Non-coding RNA Research 9 (2024) 759–771

Contents lists available at ScienceDirect

Non-coding RNA Research

journal homepage: www.keaipublishing.com/en/journals/non-coding-rna-research

Original Research Article

"Crosstalk between non-coding RNAs and transcription factor LRF in non-small cell lung cancer"

Magda Spella [a,b], Eleftherios Bochalis [a], Katerina Athanasopoulou [a], Argyri Chroni [a], Irene Dereki [a], Giannoula Ntaliarda [b], Ifigeneia Makariti [a], Georgios Psarias [a], Caterina Constantinou [a], Vasiliki Chondrou [a], Argyro Sgourou [a,*]

[a] Biology Laboratory, School of Science and Technology, Hellenic Open University, 26335 Patras, Greece
[b] Department of Physiology, Faculty of Medicine, University of Patras, Rio, 26504, Greece

ARTICLE INFO

Keywords:
Non-small cell lung cancer
Epigenetic deregulation
Micro-RNAs
Long non-coding RNAs
Transcription factor LRF
Epigenetic networks

ABSTRACT

Epigenetic approaches in direct correlation with assessment of critical genetic mutations in non-small cell lung cancer (NSCLC) are currently very intensive, as the epigenetic components underlying NSCLC development and progression have attained high recognition. In this level of research, established human NSCLC cell lines as well as experimental animals are widely used to detect novel biomarkers and pharmacological targets to treat NSCLC. The epigenetic background holds a great potential for the identification of epi-biomarkers for treatment response however, it is highly complex and requires precise definition as these phenomena are variable between NSCLC subtypes and systems origin.

We engaged an in-depth characterization of non-coding (nc)RNAs prevalent in human *KRAS*-mutant NSCLC cell lines A549 and H460 and mouse *KRAS*-mutant NSCLC tissue by Next Generation Sequencing (NGS) and quantitative Real Time PCRs (qPCRs). Also, the transcription factor (TF) LRF, a known epigenetic silencer, was examined as a modulator of non-coding RNAs expression. Finally, interacting networks underlying epigenetic variations in NSCLC subtypes were created. Data derived from our study highlights the divergent epigenetic profiles of NSCLC of human and mouse origin, as well as the significant contribution of 12qf1: 109,709,060–109,747,960 mouse chromosomal region to micro-RNA upregulated species. Furthermore, the novel epigenetic miR-148b-3p/lnc*PVT1/ZBTB7A* axis was identified, which differentiates human cell line of lung adenocarcinoma from large cell lung carcinoma, two characteristic NSCLC subtypes.

The detailed recording of epigenetic events in NSCLC and combinational studies including networking between ncRNAs and TFs validate the identification of significant epigenetic features, prevailing in NSCLC subtypes and among experimental models. Our results enrich knowledge in the field and empower research on the epigenetic prognostic biomarkers of the disease progression, NSCLC subtypes discrimination and advancement to patient-tailored treatments.

1. Introduction

Non-small cell lung cancer (NSCLC) is a leading cause for high morbidity and mortality rates worldwide. NSCLC can be further divided into three major histological subtypes: lung adenocarcinoma (LUAD), large cell carcinoma, and squamous cell carcinoma with distinct phenotypes [1]. On the genetic level, coexistence of activating mutations of the Kirsten rat sarcoma viral oncogene homologue (*KRAS*) is a common NSCLC hallmark [2]. The molecular signatures of candidate genes, epigenetically regulated by differential DNA methylation profiles have also been implicated in NSCLC as complements to already known driver mutations. Among reported epigenetically modified DNA gene loci, genes encoding non-coding RNAs (ncRNAs) are also included and represent features of high importance for the NSCLC progression and

* Corresponding author.
E-mail addresses: magsp@upatras.gr, spella.magdalini@ac.eap.gr (M. Spella), left.bochalis@gmail.com (E. Bochalis), athanasopoulou.aikaterini@ac.eap.gr (K. Athanasopoulou), argychroni@gmail.com (A. Chroni), irenenter@gmail.com (I. Dereki), ntaliarda@upatras.gr (G. Ntaliarda), i.makariti@sbbio.gr (I. Makariti), geo.psarias@gmail.com (G. Psarias), konstantinou.aikaterini@ac.eap.gr (C. Constantinou), vchondrou@eap.gr (V. Chondrou), sgourou@eap.gr (A. Sgourou).

https://doi.org/10.1016/j.ncrna.2024.03.009
Received 22 December 2023; Received in revised form 23 February 2024; Accepted 20 March 2024
Available online 23 March 2024

response to drugs [3]. Available therapies are challenging and upon administration offer extended life just by a few months. Platinum-containing drugs such as cisplatin and carboplatin are typically used as the first-line chemotherapeutic agents for the treatment of NSCLC, sometimes in combination with surgical excision and chest radiotherapy [4]. Nevertheless, advanced stages of the disease often evolve resistance to chemotherapeutics and within this frame strong research efforts and innovative approaches are necessary to improve patients' quality of life and survival, since the efficacies of current therapies are limited.

Apart from driver mutations characterized as a leading cause of NSCLC, subtle combinations of epigenetic events with hereditary potential are also associated with NSCLC progression and chemotherapeutic resistance development. Aberrant epigenetic signatures have a realistic prospect and can serve as a pragmatic approach towards the identification of functional epigenetic events and potential circuits guiding cancer development and progression. Epigenetic marks include histone code and DNA modifications simultaneously with non-coding RNA (ncRNA) tissue-specific production, accounting for the active epigenome of human tissues and playing fundamental roles in cancer [5]. These biochemical events, contrary to the established genetic mutations, sometimes represent only a snapshot of cellular life during the cell cycle and can be reversible, justifying their importance as promising therapeutic targets.

Epigenetically altered chromatin supports positively or negatively ncRNAs' transcription depending on the conditions prevailing across genetic loci coding for the various ncRNA species. Aberrant micro-RNA (miRNA) expression has been extensively studied and miRNAs are proposed as lung cancer subtype classifiers [6], prognostic and drug-response biomarkers [7,8] and as targets for novel miRNA-based treatment approaches [9]. Clinical implementation strategies and potential of lncRNAs in NSCLC have also been extensively reported [10]. MiRNA clusters and lncRNAs are considered mainly as products of RNA polymerase II and therefore, the transcription of these primary transcripts is regulated by general TFs and signal transduction pathways. Deregulation of TFs' expression and consequent tissue-imbalanced lncRNA expression is often sufficient for the development of several human disorders and for malignant cell transformation [11]. In addition to sharing certain characteristics with "ordinary" genes, such as splicing, polyadenylation etc., many genetic loci transcribing miRNAs and lncRNAs include CpG islands with conditional regulatory role on their expression. Also, genetic loci of lncRNAs encompass miRNA genes with contingent coregulation of both species by the same elements. A well-known epigenetically motivated TF, the lymphoma/Leukemia related factor (LRF) belongs to an evolutionarily conserved family of TFs with zinc finger domains, which have been assigned with regulatory properties concerning gene expression and binding preference at CG-rich and CpG island containing promoters. LRF derives from *ZBTB7A* gene and recently we investigated its capability of regulating lncRNAs [12,13] on the transcriptional level during erythropoiesis while, it has been also involved in various cancer types development [14]. Herein, all above mentioned epigenetic components constitute a network underlying critical alterations that deceitfully function in NSCLC. So far, extensive research associated with the NSCLC therapeutics has employed human samples, established human malignant cell lines and mainly mice as experimental system models. However, investigation in living cells (cell lines and mouse models) that display unique genetic/epigenetic characteristics should be cautiously interpreted for the specificity of imprinted results and efficacy of recommended treatments for humans.

In this study, we have analyzed and compared epigenetic features exhibited by different NSCLC experimental models, to determine and characterize significant epigenetic signatures. We have comprehensively studied miRNA profiles in two *KRAS*-mutant NSCLC human cell lines (A549 and H460) versus BEAS-2B bronchial epithelial cells, and in *Kras*-mutant NSCLC mouse lung malignant versus adjacent normal tissue. Additionally, we have assessed methylation and expression profiles of human lncRNAs with known altered patterns in NSCLC. Our results strongly point out different mechanisms of carcinogenesis evolution between mouse NSCLC tissue and human NSCLC cell lines (A549 and H460), simultaneously with considerable epigenetic diversity among the human cell lines, reflecting distinct malignant histological subtype-dependent behaviors. Also, our results support the existence of a miRNA/lncRNA/*ZBTB7A* epigenetic network in human large-cell lung carcinoma H460 cell line, leading to the suppression of LRF/*ZBTB7A* expression. The present study provides comprehensive and elaborate information on the epigenetic features prevailing in each tested experimental model defined as NSCLC. This kind of information is considered critical during selection of appropriate biological models for studies and drug discovery and offers new knowledge in the field of biomarkers investigation.

2. Materials and methods

2.1. Cell culture

The A549 ($KRAS^{G12S}$) and H460 ($KRAS^{Q61H}$) NSCLC cell lines of human origin as well as HEK293 (human embryonic kidney cells) were cultured in a humidified atmosphere with 5% CO2 at 37 °C in Dulbecco's Modified Eagle Medium (DMEM, Sigma). The medium was supplemented with 10% fetal bovine serum (FBS, Gibco—Thermo Fisher Scientific Inc., Grand Island, NE, USA), 100 units/mL penicillin, and 100 μg/mL streptomycin (Sigma -Aldrich Pty Ltd, An affiliate of Merck KGaA, Darmstadt, Germany). The cells were passaged every two days. A549 and H460 cell lines have the KRAS mutations c.34G > A (ClinVar = VCV000012584) and c.183A > T (ClinVar = VCV000045117) respectively, of the most prevalent mutations verified in NSCLC [15].

Human bronchial epithelial cell line BEAS-2B was cultured in Bronchial Epithelial Cell Growth Medium (BEGM) consisting of Bronchial Epithelial Cell Basal Medium (BEBM, Lonza Group AG, Basel, Switzerland Catalog No.: CC-3171), supplemented with 0.4% Bovine Pituitary Extract (BPE), 0.1% Insulin, 0.1% Hydrocortisone, 0.1% GA-1000, 0.1% Retinoic Acid, 0.1% Transferrin, 0.1% Triiodothyronine, 0.1% Epinephrine, 0.1% human Epidermal Growth Factor (hEGF) (BEGM SingleQuots kit, Lonza Group AG, Basel, Switzerland, Catalog No.: CC-4175) and 100 units/mL penicillin and 100 μg/mL streptomycin (Sigma -Aldrich Pty Ltd, An affiliate of Merck KGaA, Darmstadt, Germany). Gentamicin/Amphotericin-B, present as GA-1000 in the BEGM SingleQuots kit, was added 3 days after culture starting date, to minimize excessive cell stress originating from cryopreservation. The cells were passaged every 2 days. Upon confluency, genomic DNA and total RNA were extracted for downstream analysis.

