of the role of *ITF2* and the Wnt signaling pathway in tumor response to chemotherapy is the key to keep finding new ways to fight resistance to this widespread treatment. Here we report the downregulation of *ITF2* in NSCLC patients to be frequently involved in cancer development by affecting *RIOX1* through the Wnt/B-catenin pathway as well as to present *HOXD9* as a potential therapeutic target.

MATERIALS AND METHODS

Cell culture and cell-viability assays

The NSCLC and ovarian cancer cell lines H23, H460, OVCAR3 and A2780 were purchased from the ATCC (Manassas, USA) and ECACC (Sigma-Aldrich, Spain) and cultured as recommended. Their CDDP-resistant variants H23R, H460R, OVCAR3R and A2780R were previously established in our laboratory (10,18). Cisplatin (Farma Ferrer, Spain) was used for CDDP-viability assays. Cells were seeded in 24-well dishes at 40,000 cells/well, treated with increasing doses of CDDP (0, 0.5, 1, 1.5, 2 and 3μg/ml) for an additional 72 or 48 hours and stained as described (19). Cell viability comparing sensitive vs. resistant cell lines was estimated relative to the density recorded over the same experimental group without drug exposure at same period of time. Cell authentication is included in Supplementary Table 1.

Clinical sample and data collection

We selected a representative number of fresh frozen surgical specimens from University Hospital La Paz-Hospital (HULP) Biobank, totaling 25 NSCLC and 10 ovarian cancer samples, belonging to previously reported cohorts of patients (10). Three normal ovarian samples obtained from sex reassignment surgery patients and two lung tissues from non-neoplastic origin were used as negative controls. Follow-up was conducted according to the criteria of the medical oncology divisions from each institution. The samples were processed following the standard operating procedures with the appropriate approval of the Human Research Ethics Committees at each contributing center, including informed consent within the context of research. Clinical, pathological and therapeutic data were recorded by an independent observer, and a blind statistical analysis was performed on these data.

DNA extraction and array of Comparative Genome Hybridyzation

DNA from cell lines was isolated as previously described (20) and was used to analyze copy number variations by the Array-CGH SurePrint G3 Human CGH Microarray 1x1M (Agilent, USA). Array experiments were performed as recommended by the manufacturer (Agilent, USA). Briefly, DNA (500 ng) from the resistant and the sensitive (used as reference) subtypes of each cell line were double-digested with RsaI and AluI for 2h at 37ºC. After heat inactivation of the enzymes at 65ºC for 20 min, each digested sample was labeled by random priming (Genomic DNA Enzymatic Labeling Kit Agilent) for 2 h using Cy5-dUTP for patient DNAs and Cy3-dUTP for reference DNAs. Labeled products were column purified (Microcon Ym-30 filters, Millipore Corporation). After probe denaturation and pre-annealing with Cot-1 DNA, hybridization was performed at 65ºC with rotation for 24 h. After two washing steps, the microarray was then scanned by the Agilent Microarray Scanner and analyzed with Feature Extraction software (v9.1 Agilent Technologies) and Agilent Cytogenomic 3.0.1.1. The analysis and visualization of the microarray data was performed using Agilent Cytogenomic 3.0.1.1. Comprehensive description of the statistical algorithms is available in the user's manual provided by Agilent Technologies. The Aberration Detection Method 2 (ADM-2) quality weighted interval score algorithm identifies aberrant intervals in samples that have consistently high or low log ratios based on their statistical score. The score represents the deviation of the weighted average of the normalized log ratios from its expected value of zero calculated with Derivative Log2 Ratio Standard Deviation algorithm. A Fuzzy Zero algorithm is applied to incorporate quality information about each probe measurement. Our threshold settings for the CGH analytics software to make a positive call were 6.0 for sensitivity, 0.45 for minimum absolute average log ratio per region, and 5 consecutive probes with the same polarity were required for the minimum number of probes per region.

RNA extraction, RT-PCR, qRT-PCR

Total RNA from human cancer cell lines and surgical samples was isolated, reverse transcribed and quantitative RT-PCR analysis was performed as previously described (10). Briefly, 500 ng of total RNA was used for RT reaction using PrimeScript™ RT Master (Clontech-Takara, USA). For RT-PCR, 2 µl of the RT product (diluted 1:5) was used for semi-quantitative PCR or qPCR reactions with Promega PCR Mix (Promega, USA) and SYBR Green PCR Mix (Applied Biosystems, USA), respectively. RT-PCR was performed under the following conditions: (a) 1 cycle of 95ºC for 2 min; (b) 30–40 cycles of 95ºC for 1 min, 56°C–60ºC for 1 min, 72ºC for 1 min; (c) an extension of 5 min at 72ºC. RT-PCR products were run on a 1.5% agarose gel, using the 100 bp DNA ladder as a molecular weight marker (New England Biolabs) for appropriate identification of band size. qRT-PCR absolute quantification was calculated according to the $2^{-\Delta Ct}$ method using GAPDH as endogenous control, whereas relative quantification was calculated with the $2^{-\Delta\Delta Ct}$ using *GAPDH* as endogenous control and the sensitive-parental cell line as a calibrator. Samples were analyzed in triplicate using the HT7900 Real-Time PCR system (Applied Biosystems, USA). Primers and probes for qRT-PCR expression analysis were purchased from Applied Biosystems (*TCF4*: Hs00162613_m1; *LRP1B*: Hs01069120_m1; *DKK1*: Hs00183740_m1; *GADPH*: Hs03929097_g1). Primers for RT-PCR of *Cyclin D1*, *DKK1*, *LEF*, *TCF3* genes were kindly donated by Pilar Santisteban Laboratory. Primers for *HOXD9*, *CLDN6*, *XIST* and *RIOX1* were designed, when possible, to analyze the specific transcript that significantly showed changes in the RNA-seq; all primers and specific amplification conditions are listed in Supplementary Table 2.

NGS (RNA-seq) and Wnt signaling pathways analysis

Total RNA from nine tumor tissues, three Non-tumor tissues from NSCLC patients and two normal lung samples from healthy donors were sent to Sistemas Genómicos Company (Valencia, Spain) for RNA-sequencing. 100-200 ng of RNA were analyzed by Bioanalyzer 2100 and Qubit 2.0 to determine the quality and quantity of the RNA.

mRNA poli (A)+ fraction was isolated from total RNA to obtain the cDNA libraries, following Illumina recommendations. Briefly, mRNA poli (A)+ was isolated using magnetic beads bound to oligo-poly T following subsequent chemical fragmentation prior to retrotranscription. cDNA fragments were subjected to a final repair process through the addition of a unique Adenine base to the 3' end, followed by adaptors ligation. Finally, products were purified and enriched by PCR to generate the double strand cDNA indexed library. Libraries quality was analyzed in bioanalyzer 2100, High Sensitivity Assay. Quantity of libraries was determined by Real Time PCR in LightCycler 480 (Roche). Libraries were sequenced by end-paired method (100 x 2) in the Illumina HiSeq2500 sequencer.

The bioinformatic analysis performed in the University Hospital la Paz, was based on the FastQC flow-work, to ensure the quality of the reads, and TRIMMOMATIC to eliminate the remnants from the adaptors and discard those sequences shorter than 750 bp. Reads were analyzed to quantify genes and isoforms through the RSEM v1.2.3 methodology (RNA-seq by Expectation Maximization) (21) and using the hg19 versions as reference for annotation. The differential expression was carried out with edgeR, which can estimate the common and individual dispersion (CMN and TGW, respectively) to obtain the variability of the data (22). p-values and FDR statistical analysis were performed by Cmn and Twg models and the statistical cut-off point was set as FDR < 0.05. Normalization was performed by the TMM method (Trimmed mean of M-values) (23). The bioinformatics analysis also included an efficiency analysis for every sample, considering the total efficiency as the percentage of reads annotated belonging to a transcript regarding the total fragments initially read. The Principal Component Analysis (PCA) was used to determine if the NT samples were too different from the lung control samples and no differences were observed, so these 5 samples were considered as controls. We analyzed three different contrasts including 1) comparison between Tumor samples with upregulated *DKK1* versus Tumor samples

with downregulated *DKK1* expression, 2) Tumor samples with upregulated *DKK1* versus control samples; 3) All tumor samples versus all control samples. The bioinformatics analysis included a study of all the genes annotated in the literature that have shown an unconfirmed relation with the Wnt-signaling pathway.

Transfection assays: top-fop, B-cat and TCF4 Overexpression

A Myc-DDK-tagged ORF clone of *TCF4* and the negative control pCMV6 were used for in transient transfection (RC224345; OriGene, USA) using the previously described methodology (10). Briefly, H23 and A2780 cells were plated onto 60-mm dishes at 600,000 cells/dish and transfected with a negative control or *TCF4* vectors, using jet-PEI DNA Transfection Reagent (PolyPlus Transfection, USA). For the Wnt reporter assay A2780S/R cells were plated at a concentration of 600,000 cells/well in 6-MW plates. Cells were serum starved overnight and co-transfected with 0.2 µg of either Super8xTopFlash (2 sets of 3 copies, the second set in the reverse orientation) of the TCF binding site) or Super8xFopFlash (2 full and one incomplete copy of the TCF binding site mutated followed by 3 copies in the reverse orientation) expression plasmids, kindly donated by Pilar Santisteban Laboratory, and 0.1µg pRL-TK (Renilla-TK-luciferase vector, Promega) as a control, using Lipofectamine 2000. Cells were subsequently treated with LiCl 10mM or transfected with S33Y B catenin for 48 hours prior to luciferase activities being measured using a Glomax 96 Microplate Luminometer (Turner Biosystems Instrument, Sunnyvale, USA). After 48 hours of transfection cells were cultured and collected for quantification of Luciferase and Renilla activity. Firefly luciferase activity was calculated as light units normalized with the Renilla activity generated by the pRL-CMV vector. B-catenin/TCF specific activity was calculated by obtaining the ratio of the activities of the Top/Fop promoters and is expressed in relative terms as the number of times of activation of the levels of untreated cells (= 1).

