

**CircRNAs from *LSD1* regulate alternative splicing of its  
parental gene by R-loop formation**

Inaugural Dissertation  
zur  
Erlangung des Doktorgrades  
Dr. nat. med.  
der Medizinischen Fakultät  
und  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

**Yefeng Shen**  
aus Hubei, China

(Hundt Druck GmbH)

Köln, 2023

Betreuer/in: Prof. Dr. Margarete Odenthal

Referent/in: Prof. Dr. Niels Gehring

Prof. Dr. Dr. Michal Ruth Schweiger

Datum der mündlichen Prüfung: 10.01.2023

## Abstract

### Abstract

Lung cancer is one of the most common malignancies worldwide, with lung adenocarcinoma (LUAD) being the most common subtype. There is growing evidence that epigenetics plays a critical role in cancer initiation and development, including the key regulator lysine-specific demethylase 1 (LSD1). LSD1 regulates gene expression by demethylating histone 3 lysine 4 and lysine 9, and its high expression correlates with poor prognosis in cancer patients. Several alternative LSD1 splice variants of exon 2a and 8a are expressed, but the regulatory mechanisms are unclear. Because splicing may compete with back-splicing processes that generate circular RNAs (circRNAs), in my thesis I focused on LSD1 splice variants and circRNAs from the LSD1 gene in LUAD.

According to data from the TCGA database, the levels of LSD1+2a variants containing exon 2a differed but compared with total LSD1 transcripts - only slightly between LUAD and adjacent non-tumor tissues. Among the most highly regulated circRNAs from LUAD cells PC9 and normal lung epithelial cells PSAE, two circRNAs originating from the LSD1 gene were identified. PCR studies using different primers of exon 2 identified circRNAs derived from LSD1 (circRNA 3\_2\_2a, circRNA 3\_2, and circRNA 2a\_2). Fluorescence in situ hybridization (FISH) and qPCR revealed that circRNA 3\_2\_2a, circRNA 3\_2, and circRNA 2a\_2 were presented in the nucleus and cytoplasm. Silencing of circRNA 2a\_2 by siRNA but not circRNA 3\_2\_2a and circRNA 3\_2 resulted in a linear LSD1+2a mRNA decrease in A549 and PC9 cells. In addition, overexpression of circRNA 3\_2\_2a and circRNA 3\_2 resulted in upregulation of LSD1+2a in PC9 cells, and overexpression of circRNA 3\_2\_2a and circRNA 2a\_2 increased LSD1+2a levels in PSAE cells. Pull-down experiments with S9.6 antibody recognizing R-loops showed that R-loops are formed in response to circRNA 3\_2\_2a and circRNA 2a\_2 overexpression and that they bind to the paternal genomic DNA locus of LSD1. RNase H/R treatment of the purified RNA confirmed our findings on circRNA interaction. In particular, knockdown and overexpression of circRNA 8a\_2, newly identified in the present study, affected the expression of the LSD1+8a alternative splice variant carrying exon 8a. The neurospecific LSD1+8a isoform was found to be specifically expressed in small cell lung cancer (SCLC) cells and its suppression impaired SCLC cell viability. Importantly, overexpression of circRNA 8a\_2 induced LSD1+8a expression even in normally LSD1+8a-negative LUAD cells.

## **Abstract**

In summary, these results show that circRNA originating from the exon 2/2a region or from the 8/8a region affect the alternative splicing of the respective exons in the linear LSD1 transcript by R-loop formation. Because the alternative splice variants differ in LSD1 demethylation activity, circRNA expression is hypothesized to affect LSD1 function in cancer.

## Zusammenfassung

### Zusammenfassung

Lungenkrebs ist eine der häufigsten bösartigen Erkrankungen weltweit, wobei das Adenokarzinom der Lunge (LUAD) der häufigste Subtyp ist. Es gibt immer mehr Belege dafür, dass die Epigenetik eine entscheidende Rolle bei der Entstehung und Entwicklung von Krebs spielt, darunter auch der wichtige Regulator Lysin-spezifische Demethylase 1 (LSD1). LSD1 reguliert die Genexpression durch Demethylierung von Histon 3 Lysin 4 und Lysin 9. Seine hohe Expression korreliert mit einer schlechten Prognose bei Krebspatienten. Es werden verschiedene alternative LSD1-Spleißvarianten von Exon 2a und 8a exprimiert, aber die Regulationsmechanismen sind unklar. Da das Spleißen mit Back-Splicing-Prozessen konkurrieren kann, die zirkuläre RNAs (circRNAs) erzeugen, habe ich mich auf LSD1-Spleißvarianten und circRNAs aus dem LSD1-Gen bei LUAD konzentriert.

Den TCGA-Daten zufolge unterschieden sich die Spiegel der LSD1+2a-Varianten, die das Exon 2a enthalten, im Vergleich zu den gesamten LSD1-Transkripten zwischen LUAD und angrenzenden Nicht-Tumorgewebe nur geringfügig. Unter den am stärksten regulierten circRNAs von LUAD-Zellen PC9 und normalen Lungenepithelzellen PSAE wurden zwei circRNAs identifiziert, die vom LSD1-Gen stammen. PCR-Studien mit divergenten Primer von Exon 2 identifizierten circRNAs, die von LSD1 abgeleitet sind (circRNA 3\_2\_2a, circRNA 3\_2 und circRNA 2a\_2). Durch Fluoreszenz-in-situ-Hybridisierung (FISH) und qPCR wurde festgestellt, dass circRNA 3\_2\_2a, circRNA 3\_2 und circRNA 2a\_2 im Zellkern und im Zytoplasma exprimiert wurden. Das Silencing von circRNA 2a\_2 durch siRNA, aber nicht von circRNA 3\_2\_2a und circRNA 3\_2 führte zu einer linearen LSD1+2a mRNA-Abnahme in A549- und PC9-Zellen. Darüber hinaus resultierte die Überexpression von circRNA 2a\_2 und circRNA 3\_2 zu einer Hochregulierung von LSD1+2a in PC9-Zellen und die Überexpression von circRNA 3\_2\_2a und circRNA 2a\_2 erhöhte die LSD1+2a-Spiegel in PSAE-Zellen. Pull-Down-Experimente mit dem S9.6-Antikörper, der R-Loop-Bildungen erkennt, zeigten, dass als Reaktion auf die circRNA 3\_2\_2a- und circRNA 2a\_2-Überexpression R-Loops gebildet werden und dass sie an den paternalem genomischen DNA-Locus des LSD1 binden. Die RNase H/R-Behandlung der gereinigten RNA bestätigte unsere Ergebnisse zur circRNA-Interaktion. Insbesondere der knockdown und die Überexpression der circRNA 8a\_2, die in der vorliegenden Studie neu identifiziert wurde, beeinflusste die Expression der alternativen LSD1+8a-Spleißvariante, die Exon 8a trägt. Es wurde festgestellt,

## Zusammenfassung

dass die neurospezifische LSD1+8a-Isoform spezifisch in kleinzelligen Lungenkrebszellen (SCLC) exprimiert wird und dass ihre Unterdrückung die Lebensfähigkeit von SCLC-Zellen beeinträchtigt. Wichtig ist, dass die Überexpression der circRNA 8a\_2 die Expression von LSD1+8a sogar in normalerweise LSD1+8a-negativen LUAD-Zellen induzierte.

Zusammenfassend zeigen diese Ergebnisse, dass circRNA, die aus der Exon 2/2a-Region oder aus der 8/8a-Region stammen, das alternative Spleißen der jeweiligen Exons im linearen LSD1-Transkript durch R-Schleifenbildung beeinflussen. Da sich die alternativen Spleißvarianten in der LSD1-Demethylierungsaktivität unterscheiden, wird angenommen, dass die circRNA-Expression die LSD1-Funktion bei Krebs beeinflusst.

## Table of Content

### Table of Content

<b>1.</b>	<b>INTRODUCTION.....</b>	<b>1</b>
<b>1.1</b>	<b>LUNG CANCER .....</b>	<b>1</b>
1.1.1	EPIDEMIOLOGY OF LUNG CANCER.....	1
1.1.2	PATHOPHYSIOLOGY .....	2
1.1.3	GENETIC MUTATIONS AND PATHWAYS .....	3
1.1.4	ETIOLOGY .....	4
1.1.5	THERAPY OF LUAD .....	4
<b>1.2</b>	<b>LSD1 .....</b>	<b>5</b>
1.2.1.	VARIANTS OF LSD1 .....	5
1.2.2.	ROLES OF LSD1 IN CELLULAR PROCESSES.....	6
1.2.3.	THERAPEUTIC OPTION IN CANCER.....	9
<b>1.3</b>	<b>CIRCULAR RNAS (CIRC RNA).....</b>	<b>10</b>
1.3.1	METABOLISM OF CIRC RNAS.....	10
1.3.2	LOCALIZATION .....	12
1.3.3	FUNCTION OF CIRC RNAS .....	12
1.3.4	POTENTIAL THERAPEUTIC TARGETS .....	15
<b>1.4</b>	<b>AIMS OF THE STUDY .....</b>	<b>16</b>
<b>3.</b>	<b>MATERIALS AND METHODS.....</b>	<b>17</b>
<b>2.1</b>	<b>MATERIALS.....</b>	<b>17</b>
2.1.1	LIST OF SOFTWARE .....	17
2.1.2	LIST OF DEVICES .....	17
2.1.3	LIST OF KITS .....	18
2.1.4	PLASTIC MATERIALS .....	18
2.1.5	LIST OF CELL LINES.....	19
2.1.6	ANTIBODIES.....	20
2.1.7	ENZYMES.....	20
2.1.8	PRIMER LIST .....	20
2.1.9	PROBE LIST .....	21
2.1.10	SMALL INTERFERING RNA (SIRNA) LIST.....	21
2.1.11	SOLUTIONS .....	21
<b>2.2</b>	<b>METHODS.....</b>	<b>28</b>
2.2.1	CELL CULTURE .....	28
2.2.2	PASSAGING CELLS .....	28
2.2.3	DETERMINATION OF CELL NUMBERS.....	28
2.2.4	CRYOPRESERVATION AND THAWING OF CELLS.....	28
2.2.5	MYCOPLASMA TESTING.....	29

## Table of Content

2.2.6	DATA ACQUISITION AND ANALYSIS FROM TCGA .....	29
2.2.7	EXTRACTION OF TOTAL RNA PREPARATION FROM CLINICAL SAMPLES.....	29
2.2.8	COMPOUNDS TREATMENT .....	29
2.2.9	SIRNA TRANSFECTION.....	29
2.2.10	QUANTITATIVE ANALYSIS OF MRNA EXPRESSION LEVELS .....	30
2.2.11	FLUORESCENCE IN SITU HYBRIDIZATION (FISH).....	31
2.2.12	CYTOPLASMIC LYSIS .....	32
2.2.13	WESTERN BLOT .....	32
2.2.14	IMMUNOPRECIPITATION (IP) OF R-LOOPS .....	33
2.2.15	MTT ASSAY.....	34
2.2.16	RNA PURIFICATION.....	34
2.2.17	STATISTICAL ANALYSIS .....	34
4.	RESULTS.....	35
3.1	LSD1 SPLICE VARIANT 2A EXPRESSION PATTERN IN CANCER AND NON-CANCER	35
3.1.1	LSD1 +2A EXPRESSION IN CANCER AND NON-CANCER CELL LINES .....	35
3.1.2	LSD1 +2A EXPRESSION IN DIFFERENT CANCER TYPES.....	36
3.2	CIRCRNAS FROM LSD1 GENE.....	40
3.2.1	IDENTIFICATION OF CIRCRNAS GENERATED FROM LSD1 GENE .....	40
3.2.2	LEVELS OF CIRCRNAS IN CANCER AND NON-CANCER CELL LINES.....	43
3.2.3	LEVELS OF CIRCRNAS IN LUAD AND NON-TUMOR TISSUES .....	44
3.2.4	LOCALIZATION OF CIRCRNAS IN CELLS.....	45
3.3	CIRCRNAS REGULATE ALTERNATIVE SPLICING OF LSD1 BY R-LOOP FORMATION	49
3.3.1	KNOCKDOWN OF CIRCRNA 2A_2 INHIBITED THE EXPRESSION OF LSD1+2A ....	49
3.3.2	DECREASE OF LSD1+2A WAS REGULATED BY CIRCRNA 2A_2 KNOCKDOWN NOT BY SIRNA FOR CIRCRNA 2A_2 .....	50
3.3.3	OVEREXPRESSION OF CIRCRNAS FROM LSD1 UPREGULATED THE LEVELS OF LSD1+2A.....	51
3.4	CIRCRNA 3_2_2A FORMED R-LOOPS WITH THE PARENTAL GENE .....	52
3.5	LSD1+8A .....	56
3.5.1	LSD1+8A IN SCLC .....	56
3.5.2	LSD1+8A EXPRESSION IN SCLC CELL LINES .....	56
3.5.3	SUPPRESSION OF LSD1+8A REPRESSED THE CELL VIABILITY IN SCLC .....	58
3.5.4	CORRELATION BETWEEN LSD1+8A AND NEUROENDOCRINE MARKERS.....	59
3.5.5	CIRCRNA 8A_2 REGULATED THE LEVELS OF LSD1+8A .....	60
5.	DISCUSSION .....	62
4.1	IDENTIFICATION OF CIRCRNAS FROM LSD1 ISOFORMS .....	62
4.2	EXPRESSION OF CIRCRNAS, DERIVING FROM THE LSD1 GENE LOCUS, IN CANCER .....	63



## Table of Content

<b>4.3</b>	<b>CIRCRNAS FROM LSD1 REGULATE THE ALTERNATIVE SPLICING OF THE PARENTAL GENE LSD1.....</b>	<b>66</b>
<b>4.4</b>	<b>CIRCRNAS FROM LSD1 FORM R-LOOPS WITH ITS PARENTAL GENE.....</b>	<b>67</b>
<b>6.</b>	<b>REFERENCES.....</b>	<b>71</b>
<b>7.</b>	<b>ABBREVIATIONS .....</b>	<b>88</b>
<b>8.</b>	<b>SUPPLEMENTARY INFORMATION .....</b>	<b>91</b>
<b>9.</b>	<b>ACKNOWLEDGEMENTS .....</b>	<b>93</b>
	<b>CURRICULUM VITAE.....</b>	<b>94</b>
	<b>ERKLÄRUNG .....</b>	<b>96</b>

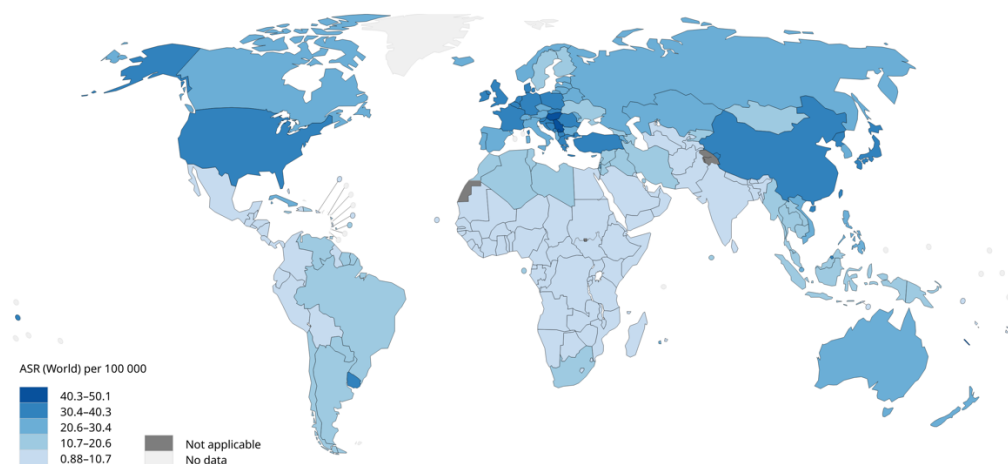
## 1. Introduction

### 1.1 Lung cancer

Lung cancer is among the most prevailing malignancy and the primary reason for fatalities from cancer in all countries, with an estimated 2.1 million new cases and 1.8 million deaths in 2018 [1, 2]. In Germany, 57,000 new lung cancer cases were detected in 2018 [3]. In China, it was anticipated that there will be 870,982 new cases of lung cancer in 2022 with 766,898 estimated deaths [4]. While it was projected that there will be approximately 238,032 new cases with 144,913 estimated deaths in the USA [4]. Thus, understanding the mechanisms of the initiation and development of lung cancer is urgently needed.

#### 1.1.1 Epidemiology of lung cancer

Uruguay, Argentina, eastern Asia, Europe, and North America have the greatest incidence rates of lung cancer in males, whereas sub-Saharan Africa has the least incidence rates [2] (Figure 1). Indeed, the incidence and fatality rate of lung cancer are consistently lower in women compared with men [5]. The incidence of lung cancer is falling at a rate that is twice as rapid in males as it is in females, which reflects the historically slower rate of tobacco use and quitting among females [6]. Lung cancer is responsible for nearly a quarter of all cancer-related deaths [7].



**Figure 1: Lung cancer incidence per 100,000 individuals in 2020.** Figure and Data from GLOBOCAN 2020 (<http://globocan.iarc.fr/Default.aspx>). ASR, age-standardized incidence rates.

## Introduction

### 1.1.2 Pathophysiology

Lung cancer is classified broadly as non-small cell lung cancer (NSCLC) (85% of total diagnoses) [8] or small cell lung cancer (SCLC) (15% of total diagnoses) [9, 10]. Adenocarcinomas are the most predominant subtype of lung cancer under the NSCLC category, followed by squamous cell carcinoma and large-cell carcinoma [11]. The lung is an intricate but delicate organ that is made up of several different kinds of cells, each of which performs a specific role that enables gas exchange [12]. This intricate collection of cells may experience a buildup of cell-autonomous and microenvironmental adaptive responses, which might shift the equilibrium of cell division as well as death and allow cancer to progress by avoiding immune detection [13]. Exposure to smoke may result in a well-defined sequence of morphological alterations to the bronchial epithelium, beginning with basal cell hyperplasia and advancing through metaplasia, severe dysplasia, carcinoma in situ, and eventually frank carcinoma [14, 15]. Such progression of alterations is most often related to the squamous subtype of NSCLC [16]. However, adenocarcinomas are widely accepted to be the predominant subset among never-smokers who have minimal exposure to carcinogens even though they may also develop under severe carcinogenic exposure long with established lung injury [17]. The onset and advancement of adenocarcinomas are linked to precancerous lesions known as atypical adenomatous hyperplasia, which is less well-elucidated [18]. SCLCs likewise frequently emerge in presence of substantial carcinogenic exposure [19, 20]; however, they originate from uncommon pulmonary neuroendocrine cells and lack well-defined preneoplastic lesions [21].

Due to the lack of clinical signs and efficient screening systems, the vast majority of cases of lung cancer are diagnosed at a late stage [22]. Accurate lung cancer staging is crucial because it determines available treatment choices and patient outcomes. The International Association for the Study of Lung Cancer staging committee revised the Tumor-Node-Metastasis (TNM) classification framework used as the foundation for cancer staging after reviewing the findings derived from a global patient database [23]. This histological categorization of lung cancer is constantly evolving, and precise terminologies and parameters are employed to differentiate squamous cell carcinoma and adenocarcinoma, especially in tumors

## Introduction

with poor differentiation. In the most recent few years, determining the specific histopathological subtypes of lung cancer is becoming increasingly significant due to the advent of an overwhelming number of therapeutic drugs that target certain variants.

### 1.1.3 Genetic mutations and pathways

There are diverse types of lung cancer with its distinct features, spanning from surgically resectable NSCLCs to highly aggressive metastatic SCLCs. The identification of mutations in the driver gene in a subgroup of lung malignancies was a major advance toward comprehending these differences. These inherited genetic mutations in kinases cause constitutive signaling, resulting in the neoplastic transformation that is almost completely independent effects of other modifications. Oncogenes associated with the initiation and development of NSCLC include activating mutations in the Kirsten rat sarcoma viral oncogene homologue (KRAS) gene and epidermal growth factor receptor (EGFR) genes. There is marked geographical EGFR mutation in prevalence, ranging from 15% in Europe to 62% in Asia [24]. Exon 19 deletion ( $\Delta$ E746–A750), L858R mutation in exon 21, as well as mutations in codon 719 of exon 18 (G719A, G719S, G719C), are the most prevalent EGFR mutations, which are all responsive to therapy [25]. KRAS mutations occur in about one-quarter of all LUAD cases and is the most frequent gain-of-function mutations in Europe. The majority of KRA genetic mutations are found in codon 12 of exon 2, with some also being found at codon 13 and hardly ever found at codon 61 of exon 3 [26]. The mutations of KRAS compromise the GTPase activity, blocking the protein in a persistent GTP-bound state and initiating LUAD. Deletion of p53 and RB1 is a common variation in SCLC, which has been recognized for many decades [27]. Additional previous investigations indicated the amplification of genes belonging to the MYC family (MYCN, MYCL, and MYC) in a subtype of SCLC [28]. Loss-of-function occurrences in RB family member p107 and p130 (encoded respectively by RBL1 and RBL2) [29], the tumor suppressor phosphatase and tensin homolog (PTEN) [30], neurogenic locus notch homolog protein (NOTCH) receptors [31] and the chromatin regulator CREB binding protein (CREBBP) [32] were among the few factors that have been experimentally verified in vivo or in vitro.

### 1.1.4 Etiology

The key factor in the onset of lung cancer is long-term cigarette smoking behavior [33]. The continuous inhalation of tobacco smoke is responsible for the onset and progression of a substantial proportion (80%) of lung cancer cases [8]. The chance of developing lung cancer among nonsmokers who were exposed to tobacco smoke was much greater than never smokers [34]. A modest contribution is made by secondhand smoking to the etiology of lung cancer [35]; a more extensive study may be necessary to establish a clear causal link between the two factors.

Even though smoking is by far the most prominent risk indicator for developing lung cancer, there is mounting evidence to suggest that genetic predisposition could also cause susceptibility to lung cancer onset as well as its progression [36]. As I described in the former part, mutations and single nucleotide polymorphisms (SNP) enhance the risk of lung cancer [37]. In SCLC, the deletion of p53 and RB1 is a common phenomenon, which has been recognized for many years [38]. Several early investigations reported the amplified expression of genes belonging to the MYC family (MYCN, MYCL, and MYC) in a subset of SCLC tumors [39]. A history of respiratory disease, viral infections, occupational and environmental exposure to carcinogens including radon, asbestos, and arsenic, as well as air pollution, all perform an extremely essential part in the development of lung cancer, either on their own or in combination with other factors whose impacts are additive or synergistic [40-43].