2.2. Mouse models of NSCLC

In our experimental design 6-week-old FVB/NJ (FVB; #001800) mice from Jackson Laboratories (Bar Harbor, MN) were introduced. Mice were bred at the Center for Animal Models of Disease of the University of Patras, maintained on a regular light-dark cycle (lights on from 8:00 a.m. to 8:00 p.m.) at an ambient temperature of 22 ± 1 °C and allowed free access to standard rodent chow diet and water.

All mice received 1 g/kg BW urethane via one intraperitoneal injection and were euthanized 6 months after urethane administration. Urethane is a chemical carcinogen contained in tobacco smoke and this regime is successful in triggering NSCLC (notably LUAD with $KRAS^{Q61R}$ mutations) in mice, as previously described [16,17]. After euthanasia, tissues of interest (liver, spleen and lung, including both lung tumors and matching adjacent normal tissue) were collected, washed in cold PBS, snap frozen in liquid nitrogen and stored for subsequent analysis. For histological confirmation of malignancy, lung samples, both malignant and normal were inflated and fixed with 10% formalin overnight and were initially inspected macroscopically under a Stemi DV4 stereoscope

equipped with an AxiocamERc 5s camera (Zeiss, Jena, Germany) in *trans*-illumination mode. Lung samples were then embedded in paraffin, and five-mm-thick paraffin sections were counterstained with hematoxylin and eosin (Sigma, St. Louis, MO) and mounted with Entellan New (Merck Millipore, Darmstadt, Germany). Bright-field images were captured with an AxioLab.A1 microscope connected to an AxioCamERc 5s camera (Zeiss, Jena, Germany).

All animal experiments were approved *a priori* by the Veterinary Administration of the Prefecture of Western Greece (approval numbers 3741/16.11.2010, 60,291/3035/19.03.2012, and 118,018/578/30.04.2014) and conducted according to Directive 2010/63/EU (http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1486710385917&uri=CELEX:32010L0063).

2.3. Selection of human lncRNAs implicated in lung cancer

Human lncRNAs selected for the present study have been previously related to lung cancerous state in humans. The study was limited to human tissues because there is no evolutionary conservation between human and mice lncRNAs. Among selected lncRNAs, HOXA Antisense RNA (*HOTAIRM1*), Maternally Expressed 3 (*MEG3*), Differentiation Antagonizing Non-Protein Coding RNA (*DANCR*), Plasmacytoma Variant Translocation 1 (*PVT1*), Metastasis Associated Lung Adenocarcinoma Transcript 1 (*MALAT1*) and Imprinted Maternally Expressed Transcript *H19* [18–24] are annotated by several databases (ENA, LncBook, Ensembl, LNCipedia, Ensembl/GENCODE, GeneCards). Sequences and chromosomal locations of CpG islands, mainly at transcriptional start site (TSS) of lnc-RNA sequences, were tracked from UCSC Genome browser (https://genome-euro.ucsc.edu/cgi-bin/hgGateway) (Table 1).

2.4. CpG methylation analysis of human origin lncRNAs assessed by pyrosequencing

Pyrosequencing methylation assay was utilized to assess DNA methylation status across CpG islands located at 5′ prime ends of lncRNAs' genetic loci. Genomic DNA was extracted from human A549, H460 and BEAS-2B cell lines with phenol:chloroform:isoamyl alcohol in a 25:24:1 ratio following standard protocol (Sigma-Aldrich Pty Ltd, An affiliate of Merck KGaA, Darmstadt, Germany) [25]. 1.5 μg of isolated genomic DNA was bisulfite converted, PCR amplified and sequenced with PyroMark® technology (QIAGEN GmbH, Hilden, Germany), according to manufacturer's instructions and as previously described [26]. Briefly, the biotinylated strand of the PCR products was immobilized by streptavidin-coated sepharose beads (Streptavidin Sepharose™ High Performance, GE HealthCare), and subsequently annealed with a Pyrosequencing primer. Pyrosequencing process was performed using the PyroMark® Gold Q24 Reagents (QIAGEN GmbH, Hilden, Germany) and PyroMark Q24 instrument. Primer sets are listed in Supplementary Table S1 and every reaction was performed in duplicate.

2.5. Expression profiles of human origin lncRNAs assessed by qPCRs

Total RNA was extracted from A549, H460, HEK293 and BEAS2B cultivated cells using the miRNeasy Mini kit (74,104, Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. 1 μg of total RNA was converted to cDNA by miScript II RT kit (218,161, Qiagen GmbH, Hilden, Germany). LRF/*ZBTB7A* and lncRNAs expression levels were assessed by qPCRs with sequence specific primer sets (Supplementary Table S1), designed to support detection of the main transcripts of each lncRNA genetic loci. Reactions performed in Rotor-Gene Q instrument (QIAGEN GmbH, Hilden, Germany) using the PowerTrack SYBRGreen Master Mix (A46012, Thermo Fisher Scientific, Waltham, MA, USA). Each reaction was performed in triplicate. The $2-\Delta\Delta CT$ method or Pfaffl equation with PCR efficiency correction was used for data analysis [27,28]. Fold changes in gene expression were normalized to the *GAPDH* reference gene and calculated relatively to either BEAS-2B or HEK293 cell line calibrator sample. Undetectable expression levels were assigned to lncRNAs with Cq values above acceptable limits (above 32–33 PCR cycles). SPSS software version 20.0 was used for data analysis. One way-Anova with Tukey's multiple comparisons post hoc tests or *t*-test were applied to compare expression levels between the different groups. P-values < 0.05 were considered statistically significant.

Table 1
Human lncRNAs and CpG chromosomal locations analyzed.

lncRNA gene	CpG island	Region analyzed
DANCR	93	Chr4: 52712612-52712726
H19	36	Chr11:1996329–1996,211
HOTAIRM1 upstream	116	Chr7:27095663-27095791
HOTAIRM1 downstream	116	Chr7: 27096410-27096495
MALAT1	50	Chr11:65497584–65497709
MEG3	45	Chr14:100826116–100826,218
MEG3	78	Chr14:100827449–100827,521
PVT1	94	Chr8:127794107-127794229

2.6. Human and mouse miRNA NSCLC profiles determined by next generation sequencing (NGS)

Total RNA including microRNA species was purified from A549, H460 and BEAS-2B human cultivated cells, as well as from three NSCLC lung tumors (LUAD) and matching adjacent normal tissues from FVB mice treated with urethane. Libraries were constructed from RNA samples using the QIAseq miRNA library kit (QIAGEN GmbH, Hilden, Germany). As an initial step, adapters were ligated to the 3′ and 5′ ends of miRNAs and subsequently, universal cDNA synthesis with Unique Molecular Indices assignment, cDNA cleanup, library amplification and library cleanup were performed. Downstream RNA sequencing reactions were performed in Illumina iSeq 100 instrument (Illumina, Inc., US) in single-ended mode with 75 bp sequence length and total reads ranging from 1.0 to 1.5 million per sample. FastQC version 0.11.9 [29] was used for quality assessment of the sequenced libraries, before downstream FASTQ file analysis. UMIs present in each sample were extracted using UMI-tools version: 1.1.2 [30] and are to be later used for deduplicating counts. The extraction pattern follows a regular expression pattern (RegEx) that searches each sequence for the 3' adapter and extracts the following 12 bases as UMI and discarding any nucleotides after that, using the parameters "extract-method = regex–bc-pattern = '.+(?P<discard_1>AACTGTAGGCACCATCAAT){s ≤ 2}(?P<umi_1>.{12}) (?P<discard_2>.+)'". UMI-extracted sequencing reads with lengths lower than 18 nucleotides were discarded using Cutadapt version 4.2 [31]. Subsequently, hairpin and mature miRNAs were mapped to their corresponding genomes using Bowtie2 version 2.5.1 [32]. Homo sapiens miRNA FASTA files were obtained from miRBase release 22.1 [33]. We performed three rounds of alignment, one for precursor miRNAs, also known as hairpin miRNA, one for mature miRNA and one for genome aligned mature miRNA. Firstly, the hairpin miRNA sequencing reads were aligned to the hairpin miRNA miRBase genome using the parameters "–local –very-sensitive –no-unal –N 0 -k 1 –threads 16". Reads were later processed to obtain mature miRNA counts using the same Bowtie2 parameters as before. Unaligned mature miRNA sequencing reads were re-aligned to the human reference genome (GRCh38) with Bowtie2 using the parameters "–local –very-sensitive –no-unal –N 1 -k 1 –threads 16", allowing now for single (1) mismatch rather than the strict no mismatch used in the original alignment round.

Original count files for the miRBase aligned sequencing reads were created with the use of Samtools version 1.16 [34] using the sort and index tools. For the GRCh38 aligned reads we used the htseq-count tool from the HTSeq version 2.0.2 package [35] with the parameters "-f bam -m union -s yes -r name -t miRNA -i product". Finally, both miRBase and GRCh38 aligned count files were deduplicated, in order to minimize falsely overestimated counts, with the use of UMI-tools dedup tool and

the option "–method = unique". Mature miRNA miRBase counts were combined with the GRCh38 mature miRNA counts using custom in-house Python scripts.

The same pipeline was implemented for the analysis of the NSCLC mouse samples. The alignment of the miRNAs was performed on the GRCm38 mouse reference genome. Differentially expressed hairpin and mature miRNAs were identified with the R package DESeq2 [36], where counts lower than 1000 were discarded and a log fold-change cutoff of ±1 and an adjusted p-value cutoff of 0.05 were applied. Total miR reads displayed by NGS analysis are listed in Supplementary Table S2 A-D (human) and S3 (mouse). Significantly altered mouse miRs, either up- or down-regulated, are presented in Supplementary Table S4 and corresponding human miRs in Supplementary Table S5. The NGS data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE251697 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE251697)".

A holistic experimental outline, as well as processing workflow of the present study are given in Fig. 1.

3. Results

3.1. Differential expression profiles of LRF/ZBTB7A and methylation/expression profiles of lncRNAs

Comparisons of mouse lrf/*Zbtb7a* expression levels between lung and other tissues (liver and spleen) showed a significant upregulation of *lrf/Zbtb7a* expression in mouse lung. Similarly, LRF/*ZBTB7A* was significantly overexpressed in human BEAS-2B cells compared to HEK293 immortalized human embryonic kidney cells. When we questioned lrf/Zbtb7a expression in NSCLC samples of FVB mice treated with urethane compared to adjacent normal lung tissue, we found similar expression levels between malignant and normal lung tissue (Fig. 2A–C); this was further corroborated in human samples, as BEAS-2B and A549 LUAD cells expressed same levels of LRF/*ZBTB7A* (Fig. 2B–D). Interestingly however, H460 large cell carcinoma cell line exhibited significantly reduced levels of LRF/*ZBTB7A* expression compared to both BEAS-2B and A549 cells, possibly indicating a differential pathway of LRF/*ZBTB7A* regulation of expression in this subtype of NSCLC.