Statistical analysis

The data were compared using the chi-squared test or Fisher's exact test for qualitative variables, and Student's t test or the Wilcoxon-Mann-Whitney test (non-normal distribution) for quantitative variables. Correlation of quantitative variables was analyzed by Pearson's test. For the translational analysis, overall survival was estimated according to the Kaplan-Meier method and compared between groups by means of the Log Rank test. All the p-values were two-sided, and the type I error was set at 5 percent. Statistical analyses were performed using SPSS 20 software.

RESULTS

ITF2 is frequently downregulated in NSCLC and ovarian cancer cell lines

To identify new cytogenetic alterations after cisplatin administration, we performed Comparative Genomic Hybridization arrays with the DNA of the four paired cell lines from NSCLC, H23S/R and H460S/R, and ovarian cancer, A2780S/R and OVCAR3S/R. The cytogenetic study showed different genomic alterations in the resistant subtypes, using the sensitive cell lines as reference genome. In the NSCLC cell lines, H23 and H460, there was a common deleted region on the q22.1 part of chromosome 2 that included the gene *LRP1B.* The resistant ovarian cancer cell lines showed common deletions in the q22.33 location of chromosome 9 and the region of chromosome 18 bands q21.2-q21.31. The deletion of chromosome 9 included part of the gene *TMOD1* and the deletion in chromosome 18 affected the genes *RAB27B, CCDC68, TCF4, TXNL1, WDR7* and *BOD1P* (Supplementary Table 3). We also found two deleted regions shared by both resistant tumor types, in at least three of the four cell lines (H23, A2780 and OVCAR3) located in chromosome 18: q21.2q-21.31 and q21.32, this one including *ZNF532*, *SEC11C*, *GRP* genes (Figure 1A and Supplementary Table 3).

We selected those genes that were completely deleted in the same tumor type or in at least three of the four cell lines, focusing on *LRP1B* for NSCLC and *TCF4 (ITF2)* in H23, A2780 and OVCAR3 and quantify the expression levels of *LRPB1* and *ITF2* to validate the results obtained in the arrays CGH by qRT-PCR. The deletion of *ITF2* in H23R and A2780R cell lines resulted in a significant loss of *ITF2* expression compared to the sensitive subtypes. No changes were observed in OVCAR3 cells (Figure 1B), while the deletion of *LRPB1* in the resistant subtype of H23 and H460 did not change at the mRNA expression level (Supplementary Figure 1). These results confirm that *ITF2* is frequently downregulated in lung and ovarian CDDP-resistant subtypes.

Wnt canonical signaling pathway is increased in CDDP resistant cells with ITF2 deletions

Due to the association of *ITF2* with the Wnt/B-catenin/TCF pathway, we studied the involvement of Wnt pathway in cisplatin resistance in H23S/R and A2780S/R cells. We transfected H23S/R and A2780S/R cells with the B-catenin/Tcf4 reporter vectors Super8xTopFlash or Super8xFopFlash and induced the B-catenin activity either by LiCl treatment, which inhibits GSK3-b, or by co-transfection with the constitutively stable B-catenin-S33Y mutant. We observed higher luciferase activity in A2780R cells compared with the parental sensitive ones, indicating an increased transcriptional activity of B-catenin/TCF in response to both pharmacological and functional activation of the pathway (Figure 2A). We did not observed this result in H23 cells (Figure 2B). Expression changes of several downstream effectors of this pathway were analyzed in sensitive and resistant cells. We observed an increase in two of these effectors, *DKK1* and *Cyclin D1* expression in A2780R cell and a slightly increment of *Cyclin D1* in H23R cells (Figure 2C).

Restoration of ITF2 increases the sensitivity to CDDP by decreasing β-catenin/TCF transcriptional activity

To test the role of *ITF2* in cisplatin resistance, we transiently overexpressed *ITF2* cDNA in the resistant subtypes of A2780 cell line, in which our previous results confirmed the ability to evaluate changes in the transcriptional activity of the Wnt pathway. Overexpression of ITF2 in A2780R resulted in a significant increase in sensitivity to cisplatin (Figure 3A), showing an intermediate phenotype between the parental resistant and sensitive subtypes. The overexpression of *ITF2* in A2780R induced a dramatic decrease in cell viability ($p<0.05$) 24 hours after transfection compared with the parental resistant cells transfected with the empty vector (Figure 3B). *ITF2* overexpression at 24 and 72 hours after transfection was confirmed by qRT-PCR

(Figure 3C). *ITF2* overexpression decreased B-catenin/TCF transcriptional activity in A2780R, recovering the levels of the parental sensitive subtype (Figure 3D). Next, we analyzed the expression of β-catenin/TCF4 downstream targets *Cyclin D1* and *DKK1* after *ITF2* overexpression. RT-PCR analysis showed that *ITF2* overexpression in A2780R (R-*ITF2*) restored in part the expression of only *DKK1* at 24 and 72 hours after transfection (Figure 3E).

The expression of ITF2 is frequently downregulated in NSCLC and ovarian tumor samples

To validate our *in vitro* results and determine the possible clinical implication of *ITF2* and *DKK1* expression in NSCLC and ovarian cancer patients, we measured the relative expression of both genes in two cohorts of fresh frozen tumor samples (T) belonging to NSCLC (Table 1) and ovarian cancer patients (Supplementary Table 4) from HULP. We also obtained Adjacent Tumor Tissues (ATT) from some of the lung tumor samples.

We observed that *ITF2* expression is frequently downregulated in NSCLC and ovarian tumor samples (Figure 4), validating our *in vitro* data. We did not observe differences between ATT and normal lung samples (LC) for *ITF2* ($p = 0.177$) and *DKK1* ($p = 0.693$) in the NSCLC cohort (Figure 4A). In addition, we found that 15 out of 25 tumor samples of NSCLC patients had lower expression of *ITF2* compared to the normal lungs mean (NLM) (Figure 4A). Furthermore, 60% of NSCLC tumor samples showed the opposite expression profile between *ITF2* and *DKK1* as observed in our experimental data. However, this situation was found only in approximately 10% (1 out of 9) of the tumors on the ovarian cancer cohort of patients (Figure 4B).

The Kaplan Meier curves analyzing the patients overall survival according to the median of *ITF2* and *DKK1* expressions showed that NSCLC patients with high expression of *ITF2* tend to have better overall survival, $p=0.1$ (Figure 4C). No

differences in survival were observed for *DKK1* expression p=0.6 (Supplementary Figure 2A). After such promising results, we also analyzed the expression of *ITF2* and *DKK1* in the Kaplan Meier plotter online tool (24) and we observed that those patients with high expression of *ITF2* (Supplementary Figure 2B) and low expression of *DKK1* (Supplementary Figure 2C) had a significantly better overall survival rate ($p = 0.016$ and $p < 0.001$, respectively).

Identification of candidate genes involved in the Wnt signaling pathway through the analysis of RNA-seq in NSCLC patients

To further explore the role of the Wnt-signaling pathway in the lung cancer tumorigenesis, we performed RNA-sequencing on 14 samples; nine were NSCLC samples, six of them with an opposite expression profile between *ITF2* and *DKK1* (Pat3, Pat6, Pat10, Pat22, Pat25 and Pat26) and three of them with the same expression profile (Pat8, Pat16 and Pat18). In addition, we included three ATT (Pat9, Pat21 and Pat25) and two normal lung samples (LC1 and LC2) that together were considered as controls for comparisons. *ITF2* is mainly downregulated in NSCLC patients, while the expression levels of *DKK1* showed a more heterogeneous pattern, therefore we considered *DKK1* as the best parameter to decide on the bioinformatics analysis of the RNA-seq. This analysis focused on three main contrasts: Contrast A, the differential gene expression analysis between tumors and controls. Contrast B, comparison between tumors with high and low expression of *DKK1*. And finally, contrast C was used to compare the tumor samples with high expression of *DKK1* with the controls (Supplementary Figure 3). The bioinformatics study also included an analysis of all the genes that have been related to the Wnt-pathway in the literature. We selected those genes that showed a significant expression differences (FDR<0.05) in at least two of the three contrasts, prioritizing contrast B. Thus, we identified three

coding genes (*HOXD9*, *RIOX1* and *CLDN6*) and one long-non coding RNA (*XIST*) for further validations ((Supplementary Figure 3 and Supplementary Table 5).

Validation of the four candidate genes was initially performed by semi-qPCR (Supplementary Figure 4). *HOXD9* expression showed a heterogeneous pattern among the samples analyzed, with almost no representation in normal samples. *XIST* expression decreased in eight out of nine tumors, compared with the controls. Regarding *CLDN6* and *RIOX1*, we observed higher band intensity in the tumor samples compared with the non-tumor ones. We further confirm these results by qRT-PCR, validating the results from the RNA-sequencing (Figure 5, left panel). *HOXD9* was significantly upregulated in 44% (four out of nine) of the tumors compared to the controls and correlated with the RNA-sequencing data accurately (r = 0.83) (Figure 5A, left panel). This correlation was even higher for *CLDN6* and *XIST* genes (r = 0.97 and r = 0.97, respectively) (Figure 5C and 5D, left panel). Conversely, *RIOX1* was downregulated in the tumors compare to the controls although the correlation coefficient was less marked (r = 0.58) (Figure 5B, left panel). We next correlated the expression of these candidates with the expression of *ITF2* in the 14 samples analyzed, and we observed a negative correlation of *HOXD9* (Pearson = - 0.24) and a positive correlation of *RIOX1* (Pearson = 0.68) with *ITF2* (Figure 5A and 5B, right panel). No correlation was found for *XIST* and *CLDN6* (Pearson = -0.14 and 0.0, respectively) (Figure 5C and 5D, right panel).