### 1.1.5 Therapy of LUAD

Among patients who are in a healthy state to receive surgical treatment and have operable early-stage lung cancer, lobectomy is still the recommended therapeutic option and patients with an inadequate physiological reserve may benefit from radical radiotherapy as an alternate treatment option with a decreased risk of morbidity [44]. The advent of stereotactic ablative radiotherapy (SABR) for treating peripheral tumors in a new era for early-stage lung cancer [45]. Surgical interventions with adjuvant (postsurgical) chemotherapy or combined chemotherapy and radiotherapy are two alternative treatments that are currently recommended by NICE 28 for patients who suffered from locally advanced lung cancer with lymph

## Introduction

nodes in the hilar or mediastinal regions [46]. Immunotherapy for different immune checkpoints (cytotoxic T-lymphocyte-associated protein 4 receptor (CTLA4), programmed death (PD-1) receptor, and programmed death-ligand 1 (PD-L1)) has led to unprecedented prolonged survival for some LUAD patients [47]. Chemotherapy is the primary treatment option for individuals diagnosed with metastatic SCLC and consolidation chest radiation is a helpful treatment option for people who respond well to chemotherapy [48]. Notably, due to the critical involvement of lysine-specific demethylase 1(LSD1) in carcinogenesis, the inhibitors for LSD1 (GSK2879552) are also an emerging option for SCLC [49]. Several inhibitors of LSD1 undergo clinical evaluation for cancer treatment (NCT02875223, NCT04081220 and EudraCT no.: 2018000482-36) [50].

### 1.2 LSD1

In recent decades, epigenetics has played vital roles in the initiation and development of cancer including LUAD [51]. LSD1, commonly referred to as KDM1A is one of the key regulators in epigenetics [52]. It was the first histone lysine demethylase observed, having the capability of demethylating H3K4me<sub>1/2</sub> and H3K9me<sub>1/2</sub> at target loci in a way that is reliant on the context [52]. LSD1 is very well conserved and has three protein domains, which are as follows: a C-terminal amine oxidase (AO) domain, a central protruding tower domain, and an N-terminal SWIRM (Swi3p/Rsc8p/Moira) structural domain [53]. Interaction between the SWIRM and AO domains results in the formation of a core structure that functions as the enzymatic domain and binds FAD in a non-covalent manner; a surface platform for engagement with partners is provided by the tower domain [54].

#### 1.2.1. Variants of LSD1

Among the splice variants of LSD1, identified in the human genome database on the Ensembl Genome Server (<https://www.ensembl.org>) a full-length human LSD1 variant (KDM1A-202) (transcript ID: ENST00000400181.9) is a LSD1 isoform that contains two additional peptides from exon 2a and 8a, whereas KDM1A-201 (ENST00000356634.7) contains no additional peptides (Figure 2). Considering the

## Introduction

combinatorial retention of exons 2a and 8a, there are four main alternative splicing variants of LSD1 in humans.



**Figure 2: Schematic image of human LSD1 splicing variants.** SWIRM, TOWER and AOD are three domains of LSD1. Exon 2a was shown in the red rectangle, and exon 8a was revealed in the green one. It is primarily composed of 3 main domains: a C-terminal amine oxidase like domain (AOD), an N-terminal SWIRM domain, and a protruding tower domain.

### 1.2.2. Roles of LSD1 in cellular processes

In both normal and malignant cells, LSD1 exerts an essential function in various fundamental biological functions, notably metabolism, epithelial-to-mesenchymal transition, cell motility, differentiation, and stemness control [55-57]. LSD1 acts as a context-dependent coregulator capable of both co-repressor and co-activator functions by its existence in multiprotein complexes [58].

#### 1.2.2.1 LSD1 leads to repression or activation of transcription

In the early stages of embryogenesis, LSD1 is responsible for regulating the expression of critical developmental modulators and the right timing of their actions. Silencing of LSD1 led to the death of mouse embryos on or before the sixth day of embryogenesis [59]. As a result of enhanced cell death, delayed cell cycle progression, and abnormalities in differentiating, embryonic stem cells (ESCs) generated from LSD1 knockout mice experienced severe growth limitation [59]. In the differentiation process, the levels of LSD1 are reduced from the high levels in undifferentiated human ESCs [60]. By controlling the fine-tuned equilibrium between H3K27 and H3K4 methylation in the regulatory regions of key genes including Forkhead Box A2 (FOXA2) and Bone morphogenetic protein (BMP2), LSD1

## Introduction

contributes to the suppression of lineage-specific developmental events and is crucial for the maintenance of pluripotency [61]. The process of epithelial–mesenchymal transition (EMT), which is the acquisition process by epithelial cells of the phenotype of mesenchymal cells, is necessary for cancer cell invasion and metastasis [62]. LSD1's role in controlling EMT has been confirmed in breast cancer [63]. LSD1 binds to SNAI1, which in turn suppresses the production of E-cadherin, a key EMT marker [64]. In addition, silencing LSD1 reduces cancer cell motility and invasiveness regardless of tumor types [64].

Through demethylation of repression-related monomethyl- and dimethyl-histone H3K9 (H3K9me1/2) markers, LSD1 may also perform a remarkable function as a co-activator in androgen receptor (AR) and estrogen receptor (ER)-regulated transcription [65]. Proline glutamate and leucine rich protein 1 (PELP1), a protein that interacts with the ER, changes the substrate selectivity of LSD1, shifting it from H3K4 to H3K9 [66]. The demethylation of histone H3K9 by LSD1 was discovered to be crucial for AR target gene activation in a study by Metzger et al [67]. In response to hormonal treatment, promoters for AR and LSD1 are colocalized in both the normal and tumorous human prostate, where they activate H3K9 demethylation without altering the enrichment of H3K4 methylation and stimulating the transcription of AR target genes by ligands, causing increased growth of tumor cells [67]. Both the downregulation of the protein levels of LSD1 and the silencing of LSD1 inhibits the transcriptional activation and cell proliferation stimulated by AR [67].

### 1.2.2.2 LSD1 in cancer

LSD1 has various oncogenic functions since it is expressed at a high level in many different types of cancer cells [68, 69]. In our previous studies, the levels of LSD1 from 182 cases of LUAD samples revealed the correlation between LSD1 overexpression and the progression and metastasis of LUAD. LSD1 has a unique function in an invasive phenotype of NSCLC cells by modulating the non-canonical integrin pathway [70]. Self-renewal is promoted by LSD1 via the regulation of the integrin 3-KRAS-NF-B pathway [70]. Suppression of LSD1 led to the arrest of tumor growth in LUAD independently of driver mutations [71]. In other solid tumors like gastric cancer, it is determined that high expression of LSD1 is implicated in various



## Introduction

pathogenic mechanisms behind the development of proliferation, apoptosis, and metastasis of gastric cancer cells [72]. Furthermore, the presence of aggressive and undifferentiated neuroblastoma is associated with overexpression of LSD1. Suppression and ablation of LSD1 led to an enhancement in global H3K4 methylation, resulting in the stimulation of genes related to differentiation and as a direct consequence, neuroblastoma cell development was slowed down both in vitro and in vivo [73].

LSD1 is upregulated in acute myeloid leukemia (AML) subgroups with less differentiation (including the M1 subtype, based on FAB classification) as opposed to other subgroups hallmarked by a greater extent of morphological differentiation [74]. A substantial correlation was found between LSD1 silencing and the absence of LSC capacity in AML cells, which was achieved by the stimulation of differentiation and apoptosis [75]. A deficiency of LSD1 in leukemic cells renders them incapable of forming colonies in vitro (AML-CFC) or transferring leukemia into subsequent mouse recipients [76]. Collectively, these findings provided evidence that LSD1 has the tumorigenic potential of AMLs, indicating a viable target for cancer therapy [76]. In Notch-mediated T-ALLs, LSD1 can serve as a repressor or an activator. In the absence of Notch, LSD1 functions as a co-repressor in combination with the CSL-repressor complex by removing histone H3 lysine 4 demethylation (H3K4me2) markers from Notch targets [77]. Nonetheless, upon Notch activation, LSD1 acts as a co-activator of NOTCH1 by facilitating efficient demethylation of H3K9me2 [77].

One important variant LSD1+8a, which includes the exon 8a, is specifically expressed throughout the neural development of mammals [78]. There is no innate demethylation capacity in this neuronal isoform for H3K4me2 although it demethylates the restrictive mark H3K9me2 in conjunction with its partner SVIL/supervillin, thereby promoting the transcription of neuronal-specific genes upon differentiation [79]. In vitro neuronal development may be hindered by targeting precisely the LSD1+8a isoform and knocking it down and, accordingly mice lacking LSD1+8a exhibit cognitive impairments and have trouble learning their environments spatially [80]. Strikingly, LSD1+8a is abundantly expressed in various SCLC cell lines, which is consistent with the feature that SCLC is an invasive neuroendocrine tumor [81]. Enhanced expression of LSD1+8a was associated with chemoresistance

## Introduction

to cisplatin (CDDP) and LSD1 inhibitors [82]. The inhibition of LSD1+8a was shown to suppress cell proliferation, which suggests that it exerts an important function in SCLC [82].

### 1.2.3. Therapeutic option in cancer

Targeting demethylase is becoming a desirable alternative as a treatment for cancer patients because of the essential role performed by LSD1 in the process of tumorigenesis and because of the numerous ways where it interacts with a multitude of signaling pathways [31, 83]. Some of the LSD1 inhibitors that have been identified are the following: ORY-2001 (vafidemstat), INCB059872, CC-90011, IMG-7289 (bomedemstat), GSK-2879552, ORY-1001 (iadademstat), and tranlycypromine (TCP or PCPA) [84]. As a new therapeutic anticancer target, pharmacological inhibition of LSD1 was proven to suppress the capacity of cancer cells to differentiate, proliferate, migrate, and invade, especially in acute myeloid leukemia [85] and SCLC [86]. In recent years, a multitude of distinct LSD1 inhibitors has been combined with a variety of other substances to test various combinatorial treatments [55]. According to the findings of Fiskus and his colleagues, coordination toxicity in cultured primary AML blasts was observed when the reversible LSD1 antagonist SP2509 was used together with the pan-HDAC antagonist panobinostat (synergistic effect) [87]. Presently, combination treatments of LSD1 inhibitors (CC-90011, IMG-7289, INCB059872, and TCP) with chemotherapy (etoposide, cisplatin, azacitidine, cytarabine, and ATRA), histone deacetylase (HDAC) inhibitors, the inhibitor of indoleamine 2,3-dioxygenase-1 (IDO1) (epacadostat), monoclonal antibody (pembrolizumab) [88], and the NEDD8-activating enzyme (NAE) inhibitor (pevonedistat) [71] are now being researched for potential application in the treatment of cancer. Consequently, LSD1 is proposed as a viable target for the treatment of cancer as a result of its prominent function in cancer.

### 1.3 Circular RNAs (circRNA)

Since LSD1 is alternatively spliced and alternative splicing is often accompanied with back splicing mechanisms, my presented work focused on the roles of circRNAs derived from the alternatively spliced LSD1 locus. In recent decades, it has been established that circRNAs function as crucial regulators in epigenetics. CircRNAs were originally identified in 1976 in a virus isolated from a plant [89], and in 1979, electron microscopy revealed their presence in human HeLa cells [90]. CircRNAs are single-stranded transcripts that have a covalently closed loop and are generated via the alternative splicing process of back splicing from a precursor mRNA (pre-mRNA) [91, 92], giving them an innate defense against degradation by exonucleases of their RNA. Nonetheless, with the advent of bioinformatics and high-throughput sequencing, circRNAs are found to be highly expressed and evolutionarily conserved across species [93-95] that are known to exhibit tissue/development-specific expression. There is mounting evidence to suggest that circRNAs are intimately implicated in the pathogenesis of a myriad of distinct illnesses, which include Alzheimer's disease [96], osteoarthritis [97], diabetes [98], cardiovascular disease [99] and cancer [100].

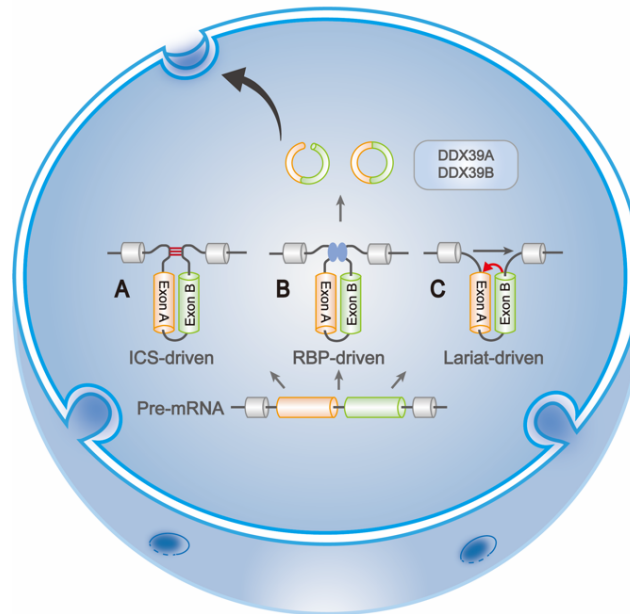
#### 1.3.1 Metabolism of circRNAs

##### 1.3.1.1 Biogenesis

Even though back splicing is deemed to be a kind of alternative splicing, the molecular process that underlies it is different from that of linear alternative splicing [101]; Nevertheless, the processes that underlie the formation of circRNA are not yet completely elucidated. The 3' untranslated region (UTR), introns, exons, the 5' UTR, or even antisense sequences may all be potential sources of circRNAs [102]. CircRNAs have been classified into the following four groups: tRNA intronic circRNAs (trircRNAs), exon-intron RNAs (elciRNAs), circular intronic RNAs (ciRNAs), and exonic circRNAs (ecircRNA) (Figure 3) [103]. It is hypothesized that an RNA-binding protein (RBP) or an intronic complementary sequence is the driving molecular regulating this back splicing process [104]. Another mechanism of their formation is through a model of biogenesis based on lariat-driven back-splicing involving the cutting off of the RNA sequence after the pre-mRNA is spliced,

## Introduction

resulting in a lariat structure that can be subsequently spliced to form a circular structure [105].



**Figure 3: Biogenesis of circRNAs.** (A) Intronic complementary sequences (ICSs) are paired and closely connected sequences that stimulate back-splicing to generate a circular morphology in the ICS-driven back-splicing biogenesis model. (B) The splicing domains at both terminals of the exon are directly coupled and stimulate back-splicing in the back-splicing biogenesis model mediated by RNA-binding protein (RBP). (C) The RNA sequence that is discarded in the linear RNA splicing generates a lariat structure in the lariat-driven back-splicing biogenesis model before being back-spliced into a circular morphology. Mature circRNAs are transported from the nucleus to the cytoplasm by the proteins DDX39A and DDX39B.

### 1.3.1.2 Regulation

CircRNA biogenesis depends on canonical splicing machinery, including splice signal sites and spliceosomes [106]. Nevertheless, blocking the pre-mRNA processing machinery redirects gene output to circRNAs [107], which suggests that circRNAs and the linear versions of their counterparts engage in a competition. Recently, it was discovered that N6-methyladenosine (m6A) plays a role in affecting the biogenesis of circular RNA [108]. Depletion of either YTH domain-containing 1 (YTHDC1) or methyltransferase-like 3 (METTL3) regulates around 20 percent of a subset of circRNAs, but there are no considerable variations in the linear isoforms [108]. YTHDF3, in the meantime, is the factor that identifies the m6A-modified start codon and facilitates the beginning of translation [109]. Nonetheless, how m6A

deposition regulates the choice between back and canonical splicing is not well known.

### 1.3.2 Localization

Visualizing circRNAs in cells is vital for discovering their functions [110]. The majority of circRNAs are normally transported from the nucleus into the cytoplasm shortly following their biogenesis, except for those circRNAs that contain introns [111]. It has been found that the nucleus contains a significant amount of both intronic elciRNAs and circRNAs [112]. Accumulating literature data suggested that circRNA shuttling is controlled by RNA length as well as modifications [113, 114]. Huang and colleagues discovered a length-dependent evolutionarily conserved pathway to control nuclear export of circRNAs [113].

### 1.3.3 Function of circRNAs

It is well determined that circRNAs serve diverse activities, which included functioning as sponges for miRNA, communication, producing pseudogenes, influencing alternative splicing and transcription, translation, transportation, and interfacing with RBPs [110].

#### 1.3.3.1 MiRNA sponges

An extensive research has proven that circRNAs exert important biological functions as miRNA or ceRNAs sponges [115, 116]. In addition, microRNAs are a kind of small RNAs that are below 20 nucleotides in length [117]. Because highly abundant circRNAs with numerous competing binding domains have a greater likelihood of possessing competing endogenous RNA functions, the stoichiometric connection between the miRNA binding domains of the circRNA must be taken into consideration when estimating miRNA sponge functionalities to circRNAs [116]. For example, the antisense transcript of cerebellar degeneration-related protein 1 (cdR1), also known as cIRS7, which is the first reported miRNA sponge produced from the cdR1 gene, can regulate miR7 in a negative manner [111], found in hepatocellular carcinoma (HCC) and gastric cancer [118]. Moreover, circFAT1 is overexpressed in papillary thyroid carcinoma (PTc) cells and tissues and acts as a miRNA sponge to remarkably lower the expression level of miR-873. Consequently, the activity of the

## Introduction

zinc finger E-box binding homeobox 1 was enhanced, which eventually led to the stimulation of PTC cells' ability to proliferate, migrate, and invade [119]. Wang and colleagues illustrated that circPTK2 (hsa\_circ\_0008305) acted as a miR-429 sponge, resulting in the promotion of the cell invasion of NSCLC [120].

### 1.3.3.2 Protein interaction

In addition to acting as a sponge for miRNA, one of the most significant functions that circRNAs perform is interacting with proteins [121]. RNA binding proteins (RBPs) are some of the most well-known proteins because of their roles in the control of RNA metabolism in a variety of disciplines, particularly formation, transportation, translation, and localization [122]. The cMYc protein, which shares consensus sequences with other RNA binding proteins, may interface with circAmotl1, facilitating nuclear translocation of cMYc, thus improving its stability and increasing the expression cMYc targets [123]. Additionally, a similar connection exists between circ-Amotl1 and the nuclear translocation of signal transducer and transcription 3 activators [124]. It has been shown that circRNA\_102171, which is expressed at a high level in PTC tissues, binds precisely with  $\beta$ -catenin interacting protein 1 (cTNNBIP1) and impedes its link to the  $\beta$ -catenin/T cell factor (TCF)3/TCF4/lymphoid enhancer factor 1 complex, hence promoting the advancement of PTC [125]. The research demonstrates that certain circRNAs are capable of interacting with a variety of proteins [126], while some proteins are also capable of dynamically binding to distinct circular RNAs [127].

### 1.3.3.3 Encoding proteins/ peptides

Despite the common assumption that circRNAs do not code for proteins, numerous research reports have proven that circRNAs are translated into proteins [128]. Researches have illustrated that circMbl3 [129] and circZNF609 [130] may be translated into proteins in a cap-independent and internal ribosome entry site (IRES)-dependent way. N6-methyladenosine (m6A) was also discovered to be a critical RNA base modification for efficient circRNAs translation initiation [131]. Even though some circRNA-derived proteins have been associated with malignancies, a variety of circRNA-derived proteins' functions are yet unknown [132]. CircSHPRH (hsa\_circ\_0001649) was found to encode a peptide consisting of 146 amino acids (SHPRH-

## Introduction

146aa) [133]. The E3 ubiquitin ligase DTL may be implicated in the interaction with both the full-length SHPRH protein and SHPRH-146aa. Stabilized SHPRH functions as an E3 ubiquitin ligase, resulting in successive ubiquitination of proliferating cell nuclear antigen (PCNA), thus contributing to the attenuated proliferation of cells as well as tumorigenicity [134].

### 1.3.3.4 Regulation of the expression or the alternative splicing

Prior research has demonstrated that elciRNA is capable of interacting with U1 snRNA, which results in the formation of the complex clciRNA-U1 snRNP [135]. This complex facilitates the expression of the host gene via its interactions with the PolII transcription complex [135]. Lu et al. discovered in their study of plants that circular RNAs as well as their linear isomers were capable of suppressing the host gene post-transcription expression via the mechanism of building a genetic transformation system of rice that resulted in upregulation [136]. CircRNAs may modulate the alternate splicing that is present in host genes [137]. CircRNA can tightly bind to the host gene DNA locus in *Arabidopsis thaliana*, which results in the formation of an R-loop configuration of RNA: DNA heterozygote [138]. The structure of an R-loop may block transcription from occurring in this region. The flowering phenotype is altered because of the presence of a cross-exon alternative splicing event, which in turn favors the development of an alternative splicing transcript variation known as SEP3.3 [138].

### 1.3.3.5 Oncologic biomarkers

The clinical application of biological markers is essential throughout all stages of carcinogenesis, and it has emerged as one of the primary strategies for diagnosing and determining the prognosis of cancer [139]. CircRNAs' patterns of expression and their individual characteristics (specific expression, highly conserved, high stability, and high and selective abundance) might help explain, at least in part, why circRNAs have such promise as prospective biological markers and therapeutic targets [140]. The existence of circRNAs may be identified by the analysis of tissue cells [141], exosomes [142], plasma [143], serum [114], cerebrospinal fluid (CSF) [144], urine [143], saliva [145], along with a variety of other bio-specimens. When pathological circumstances are present, the normally occurring expression of a vast number of

## Introduction

circRNAs alters in a significantly aberrant way [146]. In HCC tissues, several studies have illustrated the upregulation of hsa\_circ\_0000798 [147], and hsa\_circ\_0058124 [148] and the attenuated expression of circRNAs such as circSMARCA5 [149] and hsa\_circ\_0076251 [150]. Such circRNAs are being evaluated for their use as possible biological indicators of HCC [147, 148]. Moreover, the TNM stage in gastric cancer was shown to correlate with the hsa\_circ\_0000467 expression levels, which were observed to be elevated in the gastric cancer cells and tissues as well as plasma, in contrast with the healthy controls [151]. When compared to other established biological markers, the hsa\_circ\_0000467-related area under the curve (AUC) of 0.799 was remarkably higher [151]. The specificity and sensitivity of the hsa\_circ\_0000467 test were, correspondingly, 64.8 and 70.5 percent [151]. In patients with colorectal cancer, some circRNAs may also act as biological markers. Elevated plasma levels of hsa\_circ\_0000370 [152] were substantially linked to lymph node metastasis, whereas the overexpression of hsa\_circ\_0004585 was linked to the patient's tumor size [153].

### 1.3.4 Potential therapeutic targets

It has been revealed that circRNA 100146 is overexpressed in NSCLC samples [154]. SiRNA against circRNA 100146 suppressed tumor growth in vivo and inhibited the capacity to proliferate, migrate, and invade in vitro, and enhanced the rate of apoptosis [154]. Meanwhile, NSCLC tissues showed downregulation of circPTPRA, which served as a miR-96-5p sponge and had tumor-suppressing properties [155]. According to research done by Liu and his colleagues, a synthetic circRNA that contains many carcinogenic miR-21 binding sites is capable of efficiently suppressing the capacity of gastric cancer cells to proliferate by functioning as a miR-21 sponge [156]. Therefore, numerous studies confirmed the potential therapeutic values of circRNAs in cancer therapy.