Methylation and expression assays were performed in six (6) selected human lncRNAs (Table 1) to reveal correlations between methylation and expression levels (heatmaps of Fig. 2E). CpGs 93 and 50 of *DANCR*- and *MALAT1*-lncRNAs respectively, were unmethylated representing full accessibility for the transcriptional machinery although, A549 cells displayed 70% reduced levels of expression for *DANCR*-lncRNA. *MALAT1*-lncRNA was not detected in any of the cell lines tested with qPCR, as well as *HOTAIRM1*-lncRNA. Among the rest of lncRNAs analyzed, decreased expression levels of each lncRNA were analogous to the elevation of methylation status across CpG island. CpGs 36, 116, 45, 78 and 94 of *H19*-, *HOTAIRM1*, *MEG3*-, *MEG-3* and *PVT1*-lncRNAs respectively, were mainly hypermethylated in H460 cell line (Fig. 2E). Methylation status of *MEG3* CpG 45 directly upstream CpG 78, exhibited half values of methylation (45%) in BEAS-2B and A549 cells compared to CpG 78, which can be attributed to the CTCF insulator protein binding site, residing between these two rich CpG sequences [UCSC genome browser (https://genome-euro.ucsc.edu/cgi-bin/hgGateway)] and obstructing hypermethylation spreading. LncRNA methylation and expression levels are further discussed in paragraph 3.5 in direct correlation with results from miRNA expression and within the frame of potential interconnection of their functions.

3.2. Sets of mouse miRs differentially expressed between NSCLC and normal lung tissue

To identify miRs of mouse origin (mmu-miRs) that are differentially expressed in NSCLC, RNA samples from six different NSCLC tumors (experimentally developed according to Ref. [17]) and four RNA samples from adjacent normal mouse lung tissues, were utilized for miR

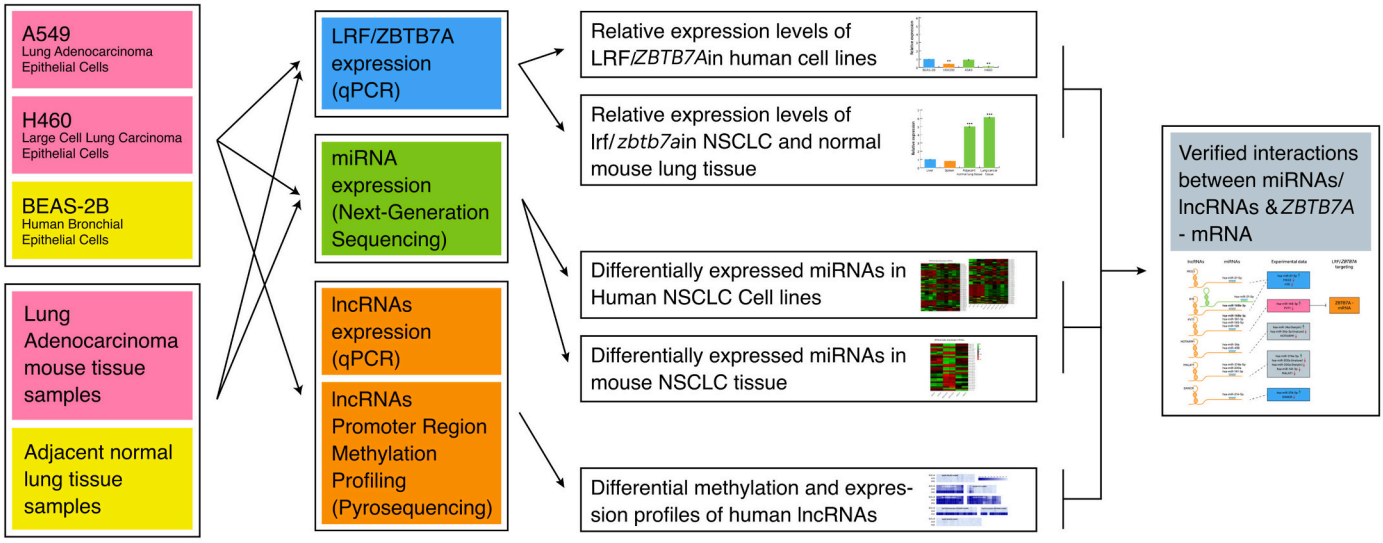

Fig. 1. Flow diagram of experimental processes performed with NSCLC experimental models involving human cell lines and mouse tissues. To investigate the epigenetic landscape of human and mouse NSCLC, we have initially assessed the differential expression patterns of *ZBTB7A* human and *Zbtb7a* mouse genes in a variety of human cells and mouse tissues and in *KRAS*-mutant NSCLC (human cell lines and mouse tissues). *ZBTB7A* encodes LRF which has been characterized as an epigenetically motivated TF acting towards transcriptional regulation of a variety of ncRNA molecules during cell differentiation [12,13]. To highlight valid ncRNA networks underlying the NSCLC epigenetic profile, which involve LRF interaction with ncRNAs, we subsequently performed analysis of expression levels of six lncRNAs confirmed by recent literature [18–23] for their implication in human lung cancer. Also, methylation screening in CpG islands flanking those six selected lncRNAs was determined to further correlate DNA methylation at 5′ site of TSS with lncRNA expression levels. LncRNAs are not conserved between humans and mice contrary to miRNA families and thus, mouse samples were subjected only to miRNA qualitative and quantitative analysis. Both mouse and human RNA samples were subjected to NGS analysis to determine miRNAs differentiated in NSCLC and normal tissue to yield the complete epigenetic field guided by aberrant miR expression levels. Finally, a comprehensive screening to assess miR/lncRNA/LRF networks differentiating NSCLC subtypes studied was conducted and a network of potential interrelationships was built.

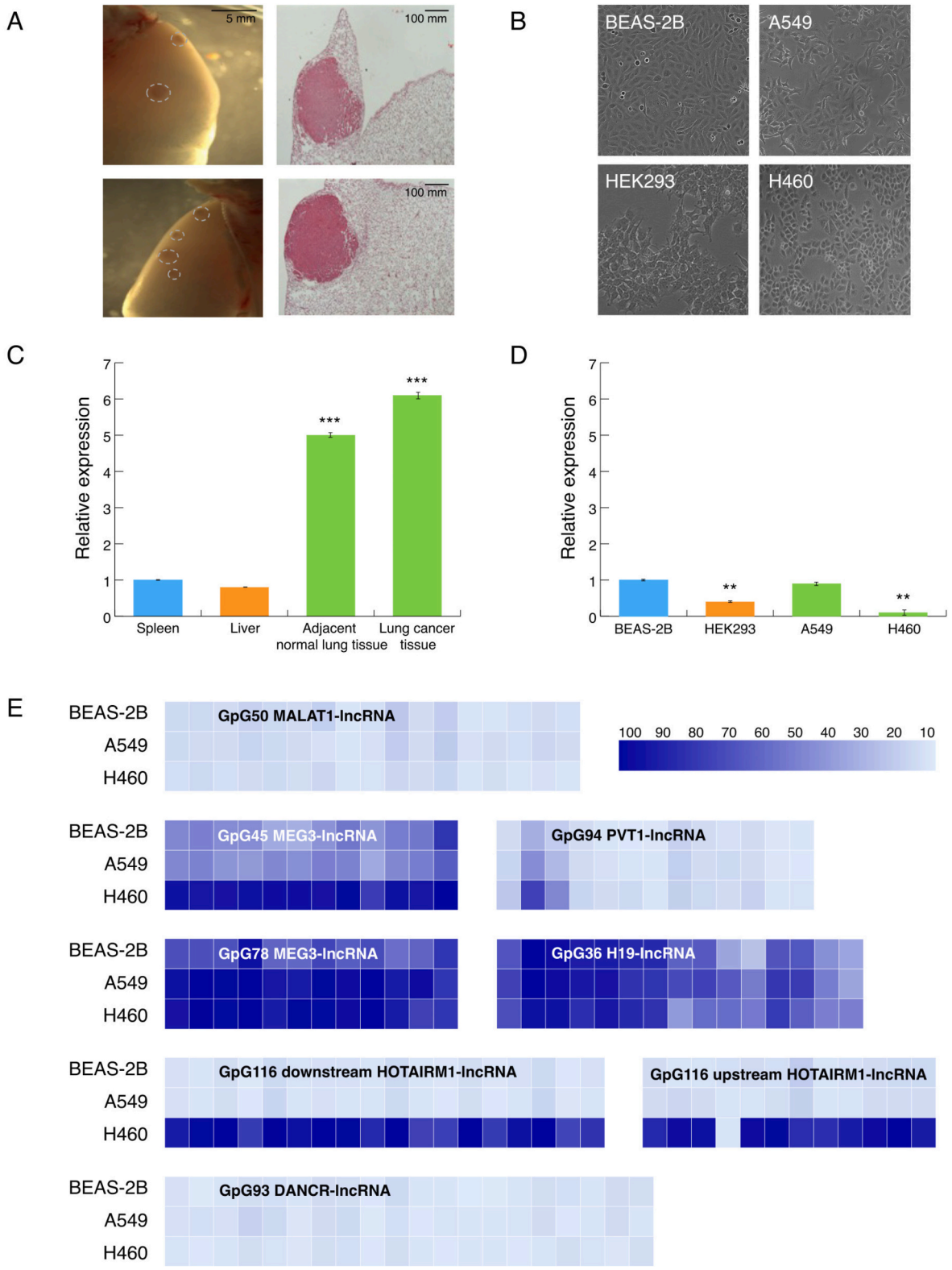

Fig. 2. (A) Representative stereoscopic images of whole lungs (left panel) and microscopic images of hematoxylin/eosin-stained sections (right panel) of lung adenocarcinomas (LUAD) induced in FVB mice by 1 intraperitoneal injection of 1 g/kg urethane (six months latency). Dashed circles outline tumors. (B) Human cell lines subjected to analysis (C) lrf/*Zbtb7a* was significantly overexpressed in normal mouse lung tissue compared to other physiological tissues (liver and spleen) and comparable between mouse NSCLC and adjacent normal lung tissue. ***$p < 0.001$ (D) LRF/*ZBTB7A* levels were significantly lower in human HEK293 compared to BEAS-2B cells. Human BEAS-2B and A549 LUAD cells exhibited equal LRF/*ZBTB7A* expression levels, in contrast to H460 large cell carcinoma human cell line which presented significantly lower expression of LRF/*ZBTB7A*. **$p < 0.01$ (E) Methylation profiles of CpG islands flanking TSS of lncRNAs. DNA samples were isolated from A549, H460 and BEAS-2B human cell lines and were bisulfite treated before subjected to methylation analysis. Each box corresponds to a Cytosine followed by Guanosine across the CpG island sequence. Data are presented as mean percentage methylation.

library development, next generation sequencing and subsequent bioinformatics analysis.