In order to gain insight into the role of *ITF2* regulating the expression of the selected candidates in NSCLC, we overexpressed *ITF2* in H23 cell line. Validation of the transfection efficiency was done by qRT-PCR at 24 and 72h after transfection (Figure 6A). Changes in the expression of the candidate genes were determined by semi-qPCR and qRT-PCR. We observed that overexpression of *ITF2* induced a decrease in the expression of *HOXD9* at 72h after transfection (p = 0.04) and, as expected, increased the expression of *RIOX1* also at 72h (p = 0.08) (Supplementary Figure 5 and

Figure 6B). No changes were observed for *CLDN6* ($p = 0.1$) or *XIST* ($p = 0.59$) expression after *ITF2* overexpression, but we found significant differences in the expression of these genes between the sensitive and the resistant subtypes ($p < 0.01$). Having identified *HOXD9* and *RIOX1* as possible candidate genes of the Wnt pathway regulated by *ITF2*, we studied their translational application in terms of treatment response in the cohort of NSCLC patients used for the RNA-seq analysis. We observed a negative correlation between *HOXD9* and the overall survival of the patients (Figure 6C, top panel) and a positive correlation between *RIOX1* and same clinical feature (Figure 6D, top panel). The survival analysis performed by the Kaplan Meier method and stratifying the patients according to the median of the expression of the genes, showed that patients with lower expression of *HOXD9* present a significant better overall survival rate ($p = 0.046$) (Figure 6C, bottom panel) whereas patients with low *RIOX1* expression tend to live less, although no statistical significance was found in this case (Figure 6D, bottom panel).

DISCUSSION

The search for new biomarkers for cancer progression and response to treatment in cancer has been a wide open field for research in the past years. In this work, we studied the involvement of the Wnt signaling pathway in tumorigenesis through a combined experimental approach by using both comparative genome hybridization arrays and RNA-sequencing, while previous genomic alteration studies on platinum resistance in different types of cancer cell lines, have used only Comparative Genomic Hybridization array methodology (7,8,25,26). We explored the deleted genes after cisplatin administration, as they could have a role in the response to the drug. We identified a common deletion in chromosome 18 involving the region q21.2-q21.31, in H23R, OVCAR3R and A2780R cells. Previous studies analyzing the cytogenetic profile of A2780S/R cells showed different profile probably due to the lower specificity of the CGH-array used to perform those experiments (7,8). From the two genes selected for further validation, *LRP1B* deletion was shown in H23 and H460 cells, although we could not find changes in expression by using an alternative technique. This is not the first time that the deletion of *LRP1B* is not validated by gene expression assays (27). In our case, it could be due to the mosaicism observed in this region that occurs in less than 20% of the resistant cells. The level of mosaicism that can be detected is dependent on the sensitivity and spatial resolution of the clones and rearrangements present (28). Nevertheless, *LRP1B* could still play a role in tumor progression as several studies link its downregulation through deletion and the development of esophageal, renal and NSCL cancers (29-31). The downregulation of the second gene, *TCF4* was confirmed by qRT-PCR assays in H23R and A2780R cells, whereas no changes were observed in OVCAR3R, again probably due in part to the low level of mosaicism (36% cells harboring the deletion). Array CGH drawbacks are its limited ability to detect mosaicism and its inability to detect aberrations that do not result in copy number changes. Our results indicate that low levels of mosaicisms would make

the validations of expression changes by another quantitative technique difficult, probably because the alteration at expression levels are not significant enough to be detected. The fact that *ITF2* is deleted and downregulated after platinum treatment, provides us with a new insight regarding its importance in resistance to chemotherapy. *ITF2* is a transcription factor belonging to the basic Helix Loop Helix (bHLH) family, also referred to as an E-protein due to its ability to bind to the Ephrussi-box in DNA. It can act as a transcriptional activator or repressor (32,33) and its regulation still is unknown. It is important to note that *ITF2*, whose real name is *TCF4*, is not the same as the T-Cell Factor 4 (*TCF7L2*), also known as *TCF4,* which is a downstream transcription factor of the Wnt/B-catenin/TCF pathway (34). In fact, *ITF2* is transcribed by the B-catenin/TCF axis and acts as a molecular repressor of the transcription mediated by this complex through interfering the binding of B-catenin with *TCF4*, leading to the repression of cell proliferation (17). Consistent with these studies, we have observed that the resistant A2780 cells have an increased activity of the B-catenin/TCF transcription, probably due to the absence of *ITF2*. In contrast, we did not observe differences in H23 cells, which could be explained by previous observations that H23 has a higher basal activity of the Wnt signaling pathway and no exogenous activation could imply a difference (35). Moreover, we have observed that the overexpression of *ITF2* in A2780R cells leads to a decrease in cell viability, rescuing the sensitive phenotype, maybe through the inhibition of the excessive proliferation and the activity levels of the B-catenin/TCF transcription. In fact, our results indicate that resistant cells respond better to the activation of the Wnt pathway and that Wnt signaling pathway plays an important role in the resistance to cisplatin through *ITF2* in cancer cell lines.

The translational analysis, based on the expression levels of *ITF2* and *DKK1*, a downstream effector of Wnt pathway in two different cohorts of patients was aimed to elucidate the role of this pathway in tumor progression and chemotherapy response.

ITF2 expression was frequently downregulated in NSCLC and ovarian tumor samples, validating our *in vitro* data. The expression levels of *DKK1*, however, showed a more heterogeneous pattern in the NSCLC tumor samples, while no differences were observed in the ovarian tumors. These results suggest an aberrant activation of the Wnt signaling pathway in lung cancer. In fact, our *in silico* analysis of 1926 NSCLC patients indicates a significant increased overall survival when patients present high expression levels of *ITF2* and low expression of *DKK1.* We observed the same findings from our own data although without statistical significance, probably due to the sample size. However, one of the strengthens of our cohort is that is comprised by fresh frozen samples, allowing to perform high quality RNA-sequencing in a group of NSCLC patients, in order to determine the involvement of the Wnt-pathway components in lung cancer development. The differential expression of *DKK1* within the tumor samples allowed us to performed three different bioinformatics contrasts in order to explore all the possibilities focused on tumor development and Wnt signaling pathway. Contrast A aimed to identify genes with a possible involvement in lung cancer development by comparing differential expression in tumors versus control samples. Contrast B was made to identify alterations in the Wnt pathway in NSCLC tumors and those which could be used as potential therapeutic targets. Finally, contrast C was able to identify genes regulated by the Wnt pathway and others involved in NSCLC development. Using this approach with RNA-sequencing allowed us to identify new coding, non-coding genes and transcripts that have not been functionally characterized previously (36,37). Indeed, in this study we have identified three coding genes, *HOXD9*, *CLDN6* and *RIOX1*, and one non-coding gene, *XIST*, which can be potentially involved in NSCLC progression through the Wnt signaling pathway. Our data was validated by two alternative methodologies, finding a strong positive correlation among these technologies for three of the candidates *HOXD9*, *CLDN6* and *XIST* and slightly weaker correlation for *RIOX1*.

HOXD9 and *CLDN6* were significantly downregulated in tumors with high expression levels of *DKK1* and upregulated in tumors compared with controls, indicating that these genes could be involved in tumor progression through an aberrant activation of the Wnt signaling pathway. For *HOXD9*, the tumor samples had higher expression levels than the controls, as previously reported (38). We also observed a negative correlation of *HOXD9* and *ITF2* expression levels. In addition, patients with lower levels of *HOXD9* had better overall survival than those with upregulated expression of this gene. These results are consistent with previous studies linking a high expression of *HOXD9* with glioblastoma and hepatocarcinoma (39,40). Our functional analysis showed that *ITF2* overexpression in H23R cells decreased the expression of *HOXD9*. However, these results do not indicate a direct regulation of *HOXD9* by *ITF2*, as we expected a recovery of its expression to the sensitive subtype levels. We believe, that an alternative regulatory mechanism affected by *ITF2* is modulating the expression of *HOXD9*. It has been shown that the expression of *HOXD9* is also regulated by epigenetic mechanisms, such as the long-non coding RNA *HOTAIR* (41,42), which has been linked with cisplatin resistance (43). Although *CLDND6* has been previously related with carcinogenesis of different tumor types (44,45), we did not find any correlation with *ITF2* expression levels. In addition, the overexpression of *ITF2* in H23R cells did not change the expression of this gene, altogether suggesting that *CLDN6* is not involved in NSCLC development or Wnt signaling pathway.

The coding gene *RIOX1* and the non-coding RNA *XIST*, showed expression inhibition in tumor samples with high levels of *DKK1* and controls. Two different bioinformatic contrasts imply both genes in tumor progression through their regulation by the Wnt signaling pathway. Although an upregulation of *XIST*, which is involved in X-chromosome inactivation (46), has been shown to be involved in tumor development through the Wnt pathway (47), our results do not show a clear correlation with *ITF2* expression or a relationship with clinical features of NSCLC patients. Conversely,

RIOX1 showed increased expression levels in the tumor samples and a positive correlation of its expression with *ITF2*. In addition, patients with higher levels of *RIOX1* tended to have better overall survival than the patients with lower expression levels. Moreover, overexpression of *ITF2* in the resistant subtype of H23 cell line increased the expression level of *RIOX1* up to the sensitive subtype. Although little is known about *RIOX1*, it has recently been shown to be involved in renal and colorectal carcinogenesis (48,49).

In summary, we have identified that *ITF2* is frequently downregulated in cisplatin-resistant cancer cells as well as in NSCLC and ovarian cancer patients. Moreover, our identification of the Wnt-signaling pathway as a molecular mechanism behind the development of cisplatin resistance in cancer cells through the action of *ITF2* provides novel insights into the molecular biology of the cellular mechanisms involved in the acquired resistance to the most widely use chemotherapy agent, as well as to present two potential therapeutic targets for further studies, *ITF2* and *HOXD9*.