### 1.4 Aims of the study

LSD1 regulates gene expression by demethylating histone 3 lysine 4 and lysine 9, and its high expression correlates with poor prognosis in cancer patients. Several alternative LSD1 splice variants of exon 2a and 8a are expressed, but the regulatory mechanisms are unclear. Because splicing may compete with back-splicing processes that generate circRNAs, in the presented thesis I focused on LSD1 splice variants and circRNAs from the LSD1 gene in LUAD.

Firstly I aimed to study the splice variant of LSD1 including exon 2a (LSD1+2a) in different lung cancer types using TCGA datasets. Furthermore, LSD1+2a and circRNA originating from the LSD1 gene 2a locus should be investigated in the lung adeno cancer and non-cancer cells as well as in clinical samples with lung adenocarcinoma (LUAD). Finally, circLSD1-RNA, the function of identified in lung cancer cells in splicing of the paternal gene locus should be analyzed. To this end the expression of circLSD1-RNA including the alternative spliced exons should be manipulated by siRNA technology and transgenic overexpression using circRNA encoding plasmids.

### 3. Materials and Methods

#### 2.1 Materials

##### 2.1.1 List of software

Listed below are the software used during the study.

Application	Name	Developer
Data analysis	Excel 2010	Microsoft, Redmont, USA
	GraphPad Prism 8	GraphPad Software, Inc., La Jolla, USA
	CLC Sequence Viewer 8.0	QIAGEN
	Chromas	Technelysium Pty. Ltd.
Real-Time PCR analysis	BioRad IQ5	BioRad, München, GER
	Lightcycler®480 SW 1.5	Roche, Mannheim, GER
	Stratagene MxPro 3000P V4.00	Stratagene, La Jolla, USA
Imaging	Image Lab	BioRad, München, GER
	CellP	Olympus Soft Imaging Solutions, Münster, GER
	Adobe photoshop CC 2018	Adobe, Dublin, Ireland
	Image J	Open Source

##### 2.1.2 List of devices

Listed below are the mechanical devices used during this study.

Name	Manufacturer
BioRad CFX96 Real-time PCR Cycler	Bio-Rad, Munich, GER
Olympus Fluoview FV 1000 (Confocal microscope)	Olympus, Hamburg, GER
Countess II FL Automated Cell Counter	Thermo Fisher Scientific, Waltham,

## Materials and Methods

	GER
Eppendorf centrifuge Type 5417R	Eppendorf, Hamburg, GER
Nanodrop 1000 Spectrophotometer	Peqlab, Erlangen, GER
Water bath	Dr. Hirtz & Co, Cologne, GER

### 2.1.3 List of kits

Listed below are the kits used during this study.

Name	Manufacturer
CellTiter 96® AQueous One Solution Cell Proliferation Assay	Promega, Madison, USA
GoTaq® qPCR Master Mix Kit	Promega, Mannheim GER
RNA purification kit	MACHEREY-NAGEL, Düren, GER
DNA purification kit	MACHEREY-NAGEL, Düren, GER
Maxwell® LEV simplyRNA Tissue Kit AS1280	Promega, Madison, USA

### 2.1.4 Plastic materials

Name	Standard	Manufacturer
Plastic-ware	6 cm, 10 cm, 15 cm	Sarstedt, Nümbrecht, GER
Multi-well plates	6-well, 12-well, 24-well, 96-well	TPP, Hörstel, GER or Nunc, Wiesbaden, GER
Falcon tubes	15 ml, 50 ml	Greiner Bio-One
Eppendorf tubes	0.5 ml, 1.5 ml, 2 ml, 5 ml	Biozym, Oldendorf, GER
Pipette tips	10 µl, 200 µl, 1 ml	Biozym, Oldendorf, GER
Cryo-vials	2 ml	Sigma-Aldrich, Taufkirchen, GER

## Materials and Methods

### 2.1.5 List of cell lines

Cell line	Origin	Characteristics	Medium	Source
A549	Human	Lung adenocarcinoma	DMEM	Roman Thomas <sup>1</sup>
PC9	Human	Lung adenocarcinoma	DMEM	Roman Thomas <sup>1</sup>
H82	Human	Small cell lung cancer	RPMI	Roman Thomas <sup>1</sup>
GCL8	Human	Small cell lung cancer	RPMI	Roman Thomas <sup>1</sup>
H187	Human	Small cell lung cancer	RPMI	Roman Thomas <sup>1</sup>
SHP77	Human	Small cell lung cancer	RPMI	Roman Thomas <sup>1</sup>
HCT15	Human	Colorectal cancer	DMEM	DSMZ <sup>2</sup>
HCT16	Human	Colorectal cancer	DMEM	DSMZ <sup>2</sup>
LS174T	Human	Colorectal cancer	DMEM	DSMZ <sup>2</sup>
EPC1	Human	Epithelial cell	DMEM	Alex Hillmer <sup>3</sup>
EPC2	Human	Epithelial cell	DMEM	Alex Hillmer <sup>3</sup>
OE21	Human	Esophageal cancer	DMEM	Sterner-Kock <sup>4</sup>
OE33	Human	Esophageal cancer	DMEM	Sterner-Kock <sup>4</sup>
ESO26	Human	Esophageal cancer	DMEM	Alex Hillmer <sup>3</sup>
BE2C	Human	Neuroblastoma	DMEM	Mathias Fischer <sup>5</sup>
KYSE 140	Human	Esophageal squamous cell carcinoma	DMEM	Sterner-Kock <sup>4</sup>
MD-MBA231	Human	Breast cancer	DMEM	DSMZ <sup>2</sup>
Huh7	Human	Hepatocellular carcinoma	DMEM	DSMZ <sup>2</sup>
HepG2	Human	Hepatocellular carcinoma	DMEM	DSMZ <sup>2</sup>
PSAE	Human	primary small airway epithelial	SABM+SA GM	ATCC PCS-301-010
HEK293T	Human	Embryonic kidney cells	DMEM	DSMZ <sup>2</sup>

1. Prof. Dr. Roman Thomas: Department of Translational Genomics, University of Cologne
2. DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany
3. Prof. Dr. Axel Hillmer: Institute of Pathology, University Hospital of Cologne

## Materials and Methods

4. Prof. Dr. Anja Sterner-Kock: Institute for Experimental Medicine, University Hospital of Cologne
5. Prof. Dr. Mathias Fischer: Department of Experimental Pediatric Oncology, University Hospital of Cologne

### 2.1.6 Antibodies

Primary antibody	Host species	Dilution	Manufacturer
LSD1	Rabbit	WB: 1:2000	Abcam, Cambridge, UK
H3K27me3	Rabbit	WB: 1:2000	Abcam, Cambridge, UK
GAPDH	Mouse	WB: 1:2000	Abcam, Cambridge, UK
IgG	Rabbit	ChIP: 3 µg for 10 µg chromatin	Cell Signalling Technology, Massachusetts, USA

### 2.1.7 Enzymes

Name	Manufacturer
RNase R	Biozym, Oldendorf, GER
RNase H	Biozym, Oldendorf, GER
SYBR Green	Promega, Mannheim, GER
GoTaq® G2 DNA Polymerase	Promega, Mannheim GER
GoTaq® QPCR Master Mix	Promega, Mannheim GER

### 2.1.8 Primer list

Primer	Sequence	Species
HPRT-F	GACCAGTCAACAGGGGACAT	Human
HPRT-R	GTGTCAATTATATCTTCCACAATCAAG	Human
Total LSD1-F	GAAACTGGAATAGCAGAGACTCC	Human
Total LSD1-R	TTCTTCCTCAGGTGGGGCTT	Human
LSD1+2a-F	GAAACTGGAATAGCAGAGACTCC	Human
LSD1+2a-R	CGTCTCCATACCCTCCAGAA	Human
LSD1+8a-F	GGCTGTGGTCAGCAAACAAG	Human

## Materials and Methods

LSD1+8a-R	GGAACCTTGACAGTGTCAGC	Human
circRNA 2a_2-F	CAAGCATCAGGTAGAGTACAG	Human
circRNA 2a_2-R	TTCTTCCTCAGGTGGGGCTT	Human
circRNA 3_2-F	GAAACCGCACAGTAGAGTACAG	Human
circRNA 3_2-R	TTCTTCCTCAGGTGGGGCTT	Human
circRNA 8a_2-F	AAGCCAACGGACAAGCTGAC	Human
circRNA 8a_2-R	TTCTTCCTCAGGTGGGGCTT	Human
circRNA 8_2-F	GCCCACTTTATGAAGCCAACG	Human
circRNA 8_2-R	TTCTTCCTCAGGTGGGGCTT	Human

### 2.1.9 Probe list

Name	Sequence
U6	TTTGCGTGTCATCCTTGCG
circRNA 2a_2	GCCGGTTCGTAGTCGTAGA
circRNA 3_2, circRNA 3_2_2a	CTTTGGCGTGTCATCTCAT

### 2.1.10 Small interfering RNA (siRNA) list

Name	Sequence
siRNA for circRNA 2a_2	GCCGGUUCGUAGUCCAUCU
siRNA for circRNA 3_2	CUUUGGCGUGUCAUCUCAU
siRNA for circRNA 2a_2 control	GCCGGUUCGUAGUCGUAGA
siRNA for LSD1+8a	CGGACAAGCUGACACUGUCAAGGUU
siRNA for circRNA 8a_2	UGUGACAGUCCAUCUCAUGU

### 2.1.11 Solutions

Solutions for cell fractionation

HMKE buffer

Component	Amount
HEPES pH 7.2	200 ul of 1 M Hepes
KCl	100 ul of 1M KCl

## Materials and Methods

MgCl <sub>2</sub>	50 ul of 1M MgCl <sub>2</sub>
EDTA	100 ul of 100 mM EDTA
Sucrose	2,5 ml of 1 M sucrose
PIC-ROCHE	1 tablet
PMSF	100 ul of 100 mM PMSF
Digitonin	2 mg digitonin
H <sub>2</sub> O	6950 ul

### Glycerol buffer

Component	Amount
Tris-HCL pH 7.5	100 ul of 1 M Tris-HCL
NaCl	375 ul of 1 M NaCl
EDTA	25 ul of 100 mM EDTA
DTT	42,5 ul of 100 mM DTT
PMSF	6,25 ul of 100 mM PMSF
glycerol	3,617 g of 85% Glycerol (2,94 ml)
Phosphatase inhibitor	1 tablet
Protease inhibitor	1 tablets
H <sub>2</sub> O	1511,25 ul

### Nucleic lysis buffer

Component	Amount
HEPES pH7.5	100 ul of 1 M HEPES
DTT	100 ul of 100 mM DTT
MgCl <sub>2</sub>	750 ul of 100 mM
EDTA	20 ul of 100 mM EDTA
NaCl	3000 ul of 1 M NaCl
UREA	5000 ul of 1 M UREA
Triton X-100	500 ul of 20 % Triton
H <sub>2</sub> O	530 ul

### Buffer C (with/without Triton-X-100)

## Materials and Methods

Component	Amount
Tris-HCL pH7.5	200 ul of 1 M Tris-HCL
MgCl <sub>2</sub>	15 ul of 1M MgCl <sub>2</sub>
NaCl	1500 ul of 1M NaCl
Triton X-100*	250 ul of 20 % Triton
Phosphatase inhibitor	1 tablet
Protease inhibitor	1 tablet
H <sub>2</sub> O	8035 ml

Solutions for fluorescence in situ hybridization (FISH)

Fix buffer

Component	Amount
Glacial acetic acid	1 ml
16% paraformaldehyde	5 ml
1.5 M NaCl	2 ml
H <sub>2</sub> O	12 ml

K-Phosph buffer, pH 7.0

Component	Amount
1M K <sub>2</sub> HPO <sub>4</sub>	61.5 ml
1M KH <sub>2</sub> PO <sub>4</sub>	38,5 ml

Hybrid buffer

Component	Amount
Formamide from -20 C	250 µl
20 x SSC	100 µl
50x Denhard	100 µl
DNA 10 mg/ml	10 µl
1M K-Phosph buffer, pH 7.0	50 µl
0.5 M EDTA	1 µl

Solutions for agarose gel electrophoresis

TAE buffer 10X



## Materials and Methods

Reagent	Amount
EDTA (pH 8.0)	1 mM
Sodium acetate (CH <sub>3</sub> COONa)	5 mM
Tris-acetate (pH 7.8)	40 mM

Solutions for immunoblotting

Running buffer 10X

Reagent	Amount
Tris	30.3 g (250 mM)
Glycine	187.7 (2.5 M)
10% SDS	1%
ddH <sub>2</sub> O	Add up to 1000 ml

Transferring buffer 12.5X

Reagent	Amount
Tris	75 g (312.5mM)
Glycine	356 g (2.4M)
Methanol	20%
ddH <sub>2</sub> O	Add up to 2000 ml

PBS 5X

Reagent	Amount
Na <sub>2</sub> HPO <sub>4</sub>	36 g (0.05M)
NaCl	200 g (0.68M)
KH <sub>2</sub> PO <sub>4</sub>	6 g (8.8 mM)
KCl	5 g (0.013M)
ddH <sub>2</sub> O	QS to 5 L

Adjust solution to desired pH (typically pH  $\approx$  7.4)

PBST 1X

PBST was made with 0.05% (v/v) solution of Tween-20 in PBS.

## Materials and Methods

TBS 10X

Reagent	Amount
Tris	12.114 g (0.2 M)
NaCl	43.83 g (1.5 M)
ddH <sub>2</sub> O	QS to 500 ml

Adjust solution to desired pH 7.2-7.4

TBST 1X

TBST was made by 0.05% (v/v) solution of Tween-20 in PBS.

SDS Stacking gel (2X)

Reagent	Amount (%V/mass)
0.5M TRIS pH 6.8	1.26 ml
10% APS	50 µl
10% SDS	50 µl
Acrylamide (30%)	830 µl
TEMED	5 µl
ddH <sub>2</sub> O	2.74 ml

Reagents for cell culture

Reagent	Manufacturer
DMEM (Dulbecco's Modified Eagle Medium)	GibcoBRL, Karlsruhe, GER
DMSO	Sigma-Aldrich, Taufkirchen, GER
Doxycycline	Sigma-Aldrich, Taufkirchen, GER
Fetal Bovine Serum (FBS)	Pan Biotech, Aidenbach, GER
GSK2879552	Xcessbio, San diego, USA
HCI-2509	Xcessbio, San diego, USA
HCI-2577 (SP-2577)	Salarius Pharmaceuticals, Houston, USA
GSK 690	Sigma-Aldrich, Taufkirchen, GER
OptiMEM	GibcoBRL, Karlsruhe, GER
Phosphate buffered saline (PBS)	GibcoBRL, Karlsruhe, GER
Poly-L/D-Lysine	Sigma-Aldrich, Taufkirchen, GER

## Materials and Methods

Puromycin	Sigma-Aldrich, Taufkirchen, GER
-----------	---------------------------------

### Reagents for molecular experiment

Reagent	Manufacturer
Acetic acid	Sigma-Aldrich, Taufkirchen, GER
Acrylamide/bis-acrylamide, 30%	Sigma-Aldrich, Taufkirchen, GER
Agarose	Biozym, Oldendorf, GER
Ammonium peroxodisulfate (APS)	Carl Roth, Karlsruhe, GER
$\beta$ -Mercaptoethanol	Carl Roth, Karlsruhe, GER
Chloroform (99%)	Sigma-Aldrich, Taufkirchen, GER
Dipotassium hydrogenphosphate (K <sub>2</sub> HPO <sub>4</sub> )	Carl Roth, Karlsruhe, GER
Disodium hydrogenphosphate (Na <sub>2</sub> HPO <sub>4</sub> )	Carl Roth, Karlsruhe, GER
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth, Karlsruhe, GER
Ethanol (99%)	Carl Roth, Karlsruhe, GER
Formaldehyde (4%)	Carl Roth, Karlsruhe, GER
Glycine	Sigma-Aldrich, Taufkirchen, GER
Glycogen Blue	Thermo Scientific, Waltham, USA
Isopropanol (99%)	Carl Roth, Karlsruhe, GER
Laemmli buffer 2x & 4x	Biorad, Hercules, USA
Magnesium chloride (MgCl <sub>2</sub> )	Carl Roth, Karlsruhe, GER
Methanol (99%)	Carl Roth, Karlsruhe, GER
Milk powder	Carl Roth, Karlsruhe, GER
NEB cell lysis buffer 10x	New England Biolabs, Frankfurt, GER
Phenylmethanesulfonyl fluoride (PMSF)	Sigma-Aldrich, Taufkirchen, GER
Potassium chloride (KCl)	Carl Roth, Karlsruhe, GER
Potassium dihydrophosphate (KH <sub>2</sub> PO <sub>4</sub> )	Carl Roth, Karlsruhe, GER
Prestained protein standard (11–245 kDa)	New England Biolabs, Frankfurt, GER
Protease inhibitor tablets	Roche, Grenzach-Wyhlen, GER
RIPA buffer	Thermo Scientific, Waltham, USA

## Materials and Methods

Silencer Select SiRNA against LSD1 (s619)	Thermo Scientific, Waltham, USA
Sodium dodecyl sulphate (SDS)	Carl Roth, Karlsruhe, GER
Sodium chloride (NaCl)	Carl Roth, Karlsruhe, GER
Sodium acetate (NaCH <sub>3</sub> COOH)	Carl Roth, Karlsruhe, GER
Sodium hydroxide (NaOH)	Carl Roth, Karlsruhe, GER
Tetramethylethylenediamine (TEMED)	Thermo Scientific, Waltham, USA
TRIS-HCl	Carl Roth, Karlsruhe, GER
Trizol	Thermo Scientific, Waltham, USA
Tween-20	Sigma-Aldrich, Taufkirchen, GER
Random primer	Thermo Scientific, Waltham, USA
dNTPs	Thermo Scientific, Waltham, USA

### 2.2 Methods

#### 2.2.1 Cell culture

Medium used for cell culture included DMEM or RPMI 1640 with heat-inactivated 10 % (v/v) fetal calf serum (FCS) at 37°C (5% CO<sub>2</sub>). When 80% cell confluence was attained, cells were routinely passaged.

#### 2.2.2 Passaging cells

Following the withdrawal of the cell culture medium, the cell monolayers were rinsed in 1x PBS and an appropriate amount of trypsin was added to the plates or wells. Cells were kept at 37°C until they were observed to be floating in the plate. A fresh growth medium was used to stop the trypsinization and cells were gently separated by pipetting. Cells in suspension were transferred to a falcon and centrifuged at 1,000 rpm for 5 minutes. This was followed by immediate resuspension of the cell pellet in the growth medium before being divided between three plates to continue the culture.

#### 2.2.3 Determination of cell numbers

After trypsinization, cells were separated in the suspension. Ten µl of cell suspension was removed from the well of Countess chamber slides. The cell counter-related slide port was used to insert the slides. The cell counting process was carried out following the manufacturer's guidelines.

#### 2.2.4 Cryopreservation and thawing of cells

Cell cryopreservation was performed when the cells reached 90% confluence. A plate was trypsinized as described above. Cells were centrifuged and resuspended in 10% DMSO (0.9ml DMEM+0.1ml DMSO), then transferred into a cryo-tube. Cryo-tubes were stored overnight at -80°C before being placed in liquid N<sub>2</sub> for long-term storage. To cultivate cryopreserved cells, a water bath maintained at 37°C was used to thaw the cells for one minute (which was long enough to allow the ice to melt). The cell suspension was mixed with 3ml fresh medium in a falcon and then added to culture plates in 37°C incubators with 5% CO<sub>2</sub>.

### 2.2.5 Mycoplasma testing

When the confluence was 80%, 100 µl culture medium was collected and heat-inactivated for 10 min at 95°C. Real-time PCR (qPCR) performed with the aid of a Venor®GeM mycoplasma PCR Detection Kit was used to identify mycoplasma contaminations as per the recommendations of the manufacturer.

### 2.2.6 Data Acquisition and analysis from TCGA

The Cancer Genome Atlas (TCGA) database was searched for RNA expression in cancer and non-cancer tissues and the corresponding clinical characteristics of patients. The expression of all LSD1 transcripts was combined and analyzed as the levels of total LSD1, while the levels of the transcript ENST00000400181.8 was considered as LSD1+2a levels.

### 2.2.7 Extraction of total RNA preparation from clinical samples

Tissues were shredded in the Precellys 24 (PepLab) (1: 6500 – 2x 20 – 005 and five ceramic beads per sample). For isolation, RNA Tissue Kit for Maxwell 16® was used and the manufacturer's protocol was followed. The RNA concentrations were measured by spectrophotometry using a NanoDrop™ 1000. Every RNA sample was frozen at -80 °C until further use.

### 2.2.8 Compounds treatment

LSD1 inhibitors HCI-2509, SP-2577, GSK 2879552, and GSK690 were first suspended in DMSO before being stored in the refrigerator. Both the control and the experimental groups had their cells seeded at an equivalent density one day prior to the treatment, to reach a confluency level of 40 % in the following day. Following a single wash using 1xPBS, the cells were placed in a new medium containing the drugs, correspondingly and cultured for a total of 72 hours, followed by result analysis.

### 2.2.9 SiRNA transfection

The cells were passaged a day before transfection and plated on required plates to achieve a 60% confluency and the siRNA transfections were carried out with the use of Lipofectamine 2000 with the instructions provided by the manufacturer. The

## Materials and Methods

concentration of siRNA to be used was determined by performing pre-experiments using 4 different siRNA concentrations to achieve downregulation of the specific gene. Both the scrambled and gene-specific siRNAs had their respective transfection processes carried out concurrently. At 6 hours following transfection, the cells were subjected to wash once using 1x PBS and then placed in a new cell medium. Afterward, the cells were grown for 48-72 hours depending on the requirement of the experiment. Cells were harvested and RNA and protein isolation were performed depending on the experiments.

### 2.2.10 Quantitative analysis of mRNA expression levels

#### 2.2.10.1 RNA extraction

An Ambion PureLink® RNA Mini Kit (Life Technologies) was used for cellular RNA extraction as instructed by the manufacturer. Following cell lysis, DNA was digested in a series of steps using DNase. The lysate was then filtered through a filter column consisting of a silica membrane, where RNA gets bound and the residual filtrate is filtered out. Next, nuclease-free water was used to elute the RNA before being stored at -80 °C for subsequent use. For the cell line array, the RNA was isolated with the aid of the Maxwell® LEV simplyRNA Tissue Kit AS1280 kit from Promega in line with the recommendations of the manufacturer.

#### 2.2.10.2 RNA reverse transcription

By applying the TaqMan Reverse Transcription Kit from Applied Biosystems, mRNA was converted into cDNA following the guidelines stipulated by the manufacturer for real-time PCR. In all the aforementioned investigations, cDNA was prepared using 500ng of RNA in a maximum of 15µl of reaction mix for PCR analysis. On a thermal block, the reaction was carried out in two steps, primer annealing at 25°C for 10 minutes and cDNA synthesis at 37°C for 2 hours. In case of low availability of sample and less RNA concentration, the amount of RNA and reagents used were scaled down proportionally.

