

Alma Mater Studiorum – Università di Bologna

**DOTTORATO DI RICERCA IN**

**SCIENZE FARMACOLOGICHE E TOSSICOLOGICHE, DELLO SVILUPPO E DEL MOVIMENTO UMANO**

Ciclo XXVII

**Settore Concorsuale di afferenza: 05/G1**

**Settore Scientifico disciplinare: BIO/14 Farmacologico**

TITOLO TESI

**Role of Non-coding RNAs in chemotherapeutic treatments**

**Presentata da: Dott. Ivan Vannini**

**Coordinatore Dottorato**

Prof. ssa Patrizia Hrelia

**Relatore**

Prof. Giorgio Cantelli Forti

**Correlatore**

Prof. Müller Fabbri

# 1. INTRODUCTION

## 2. INTRODUCTION

### 1.1 Lung cancer

Lung cancer is a malignant tumor characterized by uncontrolled cell growth in lung tissues. Most lung cancers are carcinomas that originate from epithelial cells.

The majority (80–90%) of cases of lung cancer are caused by long-term exposure to tobacco smoke. The remaining 10–15% of cases are often caused by a combination of genetic factors and exposure to asbestos, radon gas and second-hand smoke. Lung cancers are classified according to histological type. The classification is important to decide management and predicting outcomes of the tumor. The lung cancer is distinguished in non-small-cell lung carcinoma (NSCLC) and small-cell lung carcinoma (SCLC). NSCLC is often treated with surgery but the patients are frequently treated with chemotherapy both pre-operatively (neoadjuvant chemotherapy) and post-operatively (adjuvant chemotherapy). SCLC frequently responds better to chemotherapy and radiotherapy.

The common types of NSCLC are adenocarcinoma, squamous cell carcinoma and large cell carcinoma, but there are numerous other types that occur less commonly, and all types can occur in abnormal histologic variants and in mixed cell-type combination. Adenocarcinoma is at present the most frequent type of lung cancer in never smokers. It represents approximately 40% of lung cancers. Historically, adenocarcinoma was more often observed peripherally in the lungs than small cell lung cancer and squamous cell lung cancer, both of which tended to be more often centrally located. Squamous cell carcinoma of the lung is more frequent in men than in women and accounts for about 30% of lung cancers. It is strongly correlated with the tobacco smoking. Large cell lung carcinoma is a heterogeneous class of tumor originating from transformed epithelial cells in the lung. It represents typically 10% of all NSCLC. Large cell lung

carcinoma is differentiated from SCLC primarily by the larger size of the tumor cells and a higher cytoplasmic - nuclear size ratio.

SCLC presents under the microscope the small cancer cells and are mostly filled with the nucleus dense and neurosecretory granules. This tumor grows quickly and spreads more aggressively than NSCLC during the disease. Sixty to seventy percent have metastatic tumor at presentation. This lung cancer is strongly associated with smoking.

## **1.2 Non-coding RNAs**

Over the last years, despite advances in radiotherapy, chemotherapy and surgery the death rate from lung cancer has remained unchanged, which is mainly due to metastatic tumor. Therapeutic choices in NSCLC have been mainly based on, performance status, disease stage and co-morbidities, and rarely on molecular or on histological analysis. Actually this therapeutic approach to patients may result in survival improvement of only weeks to months. So it is important to discover new markers for targeted therapy.

Cancer is a genetic disease determined by dysregulation not only of genes that codes proteins but also of non-coding RNAs (ncRNAs). Among the latter group, microRNAs (miRNAs) are the most widely considered.

MiRNAs are small ncRNAs of 19-24 nucleotides in length that cooperate in regulation of mRNAs through base pairing to complementary site, mostly in the untranslated region (UTR) of the target mRNAs [1]. MiRNA genes are often located at fragile sites (FRAs), in common breakpoint regions and in minimal regions of loss of heterozygosity, so showing that miRNAs might be a new group of genes implicated in human tumorigenesis [2]. Some miRNAs are upregulated in human tumors therefore acting as an oncogene, others are downregulated and have a tumor suppressor function [3].

MiRNAs are commonly dys-regulated in cancer tissue respect to the normal tissue counterpart in different tumors, including lung cancer. Volinia et al [4] studied 540 samples including lung, breast, prostate, stomach, colon and pancreatic tumors and observed a miRNA signature composed of numerous overexpressed miRNAs, such as *miR-17-5p*, *miR-20a*, *miR-21*, *miR-92*, *miR-106a*, and *miR-155*. The targets for these miRNAs include both tumor suppressors and oncogenes. Yanaihara et al [5] analyzed miRNA expression in lung cancers and investigated miRNA's role in lung carcinogenesis. They identified a miRNA profile able to distinguish lung cancers from normal lung tissues as well as molecular signatures that differentiate NSCLC histological sub-types. MiRNA expression profiles associated with survival of adenocarcinomas, also including those classified as stage I disease. Yanaihara et al. observed that overexpression of miR-155 and also downregulation of let-7a-2 associated with poor survival. The miRNA signature on outcome was also determined by quantitative real-time RT-PCR (qRT-PCR) analysis of miRNAs and cross-validated with an independent group of adenocarcinomas.. Moreover miR-29 family (miR-29a, 29b and 29c) is downregulated in NSCLC [6]. In *in vitro* experiments, the Fabbri et al. showed that miR-29s downregulated the expression of DNA methyltransferase 3A and -3B (Dnmt3B, Dnmt3A), which transfer a methyl group from methyl donor S-adenosyl methionine onto the 5' site on the cytosine ring, specially in CG sites that were unmethylated on the parental strands of DNA. They assessed the expression of Dnmt3A and Dnmt3B proteins in 172 paired non-neoplastic/cancerous lung tissues, and Fabbri et al. discovered that high expression of Dnmt3A protein was significantly correlated with lower overall survival, while no statistically significant association with survival was shown for Dnmt3B. The mRNA DNMT3A and -3B expressions were inversely correlated with the expressions of miR-29s in NSCLC tissues. Also, they analyzed the mRNA expressions of two tumor suppressor genes, WW domain containing oxidoreductase (WWOX) and fragile histidine triad (FHIT), which are often silenced by promoter methylation in lung cancer. WWOX is involved in the regulation of a wide variety of cellular functions for example protein degradation, RNA

splicing, transcription while FHIT is a factor of the histidine triad gene family, encodes a diadenosine 5',5'''-P1,P3-triphosphate hydrolase that participate in purine metabolism. The transfection of miR-29a, -29b and 29c precursors in A549 and H1299 cell lines reduced DNMT3A and -3B expression and caused de-methylation of silenced epigenetically genes, determining to increased FHIT and WWOX expression. The re-expressed FHIT and WWOX inhibited tumor progression both *in vitro* and *in vivo*. This study shows that miRNAs are essential to regulate the epigenetic expression of tumor suppressor genes and have a important role in human carcinogenesis.

Other groups of ncRNAs are emerging as implicated in human carcinogenesis. The ultraconserved regions (UCRs) are a group of genomic sequences completely conserved between orthologous regions of the human, murine and rat genomes [7]. Their transcriptional activity is been identified and the transcribed UCRs is named T-UCR genes [8]. 53% of the UCRs were classified as nonexonic ("N", 256/481 lacking evidence of encoding protein), whereas the other 47% were designated either exonic ("E", 111/481, that overlap mRNA of known genes), or possibly exonic ("P", 114/481,with inconclusive indication of overlap with genes). Also T-UCRs are commonly located at fragile sites and other cancer-associated genomic regions, are dysregulated in numerous types of solid and hematological malignancies compared to the normal tissue and most are not translated into proteins [8]. Dysregulation of T-UCRs has been studied in pediatric tumors [8, 9] and is associated with outcome in high-risk neuroblastomas [9, 10]. Calin *et al* observed that *uc.73A* has oncogenic function in colorectal cancer (CRC) cell lines, whereas other researchers observed *uc.388* and *uc.338* as oncogenes in CRC [11] and hepatocellular carcinoma (HCC) [12], respectively. Moreover, Single Nucleotide Polymorphisms (SNPs) in T-UCR genes have been observed to be correlated with familial breast cancer risk [13]. Calin *et al* previously confirmed that T-UCR genes are epigenetically regulated by CpG island hypermethylation [14] and miRNAs [8]. It was observed that *uc.339* is over-expressed in HCC-derived exosomes and in HCC cells, contributing to a pro-tumoral HCC

microenvironment [15]. These findings support a function of T-UCRs in human carcinogenesis. However, many mechanism and consequences of dysregulation of T-UCR in human cancers remains unidentified.

## 2. Aims

## 2. AIMS

In this study it was observed whether *uc.339* is differentially expressed in NSCLC primary tumors compared to the normal tissue counterpart, and this expression related with overall survival was assessed in the TCGA database. Moreover, it was determined whether *uc.339* functions as an oncogene in lung cancer and it was investigated a possible mechanism of action of *uc.339*. Then, it was assessed whether *TP53*, a tumor suppressor gene mutated or deleted in more than 50% of human tumors, as NSCLC [16-18], might be the cause for *uc.339* dysregulation in lung cancer. Finally, the role of *uc.339* was tested in chemotherapeutic treatments.

## 3. METHODS

## **METHODS**

### **Cell lines and patient samples**

Paired frozen tumor and adjacent normal lung from 30 NSCLC patients were provided by Istituto Scientifico Romagnolo per la Cura e lo Studio dei Tumori IRST, s.r.l, IRCCS, in Meldola (FC), Italy. Written informed consent was collected from all patients before to analyze the sample. The clinical characteristics of the NSCLC patients studied in this work are summarized in **Table 1**.

All cell lines were provided from American Type Culture Collection and were incubated as a monolayer at 37°C. LoVo (colorectal adenocarcinoma) and A549 (lung adenocarcinoma) cells were grown in F12K medium (ATCC), with 10% FBS while H1299 (lung adenocarcinoma) and H460 (large cell lung cancer) cells were maintained in RPMI 1640 (ATCC) with 10% FBS. For the packaging of the *uc.339* expressing lentivirus, 293TN cells were used (or its empty lentiviral counterpart) and were maintained in DMEM medium (ATCC), with 10% FBS. Each cell line was checked for the presence of mycoplasma every 2 months (MycoAlert™ Mycoplasma Detection Kit – Lonza)

### **Chromatin Immunoprecipitation**

The TP53 binding sites upstream of the *uc.339* were obtained by combining a published global map of the TP53 binding sites [44] with the OMGProm algorithm [45]. Chromatin immunoprecipitation (ChIP) was performed with the EZ-ChIP kit (Millipore) and 5 µg of TP53 antibody (Santa Cruz Biotechnology) on H1299 transfected with CMV-TP53/Empty. As a negative control, it was immunoprecipitated 1 sample with 5 µg of pre-immune serum (Santa Cruz Biotechnology). The samples were PCR amplified using specific primers after final elution and purification of DNA. As a positive control for PCR, the input sample was diluted to 0.1 ng/L. PCR was obtained using the AmpliTaq Gold® PCR Master Mix (Applied Biosystems).

### **Isolation of *uc.339-miRNA* complexes**

The immunoprecipitation of the *uc.339* transcript was obtained as follows: RNA was extracted from LV-*uc.339* or LV-E infected H460 and A549 cells with complete RIPA buffer (Santa Cruz Biotechnology) (following the manufacturer's instructions). Next, it was incubated the RNA with 120 pmoles of a 3'-biotinylated RNA probe (see **Supplementary Table 1**) that was complementary to the *uc.339* transcript and did not overlap with the sequence of the *ATP5G2* host gene, overnight at 4°C. Co-precipitation of the bound *uc.339-miRNA* complexes was obtained by using streptavidin-conjugated magnetic beads (Miltenyi Biotec), according to the manufacturer's instructions as previously described [28].

### **Animal experiments**

Nude mice (4 week-old females) were provided from Jackson Laboratories (total  $n=14$ ). All mice were irradiated with a 200 cGy total body irradiation at 5 weeks of age at the Animal Care Facility of the Saban Research Institute of Children's Hospital Los Angeles. The next day,  $5 \times 10^6$  A549 cells stably infected with LV-*uc.339* were injected subcutaneously in seven mice and  $5 \times 10^6$  A549 cells stably infected with LV-E. were injected subcutaneously in seven mice. Tumor diameters were quantified after 8 days from the injection and then every 3 days. Mice were euthanized at 29 days after the injection and necropsy performed; the tumors were excised, photographed and measured. Tumor volumes were obtained by using the equation  $V \text{ (mm}^3\text{)} = L \times W^2/2$ , where  $W$  is the perpendicular diameter and  $L$  is the largest diameter. Total RNA was extracted from about 1/10 of the excised tumor, and *uc.339* expression was obtained by qRT-PCR in triplicate, per each mouse and as reported above.

All procedures used in this work complied with federal guidelines and were accepted by the Institutional Animal Care and Use Committee at Children's Hospital Los Angeles.

## TCGA dataset analysis

The case from TCGA analysis is of 204 Lung squamous cell carcinoma (LUSC) samples with overall survival information for which RPKM-normalized values of *uc.339* were calculated. 125 samples were profiled on Illumina Genome Analyser for miRNA expression. Clinical and level 3 miRNA information was downloaded from the Cancer Genome Atlas Project (TCGA; <http://tcga-data.nci.nih.gov/>). Deriving values for mature forms for microRNAs were used Level 3 Illumina miRNASeq “isoform\_quantification” files. It was downloaded RNA-seq BAM files from UCSC Cancer Genomics Hub (CGHub, <https://cghub.ucsc.edu/>). TCGA BAM files were obtained based on Mapsplice2 algorithm [46] for alignment on the hg19 reference genome with default parameters. It was quantified the *uc.339* expression as RPKM (reads per kilobase per million mapped reads) [47].

## Statistical analysis

Statistical data are shown as mean  $\pm$  standard deviation (s.d.) of experiments in triplicate, unless otherwise specified. Significance was computed by Student's *t*-test or by Anova test with Bonferroni correction. A *P* value  $<0.05$  was statistically significant.

For survival curves statistical analyses were computed in R (version 3.0.1) (<http://www.r-project.org/>) and the statistical significance was identified as a *P* value  $<0.05$ .

Patients were grouped into percentiles for the Kaplan-Meier curve according to *uc.339* expression. The Log-rank test was used to determine the association between overall survival and *uc.339* expression and the Kaplan-Meyer method was performed to generate survival curves. It was recorded cut-off points to significantly split (log-rank test *P* value  $<0.05$ ) the samples into low/high *uc.339* groups. It was chosen the cut-off to optimally separate the patients in high/low (min *P* value). It was chosen thus for *uc.339* the cut-off 0.54.