Distinguished miRs either overexpressed or underexpressed in mouse NSCLC tissue are listed in Supplementary Table S4. Both hairpin and mature forms of miRs were determined to address whether they occur at equal rates or post-transcriptional regulation differentiates the final levels of the mature miRs. Results provided clear evidence for the absolute match between the production of hairpin and mature miR molecules. Reads with counts lower than 1000 were discarded. Differentially expressed miRs were considered those with an adjusted p-value <0.05 and a log2fold change cutoff of ± 1. Also, mmu-miR families indicating other miR members with seed sequence of close similarity implying same gene targeting and chromosomal locations of mmu-miR clusters (when applicable) indicating chromosomal locations with altered transcriptional regulation upon malignant development, are also listed in the S4 table.

A heatmap of deregulated miRs outlining expression profiling in mouse NSCLC tissues is illustrated in Fig. 3A. 73 hairpin forms of premature miRs distinguished NSCLC from normal lung tissue in mice, which are considered useful markers in correlation tests between human and mouse NSCLC biological models.

The vast majority of upregulated mmu-miR species were assigned to a specific area of mouse chromosome 12 [(UCSC genome browser on Mouse (GRCm38/mm10), 12qf1: 109,709,060–109,747,960] (Fig. 3C), suggesting a genomic expansion of this chromosomal region or exceptional deregulation of transcription across this genomic area. More distant regions of chromosome 12 were also found to produce higher levels of miR species (chr12: 109585789–109595388 and chr12:109611500–109618336, Table S4) in mouse NSCLC tissue (not shown in Fig. 3C). Relative quantification of miRs derived from this chromosomal location, suggest that transcription starts from miR promoter regions across the forward strand of DNA however, RNA polymerase attenuates transcriptional efficiency since miR counts of each cluster are with a descending order. This phenomenon is color-coded and presented in Fig. 3C, where miRs with the highest and significantly differentiated expression are highlighted in dark green, gradation to light green represents miRs upregulated but not statistically significant and finally black-colored miRs were not detected by NGS analysis.

3.3. Altered expression profiles of human miRs

Human miR (hsa-miR) expression levels were derived from NGS and up/downregulated species were analyzed by bioinformatics analysis as presented in materials and methods. Hairpin as well as mature miR

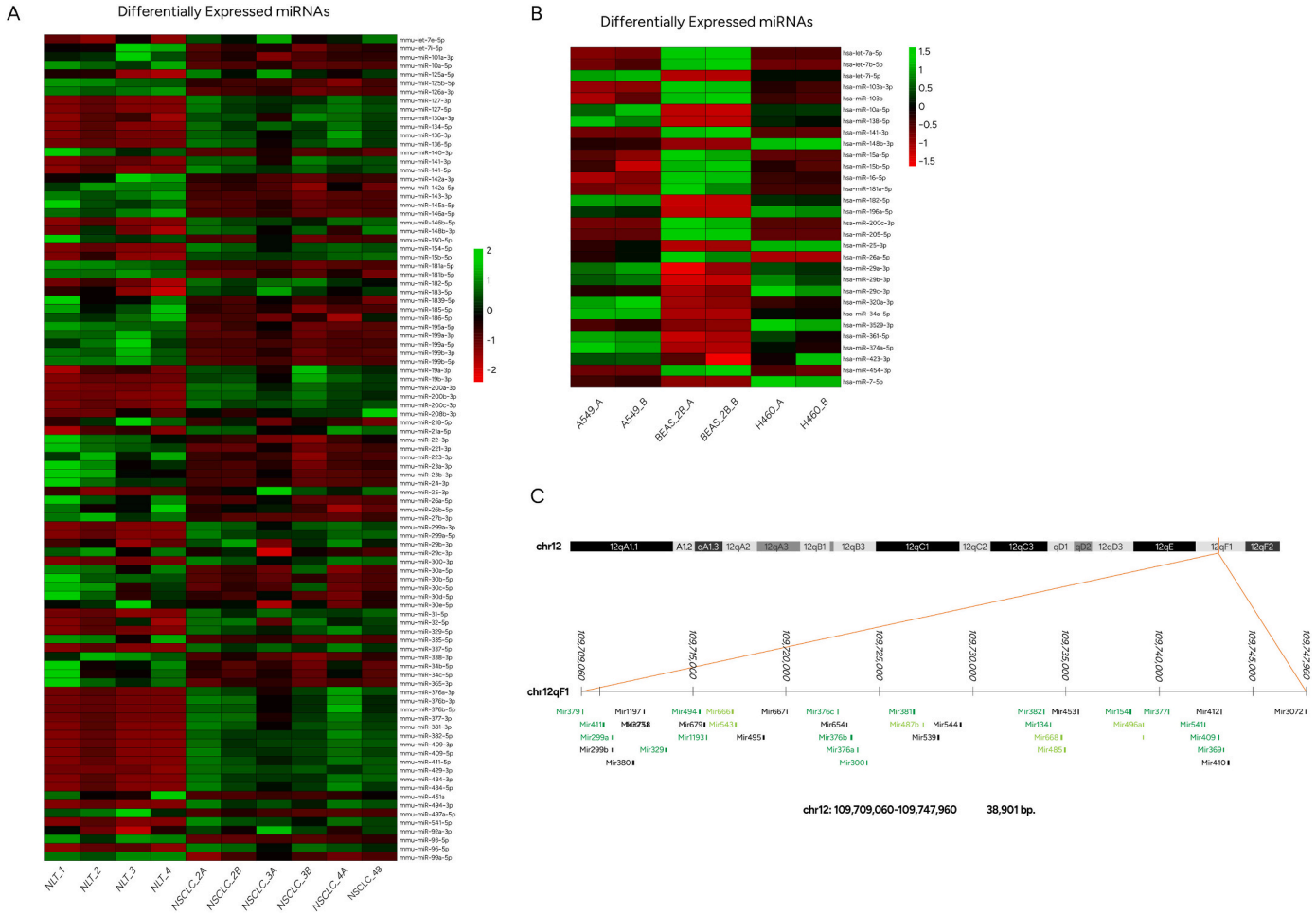

Fig. 3. (A) Heatmap of mature mmu-miR profile comparisons between mouse NSCLC and paired normal lung tissue. Each column represents an RNA sample, and each row represents a single miR. NLT: Normal Lung Tissue (Sample 1 to 4), NSCLC_2A-4B: RNA samples originating from different mice with NSCLC. (B) Heatmap presenting differentially expressed mature forms of hsa-miRs derived from human cell lines A540 and H460 compared to normal epithelial cells BEAS-2B. The A and B sample types refer to the technical replicates used in the NGS analysis. The color scale in both heatmaps shows the relative miR expression level in certain slide: green indicates high relative expression levels; red indicates low relative expression levels. (C) Mouse chromosomal location 12qf1 [UCSC genome browser on Mouse (GRCm38/mm10)] encompassing miR clusters overexpressed in NSCLC tissue (Table 2). miRs in dark green are detected in our study as significantly upregulated, miRs in light green are detected but not differentially expressed with statistical significance (compared to normal lung tissue) and black are miR species below limit of detection.

species were at similar levels, signifying a smooth maturation process from hairpin to mature miR forms and with absence of post-transcriptional regulatory intervening events. Herein, hairpin forms are not further discussed. A heatmap of the differentially expressed miRs in NSCLC cell lines (A549 and H460) is illustrated in Fig. 3B and summarized in Supplementary Table S5. Reads with counts lower than 1000 were discarded. Differentially expressed miRNAs were considered those with an adjusted p-value <0.05 and a log2fold change cutoff of ± 1. Additionally, hsa-miR families, with similar seed sequences and chromosomal locations of hsa-miR clusters lying within the same genomic region are listed in Table S5. Families of commonly upregulated miRs between A549 and H460 were miR-34, let-7 and miR-374, and accordingly downregulated families were miR-205 and miR-8. Moreover, miR-10 family members were upregulated in A549 and downregulated in H460 cells. Common members of miR-8 family, detected in both mouse and human NSCLC samples were up- or downregulated in reverse way. MiR-10 family in mice was downregulated, presenting the same trend as H460 cell line.

Significantly deregulated miRs of the human NSCLC cell lines compared to the human bronchial epithelial cell line BEAS-2B, were the following:

MiR-let-7i hairpin and mature hsa-let-7i-5p were overexpressed in NSCLC cell lines. In particular, hairpin and mature forms of hsa-let-7e and hsa-let-7e-5p were overexpressed in A549, and hsa-let-7c and hsa-let-7c-5p were overexpressed in H460. On the contrary, A549 cells were characterized by downregulation of hsa-let-7c-5p and hsa-let-7d-5p. Hairpin hsa-miR-374a and its corresponding mature form (hsa-miR-374a-5p) were also overexpressed in both A549 and H460 cell lines. Another essential miRNA, miR-34a, the first miRNA investigated in a clinical trial for solid tumors restoration [37], was upregulated as hairpin and mature (hsa-miR-34a-5p) in both A549 and H460 cells. Hairpin hsa-miR-200a, hsa-miR-200c and hsa-miR-141 along with mature forms of hsa-miR-200a-3p, hsa-miR-200c-3p and hsa-141-3p (members of mir-8 mammalian family), showed decreased levels of expression in both A549 and H460 cell lines. Finally, both human NSCLC cell lines showed a significant decrease in hsa-miR-205 (hairpin form) and hsa-miR-205-5p (mature form). Other miR molecules detected in this study either in mouse or human cells are within the list of compiled validated miRs with important diagnostic value [38].

A cluster encompassing hsa-miR-99b-5p, hsa-miR-125a-5p and hsa-let-7e-5p, located on chromosome 19, was found to be elevated only in A549 cells. These miR species derive from the same precursor RNA and although belong to the same cluster, are in separate miR families with diverse biological functions.

The same observation regarding the presence of miR species constituting a cluster applies to human samples as well as observed in mouse upregulated miR clusters of chromosome 12. All members of the same miR cluster are present (Supplementary Table S2 A-D), but only subsets of the each miR cluster reach a statistical significance of deregulation. Only miR species that were depicted by bioinformatics analysis are presented in Table S5.

3.4. Human and mouse miRs targeting LRF/ZBTB7A

Abnormal levels of Zinc-Finger/BTB-Domain-Containing Protein family members, such as LRF/*ZBTB7A* in malignant cell lines and tissues provide valuable insights into upstream signaling pathways/molecules and molecular mechanisms underlying their roles in genome integrity, cell cycle checkpoint obstruction and interacting partners. LRF/*ZBTB7A* has been characterized as an epigenetic regulator and a contingent factor orchestrating expression of non-coding RNA species [12–14]. In our analysis, *ZBTB7A* expression levels were found comparable and equally elevated in lung adenocarcinoma of mouse and human origin, while significantly lower *ZBTB7A* expression levels were obtained in H460 large cell lung carcinoma cells (Fig. 2C and D). Table 2 lists mouse and human miRs either predicted with miRDB [39] or reported in literature to target *ZBTB7A;* results obtained from our study are also included.