### 2.2.10.3 Real-time PCR

QPCR was carried out with the use of GoTaq qPCR Master Mix (Promega), respective primer sets (forward and reverse), and 2-10ng DNA as per availability of sample. A single reaction mix consisted of 2-10ng DNA, 50% by volume GoTaq qPCR Master Mix, and 10mM of each forward and reverse primer and nuclease-free water to make up the volume (10 $\mu$ l/15  $\mu$ l). New primers were always tested for their efficiency of amplification by plotting a standard curve by making a dilution series using three different DNA concentrations and observing the Ct (cycle threshold) values to change proportionally as per the used concentrations. Below are the parameters that were used for a typical PCR run:

1. 95°C 2min
2. 95°C 30sec
3. 60°C 30sec
4. 72°C 30sec
5. Steps 2 to 4 repeated for 49 cycles
6. 72°C 5min
7. Melting curve 65°C to 95°C

We used either a BioRad CFX96 Real-time PCR Cycler, an Mx3000P Cycler, or a Roche Lightcycler 480 to conduct the real-time PCR process. The primer list in the materials section includes annealing temperatures for each primer if they vary from those described above. Unless stated, the mRNA expression level for every gene quantified by real-time PCR was normalized to those of the hypoxanthine phosphoribosyltransferase 1 (HPRT) gene.

### 2.2.11 Fluorescence in situ hybridization (FISH)

Cell medium was removed, and cells were washed once with PBS. Cells were fixed and incubated for 15 min at room temperature. Ice-cold 70% ethanol was added, and the coverslips were stored overnight at -20 °C. On the second day, fixed cells were removed from -20 °C and recovered in PBS for at least 5 min at room temperature. Depending on the expected localization of the circRNA, a method is decided to be performed for the permeabilization. For preservation of the cytoplasm, a permeabilization step with 0.5% Triton X-100/0.5% saponin for 5–10 min is recommended. The coverslips were washed once with PBS before post-fixation. For



## Materials and Methods

post-fixation, 3.7% formaldehyde in PBS is added to the cells for 5 min at room temperature, followed by one 10-min wash in PBS. Next, cells are dehydrated prior to hybridization by the addition of 70, 90, and 100% ethanol (3 min every time), and then coverslips are air-dried. The probe was diluted with the hybridization mix. The probes were denatured at 90 °C for 10 min and placed immediately on ice. Fifteen  $\mu$ l of the hybridization mix was placed in the middle of a clean glass slide and a dried coverslip was inverted onto the drop using forceps. The coverslip is sealed using rubber cement, placed in a humid chamber sealed with parafilm, and incubated at 37 °C overnight. The rubber cement seal is carefully removed, and coverslips are placed back into a tissue culture dish to be washed three times for 10 min at 37 °C with 2 $\times$  SSC. On the third day, the slides were washed once with RNase-free water before DAPI staining for 5 min at room temperature. Finally, coverslips were dried by touching the edge on a clean tissue and mounted onto clean glass slides. The slides were observed by the confocal microscope.

### 2.2.12 Cytoplasmic lysis

Cells were washed twice with cold PBS and then trypsinized. Cells were resuspended with HMKE buffer and left on ice for 10min. Cells were centrifuged at 500 g for 10 min at 4°C to separate cytosol from membranes, nuclei, organelles and cytoskeleton. Pellets were washed with HMKE buffer once. Nuclei were resuspended in equal volumes of glycerol buffer and nucleic lysis buffer, vortexed 2 seconds twice and then incubated on ice for 2 min. Cells were spined at 700 g for 2 min at 4°C. The supernatant was nucleoplasmic fraction. Then chromatin fraction was resuspended and sheared 20 min sonication in biorupter (30 sec max 30",30"). Chromatin fraction was centrifuged 5 min at 16.000 g at 4°C.

### 2.2.13 Western blot

For protein analysis, western blots were performed using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). First, a separating gel and then a running gel was loaded in a cast along with a plastic comb for the creation of wells for sample loading. Protein samples (10-15 $\mu$ g) were used depending on sample availability and the requirement of the experiment. Before sample loading, protein denaturation was performed using 2X or 4X laemmli buffer supplemented with  $\beta$ mercaptoethanol and heated for ten minutes at 95 °C, after which the sample was loaded onto the gels in

## Materials and Methods

a direct manner. To estimate the sizes of the proteins, a prestained protein ladder (New England Biolabs in Frankfurt, Germany) was continuously loaded alongside the protein samples. Running the gels via the Bio-Rad Mini protein gel equipment allowed the protein samples to be separated into their constituent parts. After performing electrophoresis at 90V for the first 15 minutes to better separate the proteins, the voltage value was elevated to 120V for the remaining 60 minutes of the experiment. The Bio-Rad blotting chamber was then used to transfer the proteins that had been separated from the gel onto a 0.2-micron PVDF membrane. The transfer was carried out at 90V for 60 minutes. To observe the successful transfer of the proteins to the membrane, the proteins were first visualized by incubating the membrane with the ponceau stain. A blocking solution (5% milk in PBST) was added to incubate the membranes for an hour at RT on a rocking platform. After incubation with certain primary antibodies, various-sized proteins could be visualized after being sliced from the blots using the protein ladder as a reference. The blocking solution was used to dilute the primary antibodies, and the final concentrations of each antibody are listed in the antibody list in the Materials section. The blots were allowed to incubate with the primary antibody mixture throughout the night at 4 °C while being gently shaken. After three 10-minute washes using 1X PBST, the membranes were treated for 1 hour with HRP-labeled secondary antibodies diluted in blocking solution at RT. After another rounds of washing, the membranes were exposed to treatment for one minute with Pierce™ ECL western blotting substrate for 1 minute before being visualized in the ChemiDoc™ Imaging System. Image lab and photoshop were used to process the captured images.

### 2.2.14 Immunoprecipitation (IP) of R-loops

Lysis was performed on the cells in lysis buffer followed by sonication for 6 cycles of 20 seconds with a 10-second gap between each cycle. After that, the lysate from the cells was centrifugated for 15 minutes at 4 °C and at a rate of 12000 rpm. The antibody-linked beads were then treated with cell lysate for four hours at 4 °C. After three times of washing continuously with lysis buffer and ice-cold PBS, 100µl 2 x Laemilli buffer (Bio-Rad, GER) was added, and beads were heated at 95°C for 5min to elute proteins. RNA and DNA were isolated for subsequent experiments.

## Materials and Methods

### 2.2.15 MTT assay

To perform the MTT test, the cells were plated in a 96-well plate 24 hours before the treatment at a density of 30,000 cells per well. Inhibition or siRNA transfection was performed on cells for 72 hours and the CellTiter 96® AQueous One Solution Cell Proliferation Kit was employed to conduct the MTT test in compliance with the directions provided by the manufacturer. There were three biological replicates of each experiment. Mean absorbance was calculated for each. The background absorbance (medium only) was subtracted from both the control and inhibitor-treated cells. The percent of viable or proliferating cells was calculated as follows: For control:  $\text{mean OD control} / \text{mean OD control} * 100 = 100\%$  and inhibitor-treated:  $\text{mean OD treatment} / \text{mean OD control} * 100 = \_ \%$ .

### 2.2.16 RNA purification

Buffer RA1 and  $\beta$ -mercaptoethanol ( $\beta$ -ME) were added to the cell pellet and vortexed vigorously. A NucleoSpin® Filter was placed in a collection tube, and the mixture was added into the tube and centrifuge for 1 min at 11,000 x g. Seventy percent ethanol was added to the homogenized lysate and mixed by pipetting up and down (5 times). Pipette lysate up and down 2–3 times and load the lysate to the column and centrifuged for 30 s at 11,000 x g. (Membrane Desalting Buffer was added and centrifuged at 11,000 x g for 1 min to dry the membrane. DNase reaction mixture was directly applied onto the center of the silica membrane of the column and incubated at room temperature for 15 min. Buffer RAW2 was added to the NucleoSpin® RNA Column (centrifuged for 30 s at 11,000 x g) for the first wash and Buffer RA3 was applied (30 s at 11,000 x g) for the second wash and Buffer RA3 was added (2 min at 11,000 x g) to dry the membrane completely. RNA was eluted in RNase-free H<sub>2</sub>O, and centrifuged at 11,000 x g for 1 min.

### 2.2.17 Statistical analysis

GraphPad Prism was employed for the quantitative PCR statistical analysis. The significance level was determined using the Holm-Sidak t-test. A  $p$ -value  $< 0.05$  denoted a significant level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

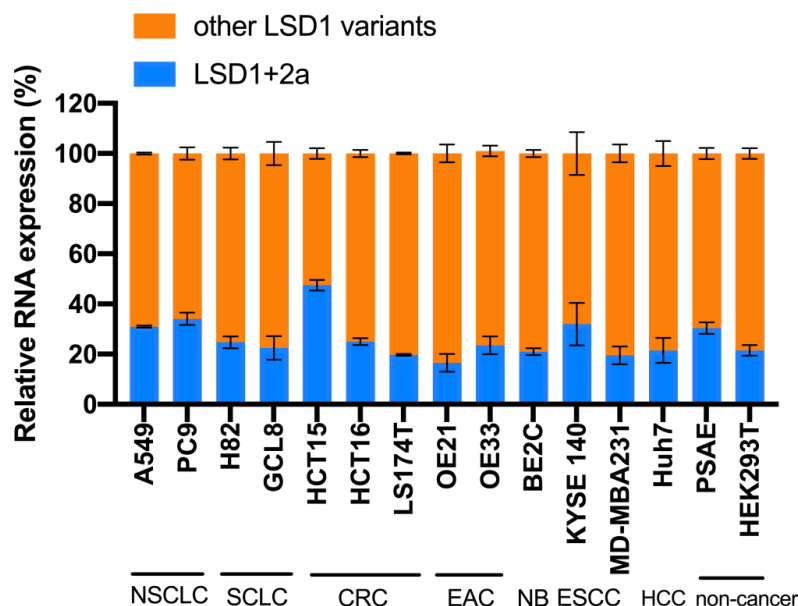
## 4. Results

### 3.1 LSD1 splice variant 2a expression pattern in cancer and non-cancer

There are four types of LSD1 isoforms (LSD1, LSD1+2a, LSD1+8a and LSD1+2a+8a) in humans (Figure 2, introduction). However, the ratios and functions of distinct variants of LSD1 remain unclear in cancers.

#### 3.1.1 LSD1 +2a expression in cancer and non-cancer cell lines

First, I explored the levels of LSD1+2a variant in several cell lines including cancer and non-cancer cells. The proportion of LSD1+2a in total LSD1 ranged from 20% - 50% in human cell lines (Figure 4). The highest proportion of LSD1+2a in total LSD1 was in HCT15, a colorectal cancer (CRC) cell line, while the lowest proportion was in OE21, which is a cell line derived from esophageal adenocarcinoma (EAC). In NSCLC cells (A549 and PC9), the proportion was about 30%. In non-cancer cells including primary small airway epithelial (PSAE) cells and human embryonic kidney (HEK293T) cells, the proportion was about 20% - 30%.



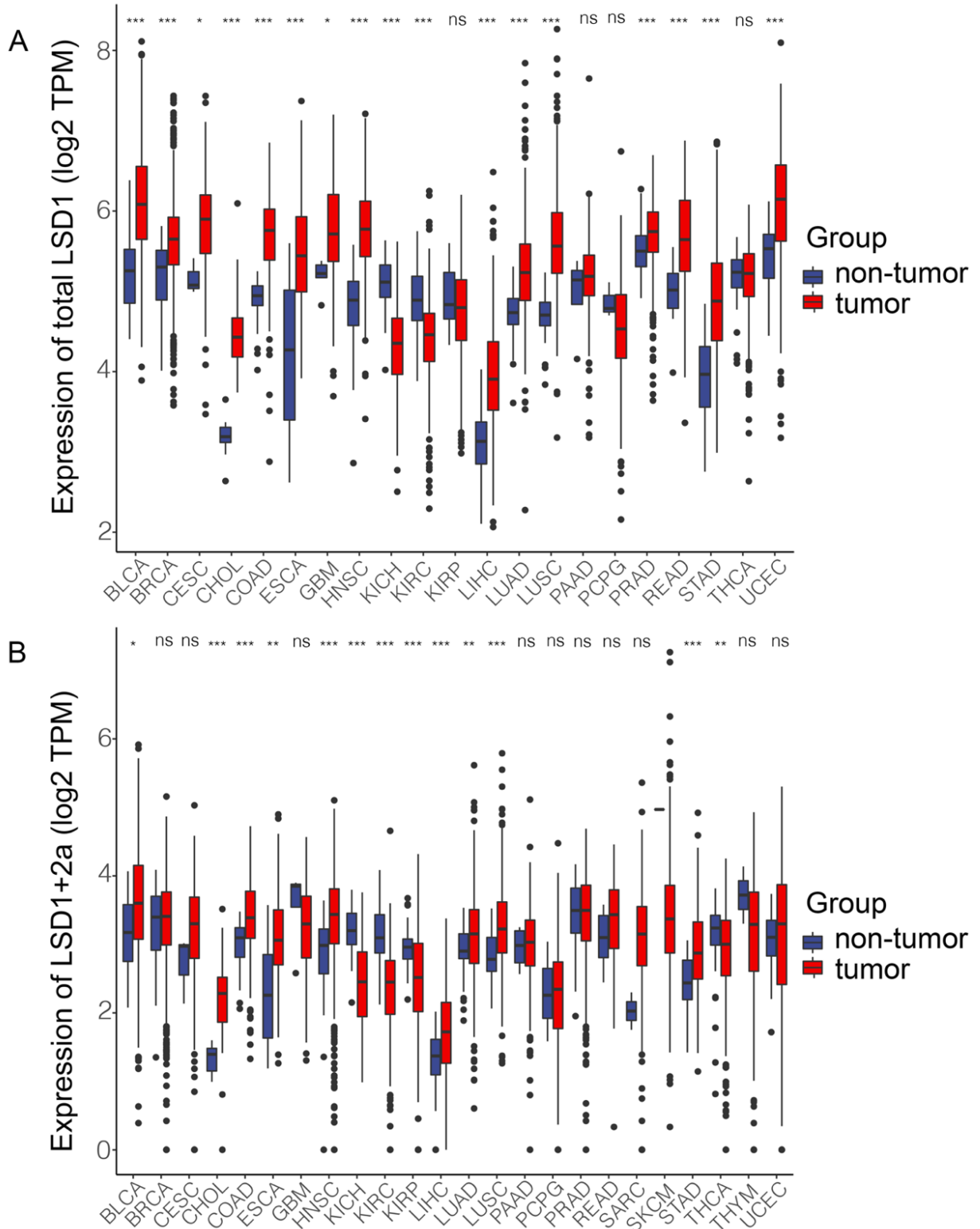
**Figure 4: Proportion of LSD1+2a in cancer and non-cancer cell lines.** NSCLC, non-small cell lung cancer. SCLC, small cell lung cancer. CRC, colorectal cancer. EAC, esophageal adenocarcinoma. NB, neuroblastoma. ESCC, esophageal squamous cell carcinoma. HCC, hepatocellular carcinoma. PSAE, primary small airway epithelial. Total LSD1, quantified by qPCR using primers recognizing all transcript isoforms, was set to 100%.

## Results

### 3.1.2 LSD1 +2a expression in different cancer types

To explore the levels of LSD1+2a in the non-tumor and LUAD tissues, I investigated the expression of total LSD1 and LSD1+2a variant from TCGA database. The expression of total LSD1 including four variants significantly differed in cancers (Figure 5A). Among 21 types of cancer, the highest expression of total LSD1 was in uterine corpus endometrial carcinoma (UCEC) and the lowest expression was in liver hepatocellular carcinoma (LIHC). In lung adenocarcinoma (LUAD), the mRNA levels of total LSD1 were much higher than those in the non-tumor tissues. To explore the roles of LSD1+2a in cancers, I analyzed the expression of LSD1+2a in TCGA cohort (Figure 5B). The highest levels of LSD1+2a was in bladder cancer (BLCA) while the lowest levels of LSD1+2a was also in LIHC. There was a significant expression difference between LUAD tumor and non-tumor samples. Collectively, I found that the levels of total LSD1 and LSD1+2a in LUAD samples were much higher than those in the non-tumor samples.

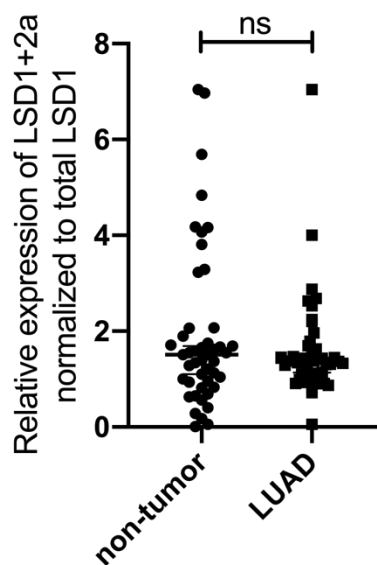
## Results



**Figure 5: Expression of total LSD1 and LSD1+ 2a in cancers from TCGA database. (A)** Expression of total LSD1 in tumor and non-tumor tissues from TCGA. **(B)** Expression of LSD1+2a in tumor and non-tumor tissues from TCGA. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ , ns, not significant. TPM, transcript per million.

## Results

I extracted from a total of 46 clinical LUAD and the matching adjacent non-tumor tissues and explored the levels of LSD1 variants. Importantly all tumors were controlled by a senior pathologist (Yuri Tolkach, Institute of Pathology, University Hospital of Cologne) and only tumors with more than 70% of tumor cells in the tumor areal were considered. RNA extracted from the macrodissected tumor areal was analyzed for expression of different LSD1 isoforms. The levels of LSD1+2a normalized to total LSD1 did not differ significantly between non-tumor and LUAD tissues (n = 46) (Figure 6).

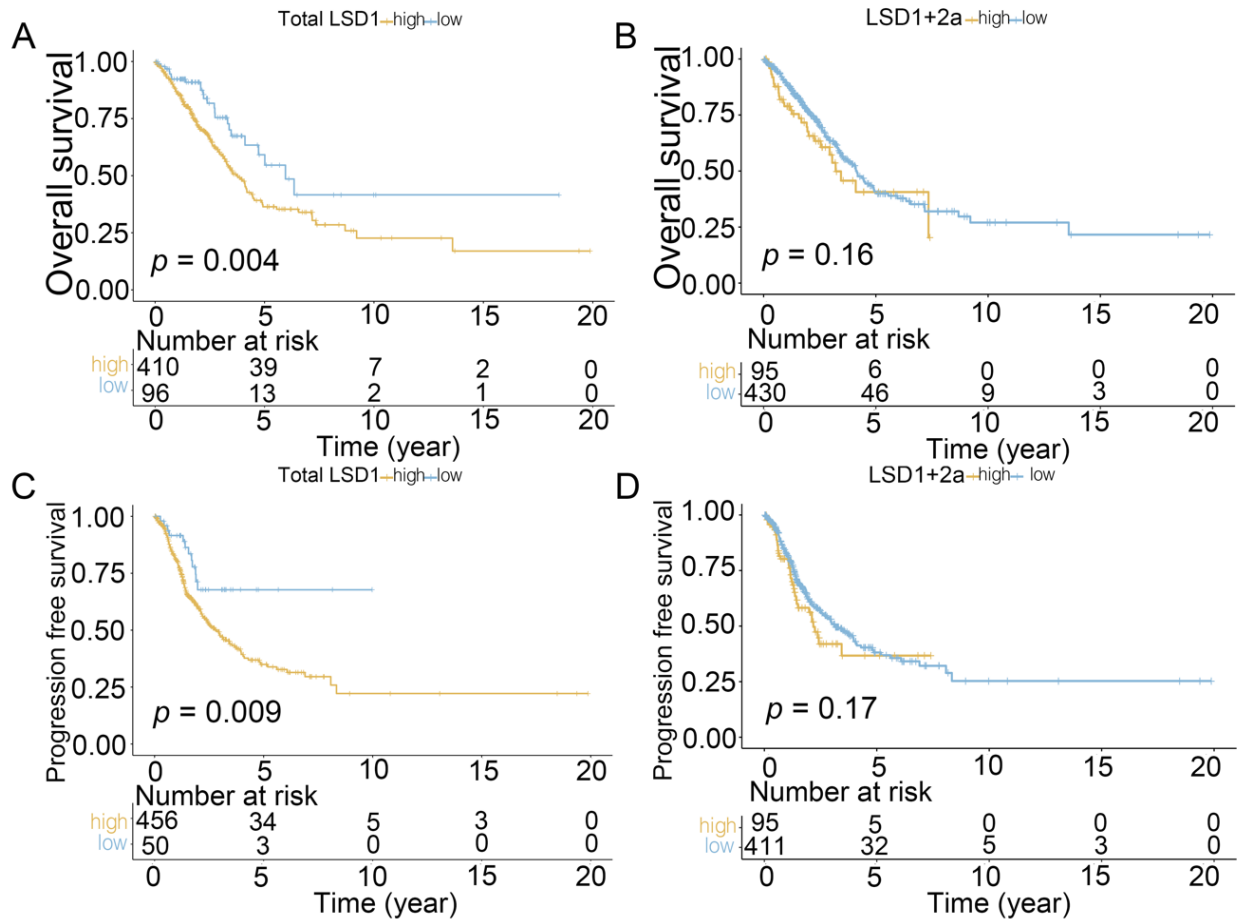


**Figure 6: Expression of LSD1+2a in paired adjacent non-tumor (n = 46, left) and tumor tissues (right) from LUAD patients.** RNA was extracted from 46 pairs of LUAD and non-tumor tissues and qPCR was performed to measure the levels of LSD1+2a which is normalized to total LSD1. ns, not significant.

I next investigated the prognostic values of total LSD1 and LSD1+2a in LUAD. Overall survival is widely used as the most significant endpoint for cancer patients. I found that patients with low total LSD1 had prolonged overall survival compared to patients with high total LSD1 expression ( $p = 0.004$ ; Figure 7A). However, I found no difference in overall survival between patients with high and low LSD1+2a levels ( $p = 0.16$ ; Figure 7B). In addition, progression-free interval (PFI), suggested to replace the phrase progression-free survival with a less ambiguous term [157], was also analyzed in our study. There was a better PFI for the patients with low total LSD1 than those in the high total LSD1 group ( $p = 0.009$ ; Figure 7C), while no difference

## Results

was found in PFI between the patients with high and low LSD1+2a levels ( $p = 0.17$ ; Figure 7D). All above data indicated that high expression of total LSD1 was highly correlated with the prognosis of LUAD and no correlation was found between the expression of LSD1+2a and the prognostic values for LUAD patients from TCGA-LUAD cohort.