ACKNOWLEDGEMENTS

The authors thank Hayley Pickett for the English language correction. The authors also acknowledge Biobank from HULP for sample processing. This study was supported by the 'Fondo de Investigación Sanitaria-Instituto de Salud Carlos III' [PI15/00186 and PI18/00050 to I.I.C.]; and the European Regional Development Fund/European Social Fund FIS [FEDER/FSE, Una Manera de Hacer Europa]. MINECO funds support C.R.A. and O.P. contracts through RTC-2016-5314-1 project.

AUTHOR CONTRIBUTIONS

IIC: Conception and design

OP, ASP, AGG, MPB, CR, RMA, VP and OVP: development of methodology

OP, ASP, CRA, AGG, RR, DSC, CR, JdC and IIC: acquisition of data

OP, ASP, AGG, RR, DSC, CR, PS, JdC, OVP and IIC: analysis and interpretation of data

CRA and MPB: bioinformatic analysis and interpretation of bioinformatic data

OVP: wrote the manuscript

All authors reviewed and/or revised the manuscript.

FIGURE LEYENDS

Figure 1. Identification of a common deletion in chromosome 18 in cisplatin-resistant cancer cell lines from NSCLC and ovarian cancer. (A) Picture extracted from the Agilent Cytogenomics 3.0.1.1 software showing the deletion of TCF4 in chromosome 18 in H23R, A2780R and OVCAR3R cell lines. (B) Relative mRNA expression levels of TCF4/ITF-2 measured by qRT-PCR. The results show the mean fold induction compare to the sensitive cells. Gene expression was normalized to GAPDH. S: sensitive; R: resistant; Data represent the relative expression levels obtained from the combination of two independent experiments measured in triplicate ± SD. *** $p < 0.001$; (Students T-test).

Figure 2. Basal status of the Wnt signaling pathway in A2780 and H23 cell lines. (A) Pharmacological activation (left) and functional activation (right) of β-catenin transcriptional activity in A2780 cells. (B) Pharmacological activation of β-catenin transcriptional activity in H23 cells. (A and B) β-catenin transcriptional activity was measure in A2780 and H23 cells after treatment with LiCl (10mM) 24 hours or transfection of bcat-S33Y, transfecting with Super8xTopFlash (Top) or Super8xFopFlash (Fop). The results show the fold induction of the Top/Fop ratio with respect to untreated cells (=1). Values represent the mean of three independent experiments measured by triplicate ± SD. *** $p < 0.001$; ** $p < 0.01$ (Students T-test). (C) Expression analysis of downstream genes regulated by *ITF2* involved in Wnt signaling pathway in H23S and H23R and A2780S and A2780R. Representative images of *LEF1*, *Cyclin D1*, *TCF3*, *DKK1* and *GAPDH* measured by RT-PCR. Each assay was performed at least three times to confirm the results

Figure 3. Effect of *ITF2* on cisplatin resistance, cell viability and Wnt pathway. (A) Viability curves of A2780 cell lines transfected with pCMV6 (S-Ø and R-Ø) and with the overexpression vector (R-*ITF2*). Each experimental group was exposed to 6 different CDDP concentrations for 48 h, and data were normalized to each untreated control, set to 100%. The data represent the mean ± SD of at least three independent experiments performed in quadruplicate at each drug concentration for each cell line analyzed. The CDDP-RI (Resistant Index to CDDP) was calculated as "IC50 from the R-Ø / IC50 from the S-Ø" and "IC50 from the S-transfected with *ITF2* / IC50 from the S-Ø" ± SD. (B) Viability of A2780 cell lines transfected with pCMV6 (R-Ø) and with the overexpression vectors (R-*ITF2*). (C) Relative expression levels of *ITF2* measured by quantitative RT-PCR, in the cell line A2780, represented in Log10 scale; In each experimental group, the sensitive cell line transfected with pCMV6 plasmid was used as a calibrator. Each bar represents the combined relative expression of two independent experiments measured in triplicate. (D) β-catenin transcriptional activity was measured in A2780 cells after *ITF2* overexpression and treatment with LiCl (10mM) treatment for 24 hours, transfecting with Super8xTopFlash (Top) or

Super8xFopFlash (Fop). The results show the fold induction of the Top/Fop ratio with respect to untreated cells (=1). Values represent the mean of three independent experiments measured by triplicate ± SD. (E) Expression analysis of the downstream genes *DKK1* and *Cyclin D1* regulated by *ITF2* in A2780 cell line transfected with pCMV6 (S-Ø and R-Ø) and with the overexpression vector (R-*ITF2*) for 24 and 72 hours. Left, representative images of *DKK1*, *Cyclin D1* and *GAPDH* measured by RT-PCR. Right, expression levels of *Cyclin D1* and *DKK1* measured by qPCR. Each assay was performed at least three times to confirm the results. *** $p < 0.001$; ** $p < 0.01$ (Students T-test).

Figure 4. Expression profile of *ITF2* and *DKK1* in patients with NSCLC and ovarian cancer. (A-B) Assessment of *ITF2* and *DKK1* expression levels measured by qRT-PCR in fresh samples from a cohort of 25 patients with NSCLC (A) and ovarian cancer (B). For all the analyses, data represents expression levels in $2^{-\Delta\Delta Ct}$ using the mean of normal lungs (NLM) or ovarian (NOM) as calibrator. (C) Analysis of mRNA expression of *ITF2* our cohort with NSCLC. Survival analysis in 25 NSCLC samples according to the mean of *ITF2* expression. LogRank, Breslow and Tarone-Ware test were used for comparisons and $p<.05$ was considered as a significant change in OS. NSCLC: non-small cell lung cancer; ATT: adjacent tumor tissue; T: tumor; LC1/LC2: Lung Control; NLM: Normal Lung Mean; NOM: Normal Ovarian Mean.

Figure 5. Expression levels of candidate target genes measured by RNA-seq and qRT-PCR and correlation with *ITF2* expression levels in a selection of 14 tumor and non-tumor samples from NSCLC patients. (A-D) Left panel, correlation between RNA-seq and qRT-PCR expression levels in *HOXD9* (A), *XIST* (B), *CLDN6* (C) and *RIOX1* (D) bars represent the relative expression of each gene measured by qRT-PCR in triplicate and represented as $2^{-\Delta Ct}$ and lines represent the count per million obtained from the RNA-seq analysis. Pearson coefficient was used for linear correlation of the quantitative variables; Right panel, correlation between *ITF2* expression and *HOXD9* (A), *XIST* (B), *CLDN6* (C) and *RIOX1* (D).

Figure 6. Effect of ITF2 overexpression over its candidate target genes in vitro and in nine NSCLC patients samples. (A) Validation of the transfection efficiency of ITF2 at mRNA levels. Relative expression levels of ITF2 measured by qRT-PCR, in the cell line H23, at 24 and 72 hour after transfection represented in Log10 the $2^{-\Delta\Delta Ct}$. (B) Relative expression levels of HOXD9, CLDN6, RIOX1 and XIST measured by quantitative RT-PCR after ITF2 overexpression. For both (A) and (B) the sensitive cell line transfected with pCMV6 plasmid was used as a calibrator (S-Ø). H23R cells were transfected with same negative control (R-Ø) or with ITF2 cDNA (R-ITF2). Each bar represents the combined relative expression of two independent experiments measured in triplicate. *** $p < 0.001$; ** $p < 0.01$ (Students T-test). (C and D) Analysis of ITF2 and its candidate target genes in a cohort of nine NSCLC patients. Top, correlation between ITF2 and HOXD9 (C) expression levels and ITF2 and RIOX1 (D) with the overall survival of the nine patients that were analyzed by RNA-seq. Data represents the quantitative expression levels of the three genes measured by qRT-PCR and represented as $2^{-\Delta\Delta Ct}$ for ITF2 (referred to the NLM) and $2^{-\Delta Ct}$ for HOXD9 and RIOX1. Bottom, survival analysis in NSCLC samples according to the mean of HOXD9 and RIOX1 expression levels. LogRank, Breslow and Tarone-Ware test were used for comparisons and $p < 0.05$ was considered as a significant change in OS.

Table 1. Clinicopathological and experimental data obtained from patients with NSCLC from La Paz University Hospital. Note: OS, Overall Survival; PFS, Progression Free Survival; CDDP: cisplatin; CBDCA: carboplatin; NA: not available.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians **2011**;61:69-90
2. Lee SY, Jung DK, Choi JE, Jin CC, Hong MJ, Do SK*, et al.* PD-L1 polymorphism can predict clinical outcomes of non-small cell lung cancer patients treated with first-line paclitaxel-cisplatin chemotherapy. Scientific reports **2016**;6:25952
3. French JD, Johnatty SE, Lu Y, Beesley J, Gao B, Kalimutho M*, et al.* Germline polymorphisms in an enhancer of PSIP1 are associated with progression-free survival in epithelial ovarian cancer. Oncotarget **2016**;7:6353-68
4. Haslehurst AM, Koti M, Dharsee M, Nuin P, Evans K, Geraci J*, et al.* EMT transcription factors snail and slug directly contribute to cisplatin resistance in ovarian cancer. BMC Cancer **2012**;12:91
5. Karaca B, Atmaca H, Bozkurt E, Kisim A, Uzunoglu S, Karabulut B*, et al.* Combination of AT-101/cisplatin overcomes chemoresistance by inducing apoptosis and modulating epigenetics in human ovarian cancer cells. Mol Biol Rep;40:3925-33
6. Hiorns LR, Seckl MJ, Paradinas F, Sharp SY, Skelton LA, Brunstrom G*, et al.* A molecular cytogenetic approach to studying platinum resistance. Journal of inorganic biochemistry **1999**;77:95-104
7. Leyland-Jones B, Kelland LR, Harrap KR, Hiorns LR. Genomic imbalances associated with acquired resistance to platinum analogues. The American journal of pathology **1999**;155:77-84
8. Yasui K, Mihara S, Zhao C, Okamoto H, Saito-Ohara F, Tomida A*, et al.* Alteration in copy numbers of genes as a mechanism for acquired drug resistance. Cancer Res **2004**;64:1403-10
9. Akervall J, Guo X, Qian CN, Schoumans J, Leeser B, Kort E*, et al.* Genetic and expression profiles of squamous cell carcinoma of the head and neck correlate with cisplatin sensitivity and resistance in cell lines and patients. Clin Cancer Res **2004**;10:8204-13
10. Vera O, Jimenez J, Pernia O, Rodriguez-Antolin C, Rodriguez C, Sanchez Cabo F*, et al.* DNA Methylation of miR-7 is a Mechanism Involved in Platinum Response through MAFG Overexpression in Cancer Cells. Theranostics **2017**;7:4118-34
11. Forrest MP, Waite AJ, Martin-Rendon E, Blake DJ. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. PLoS One **2013**;8:e73169
12. Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J*, et al.* Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). American journal of human genetics **2007**;80:994-1001
13. Brockschmidt A, Todt U, Ryu S, Hoischen A, Landwehr C, Birnbaum S*, et al.* Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. Hum Mol Genet **2007**;16:1488-94
14. Amiel J, Rio M, de Pontual L, Redon R, Malan V, Boddaert N*, et al.* Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. American journal of human genetics **2007**;80:988-93
15. Appaiah H, Bhat-Nakshatri P, Mehta R, Thorat M, Badve S, Nakshatri H. ITF2 is a target of CXCR4 in MDA-MB-231 breast cancer cells and is associated with reduced survival in estrogen receptor-negative breast cancer. Cancer biology & therapy **2010**;10:600-14