Among members of the miR-10 mammalian family, our results showed downregulation of hsa-miR-100-5p (Table S5), considered as a direct suppressor of *ZBTB7A* as well as decreased *ZBTB7A*-mRNA levels in H460 cell line. In mouse NSCLC samples, mmu-miR-10a-5p (Table S4) was significantly downregulated. Some more distant miR species of the same family, namely miR-99a/b, miR-100 and miR-125, are encoded from other genetic loci than Hox clusters [46] and imply potential mutual gene targets. Human A549 showed upregulation of hsa-miR-99b-5p and hsa-miR-125a-5p. All mentioned miR species belong to miR-10 family sharing similar seed sequences and implying similar functions and gene targets. Supportive evidence comes from studies in mice which report that *Zbtb7a* is identified as a direct target for mmu-miR-125a [40].

Moreover, A549 human cell line displayed downregulation of hsa-miR-20a-5p (Table S5), member of the miR-17 mammalian family and miR17/92 cluster, which has been reported to support development and progression of lung cancer in humans [47], whereas in mouse embryonic cells mmu-miR-20a overexpression contributes to cell senescence assisted by the direct *ZBTB7A* and *E2F1* downregulation accompanied by p16 upregulation [48]. Another cooperation of miR-15a (member of miR-15 mammalian family) and *ZBTB7A* that has been predicted by miRDB [39], was also supported by our analysis, since hsa-miR-15a-5p was downregulated in both A549 and H460 cell lines (Table S5). In mouse NSCLC samples, mmu-miR-195a-5p, member of the same family

Table 2
Targeting of ZBTB7A by miRs, predicted with respective scores by miRDB or reported in literature.

Chromo-some	Family	miRNA	Gene target	interaction with 3′UTR	Score/ [REF]	Results from our study
13	miR-17	hsa-miR-20a-5p	*ZBTB7A*	at 912, 1205, 1230, 4309 bp	96	A549 downregulated
19	miR-10	hsa-miR-125a-5p	*ZBTB7A*	at 547, 803, 1667, 4210 bp	81/[40]	A549 upregulated
12	miR-148	hsa-miR-148b-3p	*ZBTB7A*	at 1877, 2071 bp	50	H460 upregulated
13	miR-15	hsa-miR-15a-5p	*ZBTB7A*	at 3305 bp	65	A549 and H460 downregulated
19	miR-10	hsa-miR-100-5p	*ZBTB7A*		[41]	H460 downregulated
4	miR-577	hsa-miR-577	*ZBTB7A*		[42]	–
7	miR-17	hsa-miR-106b	*ZBTB7A*		[43]	–
17	miR-144	hsa-miR-144-3p	*ZBTB7A*		[44]	–
19	miR-520e	hsa-miR-520e	*ZBTB7A*		[45]	–
19	miR-10	mmu-miR-10a-5p	*Zbtb7a*		68	NSCLC mouse tissue downregulated
13	miR-15	mmu-miR-195a-5p	*Zbtb7a*			NSCLC mouse tissue downregulated
1 and 2	miR-181	mmu-miR-181a-1-5p and mmu-miR-181a-2-5p	*Zbtb7a*	at 2907 bp	87	NSCLC mouse tissue downregulated

(miR-15) was also downregulated. Another downregulated miR, namely mmu-miR-181a-5p was predicted to target *Zbtb7a* and was also detected in our sets of deregulated mmu-miR species (Table S4). Hsa-miR-148b-3p, is upregulated only in H460 cells in the present study and has also been predicted to target *ZBTB7A* (Table S5). MiR-148b-3p belongs to the miR-148/-152 family members (miR-148a, miR-148b, and miR-152), which have been associated with poor overall survival in NSCLC. Drug cocktails including metformin and cisplatin promote upregulation of miR-148/-152 family suggesting these molecules as potential targets for therapy of lung cancer [49].

Other interesting associations between miRNA species and *ZBTB7A* are presented in Table 2, which are highlighted in relevant literature, but neither were apparent among significantly altered miR species in our study nor they were predicted by miRDB [39]. Hsa-miR-577 has been shown to be suppressed in NSCLC and this reduction was associated with sponge effects of the elevated circular non-coding RNA circ_0016,760, derived from the same human tissues. Circ_0016,760 was also predicted to modulate hsa-miR-577 expression and downregulation via employment of starBase 3.0 bioinformatics tool, further promoting expression of the *ZBTB7A* (target of hsa-miR-577) gene and NSCLC progression [42]. Quantitative PCR analysis from total RNA derived from NSCLC tumor and paired adjacent normal lung tissues showed increased levels of hsa-miR-520e in NSCLC. Hsa-miR-520e was suggested to target the 3′UTR of *ZBTB7A*-mRNA and to be implicated in tumor cell growth, invasion and migration via modulation of Wnt signaling in A549 cells [45]. Another mammalian miR, hsa-miR-106b was shown to directly target the 3' UTR of *ZBTB7A*-mRNA and to alleviate its expression levels in HEK-293T cells [50]. Finally, the 3'UTR of *ZBTB7A*-mRNA sequence encompasses a putative binding site for miR-144-3p. In bladder cancer miR-144-3p suppresses malignancy via targeting and decreasing *ZBTB7A*, having an inverse correlation with *HIC1/ZBTB29* (Hypermethylated in cancer 1/Zinc Finger And BTB Domain-Containing Protein 29) gene expression [44], another member of the Zinc-Finger/BTB-Domain-Containing Proteins. In lung cancer increased expression of *ZBTB7A* has been shown to inactivate the TP53 (Tumor Protein P53) by increasing the levels of MDM2 (MDM2 Proto-Oncogene, E3 Ubiquitin Protein Ligase), frequently overexpressed in this type of malignancy. In addition, a significant number of NSCLC tissues examined displayed 2-5-fold amplification of *ZBTB7A* chromosomal locus 19p13.3, accompanied by increased mRNA and protein levels strongly suggesting that the amplification process may have functional consequences [51].

3.5. Evaluation of the combinatorial use of ncRNAs identified in this study as epigenetic biomarkers of NSCLC

The constant clinical needs for reliable markers of cancer development and progression places emphasis on established high-throughput approaches and molecular networking assessment, composed of molecules that either compete or jointly cooperate in cancer state. Following this route, we assessed the role of the combinatorial expression of the previously identified ncRNAs as epigenetic markers, with the aim to define epigenetic networks underlying NSCLC. Our methylation/expression results of lncRNAs (Table 1 and Fig. 2E) and miRs (Tables S4 and S5) of human NSCLC cell lines and mouse NSCLC tissue support several previously reported results, while in some cases indicate the presence of additional regulatory transcriptional mechanisms. Our results also unravel a novel combination of lncRNA/miR molecules with regulatory control of LRF/*ZBTB7A* expression. The above results are summarized in Table 3.

Apparently, hsa-miR-148b-3p/*PVT1*-lncRNA/*ZBTB7A*-mRNA represents a novel interesting network derived from our study, which is introduced for the first time in literature. Furthermore, it describes the significantly repressed *ZBTB7A* levels in H460 (large cell lung carcinoma) cell line in contrast to all other NSCLC (adenocarcinoma) human and mouse samples tested. All reported interactions in human samples between *ZBTB7A* and non-coding RNAs (miRs and lncRNAs) are illustrated in Fig. 4.

Also, the differentially expressed miRs and lncRNAs derived from the present study could be used as diagnostic markers for LUAD, as shown in Fig. 5 using data from the KMplot pan-cancer miR dataset (http://kmplot.com) [63].

Table 3
Summary of potentially combinatory miRs/lncRNAs expression profiles obtained and within the context of relative published results.

LncRNA	miRs/lncRNAs expression profiles obtained by the present study	miRs/lncRNAs published role in cancer
MEG3	Suppressed transcription combined with hypermethylation status (above 90%) of CpG island 78 in A549 LUAD cells, leading to upregulation of hsa-miR-21 and hsa-miR-21-5p	Downregulated in NSCLC with regulatory properties during cisplatin treatment through the control of p53 and Bcl−xl expression [52]; Overexpression and simultaneous suppress of hsa-miR-21-5p and positive regulation of tumor suppressor PTEN levels [53]
H19	Upregulation of hsa-miR-21-5p only in A549 LUAD cells accompanied by the hypermethylation status of CpG36 located upstream of *H19*-lncRNA, promoting transcriptional repression of this lncRNA	Positive correlation between increased *H19*-lncRNA and hsa-miR-21 expression with increased tumor size and advanced tumor-node-metastasis stage, but not classification in human NSCLC [54]
HOTAIRM1	Undetectable levels of expression, followed by upregulation of hsa-miR-34a (harpin and mature forms) in both A549 and H460 cells; Inversely methylated *HOTAIRM1* CpG island 116 (Fig. 2E), encompassing a promoter segment. Indication of regulatory transcriptional mechanisms other than methylation status for this lncRNA in NSCLC.	Highly expressed in lenvatinib-resistant hepatocellular carcinoma leading to hsa-miR-34a downregulation [55]
MALAT1	Undetectable expression with simultaneous upregulation of hsa-miR-374a-5p, but suppression of (hairpin) hsa-miR-200a and (mature) hsa-miR-200a-3p in both human NSCLC cell lines. Indication of alternative TSSs for *MALAT1* gene that substitute the targeted promoter in NSCLC cell lines studied.	Regulation of proliferation, apoptosis, migration and invasion via the miR-374b-5p/SRSF7 and hsa-miR-200a/ZEB-1 axis in NSCLC [56,57]; Undetectable in A549 and H460 cells [58]
DANCR	Downregulated only in A549 cells and undetectable hsa-miR-214-5p levels in NSCLC cell lines; hypomethylated CpG island 93 at the 5′ end of the *DANCR* gene in both A549 and H460 NSCLC cells (Fig. 2E), indicating of additional regulatory transcriptional mechanisms for this lncRNA in NSCLC.	Upregulated in NSCLC, exhibiting oncogenic function by regulating hsa-miR-214-5p/CIZ1 axis [59].
PVT1	Upregulated in A549 cells, combined with low methylation status of CpG island 94 and undetectable levels of hsa-miR-148b-3p. Downregulated in H460 cells, combined with upregulated hsa-miR-148b-3p and high methylation of CpG island 94.	Promotes cancer proliferation, invasion and migration in NSCLC via regulating circuits, such as hsa-miR-148/RAB34, hsa-miR-145-5p/ITGB8 (Integrin Subunit Beta 8) and hsa-miR-361-3p/SOX9 (SRY-Box Transcription Factor 9) [60–62]