**Supplementary Table 1. List of primers designed in this study.**

| Procedures  | Primer name                    | Sequence   |
|---|--------------------------------|--|
| Sequencing  | F2                             | 5'-CAGCCATTCTTTTCCTGCTC-3'                             |
| Sequencing  | R11                            | 5'-AAGTGGGCCCCTACCTAGAA-3'                             |
| Sequencing  | 5F                             | 5'-CTCTTCCTGCAGTACTCCCTGC-3'                           |
| Sequencing  | 5R                             | 5'-GCCCCAGCTGCTCACCATCGCTA-3'                          |
| Sequencing  | 6F                             | 5'-GATTGCTCTTAGGTCTGGCCCTC-3'                          |
| Sequencing  | 6R                             | 5'-GGCCACTGACAACCACCCTTAACC-3'                         |
| Sequencing  | 7F                             | 5'-GTGTTGTCTCCTAGGTTGGCTCTG-3'                         |
| Sequencing  | 7R                             | 5'-CAAGTGGCTCCTGACCTGGAGTC-3'                          |
| Sequencing  | 8F                             | 5'-ACCTGATTTCTTACTGCCTCTGGC-3'                         |
| Sequencing  | 8R                             | 5'-GTCCTGCTTGCTTACCTCGTTAGT-3'                         |
| <i>uc.339</i> RT  | RT <i>uc.339</i><br>reverse    | 5'- CTCCACAGTGCCTGGCACCAC-3'                           |
| RACE  | Gene specific F                | 5'-CATTTTTATGGCCCTGAGCT-3'                             |
| RACE  | Gene specific R                | 5'-CTTCTCGCCCGTTCTCC-3'                                |
| RACE  | Nested gene<br>specific F      | 5'-GGAGAAGCGGGCGAGAAG-3'                               |
| RACE  | Nested gene<br>specific R      | 5'-AGCTCAGGGCCATAAAAATG-3'                             |
| Cloning of <i>uc.339</i> in<br>pCDH-CMV-MCS-<br>EF1-copGFP vector | <i>uc.339</i> F                | 5'- GCGAATTCACGCAGCACGAGAAAGACG - 3'                   |
| Cloning of <i>uc.339</i> in<br>pCDH-CMV-MCS-<br>EF1-copGFP vector | <i>uc.339</i> R                | 5'- TCGCGGCCGCTTTTGGTGGGAATGGA -3'                     |
| Mutagenesis of<br><i>miR-339-3p</i> CS in<br>CCNE2 3' UTR         | Del. <i>miR-339-3p</i><br>CS F | 5'-TTTGCCTTGCCATAACACATTTTTTAATAAACCTGTGCTCTAACAG-3'   |
| Mutagenesis of<br><i>miR-339-3p</i> CS in                         | Del. <i>miR-339-3p</i><br>CS R | 5'-CTGTTTAGAGCACAGTTATTAGTTAAAAATGTGTTATGGCAAGGCAAA-3' |

|  |                              |   |
|--|------------------------------|---|
| CCNE2 3' UTR   |                              |   |
| Mutagenesis of <i>miR-663b-3p</i> CS in CCNE2 3' UTR | Del. <i>miR-663b-3p</i> CS F | 5'-TAAATGCTGTGGCTCCTTCTATTTTGTATATCACAAATTTGGGTG-3' |
| Mutagenesis of <i>miR-663b-3p</i> CS in CCNE2 3' UTR | Del. <i>miR-663b-3p</i> CS R | 5'-CACCCAAATTGTGATATACAAAATAGGAAGGAGCCACAGCATTTA-3' |
| Mutagenesis of <i>miR-95-5p</i> CS in CCNE2 3' UTR   | Del. <i>miR-95-5p</i> CS F   | 5'-CTAGATTGCTAGTTTATTTTCTTCTCCCTTTGAAGAAAC-3'       |
| Mutagenesis of <i>miR-95-5p</i> CS in CCNE2 3' UTR   | Del. <i>miR-95-5p</i> CS R   | 5'-GTTTCTTCAAAGGGAGAAGAGAAAATAAACTAGCAATCTAG-3'     |
| Cloning of <i>TP53</i> CS 1 in pGL4.23(luc2/minP)    | CS1 F                        | 5'- CAGGCATGCGCCACCATGCCCG-3'                       |
| Cloning of <i>TP53</i> CS 1 in pGL4.23(luc2/minP)    | CS1 R                        | 5'-CTAGCGGGCATGGTGGCGCATGCCTGGTAC-3'                |
| Cloning of <i>TP53</i> CS 2 in pGL4.23(luc2/minP)    | CS2 F                        | 5'-CCTCCTTGTCTCCCAGACAGGACCTGCCCG-3'                |
| Cloning of <i>TP53</i> CS 2 in pGL4.23(luc2/minP)    | CS2 R                        | 5'-CTAGCGGGCAGGTCCTGTCTGGGAGACAAGGAGGGTAC-3'        |
| Cloning of <i>TP53</i> CS 3 in pGL4.23(luc2/minP)    | CS3 F                        | 5'-CGGACTTGCGTCCCCTTTCCGAGCATGCGCG-3'               |
| Cloning of <i>TP53</i> CS 3 in pGL4.23(luc2/minP)    | CS3 R                        | 5'-CTAGCGCGCATGCTCGGAAAGGGGACGCAAGTCCGGTAC-3'       |
| Cloning of <i>TP53</i> CS 4 in pGL4.23(luc2/minP)    | CS4 F                        | 5'-CGGTCAAGATCTCTGGGGGAGCTAGGTTG-3'                 |
| Cloning of <i>TP53</i> CS 4 in pGL4.23(luc2/minP)    | CS4 R                        | 5'-CTAGCAACCTAGCTCCCCAGAGATCTTGACCGGTAC-3'          |
| <i>uc.339</i> precipitation                          | 3' biotinilated RNA          | 5'-UACCUGUUGAUUGAUUCCCAA-3'.                        |

### **Analysis of the *TP53* gene mutation status in NSCLC samples**

Sequencing of *TP53* gene was obtained on exons 5, 6, 7 and 8 of the *TP53* gene for all NSCLC patients of this study.

Tumor DNA from each patient was obtained with the QIAamp DNA Mini Kit (Qiagen) and the QIAamp DNA Micro Kit (Qiagen). The fragment between exon 11 and exon 2 of the *TP53* gene was amplified with the primers R11 and F2 (listed in **Supplementary Table 1**), with an annealing temperature of 68°C using the LA Taq<sup>®</sup> DNA Polymerase (Takara). The PCR fragment was obtained with MinElute PCR Purification kit (Qiagen), as described to the manufacturer's instructions. Exons 8 to 5 were sequenced using the BigDye<sup>®</sup> Terminator v3.1 Cycle Sequencing kit (Life Technologies) with 56°C annealing temperature, with primers 5F, 5R, 6F, 6R, 7F, 7R, 8F and 8R as listed in **Supplementary Table 1**.

The sequences were obtained with the DyeEx 2.0 Spin kit (Qiagen) and analyzed in a 3130 Genetic analyzer (Applied Biosystem) with Montage Injection Solution (Millipore).

### **Quantitative Real-Time PCR**

TRizol<sup>®</sup> reagent (Invitrogen) was used to extract RNA, as described in manufacturer's instructions. RNA concentration and quality were tested with a NanoDrop<sup>®</sup> ND-1000 Spectrophotometer (Thermo Scientific). Moreover RNA was purified with the TURBO DNA-free<sup>™</sup> kit (Life Technologies).

Retrotranscription of *uc.339* was obtained using the RT *uc.339* reverse primer (**Supplementary Table 1**) that not to overlap with the sequence of the *ATP5G2* host gene (**Figure 1c**) and through the TaqMan<sup>®</sup> MicroRNA Reverse Transcription kit (Life Technologies), as described in manufacturer's instructions. *uc.339* cDNA was pre-amplified through TaqMan<sup>®</sup> PreAmp Master Mix (Life Technologies) and quantitative Real-Time PCR (qRT-PCR) was tried in triplicate through TaqMan<sup>®</sup> Universal PCR Master Mix (Life Technologies) as described in manufacturer's instructions. As a normalizer for qRT-PCR *RNU44* was used.

Retrotranscription of *miR-339-3p*, *-95-5p* and *-663-3p* was carried out through the TaqMan<sup>®</sup> MicroRNA Reverse Transcription kit (Life Technologies), as described in the manufacturer's instructions and qRT-PCRs were performed in triplicate with the TaqMan<sup>®</sup> Universal PCR Master Mix, no AmpErase<sup>®</sup> UNG reagents (Life Technologies), as described in the manufacturer's instructions. As a normalizer for qRT-PCR RNU44 was used.

*TP53* mRNA was retrotranscribed through using the TaqMan<sup>®</sup> MicroRNA Reverse Transcription kit (Life Technologies), as described in the manufacturer's instructions, and the cDNA was amplified through TaqMan<sup>®</sup> Universal PCR Master Mix (Life Technologies). The qRT-PCR reactions were done in triplicate and as a housekeeping gene the *HPRT* gene was used.

All quantitative retrotranscriptions and qRT-PCR reactions were performed in an Applied Biosystems<sup>®</sup> 7500 Real-Time PCR System (Life Technologies), as described in the operator's manual.

### **Rapid Amplification of cDNA Ends (RACE)**

To discover the 5'- and 3'-end of the transcript of *uc.339*, H1299 and LoVo cell total RNAs were purified with DNase I (RNase-free) (Life Technologies) and the SMARTer RACE cDNA Amplification Kit (Clontech) was performed, as described in the manufacturer's instructions. The cDNA ends were obtained with the Platinum Taq DNA Polymerase High Fidelity (Life Technologies) and gene-specific primers, as listed in **Supplementary Table 1**, were used. The primer for the 5'-end didn't overlap with the transcript of the *ATP5G2* host gene, so it is sure that only the *uc.339* was amplified. Furthermore, it was performed a nested PCR through the nested universal primer included in the kit and the nested gene-specific primers as indicated in **Supplementary Table 1**. As reaction controls were used Placental RNA and Transferrin receptor-specific primers included in the kit. The PCR fragments were then checked in a 1.5% agarose gel, and DNA was extracted through the QIAquick Gel Extraction Kit (Qiagen), as described in the manufacturer's instructions. Next, the RACE fragments were cloned into a TOPO<sup>®</sup> TA pCR<sup>®</sup>2.1 cloning plasmid (Invitrogen), as described in the manufacturer's

instructions, and the fragments were sequenced through using the T7 and T3 primers, and blasted using the UCSC Genome Browser website (<http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>).

### **Plasmids, Reagents and Transfection Conditions**

The TP53-expressing plasmid TrueClone pCMV6-XL5 (pCMV-TP53) and its empty control plasmid were purchased from OriGene. The identified *uc.339* RACE sequence was cloned into the pCDH-CMV-MCS-EF1-copGFP expression plasmid (pCDH-*uc.339*) (System Biosciences), through the *EcoRI* and *NotI* restriction sites and the primers listed in **Supplementary Table 1**. This vector was also used to obtain empty lentiviral vector controls (LV-E) or *uc.339* expressing lentiviral particles (LV-*uc.339*), following the Lentiviral Expression Technology (System Biosciences) and as described in the manufacturer's instruction. The *uc.339* vector and empty vector were transfected through the pPACK™ Lentivector Packaging System (System Biosciences) in the 293TN cell line. The resulted viruses were concentrated through the PEG-it™ Virus Precipitation Solution (System Biosciences), and lentiviral particle titration was made by using the Global UltraRapid Lentiviral Titer Kit™ (System Biosciences). A549, LoVo and H460 cells were infected at an multiplicity of infection (MOI) of 10 and were selected for green fluorescence signal through cytofluorimetry using a BD FACSCalibur platform (BD Biosciences). After cytofluorimetric cell selection, the infection efficiency was >90% through fluorescent microscopy and the *uc.339* expression was further validated through qRT-PCR.

The pCMV-TP53 and *uc.339* vectors were transfected in cell lines through Lipofectamine® LTX (Invitrogen), at a final concentration of 1 µg/mL as described in manufacturer's instructions. The precursors of *miR-95-5p*, *-663b-3p* and *-339-3p* (as well as the scrambled miRNA negative control #1) were provided from Ambion, and were transfected at a final concentration of 100 nM, using Lipofectamine® 2000 (Invitrogen), as described in the manufacturer's instructions.

Two siRNAs against two diverse regions of the *uc.339* transcript were designed: *si-uc.339* 5'-GGGAAUCAAUCAACAGGUATT-3' and *si-uc.339(2)* 5'-CUCCAGUUUUAGUUGUUGATT -3'.

These two siRNAs do not overlap with *ATP5G2* host gene, to avoid the silencing of the host gene which might interfere the data interpretation. Anti-*uc.339* siRNAs [si-*uc.339* and si-*uc.339(2)*] and a scrambled siRNA (si-*SCR*) were provided by Ambion. The transfections of siRNAs were performed at a final concentration of 100 nM, through Lipofectamine<sup>®</sup> RNAi Max (Invitrogen) and as described in the manufacturer's instructions.

### **Luciferase Reporter Assays**

To observe the direct targeting of *miR-339-3p*, *-95-5p* and *-663-3p* on the *CCNE2* mRNA, the LightSwitch 3'UTR plasmid containing the *CCNE2* 3'-UTR was provided from the LightSwitch 3'UTR Reporter GoClone Collection of SwitchGear Genomics.

As a control plasmid, it was performed a mutation (deletion) of the miRNA binding sites in *CCNE2* 3'-UTR plasmid with the mutagenesis primers described in **Supplementary Table 1** and using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene), as described in the manufacturer's instructions and as previously reported in literature [8].

The TP53 CS #1-4 containing sequences were cloned in the pGL4.23 [luc2/minP] plasmid (Promega), upstream of the promoter for firefly luciferase. Cloning was successfully performed through the *KpnI* and *NheI* restriction sites and the primers described in **Supplementary Table 1**.

To normalize the luciferase reporter assay experiments, it was co-transfected cells with the plasmids including the firefly luciferase gene, and with the pGL4.74 (hRluc/TK) plasmid with Lipofectamine<sup>®</sup> 2000 (Invitrogen) as described by the supplier. The luciferase signal was quantified with the Dual-Luciferase<sup>®</sup> Reporter Assay System (Promega) and was examined in a luciferase reporter assay system kit and Glomax<sup>®</sup>96 Microplate Luminometer (Promega), according to the manufacturer's instructions and as previously described [8].

## **Western blotting**

For immunoblotting assay, cells were lysed using complete RIPA buffer (Santa Cruz Biotechnologies) and proteins were denatured at 100 °C for 10 minutes. 50 µg of extracted proteins were loaded on Criterion™ XT 4-20% Precast Gels (Bio-Rad), and successively transferred on Trans-Blot®Turbo™ Midi Nitrocellulose Transfer Pack membrane (Bio-Rad) using a Trans Blot®Turbo™ Transfer System (Bio-Rad), as described in the manufacturer's instructions. For making sure that equal amounts of proteins were loaded in each lane the membrane was stained with Ponceau S (Sigma Aldrich). Membranes were maintained for 2 hours at room temperature with T-PBS 5% non fat dry milk. The membrane was incubated overnight at 4°C with the primary antibody, then it was added horseradish peroxidase-conjugated secondary antibody (Dako Corporation) at a dilution of 1:5000. These primary antibodies were used: anti-CCNE2, rabbit monoclonal antibody (Abcam) diluted 1:500, anti-TP53 Ab-2, mouse monoclonal antibody (Thermo scientific) diluted 1:400, anti-PARP rabbit, polyclonal antibody (Cell Signaling) diluted 1:1000, anti-VINCULIN, mouse monoclonal antibody (Biohit) diluted 1:1000. The bound antibodies were observed by enhanced chemiluminescence and by using the SuperSignal West Femto Chemiluminescent Substrate (Thermo Scientific), as described in the manufacturer's instructions. The quantification of the chemiluminescent bands were obtained by using the Quantity One software (Bio-Rad).