16. Kolligs FT, Nieman MT, Winer I, Hu G, Van Mater D, Feng Y*, et al.* ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with beta-catenin defects and promotes neoplastic transformation. Cancer cell **2002**;1:145-55
17. Shin HW, Choi H, So D, Kim YI, Cho K, Chung HJ*, et al.* ITF2 prevents activation of the beta-catenin-TCF4 complex in colon cancer cells and levels decrease with tumor progression. Gastroenterology **2014**;147:430-42 e8
18. Ibanez de Caceres I, Cortes-Sempere M, Moratilla C, Machado-Pinilla R, Rodriguez-Fanjul V, Manguan-Garcia C*, et al.* IGFBP-3 hypermethylation-derived deficiency mediates cisplatin resistance in non-small-cell lung cancer. Oncogene **2010**;29:1681-90
19. Chattopadhyay S, Machado-Pinilla R, Manguan-Garcia C, Belda-Iniesta C, Moratilla C, Cejas P*, et al.* MKP1/CL100 controls tumor growth and sensitivity to cisplatin in non-small-cell lung cancer. Oncogene **2006**;25:3335-45
20. Ibanez de Caceres I, Dulaimi E, Hoffman AM, Al-Saleem T, Uzzo RG, Cairns P. Identification of novel target genes by an epigenetic reactivation screen of renal cancer. Cancer Res **2006**;66:5021-8
21. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics **2011**;12:323
22. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **2010**;26:139-40
23. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology **2010**;11:R25
24. Gyorffy B, Surowiak P, Budczies J, Lanczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. PLoS One **2013**;8:e82241
25. Cui L, Fu J, Pang JC, Qiu ZK, Liu XM, Chen FR*, et al.* Overexpression of IL-7 enhances cisplatin resistance in glioma. Cancer biology & therapy **2012**;13:496-503
26. Piskareva O, Harvey H, Nolan J, Conlon R, Alcock L, Buckley P*, et al.* The development of cisplatin resistance in neuroblastoma is accompanied by epithelial to mesenchymal transition in vitro. Cancer letters **2015**;364:142-55
27. Kohno T, Otsuka A, Girard L, Sato M, Iwakawa R, Ogiwara H*, et al.* A catalog of genes homozygously deleted in human lung cancer and the candidacy of PTPRD as a tumor suppressor gene. Genes, chromosomes & cancer **2010**;49:342-52
28. Cho EK. Array-based Comparative Genomic Hybridization and Its Application to Cancer Genomes and Human Genetics. J Lung Cancer **2011**;10:77-86
29. Brown J, Bothma H, Veale R, Willem P. Genomic imbalances in esophageal carcinoma cell lines involve Wnt pathway genes. World journal of gastroenterology **2011**;17:2909-23
30. Ni S, Hu J, Duan Y, Shi S, Li R, Wu H*, et al.* Down expression of LRP1B promotes cell migration via RhoA/Cdc42 pathway and actin cytoskeleton remodeling in renal cell cancer. Cancer science **2013**;104:817-25
31. Liu CX, Musco S, Lisitsina NM, Forgacs E, Minna JD, Lisitsyn NA. LRP-DIT, a putative endocytic receptor gene, is frequently inactivated in non-small cell lung cancer cell lines. Cancer Res **2000**;60:1961-7
32. Murre C. Helix-loop-helix proteins and lymphocyte development. Nature immunology **2005**;6:1079-86
33. Sepp M, Kannike K, Eesmaa A, Urb M, Timmusk T. Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. PLoS One **2011**;6:e22138
34. Navarrete K, Pedroso I, De Jong S, Stefansson H, Steinberg S, Stefansson K*, et al.* TCF4 (e2-2; ITF2): a schizophrenia-associated gene with pleiotropic

effects on human disease. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics **2013**;162B:1-16

35. Akiri G, Cherian MM, Vijayakumar S, Liu G, Bafico A, Aaronson SA. Wnt pathway aberrations including autocrine Wnt activation occur at high frequency in human non-small-cell lung carcinoma. Oncogene **2009**;28:2163-72
36. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet **2009**;10:57-63
37. Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods **2010**;7:1009-15
38. Plowright L, Harrington KJ, Pandha HS, Morgan R. HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). Br J Cancer **2009**;100:470-5
39. Tabuse M, Ohta S, Ohashi Y, Fukaya R, Misawa A, Yoshida K*, et al.* Functional analysis of HOXD9 in human gliomas and glioma cancer stem cells. Mol Cancer **2011**;10:60
40. Lv X, Li L, Lv L, Qu X, Jin S, Li K*, et al.* HOXD9 promotes epithelial-mesenchymal transition and cancer metastasis by ZEB1 regulation in hepatocellular carcinoma. J Exp Clin Cancer Res **2015**;34:133
41. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA*, et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell **2007**;129:1311-23
42. Woo CJ, Kingston RE. HOTAIR lifts noncoding RNAs to new levels. Cell **2007**;129:1257-9
43. Liu Z, Sun M, Lu K, Liu J, Zhang M, Wu W*, et al.* The long noncoding RNA HOTAIR contributes to cisplatin resistance of human lung adenocarcinoma cells via downregualtion of p21(WAF1/CIP1) expression. PLoS One **2013**;8:e77293
44. Ikenouchi J, Matsuda M, Furuse M, Tsukita S. Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. J Cell Sci **2003**;116:1959-67
45. Lu Y, Wang L, Li H, Li Y, Ruan Y, Lin D*, et al.* SMAD2 Inactivation Inhibits CLDN6 Methylation to Suppress Migration and Invasion of Breast Cancer Cells. Int J Mol Sci **2017**;18
46. Brown SD. XIST and the mapping of the X chromosome inactivation centre. Bioessays **1991**;13:607-12
47. Sun N, Zhang G, Liu Y. Long non-coding RNA XIST sponges miR-34a to promotes colon cancer progression via Wnt/beta-catenin signaling pathway. Gene **2018**;665:141-8
48. Pires-Luís AS, Vieira-Coimbra M, Vieira FQ, Costa-Pinheiro P, Silva-Santos R, Dias PC*, et al.* Expression of histone methyltransferases as novel biomarkers for renal cell tumor diagnosis and prognostication. Epigenetics **2015**;10:1033-43
49. Nishizawa Y, Nishida N, Konno M, Kawamoto K, Asai A, Koseki J*, et al.* Clinical Significance of Histone Demethylase NO66 in Invasive Colorectal Cancer. Ann Surg Oncol **2017**;24:841-9