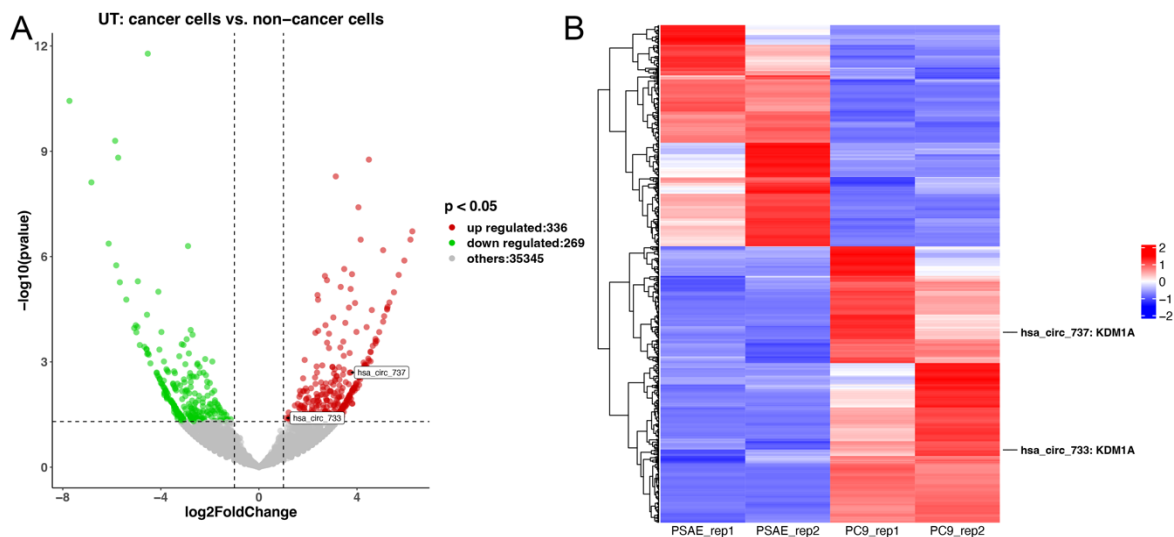
**Figure 7: Prognostic values of total LSD1 and LSD1+2a in TCGA-LUAD cohort.** (A) Kaplan-Meier curves of overall survival of patients between high and low (A) total LSD1 and (B) LSD1+2a levels in TCGA-LUAD cohort; Progression-free survival of patients between high and low (C) total LSD1 and (D) LSD1+2a levels in TCGA-LUAD cohort. The yellow curve stands for the survival of patients with high expression; the blue curve is for the survival of patients with high expression. The number under the curves are the patients at the risk.



## Results

### 3.2 CircRNAs from LSD1 gene

Splicing is often correlated with back-splicing and circRNA synthesis, therefore I analyzed how circRNAs derived from LSD1 are expressed in cancer and non-cancer cells. First, I screened circRNA sequencing data from lung adenocarcinoma cells and lung epithelial cells, kindly provided from the Odenthal team (Institute of Pathology, University Hospital of Cologne). Mr. Xinlei Zhao (Institute of Pathology, University Hospital of Cologne) kindly helped me to analyze the data. In Figure 8A the differently expressed circRNAs from PC9 lung adenocarcinoma and PSAE lung epithelial cells are shown. The top 30 upregulated and downregulated circRNAs were selected to show in Figure 8B and supplementary Table 1. Interestingly, among 605 differently expressed circRNAs, some derived from LSD1 locus, were found to be higher expressed in PC9 cells such as hsa-circ-737 and hsa-circ-733 (Figure 8B).



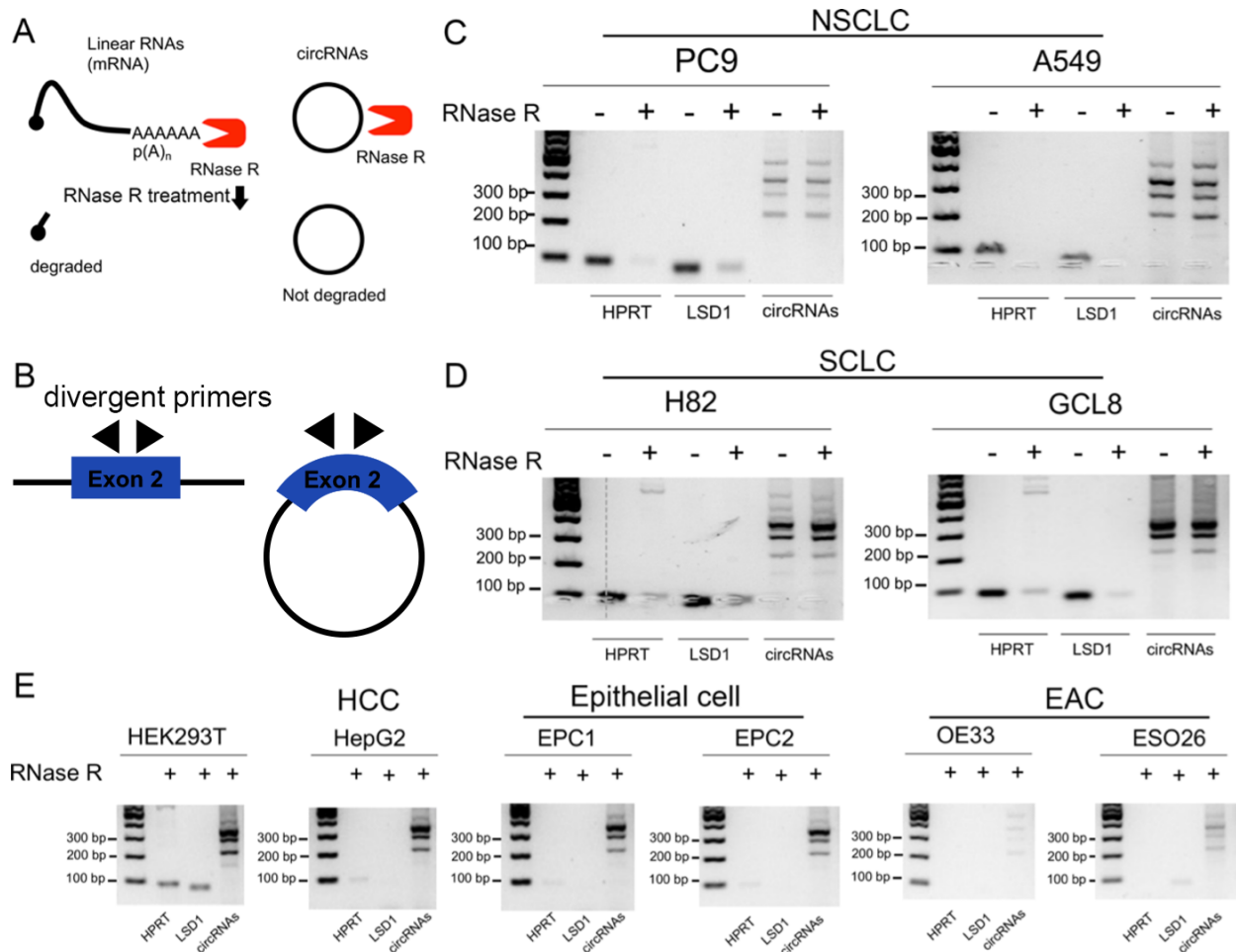
**Figure 8: Identification of differently expressed circRNAs between PC9 and PSAE cells by circRNA sequence.** (A) A volcano visualizing differently expressed circRNAs between PC9 and PSAE cells. (B) A heatmap showing the 50 upregulated and downregulated circRNAs. Each row denoted one circRNA and each column represented one sample. The color changed from red to blue indicated the dysregulation from up to down.

#### 3.2.1 Identification of circRNAs generated from LSD1 gene

I sought to uncover circRNAs arising from LSD1 gene and explore their involvement in LUAD development. To validate this RNA species harbors a cyclic structure, RNase R was used to specifically degrade linear RNAs but not circRNAs (Figure 9A). I primarily focused on circRNA deriving from the exon 2/exon2a locus. Therefore, I

## Results

used divergent primers from exon 2 to amplify circRNAs from this locus (Figure 9B). The housekeeping gene HPRT and the parental LSD1 gene were used as controls, to show the activity of RNase R and successful degradation of linear RNAs. Indeed, the gel electrophoresis of the respective amplicons revealed that HPRT and LSD1 were totally degraded with RNase R in NSCLC PC9 and A549 cells. However, putative circRNAs derived from LSD1 gene were resistant to RNase R treatment, validating these RNAs as circRNA forms (Figure 9C).



**Figure 9: Characterization of circRNAs from LSD1 in cancers.** (A) Flow chart of RNase R treatment. The red icon indicates the RNase R. Whereas linear RNA such as polyadenylated mRNA is degraded by RNase R, the circRNA is stable during RNase R treatment. (B) Verification of circRNAs using divergent primers. Northern blots showed validation and resistance of transcripts in (C) PC9 and A549, (D) H82 and GCL8 and (E) other non-cancer (HEK293T, epithelial cell) and cancer cells (HCC and EAC). HPRT and LSD1 linear mRNA was amplified with convergent primers as the negative controls for RNase R treatment. CircRNAs from LSD1 locus were amplified with divergent primers. -, no RNase R treatment. +, RNase R treatment.

## Results

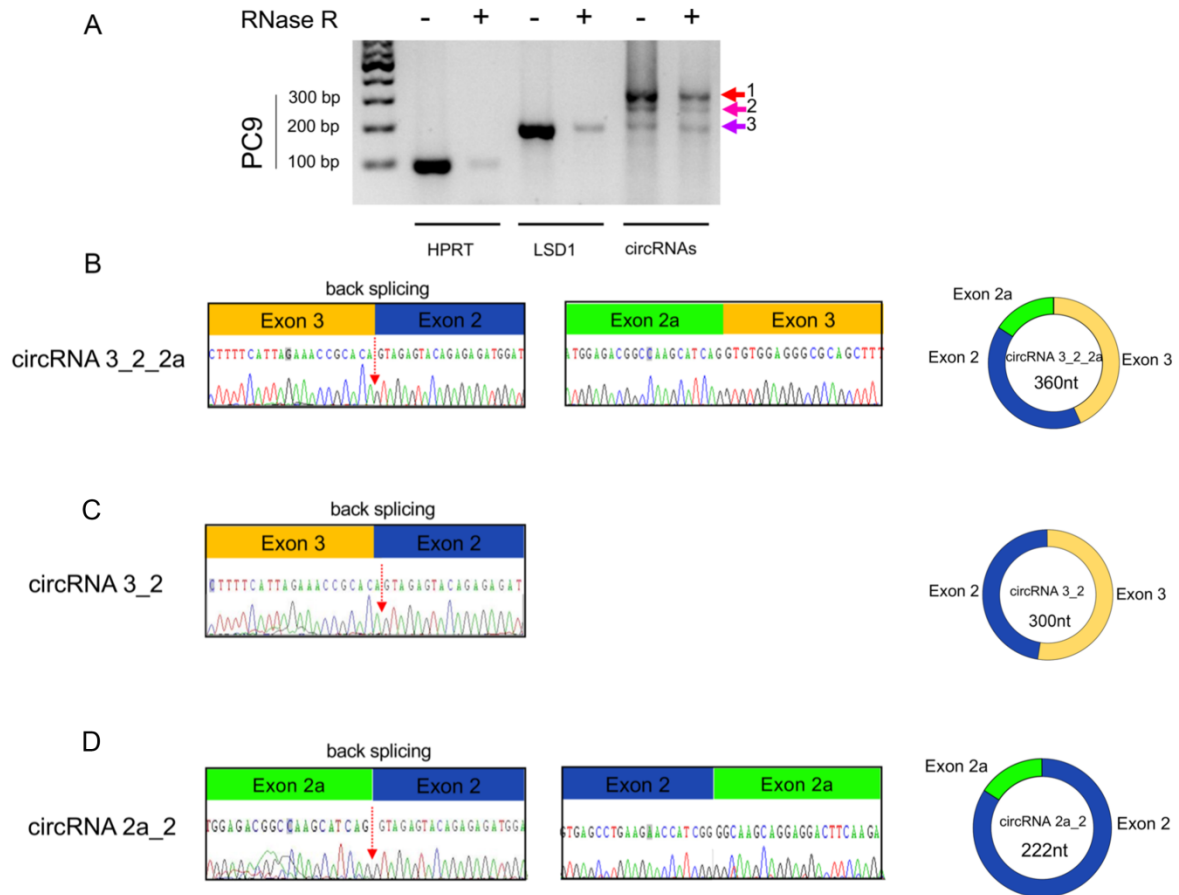
Interestingly, I found four different circRNA deriving from LSD1 2a locus. The similar results were found in SCLC H82 and GCL8 cells (Figure 9D). Moreover, the four circRNAs were also amplified in HEK293T, HCC, epithelial cells (EPC1 and EPC2) and esophageal adenocarcinoma (OE33 and ESO26) and resistant to the digestion with RNase R exonuclease (Figure 9E). Thus, four circRNAs from LSD1 locus were found in different cell lines including non-cancer and cancer cells.

To identify the circRNAs deriving from the LSD1 exon 2 locus, the amplified products of circRNA were sequenced by sanger sequencing (Figure 10A). Using the circBase database annotation (<http://www.circbase.org/>) I identified them as circRNA 3\_2\_2a, circRNA 3\_2 and circRNA 2a\_2. CircRNA 3\_2\_2a is formed by the circularization of exon 3, 2a and 2 of the LSD1 gene, which has a length of 360 nt according to circBase (Figure 10B). CircRNA 3\_2 is derived from exon 3 and 2 with a length of 300nt (Figure 10C), while circRNA 2a\_2 with 222nt is generated from the back-splicing of exon 2a and 2 of the LSD1 gene (Figure 10D). The characteristics of three circRNAs are shown in Table 1.

**Table 1: Characteristics of circRNAs from LSD1.**

	Chromatin	Isoform name	No. of circRNA	Exon count	Exon No. and size	Length (nt)
circRNA 3_2_2a	23030468-23050520	ENST00000400181	hsa_circ_733 has_circ_0009061 (circBase)	3	Exon 3 (134), Exon 2 (60), Exon 2a (134)	360
circRNA 3_2	23030468-23050520	ENST00000465864	hsa_circ_733	2	Exon 3 (134), Exon 2 (60)	300
circRNA 2a_2	23030468-23044486	ENST00000400181	hsa_circ_33057	2	Exon 2a (134), Exon 2 (60)	226

## Results



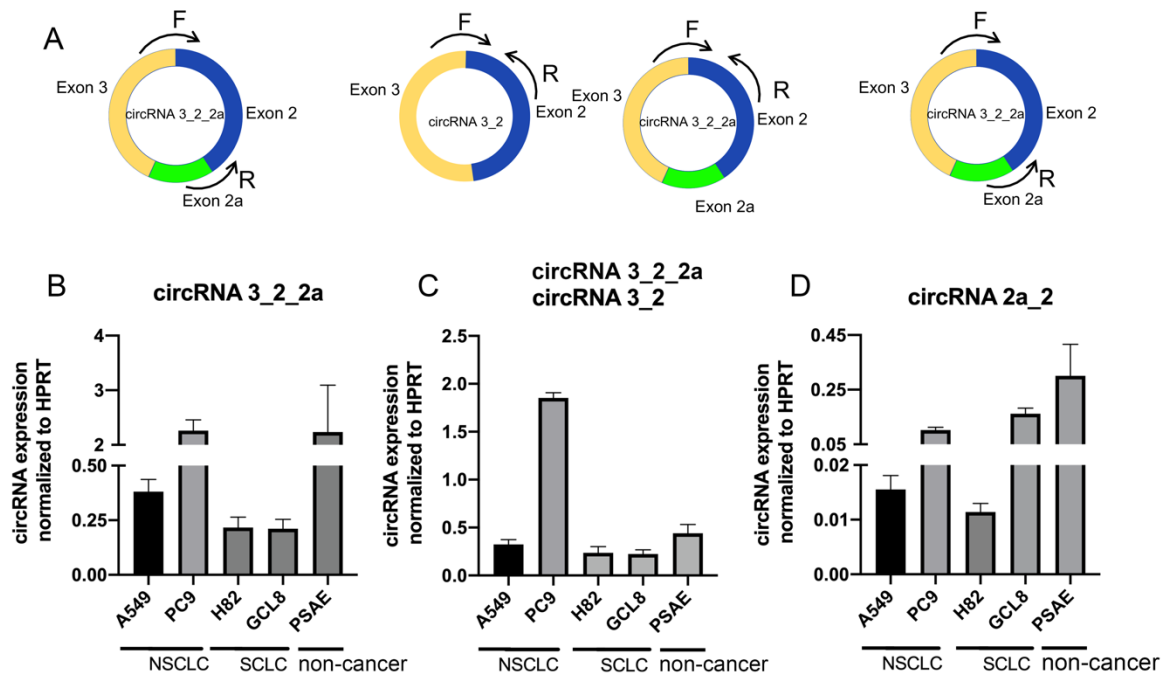
**Figure 10: Identification of circRNA from LSD1 gene.** (A) CircRNAs resistant to RNase R treatment had apparent molecular weights of circRNA from the LSD1 exon 2-2a-3 locus as shown in Table 1. The bands, which highlighted by three arrows, were taken and sequenced by means of junction flanking primers. Sequencing of the junctions of (B) circRNA 3\_2\_2a, (C) circRNA 3\_2 and (D) circRNA 2a\_2. The back splicing was determined by the red arrow. Exon 2 was revealed in dark blue; exon 2a was showed in green and exon 3 was in orange.

### 3.2.2 Levels of circRNAs in cancer and non-cancer cell lines

To determine the levels of circRNAs in cancer and non-cancer cell lines, I performed qPCR in NSCLC (A549 and PC9), SCLC (H82 and GCL8) and non-cancer cells PSAE. The primers for circRNA 2a\_2 and circRNA 3\_2\_2a are very specific while the primers covering the junction exon 3 and exon 2 can recognize both circRNA 3\_2 and circRNA 3\_2\_2a (Figure 11A). CircRNA 3\_2\_2a was much highly expressed in PC9 cells (Figure 11B) while the levels of combination of circRNA 3\_2\_2a and circRNA 3\_2 were most elevated in PC9 cells (Figure 11C). Additionally, the levels of

## Results

circRNA 2a\_2 in non-cancer cells PSAE were higher than those in lung cancer cells (NSCLC and SCLC) (Figure 11D).

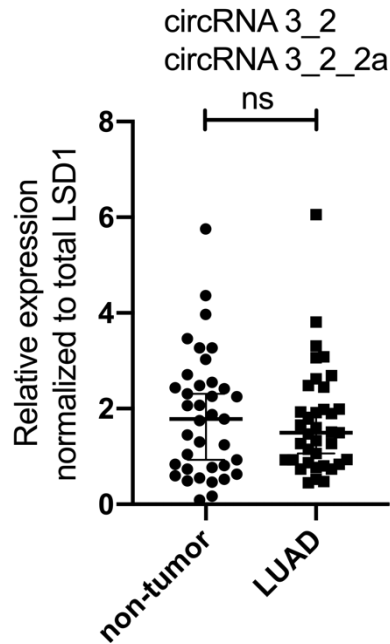


**Figure 11: Levels of circRNAs from LSD1 gene in cancer and non-cancer cell lines.** (A) Designed primers targeting the back-splicing junction for qPCR to detect the expression of circRNAs. Expression of (B) circRNA 3\_2\_2a, (C) circRNA 3\_2\_2a, circRNA 3\_2 and (D) circRNA 2a\_2 in NSCLC, SCLC and PSAE cells.

### 3.2.3 Levels of circRNAs in LUAD and non-tumor tissues

To explore the expression of circRNAs from LSD1 in clinical LUAD samples, I used the samples from a total of 46 patients as described in the part (result 3.1.2) and compared the levels of circRNAs between adjacent non-tumor and LUAD tissues by qPCR. The levels of circRNA 3\_2 and circRNA 3\_2\_2a did not differ in adjacent non-tumor and LUAD tissues (Figure 12).

## Results



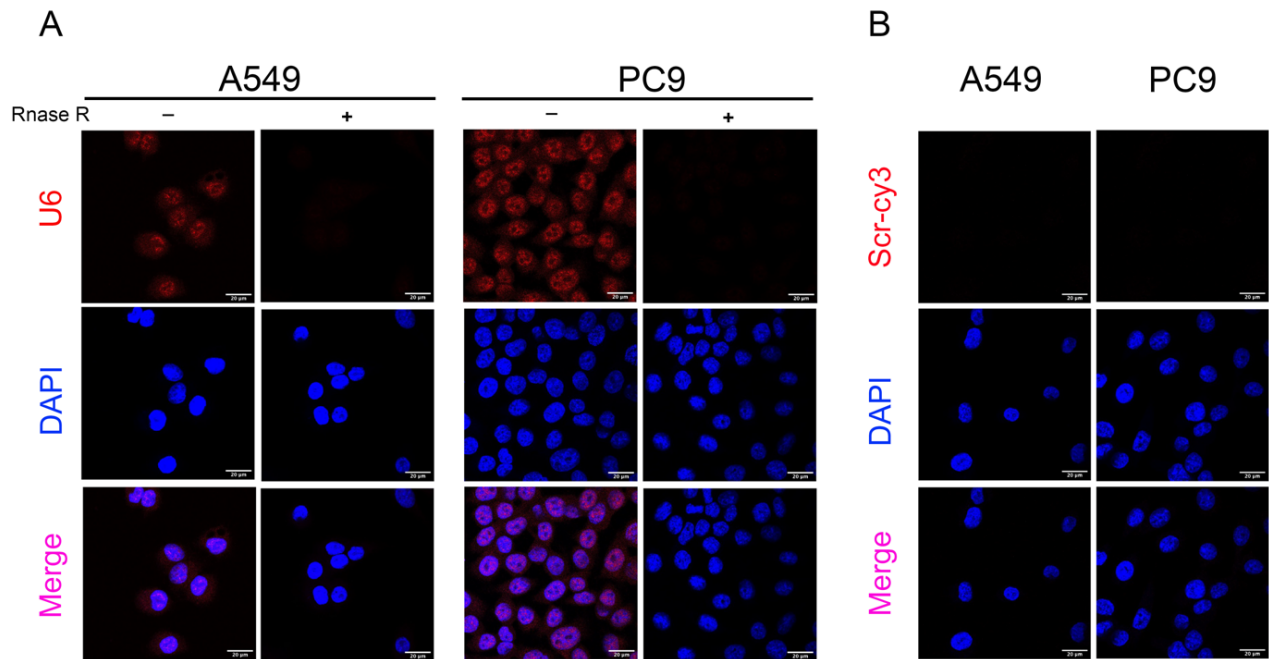
**Figure 12: Expression of circRNA 3\_2 and circRNA 3\_2\_2a between paired non-tumor and LUAD tissues.** RNA was extracted from 46 pairs of LUAD and non-tumor tissues and qPCR was performed to measure the levels of circRNAs which is normalized to total LSD1. ns, not significant.