## **Cell viability and cell cycle assays**

For cell viability assay, A549, H460, LoVo and H1299 cells were detached through trypsin after 72h from plating (LoVo and A549 stably infected with LV-E or LV-*uc.339*) or from the treatment (si-*SCR* or si-*uc.339*), washed and suspended in PBS. An aliquot of the suspension was combined with an identical volume of 0.4% Trypan Blue and maintained for 8-10 min at room temperature. Total cell numbers and proportions of viable and non viable cells were counted in a KOVA®Glasstic® Slide counting chamber (Hycor Biomedical).

For cell cycle assay, cells were collected 72h after treatment with si-*SCR* or si-*uc.339*, fixed with 70% ethanol and stained with a solution containing 10 µg/ml of propidium iodide (Sigma Aldrich), 0.01% of NP40 (Sigma Aldrich) and 10,000 units/mL of RNase (Sigma Aldrich). After 30-60 min, samples were read by flow cytometry using a BD FACSVantage™ cytofluorimeter (BD Biosciences). The obtained data (10,000 events were collected for each sample) was performed by using the BD CellQuest™ Pro software (BD Biosciences), as described in the manufacturer's instructions. Data were analyzed using the ModFit LT™ software (Verity Software House), as described in the manufacturer's instructions and indicated as fractions of cells in the different cycle cycle phases.

### **Cell migration assay**

Wound healing analysis were performed with ibidi) cell culture inserts with a distinct cell free gap(Martinsried, Germany.  $3 \times 10^5$  A549 cells infected with LV-*uc.339* or LV-E were suspended in 70 µl full media for well of the insert which had been before placed in a 12 well cell culture dish. The cells were maintained to adhere in a 37°C, 5% CO<sub>2</sub> incubator for 4 hours. The insert was removed with sterile forceps. The cells were washed with prewarmed media several times to remove the non adherent cells. Phase contrast images (4x) were taken and designated as Time 0. The position of the field was way signed. Eight and 24 hours later, images were obtained at the same location as time 0.

The images were uploaded to the ibidi company web site and analyzed through their Wim-Scratch Analysis program. The gap is 500 +/- 50 µm and the distance is quantified comparing to Time 0 (500 µm) over time that the number gets smaller. Data is reported as proportion of area of initial cell-less gap. The experiment was performed in triplicate.

## **Drug**

Cisplatin (Mayne Pharma Pty Ltd, Mulgrave, Australia) was suspended in 0.9% saline solution. Cisplatin was stored at 25°C and it was freshly diluted in culture medium before each experiment.

## **Drug exposure**

Cells were treated with cisplatin for 6 h to reproduce the clinical conditions of lung cancer treatment. Considering that the peak plasma level is 3 µg/ml for cisplatin, it was tested 3-µg/ml concentrations for cisplatin. Analysis of the cytotoxic effect was performed 48 h after the end of drug exposure.

## **4. RESULTS**

## RESULTS

### In NSCLC with poor prognosis, *uc.339* is up-regulated.

Increased expression of *uc.339* has been observed in CRC and HCC, compared to the corresponding normal tissue [8, 15]. In this work, *uc.339* expression level was quantified in 30 paired tumor and adjacent normal lung, by quantitative real-time PCR (qRT-PCR). Statistically significant *uc.339* up-regulation was shown in tumor compared to adjacent non-tumor lung ( $P < 0.0001$  **Fig. 1a**). To analyze whether *uc.339* up-regulation harbored prognostic implications the TCGA database was interrogated and observed in 204 NSCLC patients that high levels of *uc.339* significantly associate a lower overall survival ( $P = 0.0186$  **Fig. 1b**). Together these data show that in NSCLC primary tumors with poor prognosis *uc.339* is up-regulated.

### *uc.339* induces NSCLC growth and migration

The functional effects of *uc.339* dysregulation was investigated in lung cancer. As described by Bejerano et al., *uc.339* is a T-UCR with inconclusive evidence of overlap with a protein coding gene (possibly exonic T-UCR) [7], partly overlapped in the *ATP5G2* host gene (**Fig. 1c**). The length of the *uc.339* transcript was analyzed and also determined the endogenous expression level of *uc.339* in one CRC cell line (LoVo) and in three NSCLC cell lines (A549, H1299 and H460) by qRT-PCR, carefully testing the primers to amplify within the ultraconserved area as described by Bejerano et al. [7], but not to amplify with the *ATP5G2* host gene (**Fig. 1c** and **Supplementary Table 1**). The tested cell lines had a variable expression of *uc.339*, with the highest expression level in H1299 and LoVo (**Supplementary Fig. 1a**). Moreover, *uc.339* transcript length was determined through Rapid Amplification of cDNA Ends (RACE) in LoVo and H1299 cells and a prevalent *uc.339* transcript sequence of 849 nucleotides (nt) was observed, beginning 273 nt upstream and ending 324 nt downstream the sequence described by Bejerano et al. [7] (**Supplementary Fig. 1b**).

Subsequently, A549 and LoVo cells were infected with a lentiviral vector that over-expresses *uc.339* (LV-*uc.339*) or its empty vector counterpart (LV-E). LV-*uc.339*-infected cells (**Supplementary Fig. 2a**) show an increased viability (**Fig. 2a**) and increased migration in the scratch analysis (**Fig. 2b** and **c**) compared to LV-E infected cells. On the other hand, when *uc.339* endogenous expression was silenced through two different siRNAs against different regions of the *uc.339* transcript [si-*uc.339* and si-*uc.339*(2)] in A549, H460, H1299 and LoVo cells (**Supplementary Fig. 2b** and **c**), a reduction of cell viability was observed at 72h in H460, LoVo and H1299, but not in A549 (**Fig. 2d** and **Supplementary Fig. 2d**). In A549 cells, the effect of anti-*uc.339* siRNAs lacks because the A549 cells express very low endogenous levels of *uc.339* (**Supplementary Fig. 1a**). Finally, the effects of *uc.339* silencing were determined in A549, H460, LoVo, and H1299 cells by cytofluorimetry. A decreased expression level of *uc.339* was observed, also reducing of the fraction of cells in S-phase, and an increase of the fraction of cells in G<sub>0</sub>/G<sub>1</sub> phase that caused protein PARP-cleavage (**Fig. 2e, f** and **Supplementary Fig. 2e, f**). These effects were not obtained (or, for the PARP-cleavage, obtained to a lesser extent) in A549 cells, confirming the *uc.339*-mediated effect on cell viability and cell cycle. Next, the effects of *uc.339* over-expression were assessed *in vivo* in an xenograft murine model. A549 LV-*uc.339* or LV-E infected cells were injected subcutaneously in five-week old nude female mice ( $n=7$ /group). A549 cells were selected for this experiment because they express the lowest expression levels of endogenous *uc.339* (**Supplementary Fig. 1a**). Tumor volume was measured every 3 days and until day 29 in both groups. Mice injected with *uc.339* over-expressing A549 cells had a faster tumor growth than mice with empty A549 cells, and the difference between the two groups was statistically significant from day 23 (**Fig. 3a**). After euthanizing the mice, analysis of *uc.339* expression level in the xenografts observed still a significantly higher expression of *uc.339* in the LV-*uc.339* group (**Fig. 3b**) and bigger tumor growth in the LV-*uc.339* group (**Fig. 3c**). Moreover, it was tested if the *uc.339* over-expression caused the chemoresistance to Cisplatin (a drug used most in clinical to care the lung cancer

patents). The cytotoxic assay with Cisplatin treatment at peak plasma level shows that LV-*ucr.339* infected cells were more viable than LV-E infected cells (**Figure 3D**). These experiments confirm a role for *uc.339* in causing lung cancer migration and growth both *in vitro* and *in vivo*.

### ***uc.339* sequesters *miR-339-3p*, *-663-3p* and *-95-5p* and induces Cyclin E2 expression**

To discover the mechanism of action of *uc.339*, it was hypothesized that the T-UCR might function as a “decoy” for mature miRNAs. A structural analysis on *uc.339* transcript shown the presence of complementary sequences to three miRNAs (namely, *miR-339-3p*, *-663-3p* and *-95-5p*) in the *uc.339* transcript (**Fig. 4a**). Calin et al. previously observed that T-UCRs can be directly regulated by miRNAs [8]. So, to discover whether the indicated *in silico* *uc.339*-miRNA interaction determines a *uc.339* targeting by the three miRNAs, A549, H460, H1299 and LoVo cells were transfected with each of the three miRNAs (or a control as scrambled oligonucleotide), and observed *uc.339* expression after 48h. It wasn't observed any down-regulation of *uc.339* when the three miRNAs was transfected in any the tested cell lines (**Fig. 4b**), confirming that the *uc.339*-miRNA interaction indicated in **Fig. 4a** does not determine in miRNA targeting of *uc.339*. To test whether the interaction causes in *uc.339* trapping of the three miRNAs, the expression of *miR-339-3p*, *-95-5p* and *-663-3p* was determined in A549 and H460 cells (both cell lines with low levels of *uc.339* endogenous expression) stably expressing *uc.339* and a down-regulation of all three miRNAs was observed (**Fig. 4c**). Contrarily, the silencing of *uc.339* in H1299 cells (with high endogenous expression levels of *uc.339*) caused to increased expression of the three miRNAs (**Fig. 4d** and **Supplementary Fig. 3**). Moreover, the *uc.339* transcript was immunoprecipitated in A549 LV-*uc.339* and H460 LV-*uc.339* cells and obtained an enrichment of *miR-339-3p*, *-663-3p* and *-95-5p* in the immuno-precipitate (**Fig. 4e**). To observe the possible implications of such a “decoy” mechanism, an *in silico* analysis of all the target genes for *miR-339-3p*, *-663-3p* and *-95-5p* was performed [19-21] and *CCNE2*

(Cyclin E2) was identified to be a predicted target for indicated miRNAs. It was validated *CCNE2* as a direct target of these miRNAs through a luciferase reporter assay. Co-transfection of H1299 and LoVo cells was performed with a plasmid carrying the *CCNE2* 3'-UTR region downstream of the luciferase reporter gene, and *miR-339-3p*, *-95-5p*, *-663-3p*, or a scrambled miRNA as a control. Reduced luciferase activity was obtained with all three miRNAs compared to scrambled transfected cells, and this effect lacked when the three miRNA binding sites on *CCNE2* mRNA were deleted (**Fig. 5a**). It was also shown that *CCNE2* protein levels were decreased in LoVo and H1299 cells transfected with *miR-339-3p*, *-95-5p* or *-663-3p* compared to a scrambled miRNA (**Fig. 5b**). Next, it was investigated whether *uc.339* indirectly regulated *CCNE2* expression levels through inducing down-regulation of *miR-339-3p*, *-95-5p* - and *663-3p*. First, the *CCNE2* expression level was analyzed in A549, H460, and LoVo cells over-expressing *uc.339* through lentiviral infection and up-regulation of *CCNE2* protein was observed in all three cell line models (**Fig. 5c**). Then, *uc.339* expression was silenced in H460 and H1299 cells and a down-regulation of *CCNE2* protein expression was shown compared to the si-SCR control (**Fig. 5d** and **Supplementary Fig. 4**). Finally, H460 cells stably expressing *uc.339* were transfected them with *miR-339-3p*, *-95-5p*, *-663-3p*, or a scrambled miRNA. It was observed that even in presence of *uc.339*, the re-expression of the three miRNAs of interest (not of the scrambled control) reverted the protein up-regulation of *CCNE2* levels caused by *uc.339* overexpression (**Fig. 5e**). Overall, these experiments confirm that *uc.339* induces the expression of *CCNE2* by regulating the expression of *miR-339-3p*, *-95-5p* and *-663-3p*.

### ***uc.339* is transcriptionally regulated by TP53**

It was investigated how the expression level of *uc.339* is regulated. Through analysis of endogenous expression level of *uc.339* in three NSCLC cell lines and one CRC line (**Supplementary Fig. 1a**) it was noticed that while *uc.339* had the lowest levels in wild type

TP53-expressing cell lines (A549), the expression level was the highest in TP53-null H1299 cells (**Supplementary Fig. 1a** and **Supplementary Fig. 5**). This correlation between TP53 and *uc.339* expression levels was not obtained in LoVo cells (**Supplementary Fig. 1a** and **Supplementary Fig. 5**). A possible TP53 regulation of *uc.339* expression specific for NSCLC can exist. To investigate this hypothesis, A549 cells were transfected with a siRNA anti-*TP53* (or si-*SCR* control) and H1299 cells with a *TP53* expressing vector (or empty vector control) (**Supplementary Figure 6**) and a significantly induced expression of *uc.339* was obtained in A549 transfected with siRNA anti-*TP53* and a decreased expression of *uc.339* in H1299 transfected with the *TP53* expressing vector (**Fig. 6a**). Moreover in primary NSCLC patients included only if carrying a wt *TP53* status ( $n=22$ ), the TP53 protein and *uc.339* expression levels were determined both in cancerous tissues and in adjacent normal lung tissue. It was shown that normal lung tissues had higher expression level of TP53 and lower expression level of *uc.339* compared to the matched cancerous tissues, whereas lower expression of TP53 and higher levels of *uc.339* were observed in cancerous tissues (**Fig. 6b**). These data confirm a regulatory role of TP53 on *uc.339* expression. Next, it was searched for TP53 consensus sequences (CS) in the *uc.339* locus. At least three TP53 CS situated upstream of the *uc.339* sequence were identified on chromosome 12q13.13 and one TP53 CS situated downstream of the *uc.339* sequence (**Fig. 6c**). Chromatin immunoprecipitation was performed in H1299 cells transfected with a plasmid expressing TP53 (or its empty counterpart as a control) to discover whether TP53 binds to the predicted CS, and an enrichment of CS #4 (situated 2,963 bp upstream of the start of *uc.339* transcription) was observed in the TP53 immunoprecipitate, confirming that TP53 binds to this site (**Fig. 6d**). To discover whether TP53 binding to CS #4 determinates to transactivation or silencing, CS #4 was cloned in a reporter vector upstream of the luciferase gene. A decreased luciferase activity was shown when a vector over-expressing *TP53* (compared to the empty vector counterpart) was co-transfected with the vector containing CS #4 luciferase in A549, H460 and LoVo cells. This decrease was completely abolished when

CS #4 was mutated, even in presence of co-transfection with the *TP53* over-expressing vector (**Fig. 6e**). These data confirm that TP53 binds to a CS situated upstream of the *uc.339* gene and directly decreases the expression of *uc.339*.