FIGURE LEYENDS

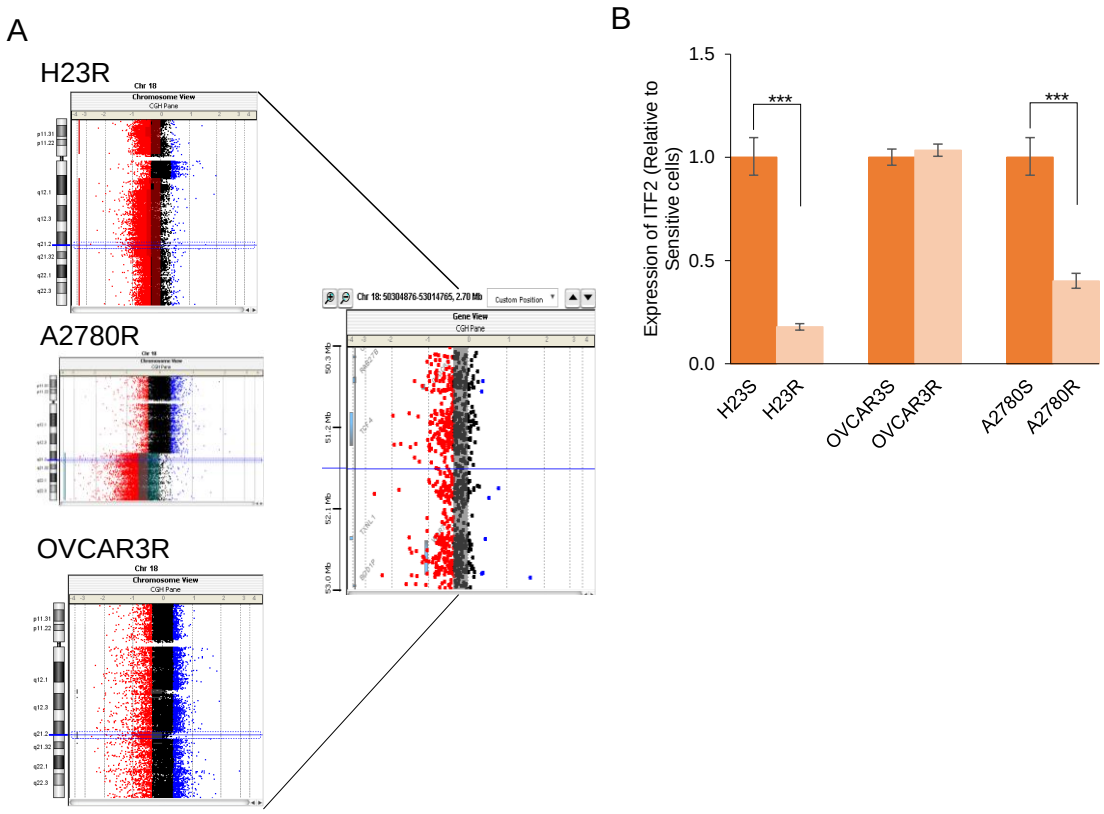

Figure 1. Identification of a common deletion in chromosome 18 in cisplatin-resistant cancer cell lines from NSCLC and ovarian cancer. (A) Picture extracted from the Agilent Cytogenomics 3.0.1.1 software showing the deletion of *TCF4* in chromosome 18 in H23R, A2780R and OVCAR3R cell lines. (B) Relative mRNA expression levels of *TCF4/ITF-2* measured by qRT-PCR. The results show the mean fold induction compare to the sensitive cells. Gene expression was normalized to *GAPDH*. S: sensitive; R: resistant; Data represent the relative expression levels obtained from the combination of two independent experiments measured in triplicate ± SD. *** $p < 0.001$; (Students T-test).

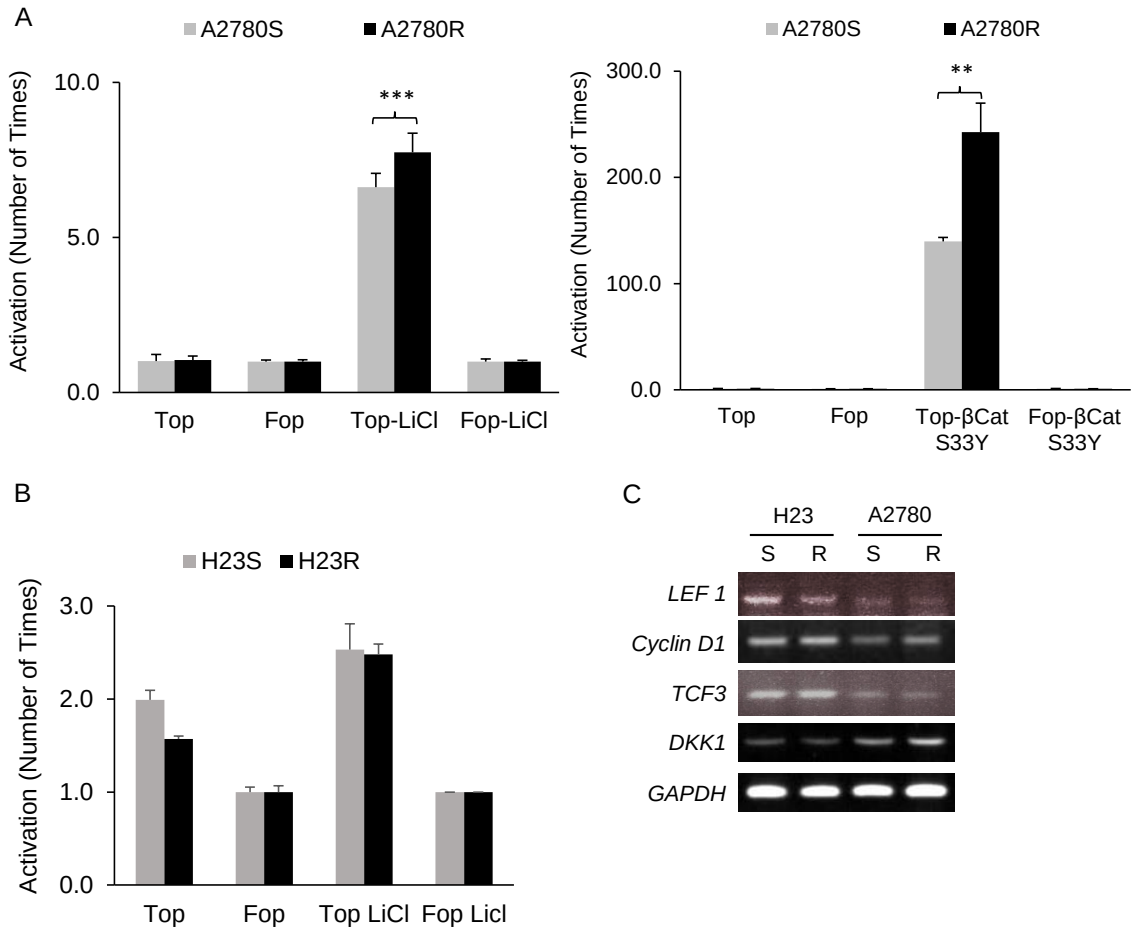

Figure 2. Basal status of the Wnt signaling pathway in A2780 and H23 cell lines. (A) Pharmacological activation (left) and functional activation (right) of β-catenin transcriptional activity in A2780 cells. (B) Pharmacological activation of β-catenin transcriptional activity in H23 cells. (A and B) β-catenin transcriptional activity was measure in A2780 and H23 cells after treatment with LiCl (10mM) 24 hours or transfection of bcat-S33Y, transfecting with Super8xTopFlash (Top) or Super8xFopFlash (Fop). The results show the fold induction of the Top/Fop ratio with respect to untreated cells (=1). Values represent the mean of three independent experiments measured by triplicate ± SD. *** $p < 0.001$; ** $p < 0.01$ (Students T-test). (C) Expression analysis of downstream genes regulated by *ITF2* involved in Wnt signaling pathway in H23S and H23R and A2780S and A2780R. Representative images of *LEF1*, *Cyclin D1*, *TCF3*, *DKK1* and *GAPDH* measured by RT-PCR. Each assay was performed at least three times to confirm the results