### 3.2.4 Localization of circRNAs in cells

To explore the localization of circRNAs in cells, I performed Fluorescence in situ hybridization (FISH) examination on one hand and studied the nuclear and cytoplasmic fractions in different cells for the presence of the circLSD1 variants. To check the efficiency of FISH, I first used a Cy3 fluorochrome labeled probe for U6 RNA, which among small nuclear RNAs (snRNAs) in the nucleus is very conserved one. The red fluorescence signals of the U6 probe showed a good probe penetration into the nucleus and the efficient detection of U6 in nuclei of both A549 and PC9 cells.

After the treatment of RNase R, the red signal was totally gone in A549 and PC9 cells (Figure 13A), indicating the red signals were from the linear RNA U6. As a negative control scramble cy3 RNA probe (scr-cy3) was used. As shown in Figure 13B, there was no red signal.

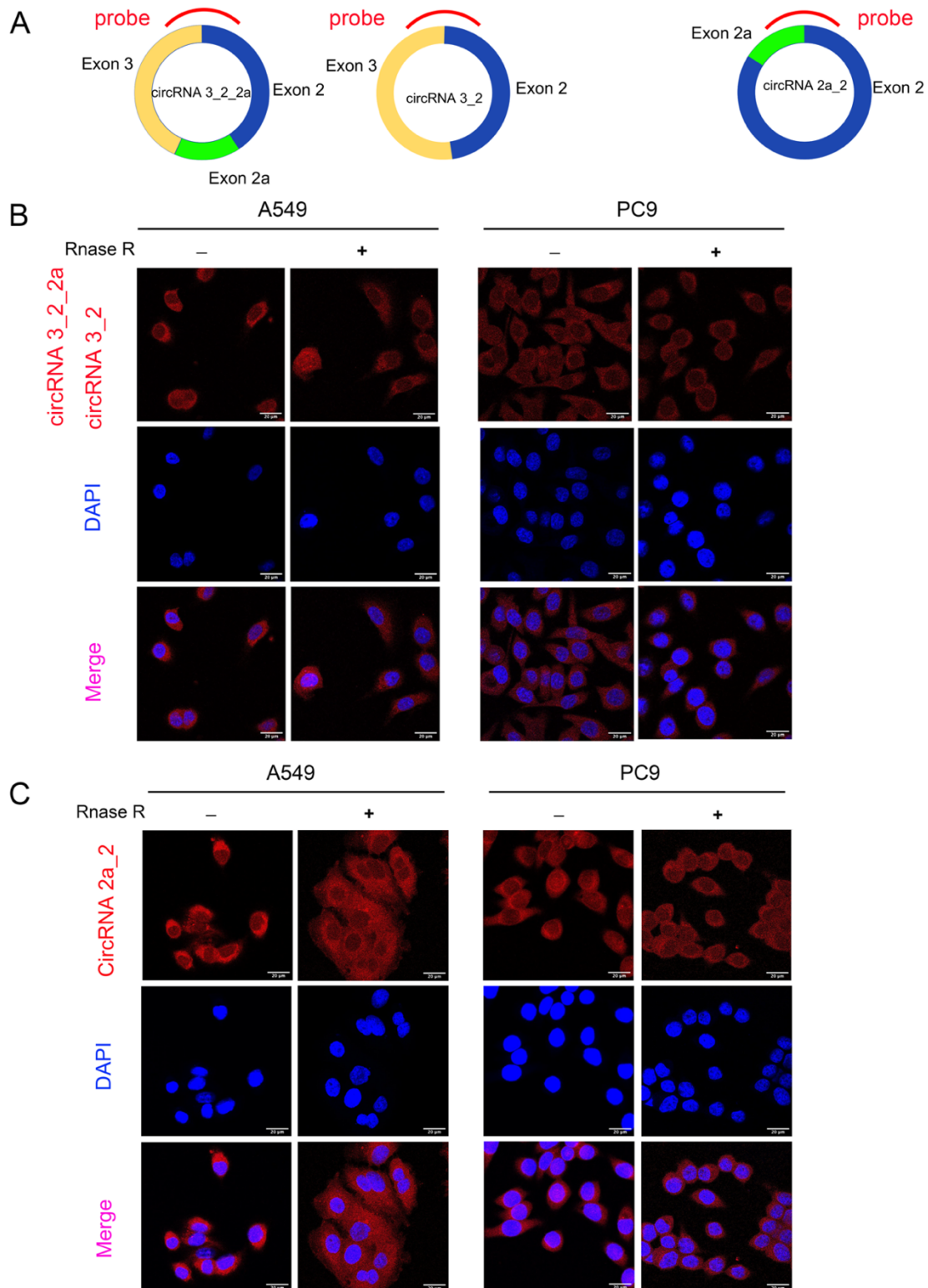
## Results



**Figure 13: RNA FISH analysis for U6 in A549 and PC9 cells.** Three independent experiments were carried out with similar results. (A) U6 (red) was stained as a positive control. -, no RNase R treatment. +, RNase R treatment. (B) Nuclei (blue) was stained with DAPI. (B) Scr-cy3 was stained as a negative control to detect the background signal in cells. Scale bar = 20  $\mu\text{m}$ .

To present the localization of circRNAs in the cells, the probes targeting the back-splicing junction of circRNAs were designed. For fluorescence microscopy they were fused to the cy3 fluorochrome. The probe, targeting the junction of exon 3 and exon 2, recognizes both the circRNA 3\_2\_2a and circRNA 3\_2 while the probe, binding the junction of exon 2a and exon 2, specifically recognizes circRNA 2a\_2 (Figure 14A). CircRNAs 3\_2 and 3\_2\_2a were predominantly located in the cytoplasm (Figure 14B). Importantly, after RNase R treatment, the red signals were not changed in A549 and PC9 cells demonstrating that circRNAs resistant to RNase R was recognized. Additionally, I also conducted FISH with the probe for circRNA 2a\_2. Similar to the results of circRNA 3\_2 and circRNA 3\_2\_2a, circRNA 2a\_2 was detected in red signal and mainly presented in cytoplasm and the signals were resistant to RNase R treatment (Figure 14C). All above data demonstrated that circRNA 3\_2, circRNA 3\_2\_2a and circRNA 2a\_2 was predominantly located in the cytoplasm.

## Results

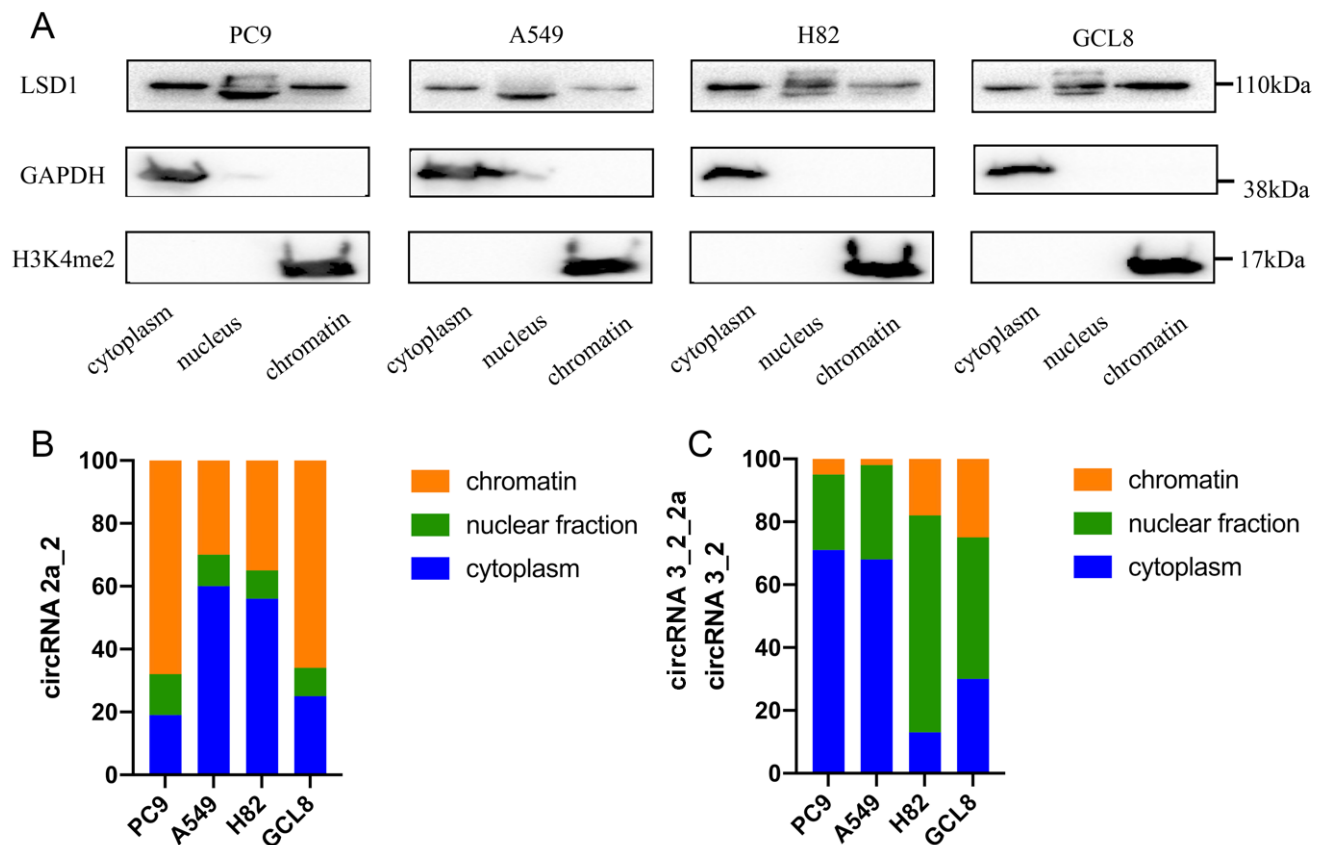


**Figure 14: Localization of circLSD1 RNAs.** (A) Design of the probes for circRNAs from LSD1, recognizing the junction region of circRNA 3\_2\_2a, circRNA 3\_2, and circRNA 2a\_2. RNA FISH analysis for (B) circRNA 3\_2\_2a, circRNA 3\_2 and (C) circRNA 2a\_2 in A549 and PC9 cells. Three independent experiments were carried out with similar results. -, no RNase R treatment. +, RNase R treatment. Recognized circRNA are stained by cy3 red fluorescence. Scale bar = 20  $\mu$ m.



## Results

To quantify the enrichment of circRNAs in the fractions of cells, we prepared the fractions of the cytoplasm, nucleus and chromatin from cells. GAPDH protein levels were taken as a marker for the cytoplasm fraction and H3K4me2 was used as a marker for the chromatin fraction. Western blot showed that the distinct fractions were well isolated (Figure 15A). Moreover, I performed qPCR to explore the levels of circRNAs after subcellular fractionation, and I found that circRNA 2a\_2 (Figure 15B) and circRNA 3\_2\_2a, circRNA 3\_2 (Figure 15C) was expressed in the nucleus and cytoplasm in NSCLC (A549 and PC9) cells. Therefore, subcellular fractionation as well as FISH examination revealed that circRNAs from LSD1 locus (circRNA 3\_2\_2a, circRNA 3\_2 and circRNA 2a\_2) were localized in the nucleus and cytoplasm.



**Figure 15: Quantitative analysis of circRNA 3\_2\_2a, circRNA 3\_2 and circRNA 2a\_2 in lung cancer cells.** (A) Protein levels of GAPDH, LSD1 and H3K4me2 in the subcellular fractions in the indicated cells. GAPDH was used as the marker for the cytoplasm and H3K4me2 was treated as the marker for the chromatin fraction. RNA levels of (B) circRNA 2a\_2 and (C) circRNA 3\_2\_2a and circRNA 3\_2 in the subcellular fractions of NSCLC (A549 and PC9) and SCLC (H82 and GCL8) cells.

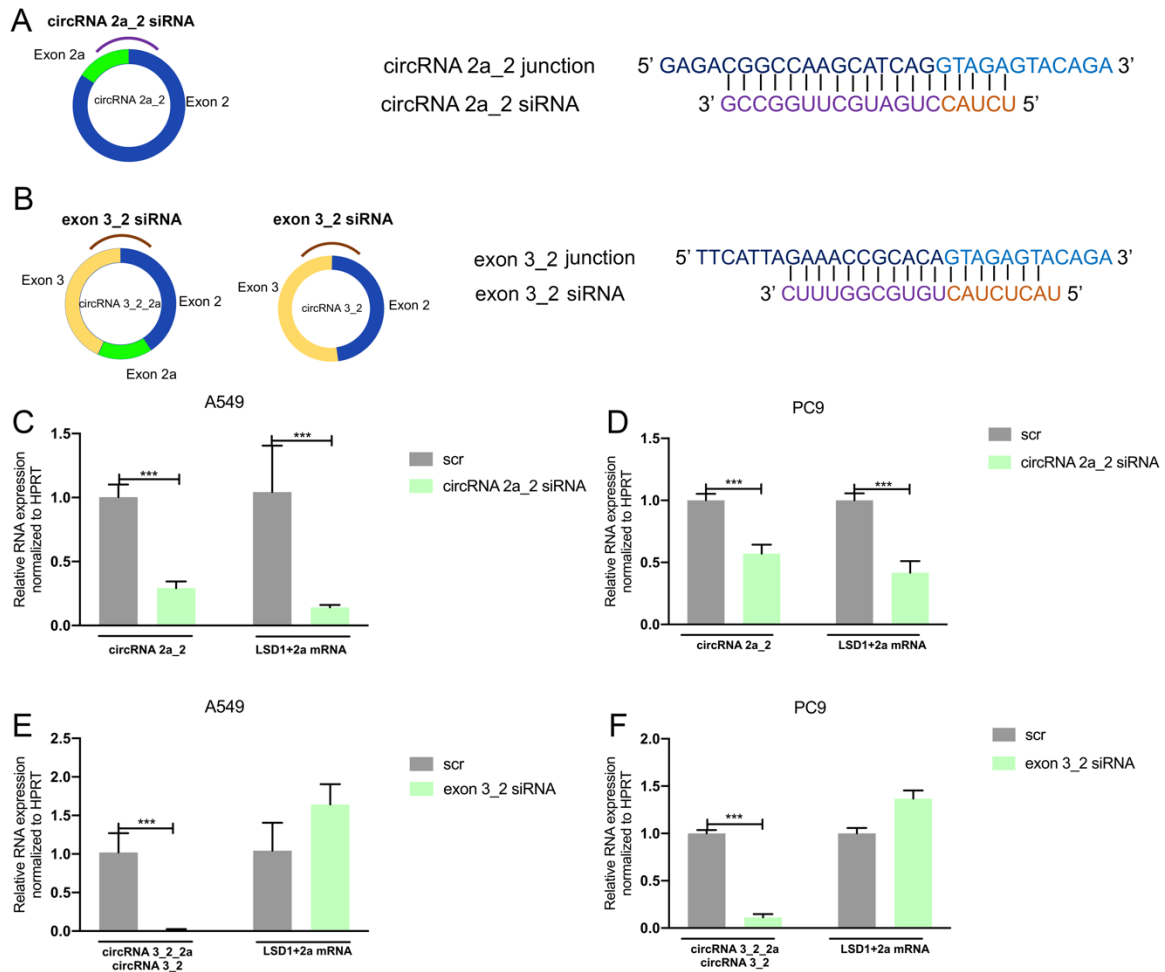
### 3.3 CircRNAs regulate alternative splicing of LSD1 by R-loop formation

Recent data provided evidence that in plants circRNA can interact with the parental genes and affect splicing [138]. However, in mammals, this was not yet observed. In order to investigate the function of circLSD1-RNA deriving from the exon 2 locus, I studied R-loop formation and splicing.

#### 3.3.1 Knockdown of circRNA 2a\_2 inhibited the expression of LSD1+2a

To explore the possible roles of circRNAs from LSD1, siRNAs were designed to knockdown the levels of circRNAs by targeting specifically the junctions of circRNAs from LSD1 (Figure 16A and B). After the transfection of siRNA, the levels of circRNA 2a\_2 and circRNA 3\_2\_2a and circRNA 3\_2 were significantly downregulated in A549 cells (Figure 16C and E) and in PC9 cells (Figure 16D and F). Interestingly, the linear LSD1+2a mRNA was also downregulated in response to circRNA 2a\_2 knockdown in A549 (Figure 16C) and PC9 cells (Figure 16D). However, the levels of LSD1+2a were not decreased after the knockdown of circRNA 3\_2\_2a and circRNA 3\_2 in A549 (Figure 16E) and PC9 cells (Figure 16F). Therefore, we suggested that the knockdown of circRNA 2a\_2 suppressed the expression of the linear LSD1+2a isoform. However, to validate this repressive interaction, it has to be validated that siRNA targeting the circRNA 2a\_2 does not also recognize the sequence of the exon 2a of the linear LSD1 2a isoform.

## Results



**Figure 16: Downregulation of circRNA 2a\_2 suppressed the expression of LSD1+2a variant.** The siRNAs targeted on (A) circRNA 2a\_2, (B) circRNA 3\_2\_2a and circRNA 3\_2 at junction sequences. The purple and orange sequences indicated the sequences forming the junction. Exon 3\_2 siRNA can target the common junction from circRNA 3\_2\_2a and circRNA 3\_2. Levels of LSD1 after knockdown of circRNA 2a\_2 in (C) A549 and (D) PC9 cells, circRNA 3\_2\_2a and circRNA 3\_2 in (E) A549 and (F) PC9 cells. \*\*\*  $p < 0.001$ .

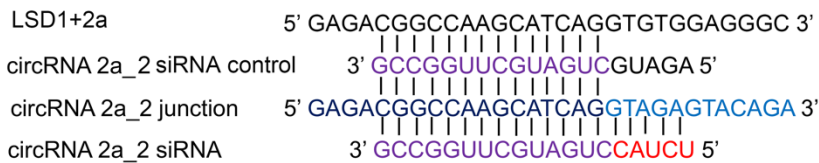
### 3.3.2 Decrease of LSD1+2a was regulated by circRNA 2a\_2 knockdown not by siRNA for circRNA 2a\_2

To confirm that the siRNA 2a\_2 does only target the circ2a\_2 and not the linear LSD1+2a isoform, the circRNA 2a\_2 control siRNA was designed. Importantly, corresponding to siRNA 2a\_2 recognizing circRNA 2a\_2, this control siRNA covered a sequence of 14 bases which was complementary to the exon 2a junction in circRNA2a\_2 and in addition to the exon 2a in linear LSD1+2a. However, the region recognizing the exon 2 of the circRNA 2a\_2 junction was replaced and transfected into A549 and PC9 cells (Figure 17A and B). Notably, after transfection of circRNA

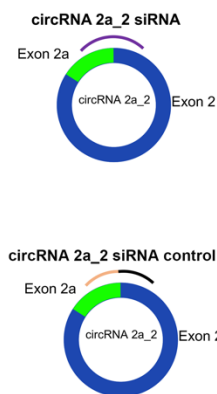
## Results

2a\_2 siRNA control into A549 and PC9 cells, the levels of circRNA 2a\_2 and LSD1+2a were not downregulated (Figure 17C). After validation of the specificity siRNA 2a\_2 for inhibition of circRNA 2a\_2 and not of the linear LSD1+2a variant, I concluded that the decrease of LSD1+2a was regulated by the knockdown of circRNA 2a\_2.

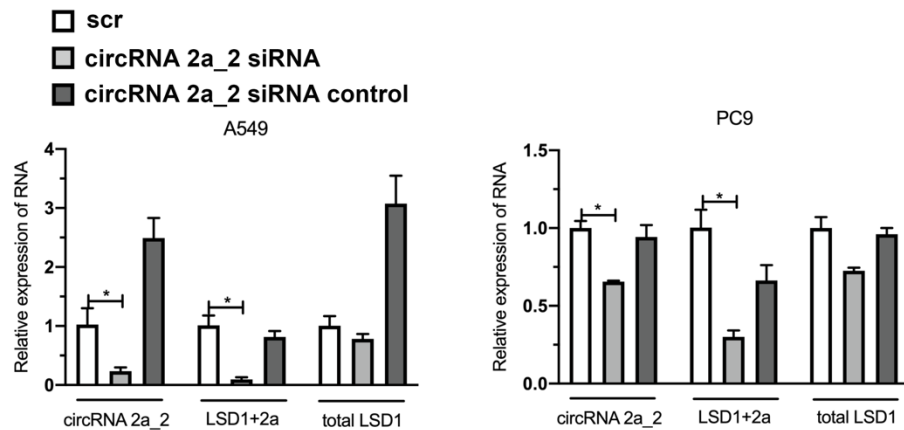
A



B



C



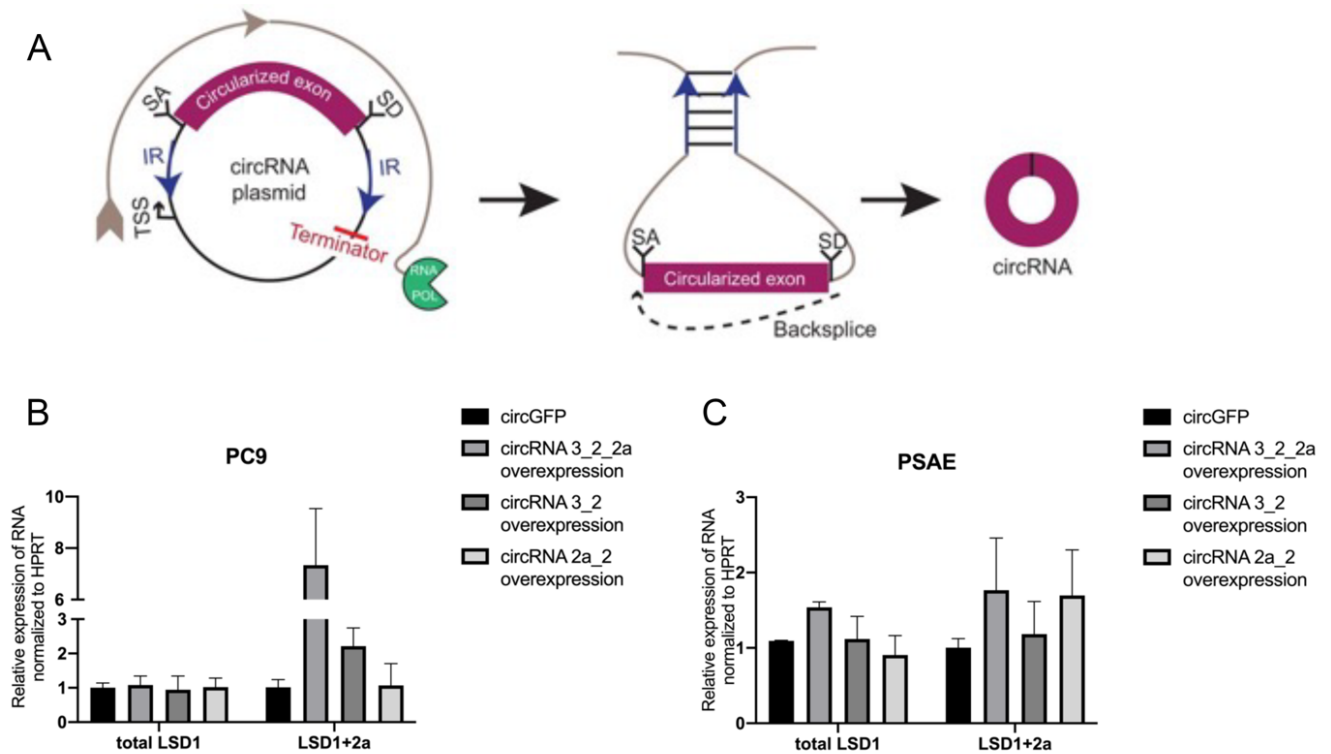
**Figure 17: siRNA 2a\_2 does not target LSD1+2a.** (A) CircRNA 2a\_2 siRNA control targeted partly the junction of circRNA. Mismatched sequences are shown in black. (B) A cartoon revealed the mismatching of siRNA control for circRNA 2a\_2. (C) Levels of LSD1+2a after transfection of circRNA 2a\_2 siRNA and siRNA control. \*  $p < 0.05$ .