### **TP53 regulates *CCNE2* expression through *uc.339***

This study shows that TP53 silenced directly *uc.339* in NSCLC (**Fig. 6**) and that *uc.339* sequesters *miR-339-3p*, *-95-5p* and *-663-3p* (**Fig. 4**) determining to up-regulation of *CCNE2* (**Fig. 5**) and increased migration and tumor growth both *in vitro* and *in vivo* (**Fig. 2** and **3**). The TP53 is mutated in about 50% of NSCLC [17-18] and in several human cancers [16], therefore it was investigated whether the functional *TP53* expression could regulate the oncogenic effects of *uc.339*.

First, *TP53*-null H1299 cells were transfected with a vector expressing wt *TP53* or with its empty vector counterpart, and analyzed the effects on *miR-339-3p*, *-95-5p* and *-663-3p* and *CCNE2* expression, through qRT-PCR and immunoblotting, respectively. Re-expression of TP53 caused up-regulation of the indicated miRNAs and down-regulation of their common target gene *CCNE2* (**Fig. 7a** and **7b, lanes 1 and 2**). Also, when H1299 cells were co-transfected with the *TP53* over-expressing vector and with a vector over-expressing *uc.339*, the expression of *CCNE2* was rescued (**Fig. 7b, lane 3**), confirming that TP53 down-regulation of *CCNE2* is, at least in part, regulated by TP53 silencing of *uc.339*. These data also confirm that re-expression of wild type TP53 can revert the effects of *uc.339*, determined by its miRNA sequestering effect and *CCNE2* over-expression in NSCLC. Finally, it was validated the *TP53-uc.339-miR-339-3p/-663-3p/-95-5p-CCNE2* pathway in 22 primary NSCLC samples, where it was obtained that when TP53 was down-expressed and *uc.339* over-expressed (**Fig. 6b**), *miR-339-3p*, *-95-5p* and *-663-3p* were down-expressed (**Fig. 7c**) and *CCNE2* protein expression was increased (**Fig. 7d**). These data indicate that *uc.339* has its oncogenic function, at least in part, through sequestering complementary mature *miR-339-3p*, *-95-5p* and *-663-3p* and

releasing oncogenic CCNE2 from its miRNA-controlled regulation in NSCLC (**Supplementary Fig. 7**).

## 5. DISCUSSION

## DISCUSSION

This study identifies *uc.339* as up-regulated in primary NSCLC tissues compared to the adjacent non-tumor lung tissues. Also it was demonstrated that *uc.339* increases lung cancer cell growth *in vitro* and in *in vivo* xenograft murine models, and increases cell migration and chemoresistance *in vitro*, therefore acting as an oncogene. Calin et al. previously observed up-regulation of *uc.339* also in CRC primary samples [8], whereas a report showed an oncogenic function for *uc.339* in HCC [15], although limited to *in vitro* experiments. Interestingly, *uc.339* has been observed also in microvesicles secreted by cancer cells [15], then it could be implicated in paracrine intercellular cross-talk mechanisms, within the cancer microenvironment as described for miRNAs [22]. The mechanism of regulation and action of *uc.339* and of T-UCRs are currently unknown. Specifically, it is unclear how T-UCR regulation occurs in cancer, and how they contribute to cancer invasiveness and cell increased growth. Many T-UCRs act as long-range enhancers during mouse development [23], this role has not been observed for all T-UCRs and it has been demonstrated that similar proportions of enhancers can be observed in less conserved sequences [24]. In this study, it was shown that *uc.339* contains three sequences complementary to the sequence of *miR-95-5p*, *-339-3p* and *-663b-3p*, and that *uc.339* acts as a “decoy” sequestering these miRNAs. As a result, the *CCNE2* mRNA, which it was demonstrated being a direct target of *miR-339-3p*, *-663-3p* and *-95-5p*, is free from the post-transcriptional regulation of the three miRNAs, and the *CCNE2* expression is increased, determining to increased cell proliferation and motility. Cyclin E2 appertains to the highly conserved cyclin family, acts as a regulatory subunit of CDK2, and has a role in cell cycle G1/S transition [25-27]. *CCNE2* is often up-regulated in human pulmonary malignancy and dysplasia and is associated with poor prognosis in lung cancer patients [28-30]. In numerous cancer types (as NSCLC and breast cancer) *CCNE2* up-regulation happens as an early event [28, 31], and over-expression of this cyclin acts as an inducer of genomic instability and polyploidy, in a different way from cyclin E1, D or A over-expression [32-33]. Even if the influence of *uc.339*-

induced up-regulation of cyclin CCNE2 was not determined on the genomic instability of the tested cell lines, this study warrants further investigation. It has also been observed that the CCNE2-CDK2 axis blocks SIRT2 through a Ser-331 phosphorylation, determining to increased cell migration [34]. In this study, over-expression of CCNE2 was observed and improved motility in cancer cells over-expressing *uc.339*. Whether this phenotype is caused by the regulation on SIRT2 phosphorylation remains to be clarified.

This study demonstrates that *CCNE2* is directly regulated by *miR-339-3p*, *-663-3p* and *-95-5p*. *MIR-95* is over-expressed in prostate, pancreatic, breast and colorectal cancer, where it increases cell proliferation by targeting Nexin-1 [35]. Moreover, it has been observed that *miR-95* increases the resistance to radiotherapy in prostate and breast cancer cells by regulating Sphingolipid Phosphatase SGPP1 [36]. In breast cancer cell lines *MIR-339-5p* is down-regulated and it inhibits cell migration and invasiveness, and is a marker of bad prognosis when down-regulated [37]. Also, it has been demonstrated that *miR-339-5p* increases resistance of glioma cells to cytotoxic T-lymphocytes, by regulating ICAM-1 [38]. In the more aggressive NSCLC, the expression of *miR-339-3p* is dysregulated [39], suggesting a role for this miRNA in lung carcinogenesis.

A double nature both as an oncogene and as a tumor suppressor gene is shown also for *miR-663*. While *miR-663* can be over-expressed by the anti-inflammatory drug resveratrol, and functions as a tumor-suppressor in human THP1 monocytic cells as well as in human blood monocytes by regulating AP-1 and inhibiting LPS-mediated up-regulation of the oncogenic *miR-155* [40], in nasopharyngeal carcinoma *miR-663* increases cancer proliferation by regulating p21 (WAF1/CIP1) [41], and in prostate cancer over-expression of *miR-663* is correlated with increased castration-resistance [42].

In this study it was observed that *uc.339* can bind to all three miRNAs in complementary sites along the *uc.339* sequence, suggesting the possibility that this T-UCR acts as a “reservoir” for the three miRNAs, regulating their disposability to target *CCNE2*. Further studies will define

whether different expression levels of *uc.339* affect the three miRNAs in different way and how this might affect the expression level of their common target *CCNE2*, determining different phenotypes. Liz et al. showed that *uc.283A* inhibits maturation of *pri-miR-195* through binding to the lower stem region of the pri-miRNA and blocking its cleavage by Drosha [43]. This study confirms a T-UCR/miRNA reciprocal interaction and regulation.

Finally, it was observed that *uc.339* is directly silenced by TP53. In this study it was identified a TP53 CS upstream of *uc.339* gene and it was shown that when transfected, TP53 binds to this sequence and silences *uc.339* expression. Also an inverse correlation between TP53 and *uc.339* expression was observed in primary NSCLC tissues, showing a negative regulation by TP53. Since mutations of the TP53 gene are in about 50% of NSCLC and are very common in several types of cancer, the identified mechanism might be a significant event in lung carcinogenesis, with implications for other types of cancer. This statement is confirmed by the observation of *uc.339* in CRC cancer cells.

In conclusion, this study observes that TP53-silenced *uc.339* functions as an oncogene in NSCLC, by sequestering *miR-339-3p*, *-663-3p* and *-95-5p* and inhibiting their targeting of oncogenic *CCNE2*, determining to increased tumor growth, migration and chemoresistance. This study identifies *uc.339* as an important molecular target for NSCLC and different types of cancer with TP53 impaired expression.

## 6. REFERENCES

## REFERENCES

1. Lagos-Quintana M., Rauhut R., Lendeckel W., Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* Oct 26; 294(5543) (2001) 853-8.
2. Calin G.A., Sevignani C., Dumitru C.D., Hyslop T., Noch E., Yendamuri S., Shimizu M., Rattan S., Bullrich F., Negrini M., Croce C.M. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. U S A* Mar 2; 101(9) (2004) 2999-3004.
3. Croce C.M. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* Oct 10(10) (2009) 704-14.
4. Volinia S., Calin G.A., Liu C.G., Ambs S., Cimmino A., Petrocca F., Visone R., Iorio M., Roldo C., Ferracin M., Prueitt R.L., Yanaihara N., Lanza G., Scarpa A., Vecchione A., Negrini M., Harris C.C., Croce C.M. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U S A*. Feb 14; 103(7) (2006) 2257-61.
5. Yanaihara N., Caplen N., Bowman E., Seike M., Kumamoto K., Yi M., Stephens R.M., Okamoto A., Yokota J., Tanaka T., Calin G.A., Liu C.G., Croce C.M., Harris C.C. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell*. Mar 9(3) (2006) 189-98.
6. Fabbri M., Garzon R., Cimmino A., Liu Z., Zanesi N., Callegari E., Liu S., Alder H., Costinean S., Fernandez-Cymering C., Volinia S., Guler G., Morrison C.D., Chan K.K., Marcucci G., Calin G.A., Huebner K., Croce C.M. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc. Natl. Acad. Sci. U S A* Oct 2; 104(40) (2007) 15805-10.
7. Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. *Science* 304, 1321-1325.
8. Calin, G.A., Liu, C.G., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E.J., Wojcik, S.E., et al. (2007). Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer cell* 12, 215-229.
9. Scaruffi, P., Stigliani, S., Moretti, S., Coco, S., De Vecchi, C., Valdora, F., Garaventa, A., Bonassi, S., and Tonini, G.P. (2009). Transcribed-Ultra Conserved Region expression is associated with outcome in high-risk neuroblastoma. *BMC cancer* 9, 441.
10. Mestdagh, P., Fredlund, E., Pattyn, F., Rihani, A., Van Maerken, T., Vermeulen, J., Kumps, C., Menten, B., De Preter, K., Schramm, A., et al. (2010). An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours. *Oncogene* 29, 3583-3592.
11. Sana, J., Hankeova, S., Svoboda, M., Kiss, I., Vyzula, R., and Slaby, O. (2012). Expression levels of transcribed ultraconserved regions uc.73 and uc.388 are altered in colorectal cancer. *Oncology* 82, 114-118.
12. Braconi, C., Valeri, N., Kogure, T., Gasparini, P., Huang, N., Nuovo, G.J., Terracciano, L., Croce, C.M., and Patel, T. (2011). Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 108, 786-791.
13. Yang, R., Frank, B., Hemminki, K., Bartram, C.R., Wappenschmidt, B., Sutter, C., Kiechle, M., Bugert, P., Schmutzler, R.K., Arnold, N., et al. (2008). SNPs in ultraconserved elements and familial breast cancer risk. *Carcinogenesis* 29, 351-355.
14. Lujambio, A., Portela, A., Liz, J., Melo, S.A., Rossi, S., Spizzo, R., Croce, C.M., Calin, G.A., and Esteller, M. (2010). CpG island hypermethylation-associated silencing of non-coding RNAs transcribed from ultraconserved regions in human cancer. *Oncogene* 29, 6390-6401.

15. Kogure, T., Yan, I.K., Lin, W.L., and Patel, T. (2013). Extracellular Vesicle-Mediated Transfer of a Novel Long Noncoding RNA TUC339: A Mechanism of Intercellular Signaling in Human Hepatocellular Cancer. *Genes Cancer* 4, 261-272.
16. Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C.C. (1991). p53 mutations in human cancers. *Science* 253, 49-53.
17. Bodner, S.M., Minna, J.D., Jensen, S.M., D'Amico, D., Carbone, D., Mitsudomi, T., Fedorko, J., Buchhagen, D.L., Nau, M.M., Gazdar, A.F., et al. (1992). Expression of mutant p53 proteins in lung cancer correlates with the class of p53 gene mutation. *Oncogene* 7, 743-749.
18. Takahashi, T., Nau, M.M., Chiba, I., Birrer, M.J., Rosenberg, R.K., Vinocour, M., Levitt, M., Pass, H., Gazdar, A.F., and Minna, J.D. (1989). p53: a frequent target for genetic abnormalities in lung cancer. *Science* 246, 491-494.
19. Alexiou, P., Vergoulis, T., Gleditzsch, M., Prekas, G., Dalamagas, T., Megraw, M., Grosse, I., Sellis, T., and Hatzigeorgiou, A.G. (2010). miRGen 2.0: a database of microRNA genomic information and regulation. *Nucleic acids research* 38, D137-141.
20. Betel, D., Koppal, A., Agius, P., Sander, C., and Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome biology* 11, R90.
21. Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., et al. (2005). Combinatorial microRNA target predictions. *Nature genetics* 37, 495-500.
22. Fabbri, M., Paone, A., Calore, F., Galli, R., Gaudio, E., Santhanam, R., Lovat, F., Fadda, P., Mao, C., Nuovo, G.J., et al. (2012). MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2110-2116.
23. Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499-502.
24. Visel, A., Prabhakar, S., Akiyama, J.A., Shoukry, M., Lewis, K.D., Holt, A., Plajzer-Frick, I., Afzal, V., Rubin, E.M., and Pennacchio, L.A. (2008). Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat Genet* 40, 158-160.
25. Gudas, J.M., Payton, M., Thukral, S., Chen, E., Bass, M., Robinson, M.O., and Coats, S. (1999). Cyclin E2, a novel G1 cyclin that binds Cdk2 and is aberrantly expressed in human cancers. *Mol Cell Biol* 19, 612-622.
26. Lauper, N., Beck, A.R., Cariou, S., Richman, L., Hofmann, K., Reith, W., Slingerland, J.M., and Amati, B. (1998). Cyclin E2: a novel CDK2 partner in the late G1 and S phases of the mammalian cell cycle. *Oncogene* 17, 2637-2643.
27. Zariwala, M., Liu, J., and Xiong, Y. (1998). Cyclin E2, a novel human G1 cyclin and activating partner of CDK2 and CDK3, is induced by viral oncoproteins. *Oncogene* 17, 2787-2798.
28. Lonardo, F., Rusch, V., Langenfeld, J., Dmitrovsky, E., and Klimstra, D.S. (1999). Overexpression of cyclins D1 and E is frequent in bronchial preneoplasia and precedes squamous cell carcinoma development. *Cancer Res* 59, 2470-2476.
29. Fukuse, T., Hirata, T., Naiki, H., Hitomi, S., and Wada, H. (2000). Prognostic significance of cyclin E overexpression in resected non-small cell lung cancer. *Cancer Res* 60, 242-244.
30. Mishina, T., Dosaka-Akita, H., Hommura, F., Nishi, M., Kojima, T., Ogura, S., Shimizu, M., Katoh, H., and Kawakami, Y. (2000). Cyclin E expression, a potential prognostic marker for non-small cell lung cancers. *Clinical cancer research : an official journal of the American Association for Cancer Research* 6, 11-16.