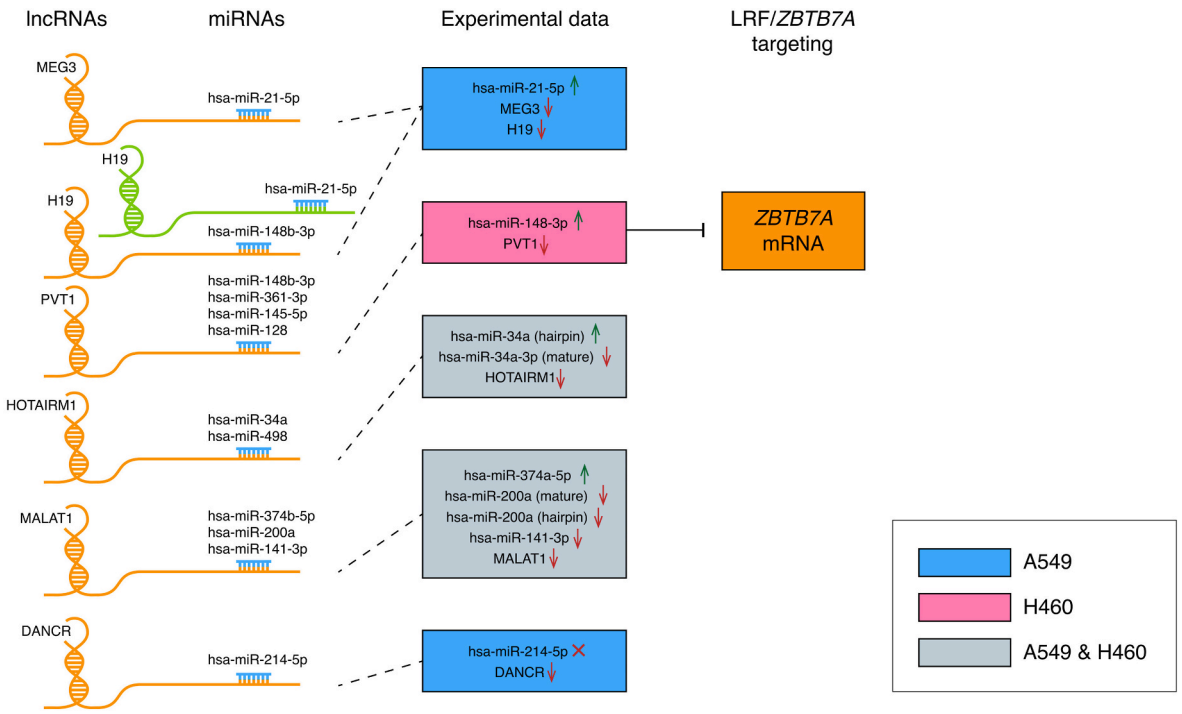

Fig. 4. Interplay of lncRNA/miRNAs and *ZBTB7A*-mRNA in human NSCLC KRAS mutant cell lines A549 and H460, derived from the present study.

4. Discussion

Our study is an attempt to deepen our understanding on epigenetic events discriminating NSCLC subtypes and inspire new directions for the design and testing of treatment agents interfering with the cancer-cell epigenome. We aimed to uncover the divergent epigenetic background of *KRAS*-mutant NSCLC in mouse and human cell lines, extensively used as biological models in relevant research. To that end, we have comprehensively mapped the functional miRNA and lncRNA landscape of human (A549 and H460) and mouse (experimentally developed LUAD) NSCLC models and correlated current results with LRF/*ZBTB7A* expression. LRF is a BTB-ZF (broad-complex, tramtrack and bric-à-brac-zinc finger) protein, strongly implicated in epigenetic regulation of human genes including ncRNAs [12–14]. Interestingly, *ZBTB7A* gene (transcribing LRF) was found to be differentially expressed in the two NSCLC subtypes that we studied: *ZBTB7A* was overexpressed in human A549 and mouse lung adenocarcinoma samples, but significantly underexpressed in H460 large cell carcinoma cell line (Fig. 2C–D), suggesting an exceptional role of *ZBTB7A* expression signature in the large cell lung carcinoma. Another important result was also the significant upregulation of hsa-miR-148b-3p only in H460 cell line, which is predicted to target *ZBTB7A*-mRNA (Table 2) among other gene targets, and it is sponged by the *H19*- and *PVT1*-lncRNAs [58,62]; these traits render hsa-miR-148b-3p an attractive candidate regarding the differential regulation of *ZBTB7A* expression in lung large cell carcinoma. Its capability of attacking also DNMTs with other members of the miR-148/-152 family makes this molecule an essential epigenetic regulator and a novel potential target for lung cancer treatment [49]. The repressed *PVT1*-lncRNA levels that were detected in H460 cells are suggested to contribute to the observed hsa-miR-148b-3p upregulation and *ZBTB7A* downregulation, signifying the regulatory axis miR-148b-3p/*PVT1*-lncRNA/*ZBTB7A*-mRNA, that merits further research (Fig. 4) and verification in human samples.

With the view of deregulated miR families associated with LRF (all listed in Table 2), our results showed downregulation of both hsa-miR-100-5p, member of the mammalian miR-10 family (Table S5) and *ZBTB7A*-mRNA levels in H460 cell line (Fig. 2D). Relative lower expression of hsa-miR-100, a direct suppressor of *ZBTB7A*, has been positively correlated with lymph node metastasis and lager tumor size in gastric cancer implying another emerging cooperation between miRs and transcription factors [41]. Also, A549 human cell line displayed downregulation of hsa-miR-20a-5p (Table S5), member of the miR-17 mammalian family and miR17/92 cluster, implicated in lung cancer in humans [47,48]. Hsa-miR-15a-5p, member of miR-15 mammalian family was downregulated in both A549 and H460 cell lines (Table S5) and mmu-miR-195a-5p, member of the same family was also downregulated (Table S4). Hsa-miR-148b-3p, was upregulated only in H460 cells, which belongs to the miR-148/-152 family members (miR-148a, miR-148b, and miR-152), and its overexpression has been associated with longer overall survival of patients with NSCLC [49,62]. Other, deregulated miR families and species either reported in research studies or predicted by miRDB [39], are discussed in 3.4 paragraph within the results section.

Expanding on this track, relevant lncRNA transcripts expressed in NSCLC human cell lines against the physiological human cells BEAS-2B, were evaluated. *MEG3*-and *PVT1*-lncRNAs levels of expression were reduced, mainly in H460 cell line. *H19*-lncRNA showed insignificant differences in expression levels, *DANCR*-lncRNA was reduced only in A549, whereas *MALAT1*-and *HOTAIRM1*-lncRNAs were undetected in all cell lines tested. Methylation levels in CpG islands flanking the 5' end of lncRNA genes were very high (over 80%) in all lncRNAs tested in H460 cells except from *MALAT1*-and *DANCR*-lncRNAs. A549 and BEAS-2B cells displayed comparable methylation levels (Fig. 3E) among CpG islands studied. Our findings in relation to established cooperation with miRs defined by the literature are presented in the corresponding paragraph *3.5* of the results section.

Significantly deregulated miR families between human and mouse NSCLC samples, detected by NGS and consequent bioinformatics analysis, were members of the mammalian family miR-21 namely, mmu-miR-21a and mmu-miR-21a-5p as well as hsa-miR-21 and hsa-miR-21-3p hairpin and mature forms, which were found upregulated in both NSCLC mouse tissue and A549 cell line (Fig. 3A and B, Tables S4 and S5). MiR-21 involvement in NSCLC tumorigenesis and drug resistance has already been documented in humans [64,65]. On the other hand,

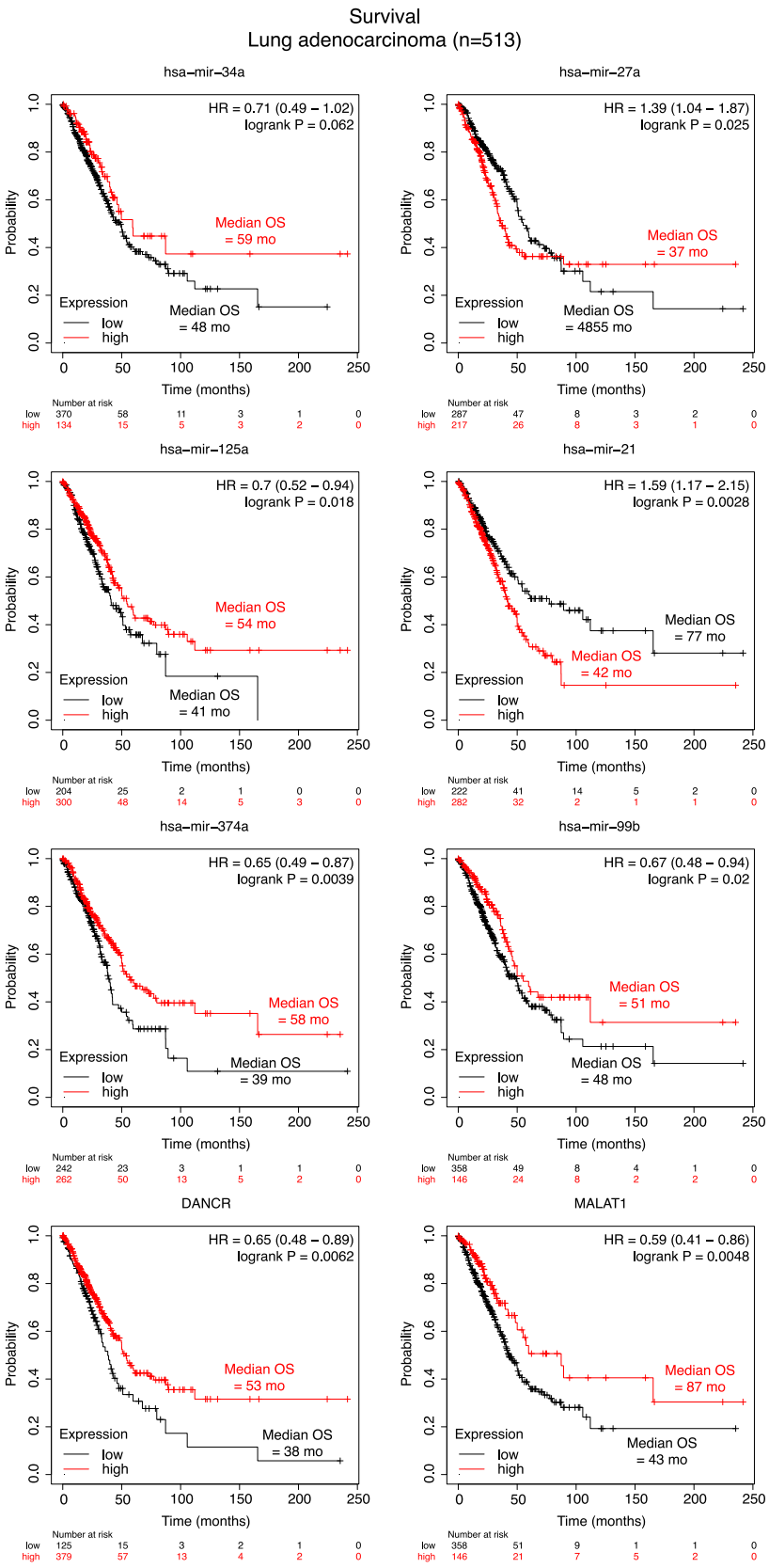

Fig. 5. miR and lncRNA species produced by this study as biomarkers of lung adenocarcinoma. Kaplan-Meier survival plots with median overall survival (OS), hazard ratios [20] with 95% confidence interval, and univariate log-rank probability values (P) from 513 patients with lung adenocarcinoma, stratified into low (black) and high (red) miRs and lncRNAs expression by the optimal cut-offs indicated. Data from Kmplot (http://kmplot.com).