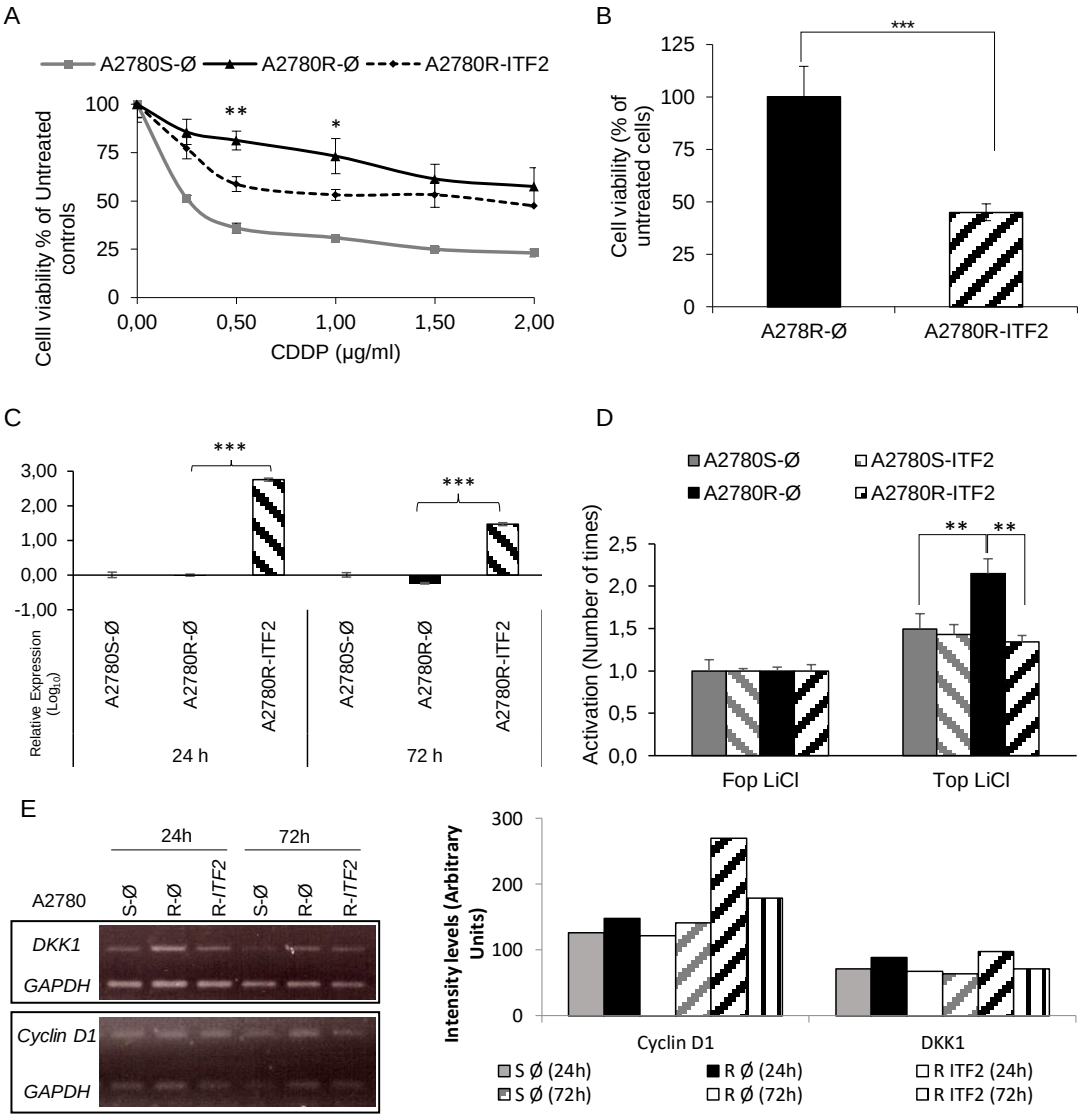

Figure 3. Effect of *ITF2* on cisplatin resistance, cell viability and Wnt pathway. (A) Viability curves of A2780 cell lines transfected with pCMV6 (S-Ø and R-Ø) and with the overexpression vector (R-*ITF2*). Each experimental group was exposed to 6 different CDDP concentrations for 48 h, and data were normalized to each untreated control, set to 100%. The data represent the mean ± SD of at least three independent experiments performed in quadruplicate at each drug concentration for each cell line analyzed. The CDDP-RI (Resistant Index to CDDP) was calculated as "IC50 from the R-Ø / IC50 from the S-Ø" and "IC50 from the S-transfected with *ITF2* / IC50 from the S-Ø" ± SD. (B) Viability of A2780 cell lines transfected with pCMV6 (R-Ø) and with the overexpression vectors (R-*ITF2*). (C) Relative expression levels of *ITF2* measured by quantitative RT-PCR, in the cell line A2780, represented in Log10 scale; In each experimental group, the sensitive cell line transfected with pCMV6 plasmid was used as a calibrator. Each bar represents the combined relative expression of two independent experiments measured in triplicate. (D) β-catenin transcriptional activity was measured in A2780 cells after *ITF2* overexpression and treatment with LiCl (10mM) treatment for 24 hours, transfecting with Super8xTopFlash (Top) or Super8xFopFlash (Fop). The results show the fold induction of the Top/Fop ratio with respect to untreated cells (=1). Values represent the mean of three independent experiments measured by triplicate ± SD. (E) Expression analysis of the downstream genes *DKK1* and *Cyclin D1* regulated by *ITF2* in A2780 cell line transfected with pCMV6 (S-Ø and R-Ø) and with the overexpression vector (R-*ITF2*) for 24 and 72 hours. Left, representative images of *DKK1*, *Cyclin D1* and *GAPDH* measured by RT-PCR. Right, expression levels of *Cyclin D1* and *DKK1* measured by qPCR. Each assay was performed at least three times to confirm the results. *** $p < 0.001$; ** $p < 0.01$ (Students T-test).

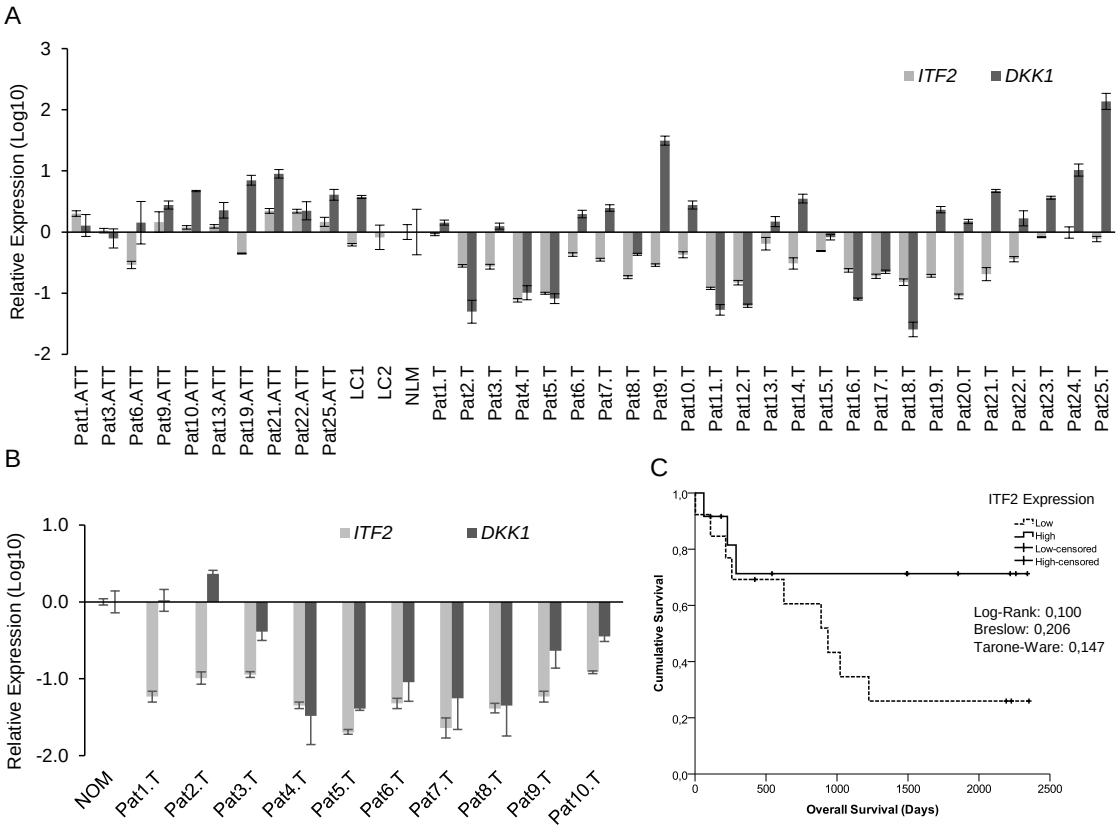

Figure 4. Expression profile of *ITF2* and *DKK1* in patients with NSCLC and ovarian cancer. (A-B) Assessment of *ITF2* and *DKK1* expression levels measured by qRT-PCR in fresh samples from a cohort of 25 patients with NSCLC (A) and ovarian cancer (B). For all the analyses, data represents expression levels in 2-ΔΔCt using the mean of normal lungs (NLM) or ovarian (NOM) as calibrator. (C) Analysis of mRNA expression of *ITF2* our cohort with NSCLC. Survival analysis in 25 NSCLC samples according to the mean of *ITF2* expression. LogRank, Breslow and Tarone-Ware test were used for comparisons and p<.05 was considered as a significant change in OS. NSCLC: non-small cell lung cancer; ATT: adjacent tumor tissue; T: tumor; LC1/LC2: Lung Control; NLM: Normal Lung Mean; NOM: Normal Ovarian Mean.

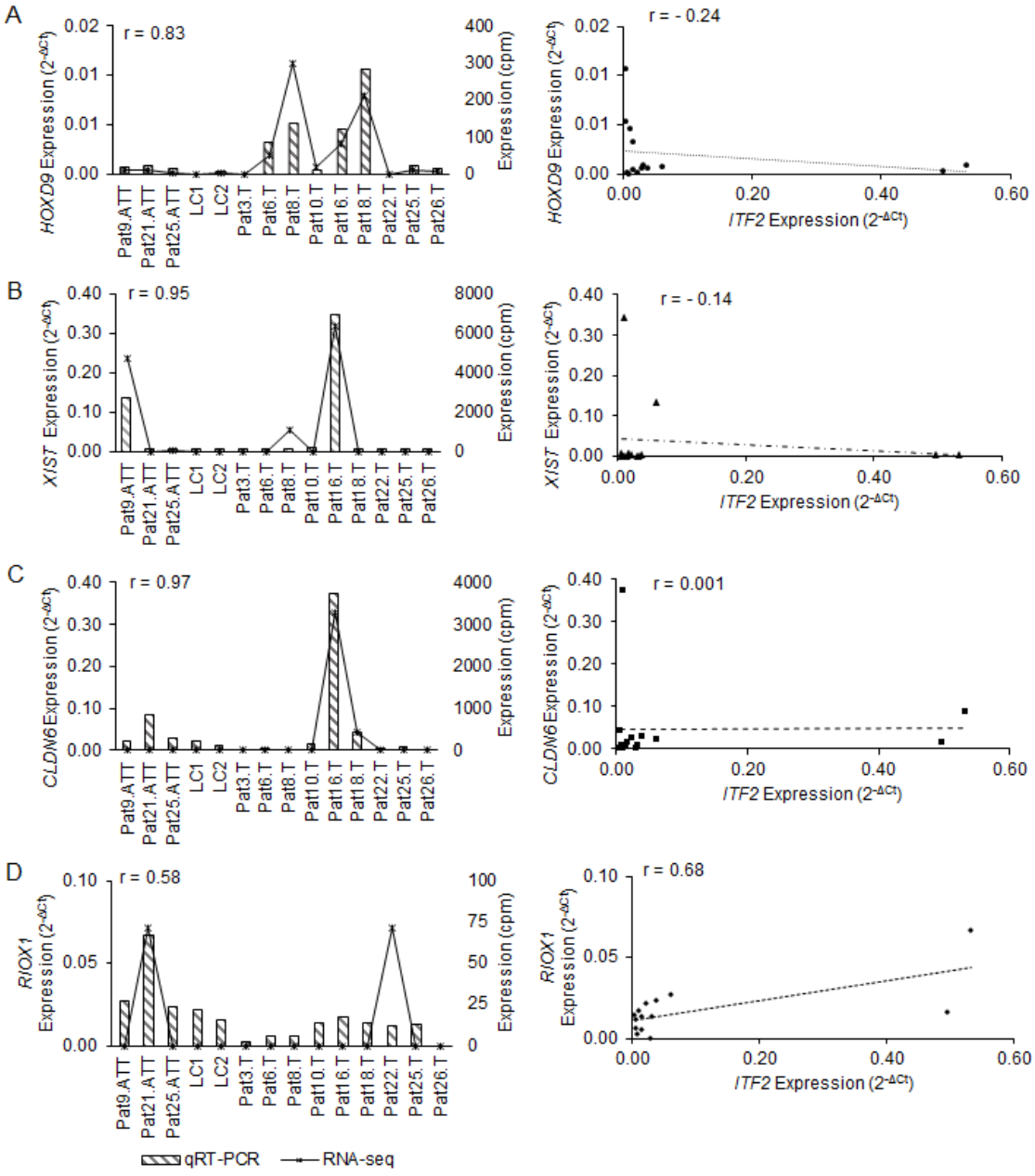

Figure 5. Expression levels of candidate target genes measured by RNA-seq and qRT-PCR and correlation with *ITF2* expression levels in a selection of 14 tumor and non-tumor samples from NSCLC patients. (A-D) Left panel, correlation between RNA-seq and qRT-PCR expression levels in *HOXD9* (A), *XIST* (B), *CLDN6* (C) and *RIOX1* (D) bars represent the relative expression of each gene measured by qRT-PCR in triplicate and represented as 2-ΔCt and lines represent the count per million obtained from the RNA-seq analysis. Pearson coefficient was used for linear correlation of the quantitative variables; Right panel, correlation between *ITF2* expression and *HOXD9* (A), *XIST* (B), *CLDN6* (C) and *RIOX1* (D).