31. de Seranno, S., and Meuwissen, R. (2010). Progress and applications of mouse models for human lung cancer. *Eur Respir J* 35, 426-443.
32. Spruck, C.H., Won, K.A., and Reed, S.I. (1999). Deregulated cyclin E induces chromosome instability. *Nature* 401, 297-300.
33. Caldon, C.E., Sergio, C.M., Burgess, A., Deans, A.J., Sutherland, R.L., and Musgrove, E.A. (2013). Cyclin E2 induces genomic instability by mechanisms distinct from cyclin E1. *Cell Cycle* 12, 606-617.
34. Pandithage, R., Lilischkis, R., Harting, K., Wolf, A., Jedamzik, B., Luscher-Firzlaff, J., Vervoorts, J., Lasonder, E., Kremmer, E., Knoll, B., et al. (2008). The regulation of SIRT2 function by cyclin-dependent kinases affects cell motility. *J Cell Biol* 180, 915-929.
35. Huang, Z., Huang, S., Wang, Q., Liang, L., Ni, S., Wang, L., Sheng, W., He, X., and Du, X. (2011). MicroRNA-95 promotes cell proliferation and targets sorting Nexin 1 in human colorectal carcinoma. *Cancer research* 71, 2582-2589.
36. Huang, X., Taeb, S., Jahangiri, S., Emmenegger, U., Tran, E., Bruce, J., Mesci, A., Korpela, E., Vesprini, D., Wong, C.S., et al. (2013). miRNA-95 Mediates Radioresistance in Tumors by Targeting the Sphingolipid Phosphatase SGPP1. *Cancer Res* 73, 6972-6986.
37. Wu, Z.S., Wu, Q., Wang, C.Q., Wang, X.N., Wang, Y., Zhao, J.J., Mao, S.S., Zhang, G.H., Zhang, N., and Xu, X.C. (2010). MiR-339-5p inhibits breast cancer cell migration and invasion in vitro and may be a potential biomarker for breast cancer prognosis. *BMC Cancer* 10, 542.
38. Ueda, R., Kohanbash, G., Sasaki, K., Fujita, M., Zhu, X., Kasthuber, E.R., McDonald, H.A., Potter, D.M., Hamilton, R.L., Lotze, M.T., et al. (2009). Dicer-regulated microRNAs 222 and 339 promote resistance of cancer cells to cytotoxic T-lymphocytes by down-regulation of ICAM-1. *Proc Natl Acad Sci U S A* 106, 10746-10751.
39. Boeri, M., Verri, C., Conte, D., Roz, L., Modena, P., Facchinetti, F., Calabro, E., Croce, C.M., Pastorino, U., and Sozzi, G. (2011). MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3713-3718.
40. Tili, E., Michaille, J.J., Adair, B., Alder, H., Limagne, E., Taccioli, C., Ferracin, M., Delmas, D., Latruffe, N., and Croce, C.M. (2010). Resveratrol decreases the levels of miR-155 by upregulating miR-663, a microRNA targeting JunB and JunD. *Carcinogenesis* 31, 1561-1566.
41. Yi, C., Wang, Q., Wang, L., Huang, Y., Li, L., Liu, L., Zhou, X., Xie, G., Kang, T., Wang, H., et al. (2012). MIR-663, a microRNA targeting p21(WAF1/CIP1), promotes the proliferation and tumorigenesis of nasopharyngeal carcinoma. *Oncogene* 31, 4421-4433.
42. Jiao, L., Deng, Z., Xu, C., Yu, Y., Li, Y., Yang, C., Chen, J., Liu, Z., Huang, G., Li, L.C., et al. (2013). MicroRNA-663 Induces Castration-Resistant Prostate Cancer Transformation and Predicts Clinical Recurrence. *J Cell Physiol*.
43. Liz, J., Portela, A., Soler, M., Gomez, A., Ling, H., Michlewski, G., Calin, G.A., Guil, S., and Esteller, M. (2014). Regulation of pri-miRNA Processing by a Long Noncoding RNA Transcribed from an Ultraconserved Region. *Molecular cell*.
44. Wei, C.L., Wu, Q., Vega, V.B., Chiu, K.P., Ng, P., Zhang, T., Shahab, A., Yong, H.C., Fu, Y., Weng, Z., et al. (2006). A global map of p53 transcription-factor binding sites in the human genome. *Cell* 124, 207-219.
45. Palaniswamy, S.K., Jin, V.X., Sun, H., and Davuluri, R.V. (2005). OMGProm: a database of orthologous mammalian gene promoters. *Bioinformatics* 21, 835-836.
46. Wang, K., Singh, D., Zeng, Z., Coleman, S.J., Huang, Y., Savich, G.L., He, X., Mieczkowski, P., Grimm, S.A., Perou, C.M., et al. (2010). MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic acids research* 38, e178.

47. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* 5, 621-628.

## 7. FIGURES

## FIGURE LEGENDS

### Figure 1. In NSCLC *uc.339* is up-regulated and associates with prognosis

- (a) qRT-PCR for *uc.339* in 30 primary NSCLC cancerous tissues (C) and the adjacent normal lung (N). The expression of *uc.339* was normalized to *RNU44* and is shown as normalized to N.
- (b) Kaplan-Meier survival analysis of 204 patients with lung squamocellular carcinoma (LUSC) obtained from the TCGA database. High expression (blue) or low expression (red) of *uc.339*; OS, overall survival; mo, months; TCGA, The Cancer Genome Atlas.
- (c) Location of *uc.339* on chromosome 12q13.13 and its relation with its host gene *ATP5G2*. Arrows specify the position of the primers used in this work for the detection of *uc.339* through qRT-PCR.

### Figure 2. *uc.339* induces cancer cell growth and migration.

- (a) Cell viability assay in LoVo and A549 cells infected with a lentiviral vector over-expressing *uc.339* (LV-*uc.339*) or its empty vector (LV-E) and detected after 72h. Data are shown as mean  $\pm$  s.d. of experiments tested in triplicate and normalized to LV-E. \* $P < 0.05$ .
- (b) Image of a scratch migration assay in A549 LV-E and A549 LV-*uc.339* cells at 0h, 8h and 24h.
- (c) Quantification of the proportion of wound area of the scratch migration assay observed in (b) and shown as mean  $\pm$  s.d. of experiments tested in triplicate. \*\* $P < 0.01$ .
- (d) Cell viability assay in A549, H460, LoVo and H1299 cells transfected with si-*uc.339* or si-*SCR* for 72h. Data are shown as mean  $\pm$  s.d. of experiments tested in triplicate and normalized to si-*SCR*. \* $P < 0.05$ .
- (e) Cell cycle analysis (shown as proportion of cells in G<sub>0</sub>/G<sub>1</sub> or G<sub>2</sub>/M or S phase of the cell cycle) tested by cytofluorimetry with propidium iodide staining in A549, H460, LoVo and H1299 cells transfected with si-*uc.339* or si-*SCR* for 72h. Data are shown as mean  $\pm$  s.d. of experiments tested in triplicate. \*\* $P < 0.01$ . \*\*\* $P < 0.001$ .

(f) Immunoblotting for cleaved PARP and Vinculin proteins in A549, H460, LoVo and H1299 cells transfected with si-*uc.339* or si-*SCR* for 72h.

**Figure 3. *In vivo* xenograft murine model *uc.339* induces NSCLC growth.**

(a) Curve of tumor volume growth in nude mice injected subcutaneously with A549 LV-E ( $n=7$ ) or A549 LV-*uc.339* ( $n=7$ ) and observed every 3 days from day 8 from the injection until day 29.

Data are shown as the mean volumes + s.d. for each animal set. \* $P<0.05$ . \*\* $P<0.01$ .

(b) qRT-PCR for analyzing *uc.339* expression in *ex vivo* xenografts from the same mice of the experiment (a). The expression of *uc.339* was normalized to *RNU44* and data are shown as mean  $\pm$  s.d. of experiments conducted in triplicate per each mouse and normalized to A549 LV-EV. \* $P<0.05$ .

(c) Images of the 14 tumors excised from the mice of experiments (a) and (b).

(D) Cell viability assay of A549 cells infected with a lentiviral plasmid over-expressing *ucr.339* (LV-*ucr.339*) or its empty plasmid counterpart (LV-E) after the Cisplatin treatment at peak plasma level. Data are shown as mean  $\pm$  s.d. of experiments conducted in triplicate and normalized to LV-E. \* $P<0.05$ .

**Figure 4. *uc.339* is a decoy for *miR-339-3p*, *95-5p* and *-663b-3p*.**

(a) *In silico* analysis of the interaction between *miR-339-3p*, *-95-5p* and *-663b-3p* (green) and *uc.339* (in red).

(b) qRT-PCR for *uc.339* in A549, H1299, H460 and LoVo cells transfected with *miR-339-3p*, *-95-5p* and *-663b-3p* or a scrambled miRNA (SCR) (after 48h from transfection). The expression of *uc.339* was normalized to *RNU44* RNA and shown as normalized to SCR. Data are shown as mean  $\pm$  s.d. of experiments in triplicate.

(c) qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in H460 and A549 cells infected with a lentiviral plasmid over-expressing *uc.339* (LV-*uc.339*) or its empty plasmid counterpart (LV-E) and analyzed after 72h. The expression of miRNAs was normalized to *RNU44* RNA and the results are shown as normalized to LV-E. Data are presented as mean  $\pm$  s.d. of experiments conducted in triplicate. \* $P < 0.05$ .

(d) qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in H1299 cells transfected with si-*SCR* or si-*uc.339* for 72h. The expression of miRNAs was normalized to *RNU44* RNA and the results are shown as normalized to si-*SCR*. Data are presented as mean  $\pm$  s.d. of experiments in triplicate. \* $P < 0.05$ .

(e) qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in H460 and A549 cells infected with LV-*uc.339* or LV-E, in which *uc.339* was immunoprecipitated. The expression of miRNAs is shown as normalized to the LV-E group. Data are presented as mean  $\pm$  s.d. of experiments in triplicate. \* $P < 0.05$ .

**Figure 5. *uc.339* induces the expression of *CCNE2* through sequestering of targeting miRNAs *miR-339-3p*, *-663b-3p* and *-95-5p*.**

(a) Luciferase reporter assay in LoVo and H1299 cells co-transfected with *miR-339-3p*, *-95-5p* and *-663b-3p* or a scrambled miRNA (SCR) and a reporter vector containing the 3'-UTR binding site (BS) for the miRNAs on the *CCNE2* mRNA (wt BS) or a mutant in which the binding site was deleted (del BS). Luciferase activity was normalized to Renilla (RLU) and the results are shown as normalized to SCR. Data are presented as mean  $\pm$  s.d. of experiments in sixtuplicate. \* $P < 0.05$ .

(b) Immunoblotting for *CCNE2* and Vinculin proteins in LoVo and H1299 cells transfected with *miR-339-3p*, *-95-5p* and *-663b-3p* or a scrambled miRNA (SCR) (after 48h from transfection).

(c) Immunoblotting for *CCNE2* and Vinculin proteins in A549, LoVo and H460 cells infected with LV-*uc.339* or LV-E.

(d) Immunoblotting for CCNE2 and Vinculin proteins in H1299 and H460 cells transfected with si-*uc.339* or si-SCR for 72h.

(e) Immunoblotting for CCNE2 and Vinculin proteins in H460 cells infected with LV-*uc.339* and transfected with *miR-339-3p*, *-95-5p* and *-663b-3p* or a scrambled miRNA (SCR) (after 48h from transfection).

The numbers above each lane indicate a quantification of the band intensity, normalized to the Vinculin band.

**Figure 6. TP53 silences directly *uc.339*.**

(a) qRT-PCR for *uc.339* in H1299 and A549 cells transfected through an anti-*TP53* siRNA (si-*TP53*) [or its anti-scrambled control (si-SCR)] or with a vector expressing *TP53* (or its empty vector counterpart) for 72h, respectively. The expression of *uc.339* was normalized to *RNU44* RNA and the results are shown as normalized to si-SCR or Empty, respectively. Data are shown as mean  $\pm$  s.d.. \* $P < 0.05$ .

(b) Expressions of TP53 protein (left panel) and of *uc.339* (right panel) in 22 paired primary NSCLC tumor tissues (C) and the adjacent normal lung (N). TP53 expression was quantified with chemiluminescence (Femto) Kit and normalized to Vinculin, while *uc.339* expression was determined by qRT-PCR and normalized to *RNU44*. Data are presented as mean  $\pm$  s.d. normalized to N. \* $P < 0.0001$ .

(c) Map showing the position of four identified consensus sequences (CS #1-4) of TP53 relative to the *uc.339* transcription starting nucleotide on chromosome 12.

(d) Chromatin Immunoprecipitation presenting the binding of TP53 to the four indicated *TP53* CS.

(e) Luciferase reporter assay in A549, LoVo and H460 cells transfected with a *TP53* expressing vector (or its empty vector counterpart) and a reporter vector containing the sequence of CS #4, or a vector in which CS #4 was deleted (del CS #4). Luciferase activity was normalized to

Renilla (RLU) and the results are shown as normalized to Empty. Data are presented as mean  $\pm$  s.d. of experiments in sixtuplicate. \* $P$ <0.05.

**Figure 7. TP53 regulates CCNE2 expression through *uc.339*.**

(a) qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in H1299 cells transfected with a vector expressing *TP53* or its empty vector counterpart (after 72h from the transfection). The expression of miRNAs was normalized to *RNU44* RNA and the results was presented as normalized to Empty. Data are presented as mean  $\pm$  s.d. of experiments in triplicate. \* $P$ <0.05.

(b) Immunoblotting for CCNE2 and Vinculin proteins in H1299 transfected with an empty vector (lane 1), the same vector expressing *TP53* (lane 2) or co-transfected with a vector expressing *TP53* and a vector expressing *uc.339* (lane 3) (after 72h from the transfection). The numbers above each lane indicate a quantification of the band intensity, normalized to the corresponding Vinculin protein band.

(c) qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in the same 22 primary NSCLC samples shown in **Fig. 6b**. The expression of miRNAs was normalized to *RNU44* RNA and the results are shown as normalized to N. Data are presented as mean  $\pm$  s.d. of experiments in triplicate per each patient. \* $P$ <0.05.

(d) Expression of CCNE2 protein in the same 22 primary NSCLC samples shown in **Fig. 6b**. CCNE2 expression was quantified with by chemiluminescence (Femto) kit, normalized to Vinculin protein, and shown as normalized to N. Data are presented as mean  $\pm$  s.d.. \* $P$ <0.05.

## SUPPLEMENTARY FIGURE LEGENDS

### Supplementary Figure 1. Endogenous expression in cell lines and Cloning of *uc.339*.

(a) *uc.339* endogenous expression in A549, H460, LoVo and H1299 cells, as shown by qRT-PCR. The expression of *uc.339* was normalized to *RNU44* and the results are shown as mean  $\pm$  s.d. of experiments in triplicate.