members of mammalian families miR-221, miR-222 and miR-27 (hairpin and mature structures) were significantly reduced in mouse NSCLC and human H460 cells. In humans, hsa-miR-221 has been proposed as a marker for early diagnosis and NSCLC screening [66], the HMGA1/miR-222 signaling axis to regulate AKT and to act as an effective tumor suppressor pathway in lung cancer cells, whereas miR-27a has been shown to negatively regulate EGFR in NSCLC [67]. Other identified miR subgroups with differential expression patterns between NSCLC samples of human and mouse origin were members of family miR-34 (mmu-miR-34b and mmu-miR-34c), which were found downregulated in mouse NSCLC, while in both human NSCLC cell lines hsa-miR-34a and hsa-miR-34a-5p species, were upregulated. Furthermore, in mouse NSCLC samples, all members of the family miR-8, namely mmu-miR-141, mmu-miR-200a, mmu-miR-200b, mmu-miR200c and mmu-miR-429 were upregulated; however, all corresponding hsa-miRs of the same family, apart from miR-200b and miR-429, were downregulated in both A549 and H460 cell lines. MiR-8 family includes highly conserved miRs and some particular species are reported to be implicated in tumor growth and progression [68,69].

Conclusively, miR collections among human and mouse samples, although belong to conserved miR families, displayed highly divergent expression profiles. These observations suggest variability in NSCLC stage and subtype, and highlight the diverse pathways disrupted in human and mouse NSCLC, further to the extended biological variations existing between origins of these mammals, reflecting also a limitation of the present study, as results obtained revealed a highly uncommon epigenetic profile between human and mouse samples, making insubstantial any further comparisons. A prominent connection between human NSCLC cell lines and mouse NSCLC tissue samples was the detection of deregulated miRs comprising the mammalian miR-8, miR-10 and miR-15 family members, encompassing miR species with identical seed sequences and mutual mRNA targets. In this context, deregulated expression levels of miR-10 and miR-15 family members represent a very interesting observation since they typically target *ZBTB7A*-mRNA (Table 2). Another important observation was that the upregulated miR species in mouse NSCLC were mainly transcribed from the 12qf1 chromosomal region encompassing a large number of miR clusters (Table 2, Fig. 3C). Quantification analysis on miRs' expression levels based on read counts per miR led to an interesting result. Mmu-miR clusters transcribed from this chromosomal area displayed gradually decreased read counts which can be attributed to the attenuation effect of RNA polymerase during transcriptional elongation but raised also intriguing research perspectives for the transcriptional regulation of this chromosomal area during the NSCLC development and progression in mice.

Direct comparisons of deregulated miR sets between human NSCLC cell lines and the physiological human bronchial epithelial cell line BEAS-2B, showed overexpressing values for miR-let-7i (hairpin and mature forms), known to have as direct target *BAG-1* (Bcl−2 associated athanogene-1)-mRNA, a member of the BAG-family of co-chaperones, with motility, differentiation and apoptosis properties. However, high levels of *XIST*-lncRNA represses hsa-miR-let-7i expression, thus preventing BAG-1 inhibition in A549 cells [70]. Other miR-let-7 family members were significantly deregulated also, which have been reported to regulate the expression of genes involved in cell proliferation and cell cycle and to modulate GTPase-KRAS expression in malignancy [71,72] but nevertheless, showed uncommon patterns of expression between A549 and H460 cell lines. Hsa-miR-374a-5p and hsa-miR-34a-5p were also overexpressed in both A549 and H460 cell lines. Hsa-miR-374a-5p has been associated with the TGFA (transforming growth factor-alpha) gene silencing, leading to significant inhibitory effects on proliferation, migration and invasion of malignancy [73], whereas high expression of hsa-miR-34a and hsa-miR-34c in plasma and tumor tissue of 196 NSCLC patients has been associated with prolonged overall survival and disease-free survival [37,74].

Decreased levels of miR expression in both A549 and H460 cell lines compared to BEAS-2B cells highlighted members of the mir-8 and mir-205 mammalian families. MiR-8 family consists of evolutionarily conserved miRNAs in two human genomic loci, miR-200b/200a/429 on chromosome 1 and miR-200c/141 on chromosome 12 and acquires significant tumor-suppressive roles in a wide array of human cancers [75]. MiR-200c/141 species deficiency promotes tumor growth and progression and also increases the metastasis frequency by recapitulating the desmoplastic lung tumor microenvironment [69]. Concerning mir-205 family members, hsa-miR-205-5p has been recently characterized as a specific prognostic biomarker, mainly for squamous cell lung carcinoma, whereas serum levels of this miR have been documented as a valuable biomarker for lung cancer diagnosis [38]. Kaplan Meier plotting tool revealed strong correlation between some of the miR and lncRNA species detected by this study and median overall survival (Fig. 5) in LUAD, results that merit further investigation and confirmation.

5. Conclusions

This study offered a unique opportunity to study the various epigenetic aspects of *KRAS*-mutant NSCLC in established biological models, comprised of immortalized human cell lines (A549 and H460) and urethane-treated FVB mice.

Our study highlighted variable epigenetic features and uncovered significant relationships between different epigenetically motivated molecular biomarkers, mainly ncRNAs and TFs. Several TFs have been proved to control ncRNAs' expression in a tissue-specific and highly-selective manner; LRF encoded by *ZBTB7A* gene holds an outstanding position previously documented. Cancer state arises from extended dysregulation of cell pathways and thus, network-based approaches for designating ncRNAs that dominate and interaction subnetworks with TFs hold a great potential for discriminating NSCLC subtype signatures and integration to biomarker identification. Results presented in this study highlight considerations and limitations arising from the variation displayed by the experimental models utilized in NSCLC research and prospects for new applications deriving from the epigenetic view.

Compliance with ethics guidelines

Mice were bred at the Center for Animal Models of Disease of the University of Patras. Experiments were designed and approved *a priori* by the Veterinary Administration of the Prefecture of Western Greece (approval numbers 3741/16.11.2010, 60,291/3035/19.03.2012, and 118,018/578/30.04.2014) and were conducted according to Directive 2010/63/EU (http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1486710385917&uri=CELEX:32010L0063).

Funding

This work was supported by the 80,250 (ELKE_EAP) grant of the Hellenic Open University (A.S.) and the General Secretariat for Research and Innovation and Hellenic Foundation for Research and Innovation grant 1853 (M.S.).

Consent for publication

All authors consented to the manuscript.

Availability of data and materials

Data and results generated or analyzed during this study are included in the manuscript. NGS data are uploaded to GEO repository (accession number GSE251697).

CRediT authorship contribution statement

Magda Spella: Investigation, Funding acquisition, Data curation. **Eleftherios Bochalis:** Validation, Software, Methodology, Data curation. **Katerina Athanasopoulou:** Validation, Investigation, Data curation. **Argyri Chroni:** Investigation, Data curation. **Irene Dereki:** Software, Investigation, Data curation. **Giannoula Ntaliarda:** Investigation, Data curation. **Ifigeneia Makariti:** Methodology, Investigation, Formal analysis. **Georgios Psarias:** Methodology, Investigation, Formal analysis. **Caterina Constantinou:** Methodology, Investigation, Formal analysis. **Vasiliki Chondrou:** Supervision, Methodology, Formal analysis. **Argyro Sgourou:** Writing – review & editing, Writing – original draft, Project administration, Funding acquisition.

Declaration of competing interest

Ifigeneia Makariti is financially supported by BioAnalytica SA, Biotechnology supplier company and also a member of the research team of biology HOU lab. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ncrna.2024.03.009.