### 3.3.3 Overexpression of circRNAs from LSD1 upregulated the levels of LSD1+2a

To determine the functions of circRNAs after overexpression, according to plasmid construction described by Barrett and the colleagues [158], Dr. Maria Anokhina (RG Odenthal, Institute for Pathology, University of Cologne) established plasmids, harboring the expression cassettes for the different circRNAs deriving from the exon 2 locus (Figure 18A). CircGFP was used as the control of transgenic circRNA expression. After the transfection of plasmids, the levels of LSD1+2a were upregulated after circRNA 3\_2\_2a and circRNA 3\_2 overexpression in PC9 cells

## Results

(Figure 18B). Moreover, the levels of LSD1+2a were increased after circRNA 3\_2\_2a and circRNA 2a\_2 overexpression in PSAE cells (Figure 18C).



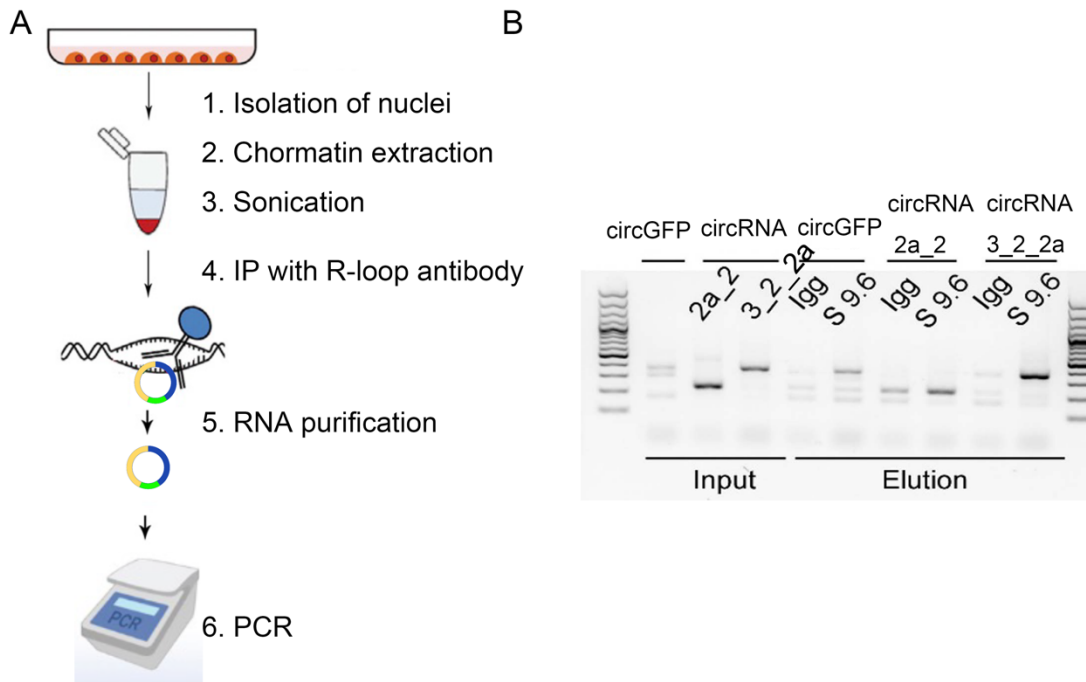
**Figure 18: Overexpression of circRNAs upregulated the levels of LSD1+2a.** (A) Scheme of structure and function of plasmids carrying the circRNA expression cassettes. (B, C) Levels of total LSD1 and LSD1+2a after transfection of plasmids in PC9 and PSAE cells.

### 3.4 CircRNA 3\_2\_2a formed R-loops with the parental gene

After showing that circRNA 3\_2\_2a and circRNA 3\_2 overexpression increased the levels of LSD1+2a variant, I analyzed the mechanisms of circRNA 3\_2\_2a and circRNA 3\_2 for the regulation. Given the circRNA and pre-mRNA are derived from the same strand and lack predicted complementary binding sites that could mask the splicing donor/acceptor sites, I hypothesized that this circRNA might bind genomic DNA to form an RNA: DNA hybrids, or R-loops. I performed Immunoprecipitation (IP) with the antibody of R-loops S9.6 recognizing R-loop formations and purified RNA from the pulled down fraction and performed PCR with divergent primers to amplify circRNAs (Figure 19A). These experiments were performed under the guidance of Dr. M. Anokhina (RG Odenthal, Institute for Pathology, University Hospital of Cologne). GAPDH was used as the control to show equal input. The gel

## Results

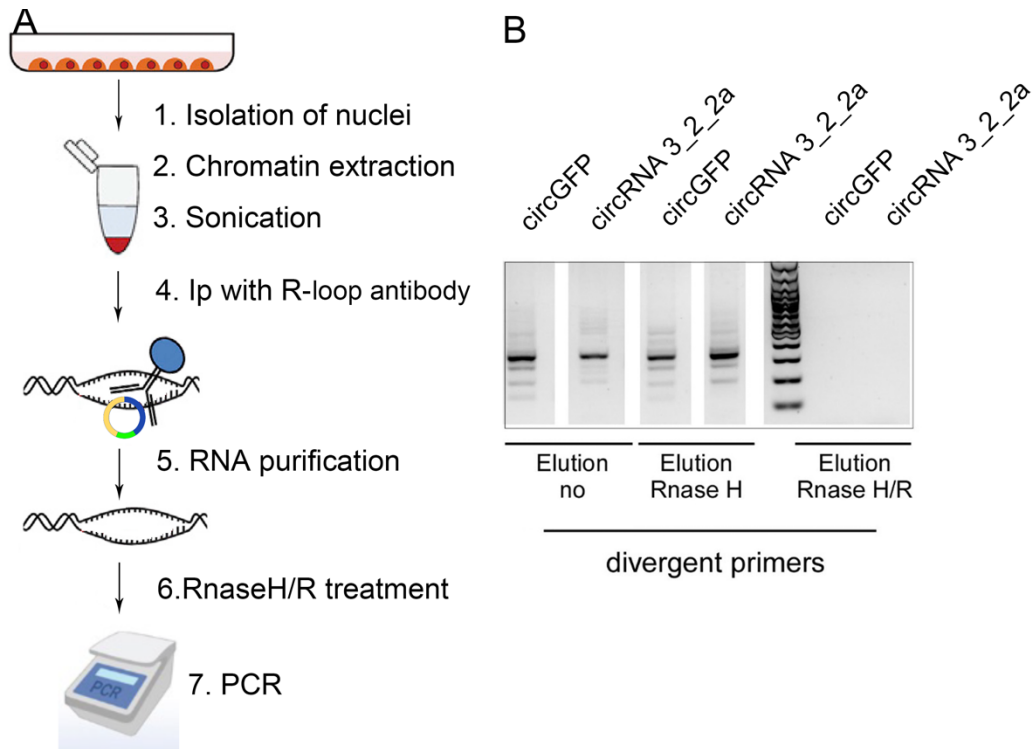
electrophoresis of the respective amplicons indicated that circRNA 3\_2\_2a was much higher enriched in the R-loop S 9.6 precipitate compared to the IgG elution fraction (Figure 19B). Importantly, this findings provided evidence that circRNA 3\_2\_2a and circRNA 2a-2 bind to the genomic DNA locus of the LSD1.



**Figure 19: CircRNA 3\_2\_2a formed R-loops.** (A) Scheme of the workflow. (B) PCR results showing the products from the RNA fraction purified from immunoprecipitation (IP) prepared from chromatin of PC9. The S 9.6 R-loop antibody was applied to IP. RNA input of the chromatin fraction was used as reference. The pull-down from the IgG was taken as a negative control, showing the background of IP.

To identify the purified RNA is circRNAs forming R-loops, the purified RNA was treated with RNase H or RNase H/R (Figure 20A). Pre-treatment with RNase H which degrades RNA in RNA: DNA hybrids and RNase R ablated the signal, suggesting the signal was R-loop-specific (Figure 20B). This finding confirmed the results that circRNA 3\_2\_2a bound to genomic DNA forming R-loops.

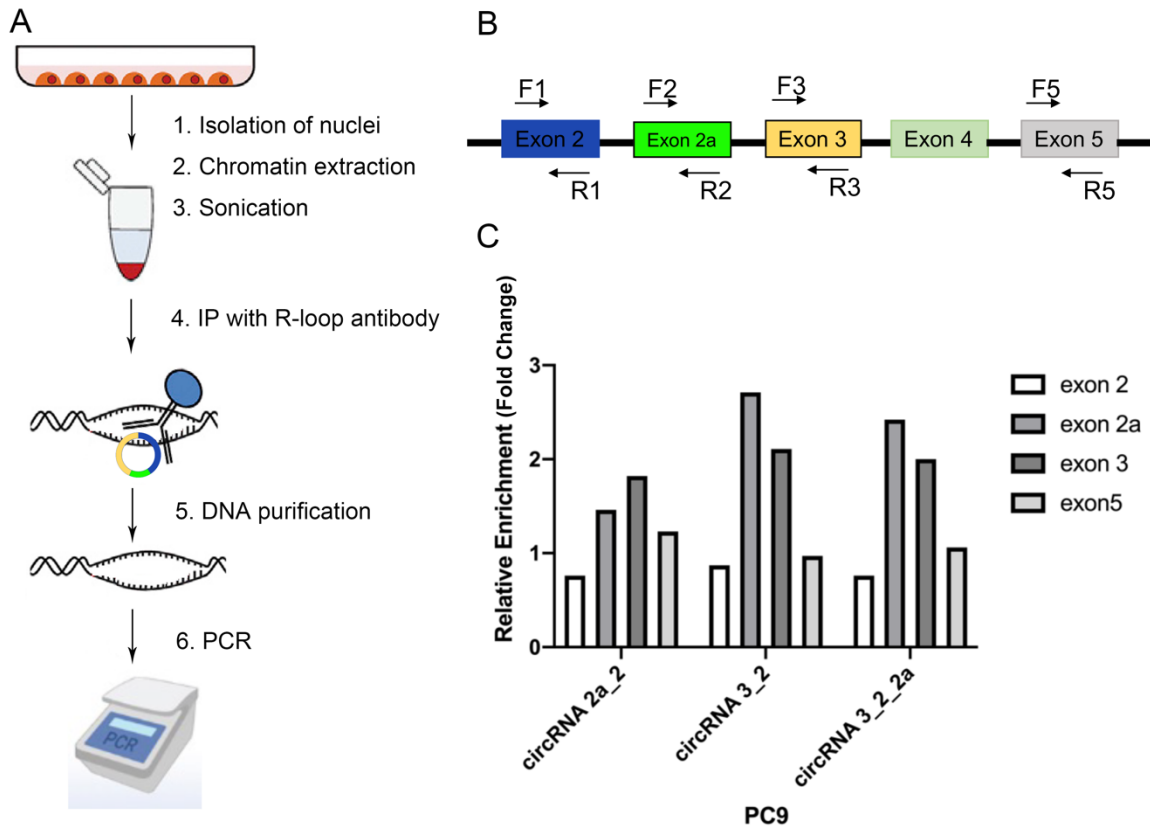
## Results



**Figure 20: Overexpression of circRNA 3\_2\_2a and R-loop formation.** (A) A scheme of workflow. (B) PCR results showing the products from the RNA fraction purified from immunoprecipitation (IP) prepared from chromatin of PC9 after circRNA overexpression. After the purification of RNA, we treated it with RNase H and R. Then the divergent primers were used to amplify circRNAs.

I suggested that the circLSD1-RNA formed the R-Loop with DNA of the parental gene at the locus from which they are deriving, because there are complementary sequences. To further confirm the R-loop with the parental genes we transfected PC9 cells with the plasmids to overexpress circRNAs, then conducted IP and purified DNA from IP and performed qPCR to measure the enrichment of LSD1 exons (Figure 21A). We designed several pairs of primers covering the regions from exon 2 to 5 of LSD1 (Figure 21B). The overexpression of circRNA 2a\_2, circRNA 3\_2 and circRNA 3\_2\_2a upregulated the enrichment of exon 2a and 3 in PC9 cells (Figure 21C). This results revealed that circRNAs interacted with genomic DNA of its parental gene to form R-loops in LUAD cells.

## Results



**Figure 21: CircRNAs interacted with its parental gene to form R-loops.** (A) Flow chart of workflow. (B) Designed primers for qPCRs targeting different exons of LSD1. (C) qPCR showing the enrichment of exons of LSD1 after circRNAs overexpression. The levels of exons of LSD1 were normalized to the input and circGFP control.



## Results

### 3.5 LSD1+8a

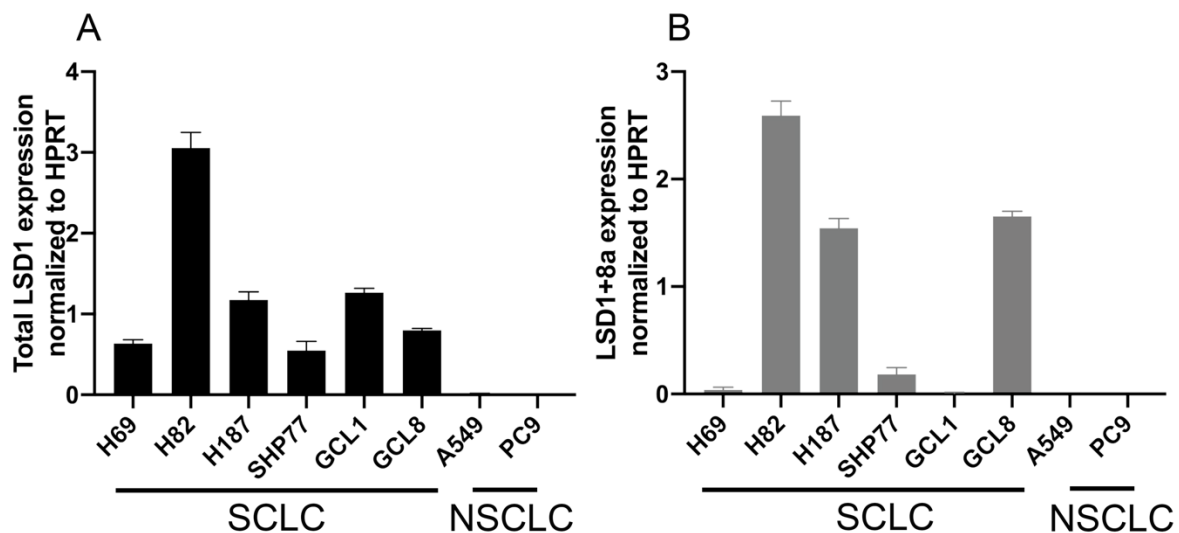
LSD1+8a, a neurospecific splice variant of LSD1, contains the additional exon 8a and expressed specifically during mammalian neuronal development [78]. The alternative splicing mechanism regulating inclusion of exon 8a is fundamental for the regulation of specific expression programs that promote neuronal differentiation and is strictly regulated during mammalian brain development.

#### 3.5.1 LSD1+8a in SCLC

Strikingly, Jotatsu et al. found that LSD1+8a is also expressed in SCLC cell lines and cells with high LSD1+2a is resistant to cisplatin (CDDP), suggesting the potential therapeutic target for treatment of SCLC chemoresistance [82].

#### 3.5.2 LSD1+8a expression in SCLC cell lines

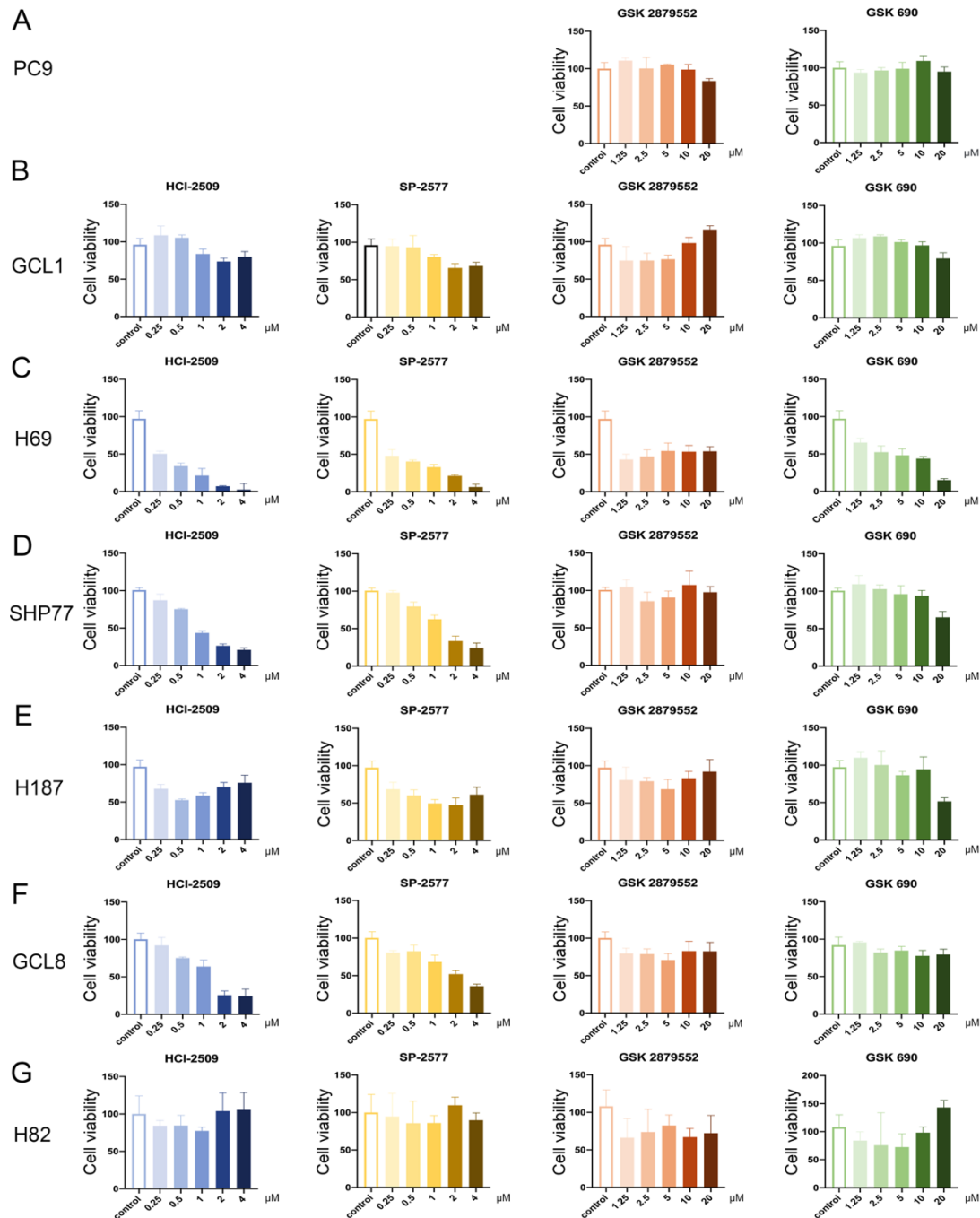
To determine the roles of LSD1+8a in SCLC, I performed qPCR for 8 human lung cancer cell lines including 6 SCLC and 2 NSCLC cell lines. The levels of total LSD1 (Figure 22A) and LSD1+8a (Figure 22B) was much higher in SCLC cells compared to that in NSCLC cells. Notably, the levels of total LSD1 (Figure 22A) and LSD1+8a (Figure 22B) were highest in H82 cells than those in other cells. It was indicated that LSD1+8a was higher expressed in SCLC cells.



**Figure 22: Levels of (A) total LSD1 and (B) LSD1+8a in SCLC and NSCLC cell lines.** Absolute quantification was used to measure the levels of total LSD1 and LSD1+8a which were normalized to HPRT.

## Results

In order to explore the possible correlation between the levels of LSD1+8a and the cell viability to LSD1 inhibitors, I performed MTT assay and measured the OD values of cells after treatment of LSD1 inhibitors. NSCLC cells PC9 were sensitive to HCI-2509 with an IC<sub>50</sub> of 1.5  $\mu\text{M}$  *in vitro*, which has been validated by the former colleague Dr. Macheleidt [71]. GCL1 (Figure 23A), H187 (Figure 23E) and H82 (Figure 23G) cells were resistant to any LSD1 inhibitors.



**Figure 23: Effect of LSD1 inhibitors on cell viability in lung cancer cells.** (A) PC9; (B) GCL1; (C) H69; (D) SHP77; (E) H187; (F) GCL8 and (G) H82. Cells were treated with LSD1 inhibitors for 48 h and cell viability was determined by MTT assay.

## Results

Meanwhile, SHP77 (Figure 23D) and GCL8 (Figure 23F) cells were sensitive to HCI-2509 and SP-2577 but resistant to GSK2897552 and GSK690. H69 cells were sensitive to HCI-2509, SP-2577 and GSK2897552. Therefore, this table showed the characteristics of cell lines and LSD1 inhibitors (Table 2). Considering the levels of total LSD1 and LSD1+8a in distinct SCLC cells, no strong correlation was found between the levels of LSD1+8a and the effect to LSD1 inhibitors.