(b) Sequence of the *uc.339* transcript, as resulted by RACE. Highlighted in grey is the sequence described by Bejerano et al. [7].

### Supplementary Figure 2. *uc.339* increases NSCLC growth in cell lines.

(a) qRT-PCR for *uc.339* in LoVo and A549 infected with a lentiviral plasmid over-expressing *uc.339* (LV-*uc.339*) or its empty plasmid counterpart (LV-E) and analyzed after 72h.

(b-c) *uc.339* expression of qRT-PCR for in A549, H1299, H460, and LoVo cells transfected with two diverse anti-*uc.339* siRNA [*si-uc.339* and *si-uc.339* (2)], or an anti-scrambled siRNA (*si-SCR*) (for 72h).

The expression of *uc.339* was normalized to *RNU44* and the results are shown as mean  $\pm$  s.d. of experiments in triplicate and normalized to *si-SCR*. \* $P < 0.05$ .

(d) Cell viability assay in A549, H460, LoVo and H1299 cells transfected with *si-uc.339* (2) or *si-SCR* (for 72h). Data are shown as mean  $\pm$  s.d. of experiments in triplicate and normalized to *si-SCR*. \* $P < 0.05$ .

(e) Cell cycle analysis (indicated as proportion of cells in G<sub>0</sub>/G<sub>1</sub> or G<sub>2</sub>/M or S phase of the cell cycle) performed by cytofluorimetry with propidium iodide staining in A549, H460, H1299 and LoVo cells transfected with *si-uc.339* (2) or *si-SCR* (for 72h). Data are shown as mean  $\pm$  s.d. of experiments in triplicate. \*\* $P < 0.01$ . \*\*\* $P < 0.001$ .

(f) Immunoblotting for cleaved PARP and Vinculin proteins in A549, H460, LoVo and H1299 cells transfected with *si-uc.339* (2) or *si-SCR* (for 72h).

**Supplementary Figure 3. Expression of *miR-339-3p*, *-95-5p* and *-663-3p* in H1299 cell line silenced with an siRNA anti-*uc.339*.**

qRT-PCR for *miR-339-3p*, *-95-5p* and *-663b-3p* in H1299 cells transfected with si-*uc.339* (2) or si-*SCR* (for 72h). The expression of miRNAs was normalized to *RNU44* RNA and the results are shown as normalized to si-*SCR*. Data are presented as mean  $\pm$  s.d. of experiments in triplicate. \* $P < 0.05$ .

**Supplementary Figure 4. Expression of protein CCNE2 in H1299 and H460 silenced with an siRNA anti-*uc.339*.**

Immunoblotting for CCNE2 and Vinculin proteins in H460 and H1299 cells transfected with si-*uc.339* (2) or si-*SCR* (for 72h).

**Supplementary Figure 5. mRNA and protein endogenous expression of *TP53* in cell lines.**

(a) qRT-PCR for *TP53* in A549, H460, LoVo and H1299 cells. The expression of *TP53* was normalized to *HPRT1* and the results are shown as mean  $\pm$  s.d. of experiments in triplicate relative to A549 cells.

(b) Immunoblotting for *TP53* and Vinculin in A549, H460, LoVo and H1299 cells. The numbers above each lane indicate a quantification of the band intensity, normalized to the corresponding Vinculin protein band.

**Supplementary Figure 6. *TP53* mRNA expression in transfected cell lines.**

qRT-PCR for *TP53* in H1299 and A549 cells transfected with anti-scrambled siRNA (si-*SCR*) or an anti-*TP53* siRNA (si-*TP53*) (A549, left panel) or with a vector expressing *TP53* or its empty vector counterpart (Empty) (H1299, right panel) (after 72h from the transfection). The expression of *TP53* was normalized to *HPRT1* and the results are shown as mean  $\pm$  s.d. of experiments in triplicate and normalized to si-*SCR* or Empty, respectively. \* $P < 0.05$ .

**Supplementary Figure 7. Summary of the newly determined TP53-*uc.339*-*miRNA*-CCNE2 network.**

TP53 directly down-regulates the expression of *uc.339* . *uc.339* acts as a decoy for CCNE2 targeting *miR-339-3p*, *95-5p* and *-663b-3p* determining to up-regulation of CCNE2 and increased migration and tumor growth.

**Table 1. Clinical characteristics of the 30 NSCLC patients analyzed in this work (TP53 status for all patients: wild-type).**

|                                 |                       | Patient # |   |
|---------------------------------|-----------------------|-----------|---|
| <b>Sex</b>                      | <i>M</i>              | 25        |   |
|                                 | <i>F</i>              | 5         |   |
| <hr/>                           |                       |           |   |
| <b>Average Age at Diagnosis</b> | 68.69 yo [54-82]      |           |   |
| <hr/>                           |                       |           |   |
| <b>Histology</b>                | <i>Squamous</i>       | 2         |   |
|                                 | <i>Adenocarcinoma</i> | 28        |   |
| <hr/>                           |                       |           |   |
| <b>TNM</b>                      | <i>T1a</i>            | 2         |   |
|                                 | <i>T1b</i>            | 7         |   |
|                                 | <i>T2a</i>            | 10        |   |
|                                 | <i>T2b</i>            | 6         |   |
|                                 | <i>T3</i>             | 4         |   |
|                                 | <i>T4</i>             | 1         |   |
|                                 | <i>N0</i>             | 13        |   |
|                                 | <i>N1</i>             | 13        |   |
|                                 | <i>N2</i>             | 4         |   |
|                                 | <i>N3</i>             | 0         |   |
|                                 | <i>M0</i>             | 29        |   |
|                                 | <i>M1</i>             | 1         |   |
|                                 | <hr/>                 |           |   |
| <b>Stage</b>                    | <i>IA</i>             | 5         |   |
|                                 | <i>IB</i>             | 3         |   |
|                                 | <i>IIA</i>            | 11        |   |
|                                 | <i>IIB</i>            | 4         |   |
|                                 | <i>IIIA</i>           | 6         |   |
|                                 | <i>IIIB</i>           | 0         |   |
|                                 | <i>IV</i>             | 1         |   |
| <hr/>                           |                       |           |   |
| <b>Treatment</b>                | <i>Neo-Adjuvant</i>   | 0         |   |
|                                 | <i>Adjuvant</i>       | 3         | CDDP+Gemcitabine (1);<br>CDDP+Vinorelbine (2) |

Figure 1.

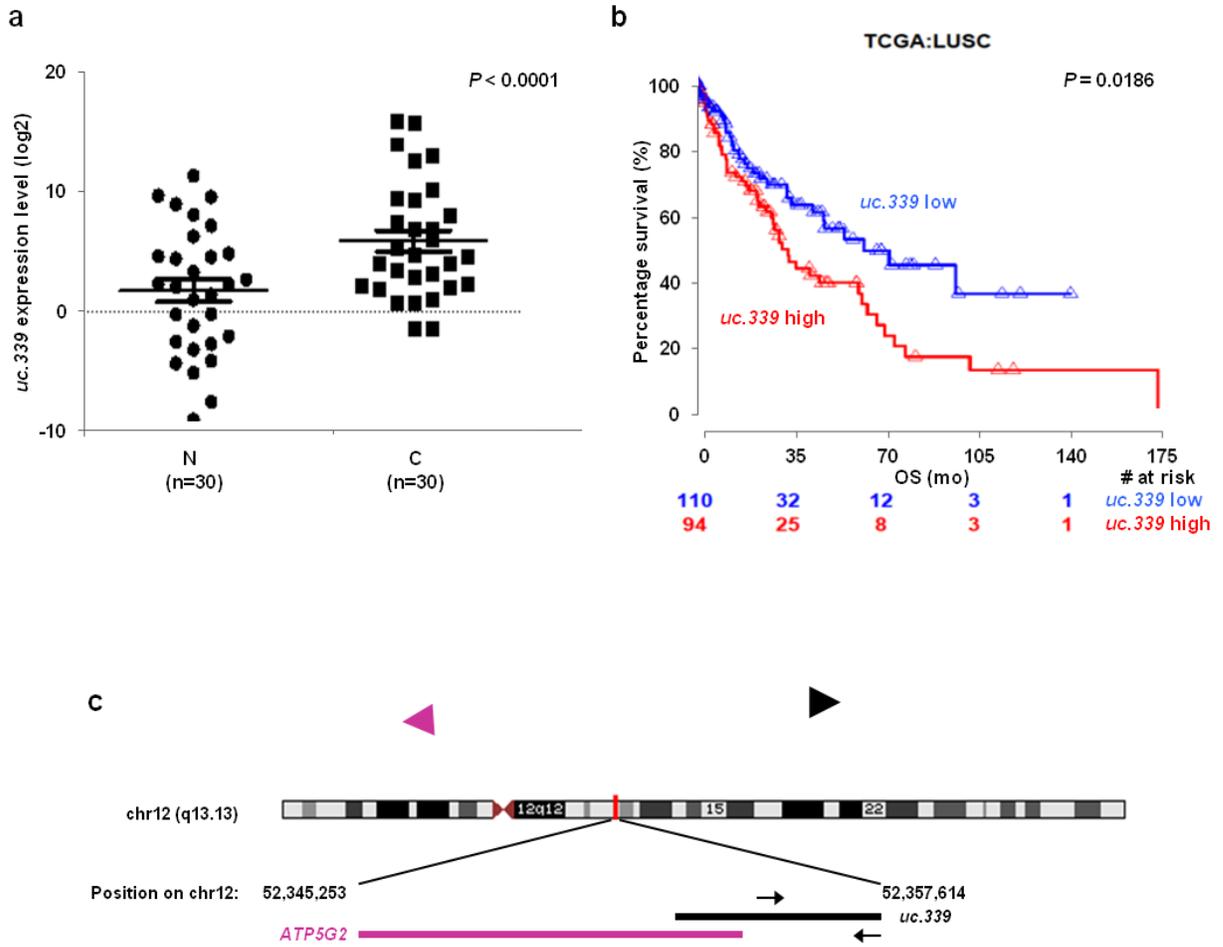


Figure 2.

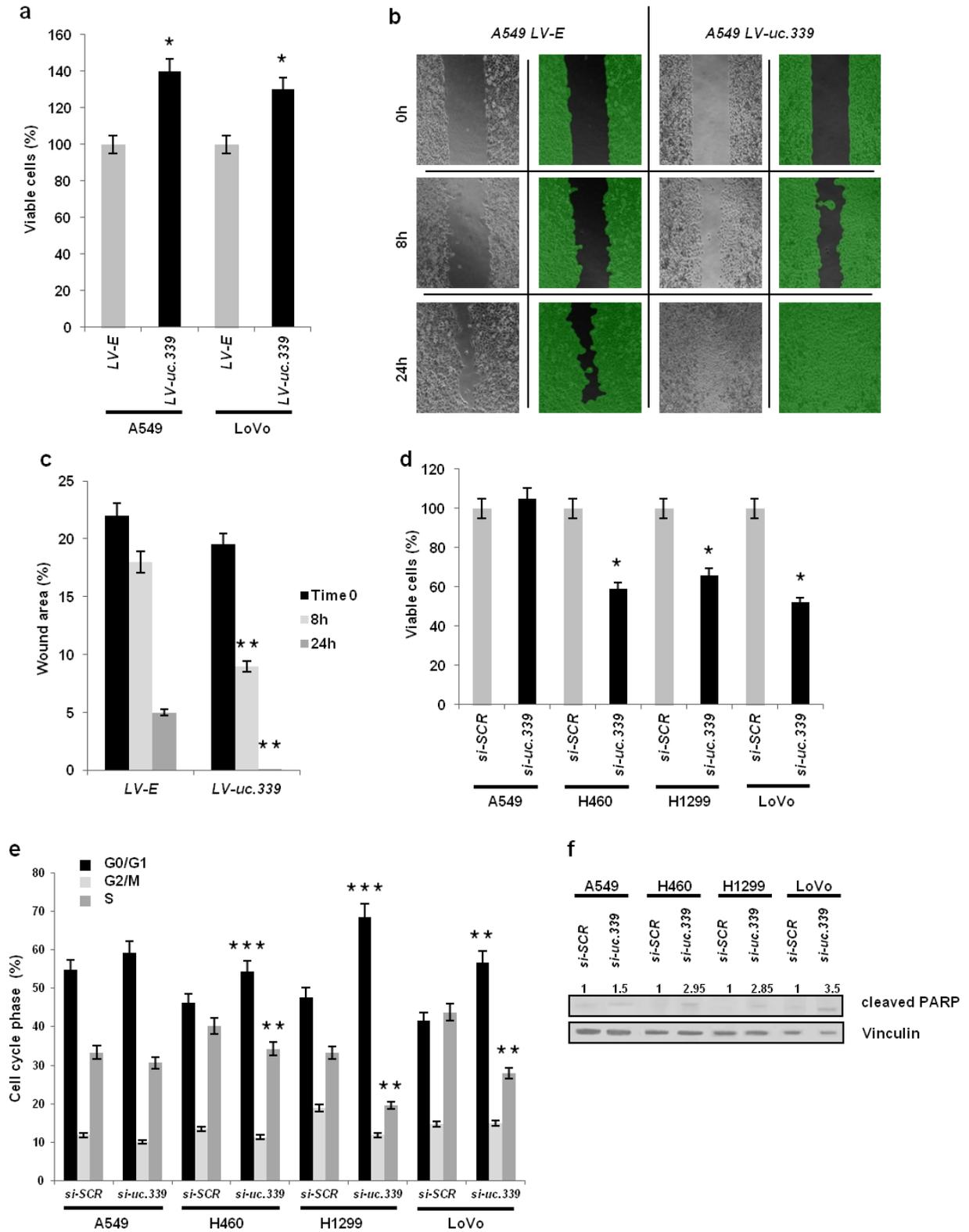


Figure 3.