References

[1] W.D. Travis, et al., The 2015 world health organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification, J. Thorac. Oncol. 10 (9) (2015) 1243–1260.
[2] P. Cascetta, et al., KRAS in NSCLC: state of the art and future perspectives, Cancers 14 (21) (2022).
[3] M. Horie, et al., Integrative CAGE and DNA methylation profiling identify epigenetically regulated genes in NSCLC, Mol. Cancer Res. 15 (10) (2017) 1354–1365.
[4] A. Chang, Chemotherapy, chemoresistance and the changing treatment landscape for NSCLC, Lung Cancer 71 (1) (2011) 3–10.
[5] P.A. Jones, S.B. Baylin, The fundamental role of epigenetic events in cancer, Nat. Rev. Genet. 3 (6) (2002) 415–428.
[6] M. Sherafatian, F. Arjmand, Decision tree-based classifiers for lung cancer diagnosis and subtyping using TCGA miRNA expression data, Oncol. Lett. 18 (2) (2019) 2125–2131.
[7] B. Chen, et al., Prognostic value of survival of MicroRNAs signatures in non-small cell lung cancer, J. Cancer 10 (23) (2019) 5793–5804.
[8] S. Hu, et al., Expression profiles of microRNAs in drug-resistant non-small cell lung cancer cell lines using microRNA sequencing, Cell. Physiol. Biochem. 51 (6) (2018) 2509–2522.
[9] M.A. Cortez, et al., Therapeutic delivery of miR-200c enhances radiosensitivity in lung cancer, Mol. Ther. 22 (8) (2014) 1494–1503.
[10] L. Ginn, et al., LncRNAs in non-small-cell lung cancer, Noncoding RNA 6 (3) (2020).
[11] Q. Guo, et al., Dynamic TF-lncRNA regulatory networks revealed prognostic signatures in the development of ovarian cancer, Front. Bioeng. Biotechnol. 8 (2020) 460.
[12] V. Chondrou, et al., LRF promotes indirectly advantageous chromatin conformation via BGLT3-lncRNA expression and switch from fetal to adult hemoglobin, Int. J. Mol. Sci. 23 (13) (2022).
[13] K. Athanasopoulou, et al., Transcriptional repression of lncRNA and miRNA subsets mediated by LRF during erythropoiesis, J. Mol. Med. (Berl.) 101 (9) (2023) 1097–1112.
[14] S. Gupta, et al., Emerging role of ZBTB7A as an oncogenic driver and transcriptional repressor, Cancer Lett. 483 (2020) 22–34.
[15] I. Ferrer, et al., KRAS-Mutant non-small cell lung cancer: from biology to therapy, Lung Cancer 124 (2018) 53–64.
[16] N.I. Kanellakis, et al., Tobacco chemical-induced mouse lung adenocarcinoma cell lines pin the prolactin orthologue proliferin as a lung tumour promoter, Carcinogenesis 40 (11) (2019) 1352–1362.
[17] M. Spella, et al., Club cells form lung adenocarcinomas and maintain the alveoli of adult mice, Elife 8 (2019).
[18] Y. Aftabi, et al., Long non-coding RNAs as potential biomarkers in the prognosis and diagnosis of lung cancer: a review and target analysis, IUBMB Life 73 (2) (2021) 307–327.
[19] Z. Chen, et al., Long non-coding RNA in lung cancer, Clin. Chim. Acta 504 (2020) 190–200.
[20] M. Entezari, et al., Long non-coding RNAs and exosomal lncRNAs: potential functions in lung cancer progression, drug resistance and tumor microenvironment remodeling, Biomed. Pharmacother. 150 (2022) 112963.
[21] C. Fang, et al., Long non-coding RNAs: how to regulate the metastasis of non-small-cell lung cancer, J. Cell Mol. Med. 24 (6) (2020) 3282–3291.
[22] T. Guo, et al., Multidimensional communication of microRNAs and long non-coding RNAs in lung cancer, J. Cancer Res. Clin. Oncol. 145 (1) (2019) 31–48.
[23] T. Liu, et al., Long non-coding RNAs in lung cancer: implications for lineage plasticity-mediated TKI resistance, Cell. Mol. Life Sci. 78 (5) (2021) 1983–2000.
[24] Y. Zhao, et al., Long noncoding RNA HOTAIRM1 in human cancers, Clin. Chim. Acta 511 (2020) 255–259.
[25] C.G. Albarino, V. Romanowski, Phenol extraction revisited: a rapid method for the isolation and preservation of human genomic DNA from whole blood, Mol. Cell. Probes 8 (5) (1994) 423–427.
[26] A.M. Vlaikou, et al., Mechanical stress affects methylation pattern of GNAS isoforms and osteogenic differentiation of hAT-MSCs, Biochim. Biophys. Acta Mol. Cell Res. 1864 (8) (2017) 1371–1381.
[27] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods 25 (4) (2001) 402–408.
[28] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (9) (2001) e45.
[29] S. Andrews, FASTQC. A quality control tool for high throughput sequence data, Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, 2010.
[30] T. Smith, A. Heger, I. Sudbery, UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy, Genome Res. 27 (3) (2017) 491–499.
[31] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads 17 (1) (2011, 2011) 3.
[32] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (4) (2012) 357–359.
[33] A. Kozomara, M. Birgaoanu, S. Griffiths-Jones, miRBase: from microRNA sequences to function, Nucleic Acids Res. 47 (D1) (2019) D155–D162.
[34] P. Danecek, et al., Twelve years of SAMtools and BCFtools, GigaScience 10 (2) (2021).
[35] S. Anders, P.T. Pyl, W. Huber, HTSeq–a Python framework to work with high-throughput sequencing data, Bioinformatics 31 (2) (2015) 166–169.
[36] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (12) (2014) 550.
[37] D.S. Hong, et al., Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours, Br. J. Cancer 122 (11) (2020) 1630–1637.
[38] S. Zhong, et al., miRNAs in lung cancer. A systematic review identifies predictive and prognostic miRNA candidates for precision medicine in lung cancer, Transl. Res. 230 (2021) 164–196.
[39] Y. Chen, X. Wang, miRDB: an online database for prediction of functional microRNA targets, Nucleic Acids Res. 48 (D1) (2020) D127–D131.
[40] N. Hojo, et al., A Zbtb7a proto-oncogene as a novel target for miR-125a, Mol. Carcinog. 55 (12) (2016) 2001–2009.
[41] D.B. Shi, et al., C/EBPalpha-induced miR-100 expression suppresses tumor metastasis and growth by targeting ZBTB7A in gastric cancer, Cancer Lett. 369 (2) (2015) 376–385.
[42] Y. Hao, et al., Circular RNA Circ_0016760 modulates non-small-cell lung cancer growth through the miR-577/ZBTB7A Axis, Cancer Manag. Res. 12 (2020) 5561–5574.
[43] D. Jiao, et al., miR-106b regulates the 5-fluorouracil resistance by targeting Zbtb7a in cholangiocarcinoma, Oncotarget 8 (32) (2017) 52913–52922.
[44] J. Liu, et al., ZBTB7A, a miR-144-3p targeted gene, accelerates bladder cancer progression via downregulating HIC1 expression, Cancer Cell Int. 22 (1) (2022) 179.
[45] Z. Zhijun, H. Jingkang, MicroRNA-520e suppresses non-small-cell lung cancer cell growth by targeting Zbtb7a-mediated Wnt signaling pathway, Biochem. Biophys. Res. Commun. 486 (1) (2017) 49–56.
[46] D. Tehler, N.M. Hoyland-Kroghsbo, A.H. Lund, The miR-10 microRNA precursor family, RNA Biol. 8 (5) (2011) 728–734.
[47] X. Zhang, et al., Biology of MiR-17-92 cluster and its progress in lung cancer, Int. J. Med. Sci. 15 (13) (2018) 1443–1448.
[48] L. Pitto, et al., microRNA(interference) networks are embedded in the gene regulatory networks, Cell Cycle 7 (16) (2008) 2458–2461.
[49] B.B. Lee, et al., Metformin regulates expression of DNA methyltransferases through the miR-148/-152 family in non-small lung cancer cells, Clin. Epigenet. 15 (1) (2023) 48.
[50] X. Liang, et al., MiR-106b regulates the apoptosis and tumorigenesis of hepatocellular carcinoma via targeting Zinc finger and BTB domain-containing protein 7A (Zbtb7a), J. Biochem. Mol. Toxicol. 32 (8) (2018) e22169.
[51] K. Apostolopoulou, et al., Gene amplification is a relatively frequent event leading to ZBTB7A (Pokemon) overexpression in non-small cell lung cancer, J. Pathol. 213 (3) (2007) 294–302.
[52] J. Liu, et al., The long noncoding RNA MEG3 contributes to cisplatin resistance of human lung adenocarcinoma, PLoS One 10 (5) (2015) e0114586.

[53] D. Lv, et al., Long non-coding RNA MEG3 inhibits cell migration and invasion of non-small cell lung cancer cells by regulating the miR-21-5p/PTEN axis, Mol. Med. Rep. 23 (3) (2021).
[54] Y. Zhou, et al., Association of long non-coding RNA H19 and microRNA-21 expression with the biological features and prognosis of non-small cell lung cancer, Cancer Gene Ther. 24 (8) (2017) 317–324.
[55] D. Gu, et al., Overexpression of the lncRNA HOTAIRM1 promotes lenvatinib resistance by downregulating miR-34a and activating autophagy in hepatocellular carcinoma, Discovery Oncol. 14 (1) (2023) 66.
[56] C. Feng, et al., LncRNA MALAT1 promotes lung cancer proliferation and Gefitinib resistance by acting as a miR-200a sponge, Arch. Bronconeumol. 55 (12) (2019) 627–633.
[57] J. Song, Z.Z. Su, Q.M. Shen, Long non-coding RNA MALAT1 regulates proliferation, apoptosis, migration and invasion via miR-374b-5p/SRSF7 axis in non-small cell lung cancer, Eur. Rev. Med. Pharmacol. Sci. 24 (4) (2020) 1853–1862.
[58] R. Esposito, et al., Multi-hallmark long noncoding RNA maps reveal non-small cell lung cancer vulnerabilities, Cell Genom. 2 (9) (2022) 100171.
[59] Y.R. Chen, et al., Upregulation of lncRNA DANCR functions as an oncogenic role in non-small lung cancer by regulating miR-214-5p/CIZ1 axis, Eur. Rev. Med. Pharmacol. Sci. 24 (5) (2020) 2539–2547.
[60] G. Qi, L. Li, Long non-coding RNA PVT1 contributes to cell growth and metastasis in non-small-cell lung cancer by regulating miR-361-3p/SOX9 axis and activating Wnt/beta-catenin signaling pathway, Biomed. Pharmacother. 126 (2020) 110100.
[61] C.M. Wei, et al., The long non-coding RNA PVT1/miR-145-5p/ITGB8 axis regulates cell proliferation, apoptosis, migration and invasion in non-small cell lung cancer cells, Neoplasma 67 (4) (2020) 802–812.
[62] Y. Xi, et al., PVT1 promotes the proliferation and migration of non-small cell lung cancer via regulating miR-148/RAB34 signal Axis, OncoTargets Ther. 13 (2020) 1819–1832.
[63] B. Gyorffy, Transcriptome-level discovery of survival-associated biomarkers and therapy targets in non-small-cell lung cancer, Br. J. Pharmacol. 181 (3) (2024) 362–374.
[64] C. Bica-Pop, et al., Overview upon miR-21 in lung cancer: focus on NSCLC, Cell. Mol. Life Sci. 75 (19) (2018) 3539–3551.
[65] F. Wang, et al., Imaging dendrimer-grafted graphene oxide mediated anti-miR-21 delivery with an activatable luciferase reporter, ACS Appl. Mater. Interfaces 8 (14) (2016) 9014–9021.
[66] S.M. Abdel Ghany, et al., Circulating miRNA-30a and miRNA-221 as novel biomarkers for the early detection of non-small-cell lung cancer, Middle East J. Cancer 11 (1) (2020) 50–58.
[67] M. Brighenti, MicroRNA and MET in lung cancer, Ann. Transl. Med. 3 (5) (2015) 68.
[68] D. Trumbach, N. Prakash, The conserved miR-8/miR-200 microRNA family and their role in invertebrate and vertebrate neurogenesis, Cell Tissue Res. 359 (1) (2015) 161–177.
[69] B. Xue, et al., miR-200 deficiency promotes lung cancer metastasis by activating Notch signaling in cancer-associated fibroblasts, Genes Dev. 35 (15–16) (2021) 1109–1122.
[70] J. Sun, et al., LncRNA XIST promotes human lung adenocarcinoma cells to cisplatin resistance via let-7i/BAG-1 axis, Cell Cycle 16 (21) (2017) 2100–2107.
[71] Z. Saridaki, et al., A let-7 microRNA-binding site polymorphism in KRAS predicts improved outcome in patients with metastatic colorectal cancer treated with salvage cetuximab/panitumumab monotherapy, Clin. Cancer Res. 20 (17) (2014) 4499–4510.
[72] X. Dai, Y. Jiang, C. Tan, Let-7 sensitizes KRAS mutant tumor cells to chemotherapy, PLoS One 10 (5) (2015) e0126653.
[73] H. Wu, et al., MiR-374a suppresses lung adenocarcinoma cell proliferation and invasion by targeting TGFA gene expression, Carcinogenesis 37 (6) (2016) 567–575.
[74] K. Zhao, et al., Circulating microRNA-34 family low expression correlates with poor prognosis in patients with non-small cell lung cancer, J. Thorac. Dis. 9 (10) (2017) 3735–3746.
[75] X. Feng, et al., MiR-200, a new star miRNA in human cancer, Cancer Lett. 344 (2) (2014) 166–173.