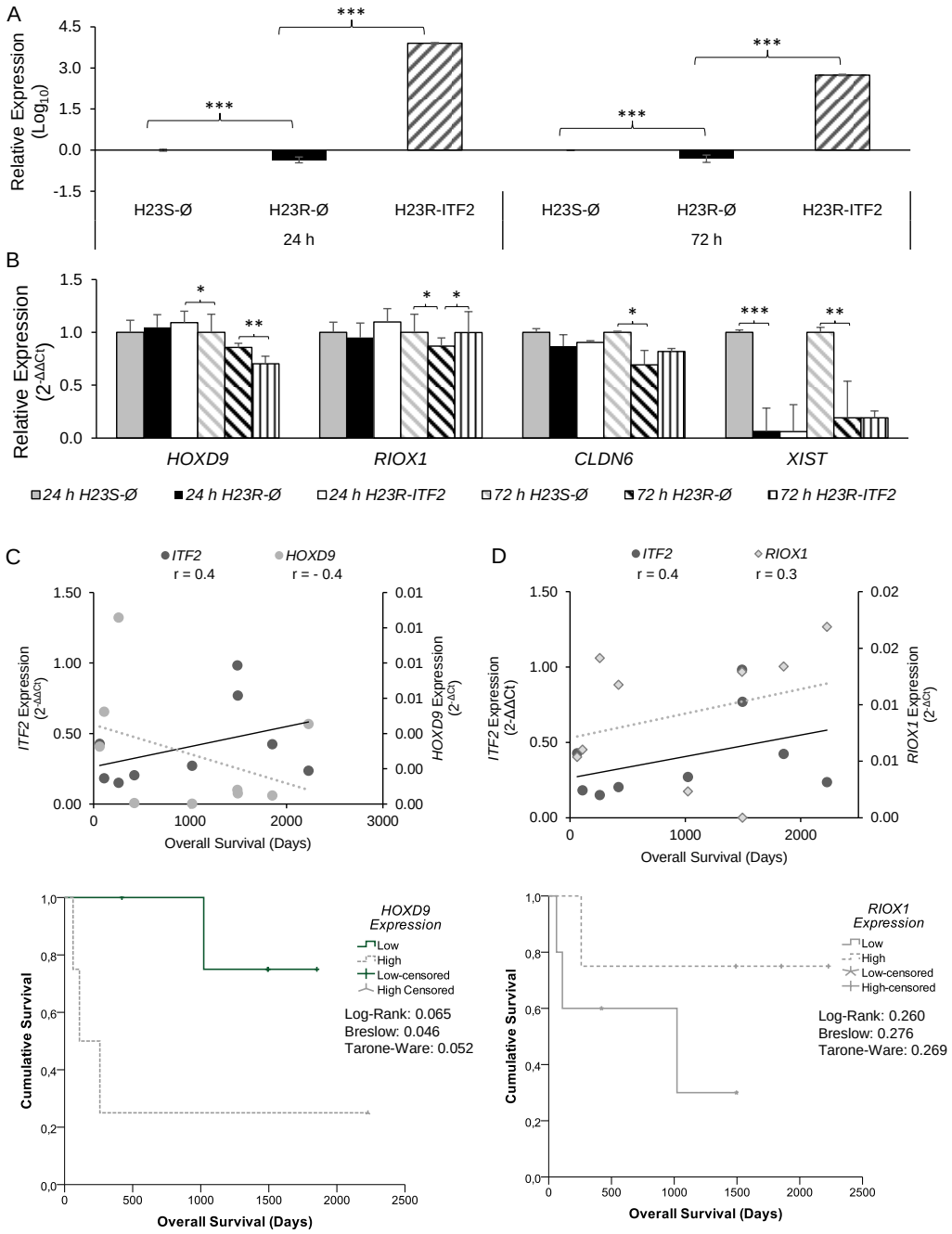

Figure 6. Effect of *ITF2* overexpression over its candidate target genes in vitro and in nine NSCLC patients samples. (A) Validation of the transfection efficiency of *ITF2* at mRNA levels. Relative expression levels of *ITF2* measured by qRT-PCR, in the cell line H23, at 24 and 72 hour after transfection represented in Log10 the 2-ΔΔCt. (B) Relative expression levels of *HOXD9*, *CLDN6*, *RIOX1* and *XIST* measured by quantitative RT-PCR after *ITF2* overexpression. For both (A) and (B) the sensitive cell line transfected with pCMV6 plasmid was used as a calibrator (S-Ø). H23R cells were transfected with same negative control (R-Ø) or with *ITF2* cDNA (R-*ITF2*). Each bar represents the combined relative expression of two independent experiments measured in triplicate. *** $p < 0.001$; ** $p < 0.01$ (Students T-test). (C and D) Analysis of *ITF2* and its candidate target genes in a cohort of nine NSCLC patients. Top, correlation between *ITF2* and *HOXD9* (C) expression levels and *ITF2* and *RIOX1* (D) with the overall survival of the nine patients that were analyzed by RNA-seq. Data represents the quantitative expression levels of the three genes measured by qRT-PCR and represented as 2-ΔΔCt for *ITF2* (referred to the NLM) and 2-ΔCt for *HOXD9* and *RIOX1*. Bottom, survival analysis in NSCLC samples according to the mean of *HOXD9* and *RIOX1* expression levels. LogRank, Breslow and Tarone-Ware test were used for comparisons and $p < 0.05$ was considered as a significant change in OS.

TABLES

Table 1. Clinicopathological and experimental data obtained from patients with NSCLC from La Paz University Hospital.

Patient	Histology	Sex	Stage	Chemotherapy	Relapse	Status	OS (days)	PFS (days)	*TCF4* ($2^{-\Delta\Delta Ct}$)	*DKK1* ($2^{-\Delta\Delta Ct}$)	*HOXD9* ($2^{-\Delta Ct}$)	*CLNDA6* ($2^{-\Delta Ct}$)	*XIST* ($2^{-\Delta Ct}$)	*RIOX1* ($2^{-\Delta Ct}$)
Pat1.T	Adenocarcinoma	Female	IA	No	Yes	Alive	2220	1490	0.91	1.43				
Pat2.T	Adenocarcinoma	Male	NA	No	Yes	Alive	2352	1860	0.28	0.05				
Pat3.T	Epidermoid	Male	IB	No	Yes	Exitus	1022	825	0.27	1.25	$2.94 \cdot 10^{-5}$	$1.18 \cdot 10^{-3}$	$2.71 \cdot 10^{-7}$	$2.33 \cdot 10^{-3}$
Pat4.T	Adenocarcinoma	Male	IB	No	No	Exitus	3	3	0.08	0.10				
Pat5.T	Adenocarcinoma	Male	NA	No	No	Exitus	626	626	0.10	0.08				
Pat6.T	Large cell	Male	IIB	No	No	Exitus	62	62	0.43	1.97	$3.27 \cdot 10^{-3}$	$7.43 \cdot 10^{-3}$	$3.92 \cdot 10^{-6}$	$5.39 \cdot 10^{-3}$
Pat7.T	Adenocarcinoma	Male	IIIA	Otro	Yes	Exitus	228	138	0.35	2.48				
Pat8.T	Epidermoid	Female	IIIB	CDDP + Others	No	Exitus	109	109	0.18	0.43	$5.24 \cdot 10^{-3}$	$3.63 \cdot 10^{-4}$	$4.16 \cdot 10^{-3}$	$6.03 \cdot 10^{-3}$
Pat9.T	Adenocarcinoma	Female	IIA	CDDP + Others	Yes	Alive	2260	2260	0.29	31.30				
Pat10.T	Epidermoid	Male	IB	No	No	Alive	1853	1853	0.42	2.78	$4.94 \cdot 10^{-4}$	$1.58 \cdot 10^{-2}$	$7.09 \cdot 10^{-3}$	$1.43 \cdot 10^{-2}$
Pat11.T	Adenocarcinoma	Male	IA	No	No	Exitus	216	216	0.12	0.05				
Pat12.T	Adenocarcinoma	Female	IIIA	CBDCA + Others	Yes	Alive	2192	2192	0.15	0.06				
Pat13.T	Epidermoid	Male	IB	CDDP + Others	No	Alive	2341	2341	0.65	1.48				
Pat14.T	Epidermoid	Male	IIA	No	NA	Exitus	289	289	0.31	3.53				
Pat15.T	Epidermoid	Male	IIA	No	NA	NA	109	109	0.49	0.83				
Pat16.T	Adenocarcinoma	Female	IIIA	CDDP + Others	NA	Alive	2228	2228	0.24	0.08	$4.55 \cdot 10^{-3}$	$3.73 \cdot 10^{-1}$	$3.45 \cdot 10^{-1}$	$1.69 \cdot 10^{-2}$
Pat17.T	Adenocarcinoma	Male	IIB	Otro	Yes	Exitus	888	443	0.19	0.22				
Pat18.T	Epidermoid	Male	IIB	CBDCA + Others	No	Exitus	259	259	0.15	0.03	$1.06 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	$2.67 \cdot 10^{-4}$	$1.41 \cdot 10^{-2}$

Note: OS, Overall Survival; PFS, Progression Free Survival; CDDP: cisplatin; CBDCA: carboplatin; NA: not available.