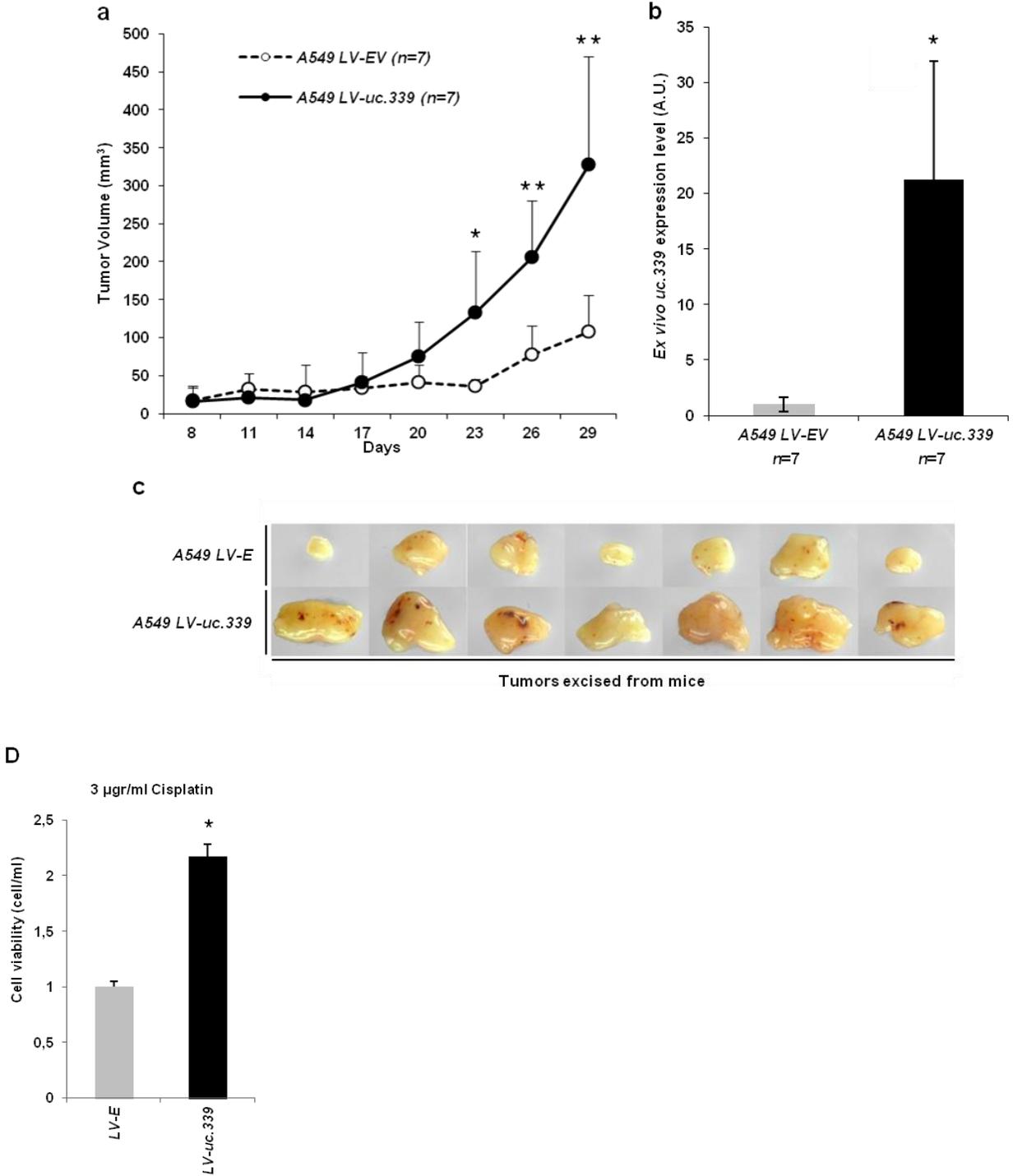


Figure 4.

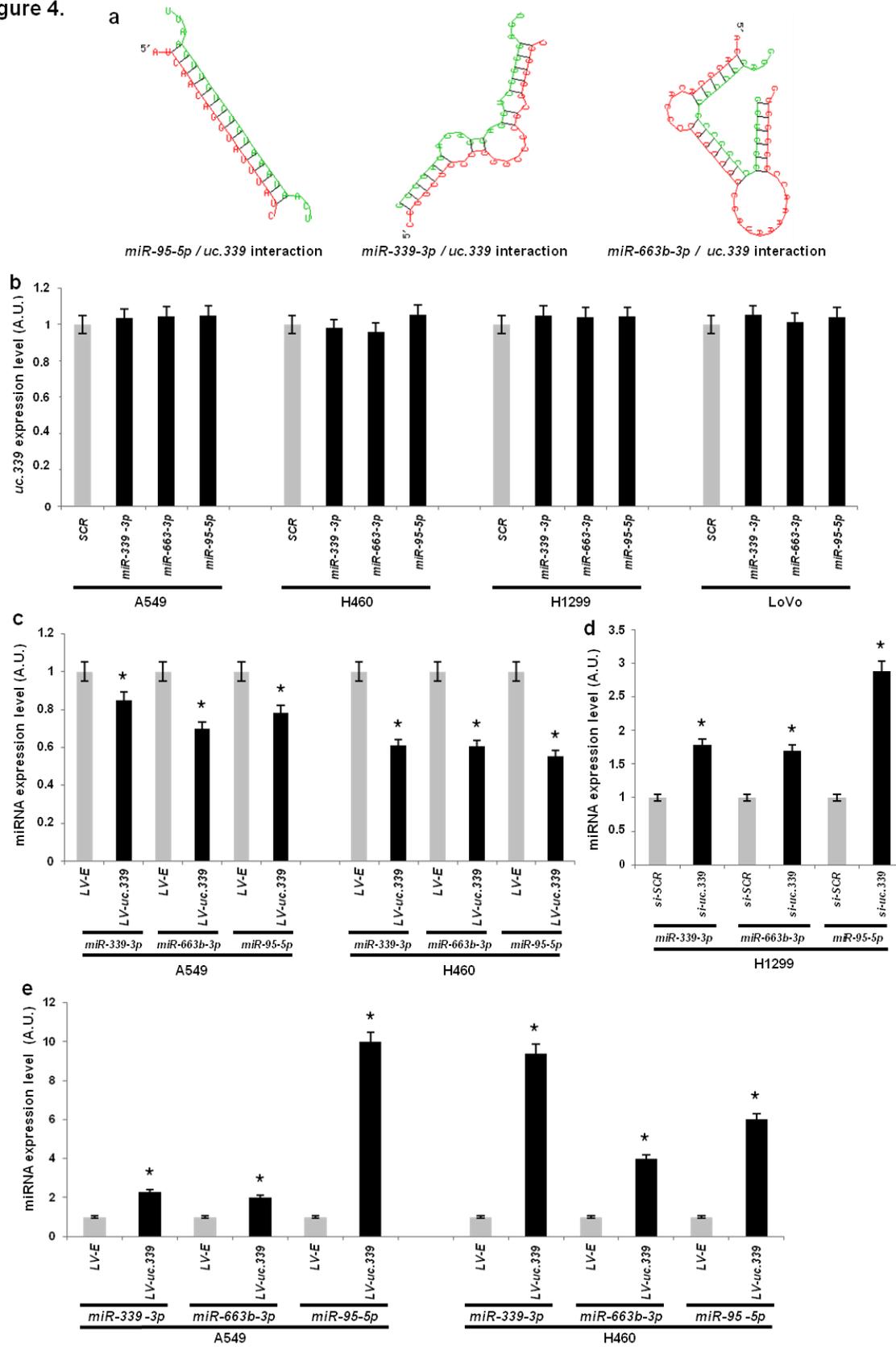


Figure 5.

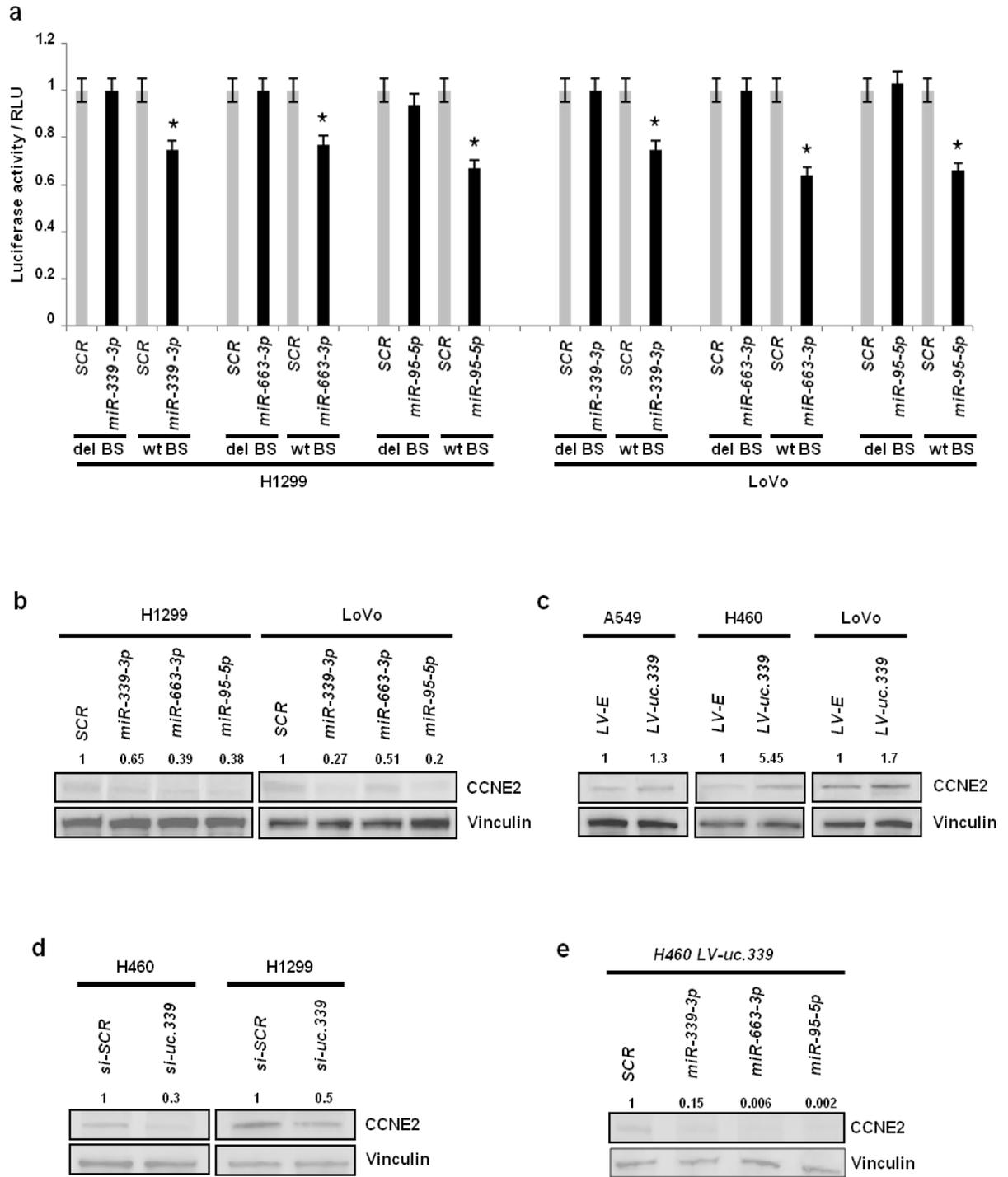


Figure 6.

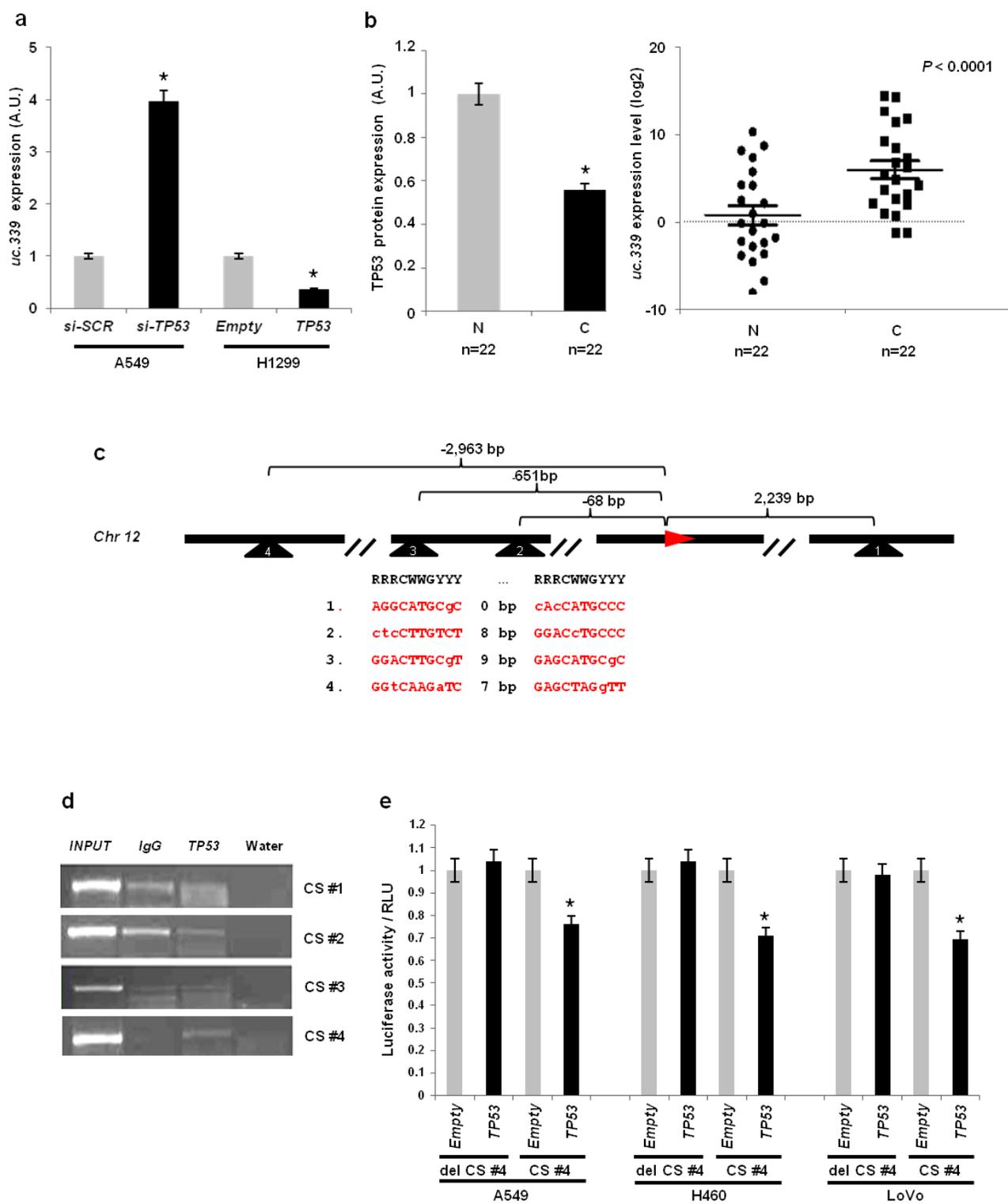
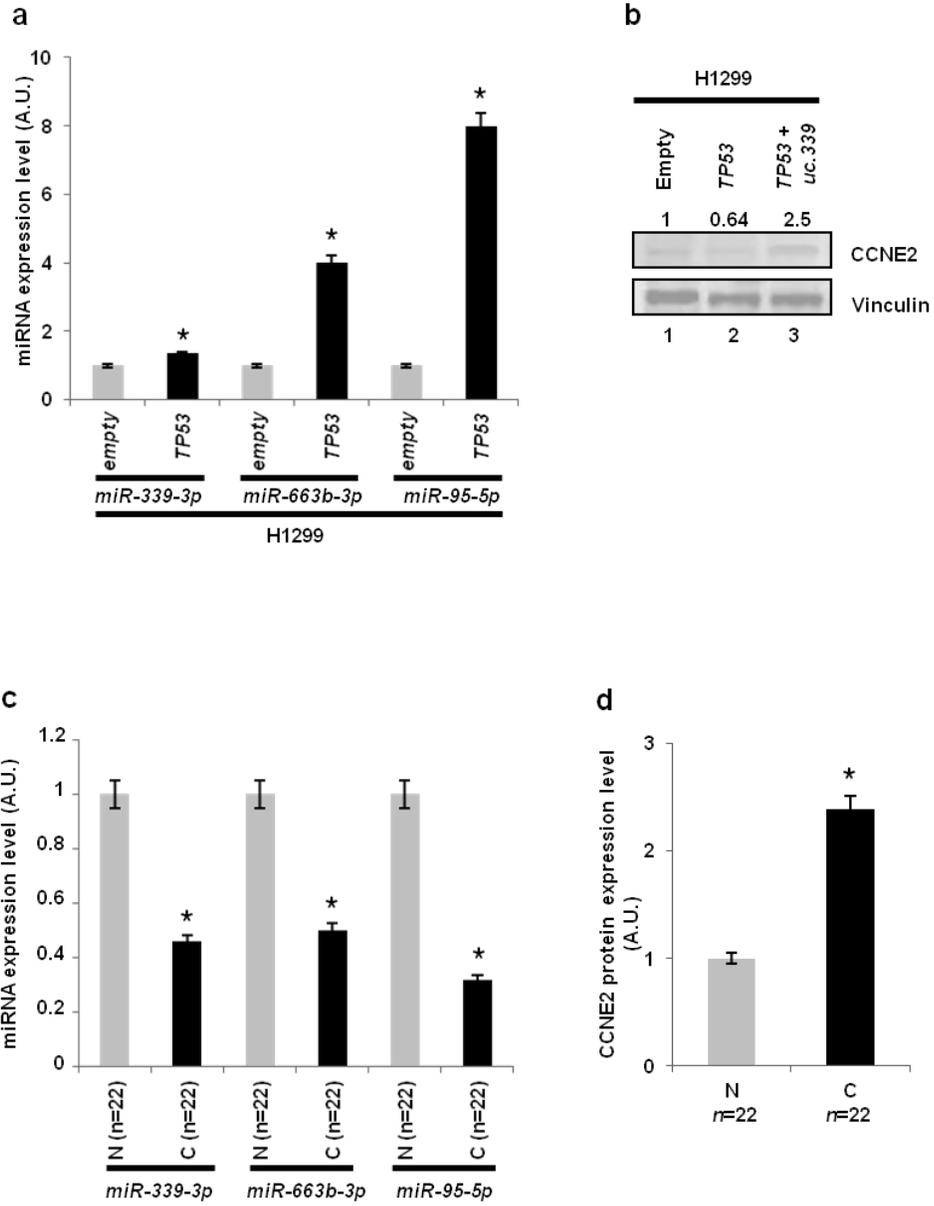


Figure 7.

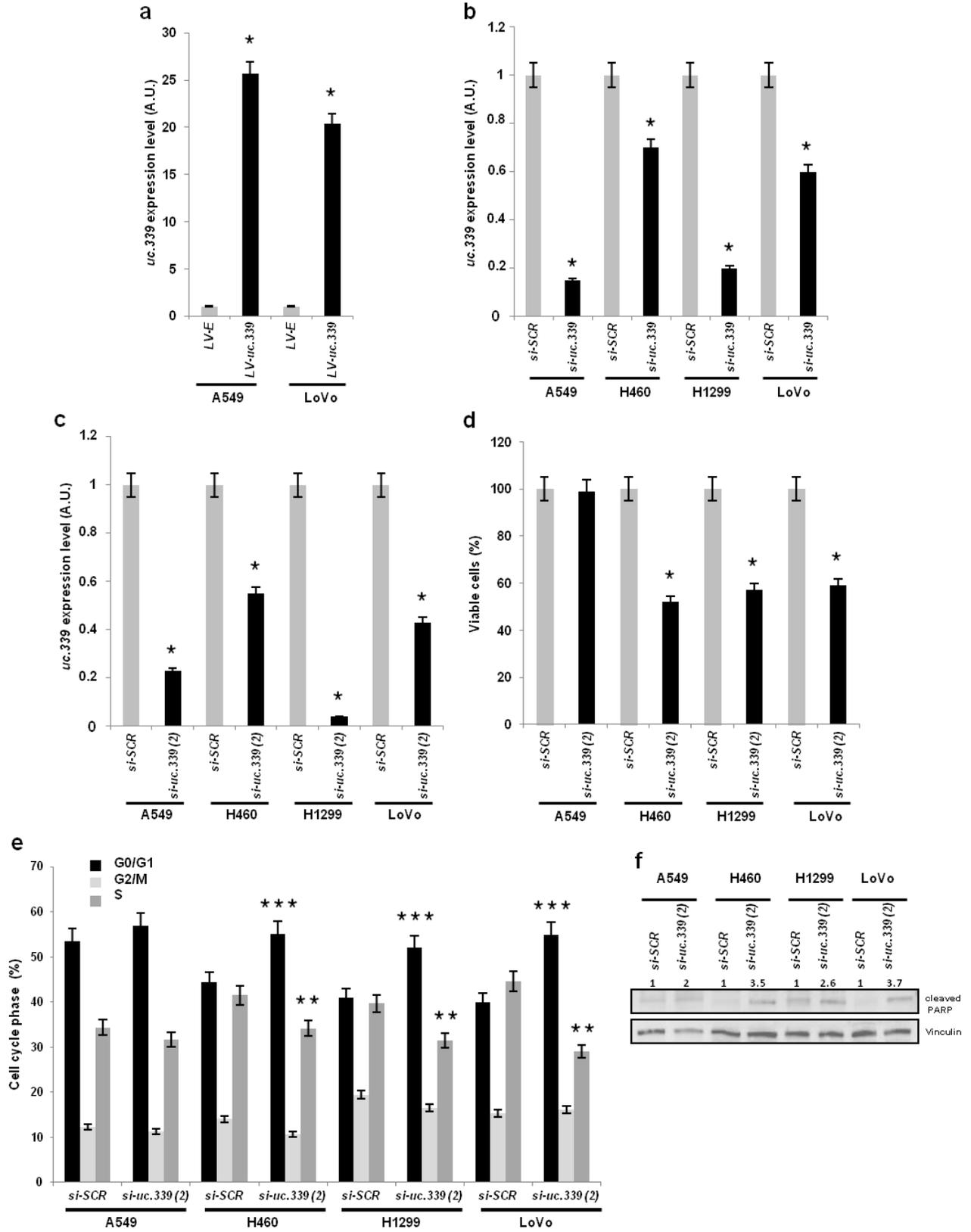


SUPPLEMENTARY FIGURES

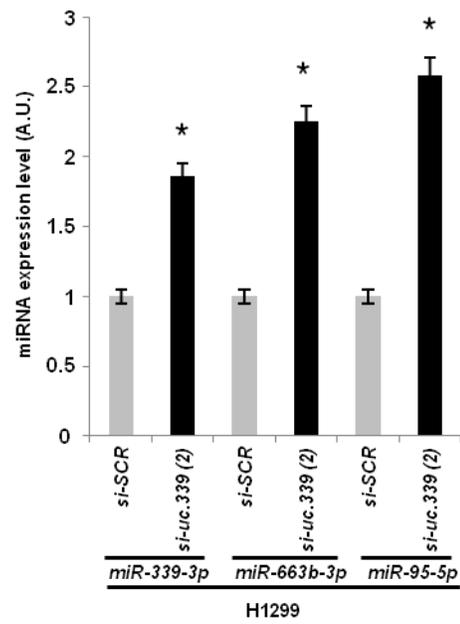
Supplementary Figure 1.



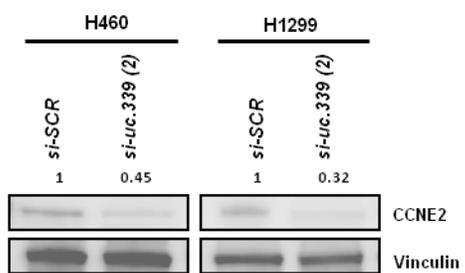
Supplementary Figure 2.



### Supplementary Figure 3.

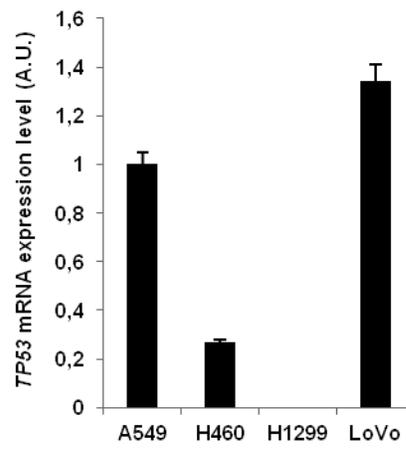


### Supplementary Figure 4.

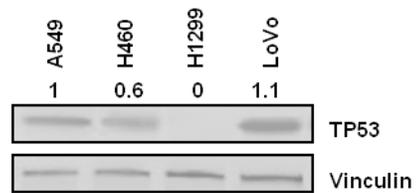


## Supplementary Figure 5.

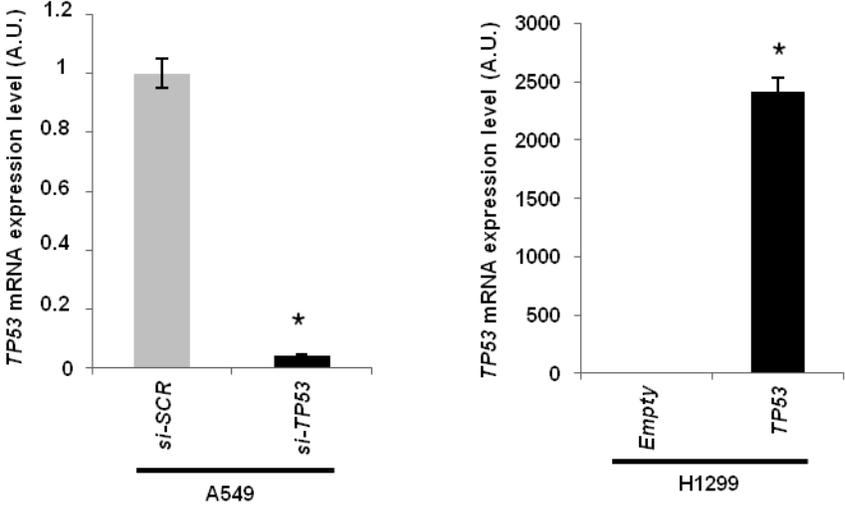
**a**



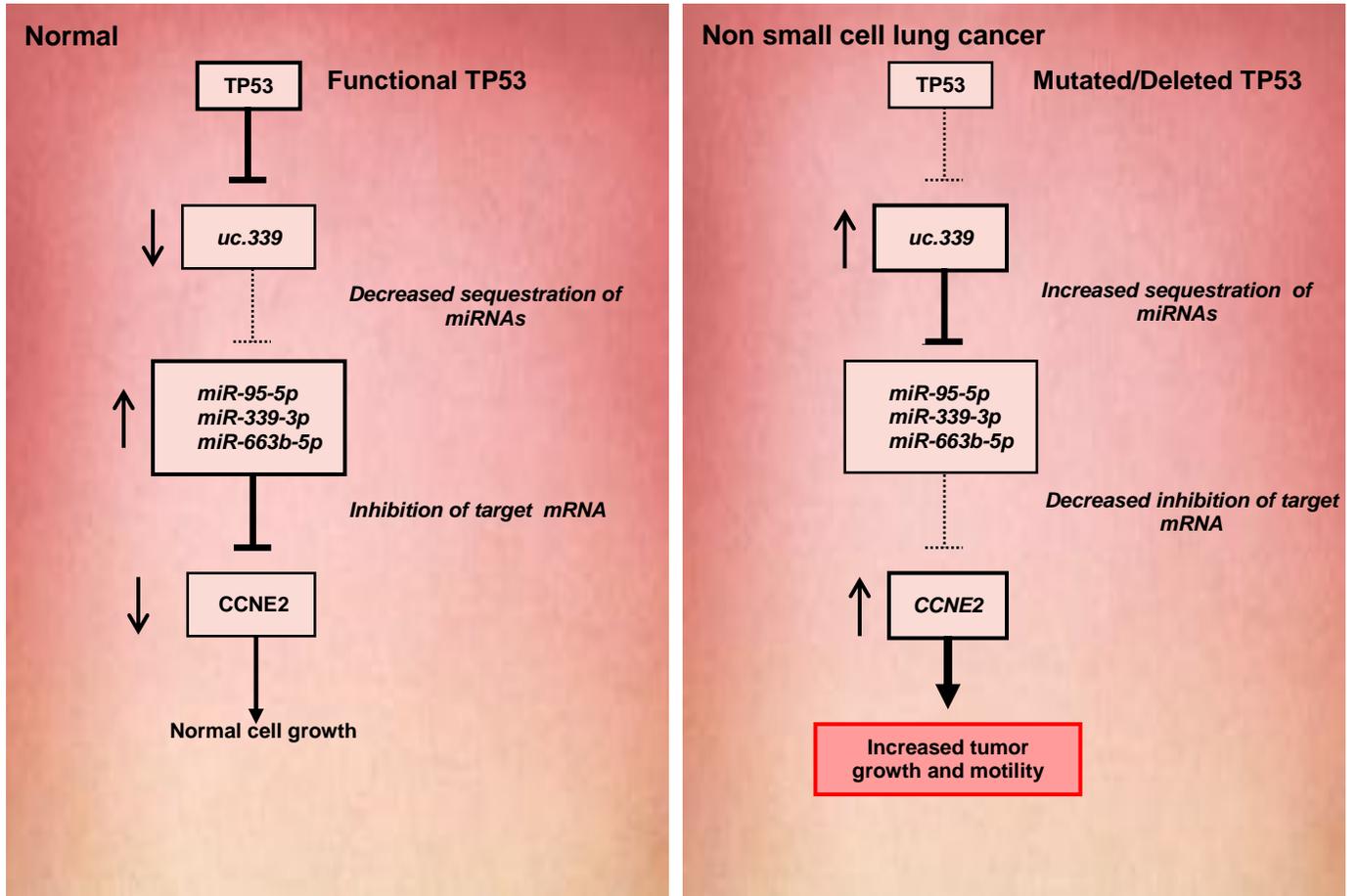
**b**



Supplementary Figure 6.



Supplementary Figure 7.



## **ACKNOWLEDGEMENTS**

**I thank Prof. Giorgio Cantelli Forti for supporting me during these years in my thesis project and Prof. Muller Fabbri and Prof. George A. Calin for the teachings in research.**

**I thank all the people that I worked in this research project.**

**A special thanks to my family.**

**Dr. Ivan Vannini**

# INDEX

➤ **Page 2, INTRODUCTION**

➤ **Page 8, AIMS**

➤ **Page 10, METHODS**

➤ **Page 23, RESULTS**

➤ **Page 31, DISCUSSION**

➤ **Page 35, REFERENCES**

➤ **Page 40, FIGURES**