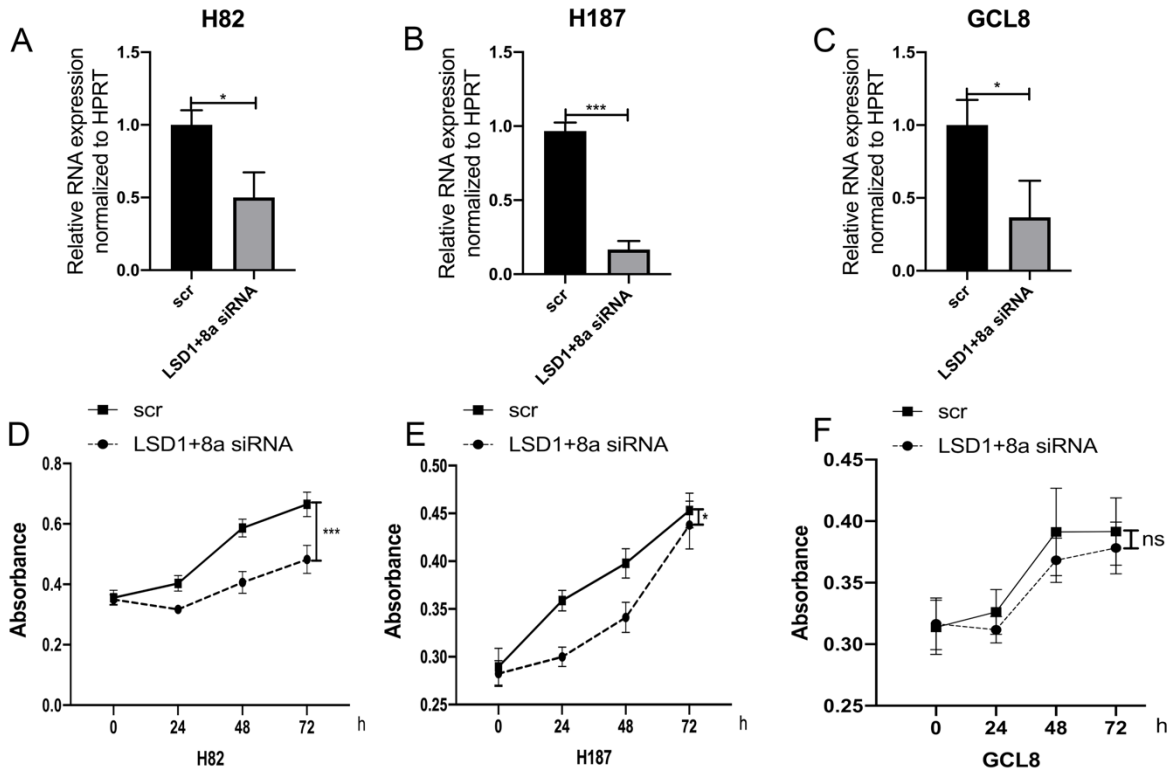
**Table 2: Characteristics of cell lines and LSD1 inhibitors.**

Cell lines	Total LSD1 levels	LSD1+8a levels	HCI-2509	SP-2577	GSK28759552	GSK690
H69	low	low	+	+	+	+
H82	very high	very high	-	-	-	-
H187	very high	high	-	-	-	-
GCL1	low	high	-	-	-	-
GCL8	very high	Low	+	+	-	-
SHP77	high	low	+	+	-	-

### 3.5.3 Suppression of LSD1+8a repressed the cell viability in SCLC

To discover whether the expression of LSD1+8a affects the cell viability, I transfected SCLC cells with scramble (scr) and siRNA for LSD1+8a and performed MTT assay. The levels of LSD1+8a were downregulated as much as 50%, 80% and 60% in H82 (Figure 24A), H187 (Figure 24B) and GCL8 (Figure 24C), respectively. It suggested the levels of LSD1+8a was significantly downregulated with high efficiency of siRNA for LSD1+8a. Meanwhile, I performed MTT assay to investigate SCLC cell viability. In H82 (Figure 24D) and H187 (Figure 24E) cells, downregulation of LSD1+8a led to the significant decrease of OD values, which was used for the analysis of cell viability. However, there was no difference of OD values between scr group and siRNA for LSD1+8a (Figure 24F). Thus, all above data revealed that suppression of LSD1+8a inhibits cell viability in SCLC.

## Results

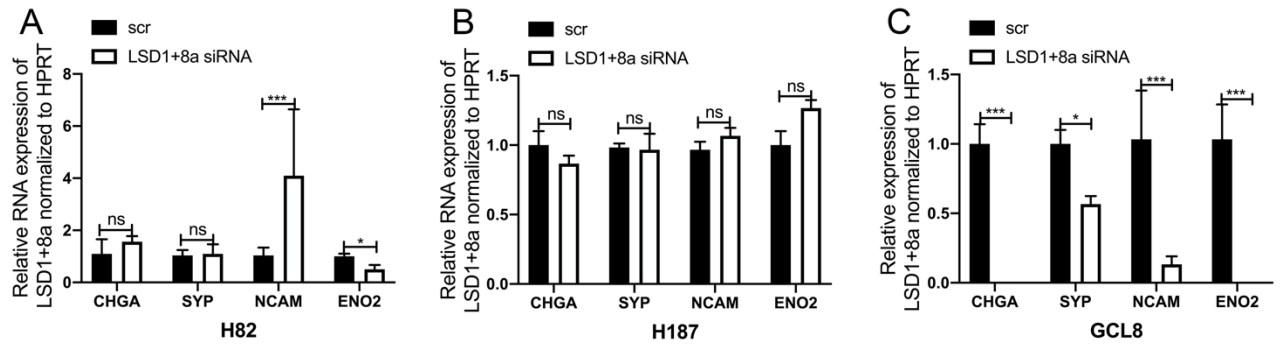


**Figure 24: Knockdown of LSD1+8a inhibited the cell viability in SCLC.** Efficiency of siRNA for LSD1+8a in (A) H82, (B) H187 and (C) GCL8 cells. MTT assay to detect cell viability of (D) H82, (E) H187 and (F) GCL8 cells transfected with scr or siRNA for LSD1+8a at 24, 48 and 72h. \*  $p < 0.05$ , \*\*\*  $p < 0.001$ , ns, not significant.

### 3.5.4 Correlation between LSD1+8a and neuroendocrine markers

A prior study found the correlation of LSD1+8a and neuroendocrine markers. To discover the regulation between LSD1+8a and several neuroendocrine makers (CHGA, SYP, NCAM and ENO2) in SCLC, I transfected H82, H187 and GCL8 cells with scr or siRNA for LSD1+8a. After the transfection of siRNA for LSD1+8a, the levels of CHGA and SYP did not change, and the levels of NCAM was increased and the levels of ENO2 was downregulated in H82 cells (Figure 25A). In H187 cells, the levels of neuroendocrine markers were not affected by knockdown of LSD1+8a (Figure 25B). Notably, the levels of neuroendocrine markers were dramatically downregulated by knockdown of LSD1+8a (Figure 25C).

## Results

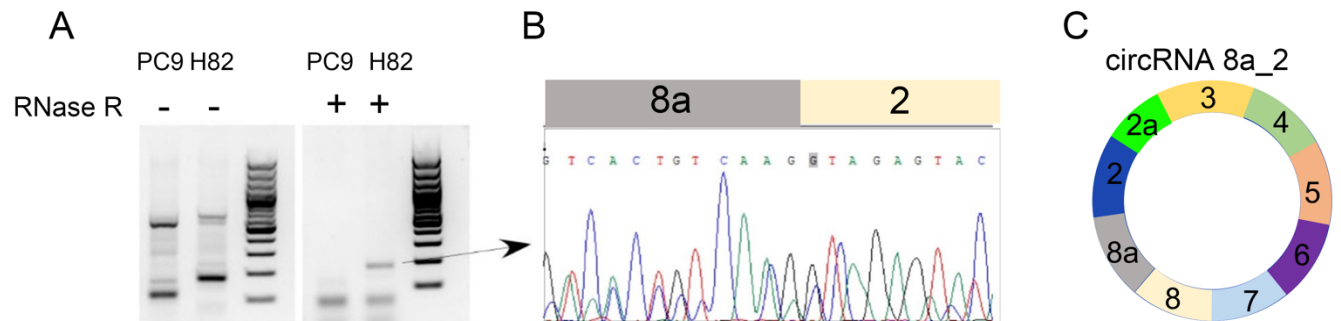


**Figure 25: Knockdown of LSD1+8a affected the expression of neuroendocrine marker genes in (A) H82, (B) H187 and (C) GCL8 cells. \*  $p < 0.05$ , \*\*\*  $p < 0.001$ , ns, not significant.**

### 3.5.5 CircRNA 8a\_2 regulated the levels of LSD1+8a

#### 3.5.5.1 Knockdown of circRNA 8a\_2 led to the downregulation of LSD1+8a

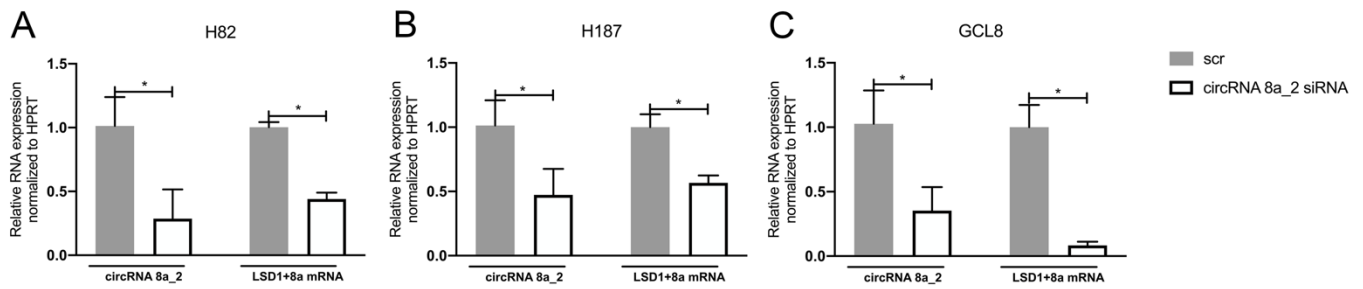
To explore whether circRNAs can regulate the levels of the novel variant LSD1+8a, we purified RNA from IP and treated it with RNase R in PC9 and H82 cells (Figure 26A). After the pre-treatment with RNase R, we found a novel circRNA containing the back-splicing junction of exon 8a and exon 2 from LSD1 (Figure 26B and C).



**Figure 26: CircRNA 8a\_2 was identified in H82 cells. (A) Agarose gels indicating the circRNAs after RNase R treatment. (B) CircRNA 8a\_2 contains the back-splicing junction of exon 8a and exon 2. (C) Cartoon showing the sequences of exons from LSD1.**

To determine the function of circRNA 8a\_2 in SCLC, we hypothesized whether circRNA 8a\_2 can regulate the levels of LSD1+8a isoform similar to circRNA 3\_2\_2a. I transfected SCLC cells with siRNA to knockdown the levels of circRNA 8a\_2. After the transfection of siRNA for circRNA 8a\_2, the levels of circRNA 8a\_2 were decreased and the levels of LSD1+8a were also significantly downregulated in H82 (Figure 27A), H187 (Figure 27B) and GCL8 (Figure 27C).

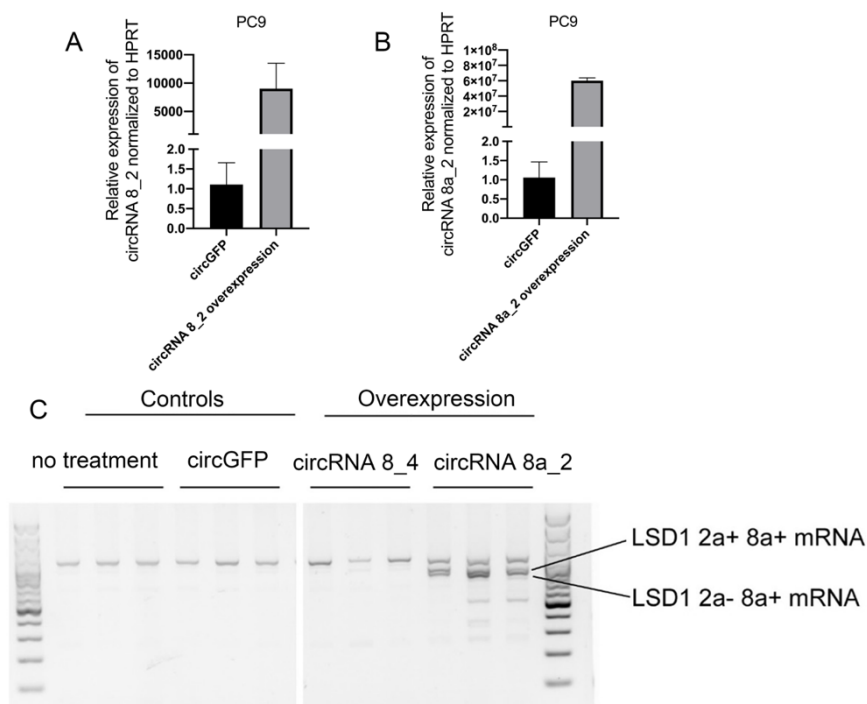
## Results



**Figure 27: Knockdown of circRNA 8a\_2 led to the downregulation of LSD1+8a in (A) H82, (B) H187 and (C) GCL8 cells. \*  $p < 0.05$ .**

### 3.5.5.2 Overexpression of circRNA 8a\_2 upregulated the levels of LSD1+8a

The plasmids for circRNA 8\_2 and 8a\_2 overexpression were transfected to PC9 cells. The levels of circRNA 8\_2 (Figure 28A) and circRNA 8a\_2 (Figure 28B) was dramatically upregulated in PC9 cells. In addition, the levels of LSD1+8a were increased following the upregulation of circRNA 8a\_2 (Figure 28C).



**Figure 28: Overexpression of circRNA 8a\_2 upregulated the levels of LSD1+8a. Levels of (A) circRNA 8\_2 and (B) circRNA 8a\_2 in PC9 cells after the transfection of plasmids. (C) Agarose gel indicating the upregulation of LSD1+8a after circRNA overexpression.**

### 5. Discussion

In addition to playing crucial functions in reprogramming throughout normal development and maintaining tissue-specific transcription patterns, epigenetic mechanisms including post-translational histone alterations and DNA methylation also modulate gene expression. As one of the most well-studied epigenetic regulators in cancers, the first identified histone demethylase, LSD1, context-dependently demethylates H3K4me1/2 and H3K9me1/2 at target loci [52]. LSD1 is expressed at a high level in many malignancies and performs a fundamental function in tumorigenesis, cell differentiation, proliferation, and self-renewal [159]. In the present thesis work, the expression of LSD1 isoforms and circRNAs has been studied.

#### 4.1 Identification of circRNAs from LSD1 isoforms

Numerous studies revealed the important roles of LSD1 in the tumorigenesis and development of tumors. In our lab, former colleagues Lim et al. found that elevated expression level of LSD1 was correlated with malignant lung tumors [70]. In another study, overexpression of LSD1 was shown to be linked to prostate cancer progression and recurrence [160]. Moreover, the elevated expression level of LSD1 was linked to unfavorable prognosis in NSCLC patients, which enhanced the capacity of tumor cells to proliferate, migrate, and invade [161]. Even though the roles of LSD1 are well-studied in cancers, the functions of LSD1 isoforms containing exon 2a and/or exon 8a remained to be explored.

There are 19 exons in the LSD1 gene, all of which are well conserved among vertebrates. Four distinct LSD1 variants may be produced by RNA alternative splicing of two extra exons, exon 2a, and exon 8a. These variants include the conventional LSD1, LSD1 plus exon 2a (LSD1+2a), exon 8a (LSD1+8a), or both (LSD1+2a+8a). These four alternative splicing variants were also found in zebrafish [162]. To investigate the roles of LSD1+2a in LUAD, I measured the proportion of LSD1+2a/total LSD1. In most cancer and non-cancer cell lines, the proportion of LSD1+2a in total LSD1 ranged from 20% - 50% in human cell lines while in NSCLC cells, it was 30%. To compare the levels of LSD1+2a in the non-tumor and LUAD

## Discussion

tissues, I analyzed the total LSD1 and LSD1+2a mRNA levels with the TCGA cohort. Among 24 types of cancer, LSD1+2a is higher expressed in more than 10 types of cancer including LUAD. However, the information about the ratio of LSD1+2a/ total LSD1 variant is missing in the TCGA dataset. In clinical samples from University Hospital of Cologne, there was no difference between LUAD tissue areas and matching non-tumor tissues, which is different from the TCGA cohort data. The possible reason might be the low quality of extracted RNA from clinical samples. In particular, we found that in the peritumor tissue the degradation of the RNA was very high. Therefore, and the linear RNA (HPRT and LSD1) was partial degradation but circRNA was more stable for the degradation. More high-quality RNAs from samples are needed to further quantify the levels of LSD1+2a in LUAD and adjacent non-tumor tissues.

To identify the prognostic values of LSD1+2a for LUAD patients, I analyzed the prognostic indicators (overall survival and progression-free survival) with the data from the TCGA database. High levels of total LSD1 indicated a worse prognosis for LUAD patients, which was similarly found in thyroid cancer [83], neuroblastoma [73], endometrial adenocarcinoma [163] and breast cancer [164]. However, there was no significant difference between high and low LSD1+2a patients on the prognosis. The possible reason for the inconsistency was that the shortest LSD1-2a-8a which does not carry the exon 2a and 8a demethylated H3K4 and H3K9 most efficiently. In contrast, in this report it was shown that the variant LSD1+2a, carrying the exon 2a had little demethylation activity for both targets H3K9 and H3K4 [165]. That might differ in distinct functions. But the molecular mechanisms are not clear. The potential reasons for the different prognostic values between total LSD1 and LSD1+2a need to be explored by further experiments.

### **4.2 Expression of circRNAs, deriving from the LSD1 gene locus, in cancer**

Due to the close association between back-splicing and circRNAs, I explored the potential of the LSD1 gene to generate circRNAs. After circRNA sequencing, it was found that two circRNAs (hsa\_circ\_737 and hsa\_circ\_733) derived from LSD1 were higher expressed in lung adenocarcinoma PC9 cells than in lung epithelial cells PSE. To proof the expression of circRNAs from LSD1 in cells, I used divergent



## Discussion

primers from exon 2 and amplified the four circRNAs. The circRNAs were resistant to the degradation of RNase R, validating that they were circular RNA. Combining the sequence of amplified products of circRNA, three circRNAs were identified from the LSD1 locus named circRNA 3\_2\_2a (has\_circ\_0009061), circRNA 3\_2 and circRNA 2a\_2. To explore the roles in tumorigenesis, I measured the levels of three circRNAs in NSCLC, SCLC and non-cancer cell lines. The combined levels of circRNA 3\_2\_2a and circRNA 3\_2 and circRNA were much higher in PC9 cells than in the non-cancer cells. However, in a recent study, it was reported that the levels of circRNA 3\_2\_2a (has\_circ\_0009061) were lower in prostate cancer tissues than in the non-tumor prostate tissues [166], which was not consistent with our results. However, the difference between non-cancer and cancer-associated expression might be tissue, cancer or mutation dependent. Indeed, my result showed a big difference between PC9 and A549 cells, carrying both NSCLC relevant mutations. These cells show a different LSD1 expression (own data, Lim et al) with high levels in EGFR mutated PC9 cells, whereas A549 cells with a KRAS mutation have relatively low levels. The circRNA 3\_2\_2a pattern is also different. Notably, a plenty of circRNAs such as circHMGB2 (hsa\_circ\_0071452 [167]), has\_circRNA\_002178 [168], circFAT1(hsa\_circ\_0001461 [169]) and circXPO1 (hsa\_circ\_0001016 [170]) were significantly upregulated in LUAD tissues. However, circGSK3B (hsa\_circ\_0066903 [171]), circDCUN1D4 (hsa\_circ\_0007928 [172]), circKEAP1 (hsa\_circ\_102442 [173]) and circFBXW7 [174] were found to be subjected to downregulation in LUAD tissues. Therefore, the expression profiles of distinct circRNAs differed in LUAD tissues.

Moreover, the diagnostic values for circRNA 3\_2\_2a were found in a study for prostate cancer. In the study of Song et al., the diagnostic values of circRNA 3\_2\_2a were measured by Receiver Operating Characteristic (ROC) curves [166]. The area under the curve (AUC) for circRNA 3\_2\_2a was 0.711, suggesting a diagnostic biological marker for the identification of prostate cancer. Combined with the clinicopathological features of patients with prostate cancer, circRNA 3\_2\_2a was found to be associated with the Gleason score and patients' pathological stages, demonstrating the clinical significance of circRNA 3\_2\_2a in this disease. In recent studies, several circRNAs were also identified as biomarkers for the prognosis and diagnosis of LUAD. It was revealed that has\_circ\_002178 existed in exosomes and the AUC of the has\_circ\_002178 was 0.9956, serving as a novel diagnostic

## Discussion

biomarker for LUAD [168]. Moreover, the link between hsa\_circ\_0056616 and the diagnosis of lymph node metastasis of LUAD was detected based on the AUC of 0.812 [175]. The levels of hsa\_circ\_0056616 were found to be negatively linked to the T stage, M stage, and TNM grade of LUAD [175]. Meanwhile, hsa\_circ\_0001715 was determined as a novel prognostic marker for LUAD [176]. LUAD patients with high hsa\_circ\_0001715 levels suffered from worse OS than those with low levels. Univariate Cox analysis and multivariate Cox analysis supported the prognostic values [176]. Interestingly, the circRNA 3\_2\_2a was found to be downregulated in LUAD cells with resistance to gefitinib [177]. Thus, circRNA 3\_2\_2a may play vital roles in the chemoresistance in LUAD, but functional analysis of therapeutic transgenic expression is missing. Reports have shown that circASK1 (hsa\_circ\_0007798) increases gefitinib sensitivity in LUAD cells while being remarkably downregulated in gefitinib-resistant cells. Thus, subsequent experiments are important to verify the chemoresistance of circRNAs from LSD1 in LUAD.

One remarkable feature of circRNA is subcellular distribution. Most circRNAs, excluding those containing introns, are transported from the nucleus to the cytoplasm shortly following their biogenesis [102, 178]. It has been shown that both intronic circRNAs and elciRNAs are abundant in the nucleus [112, 179]. In my study, three circRNAs are composed of exons from LSD1 and were visualized by FISH to determine the subcellular localization. The specific probes for the backsplicing junction of circRNAs revealed that circRNA 3\_2\_2a, circRNA 3\_2 and circRNA 2a\_2 were predominantly located in the cytoplasm. The fluorescence intensity was not changed after RNase R treatment demonstrated the specificity of the red signal for circRNAs respectively. In addition, I performed qPCR to analyze the levels of three circRNAs in the fractions of cells, and three circRNAs were found in the cytoplasm and nucleus, consistent with the findings by FISH.

The functional roles of circRNAs are important in the proliferation, invasion, migration and chemoresistance of LUAD. Thus, knockdown of circFBXW7 enhanced the proliferative capacity of LUAD cells in vitro, while the overexpression of circFBXW7 suppressed the tumor growth in vivo [174]. Moreover, the overexpression of circASPH by lentivirus enhanced the proliferative, migratory, and invasive capacities of A549 and PC9 cells, consistent with the data from in vivo experiments [180].

## Discussion

Hsa\_circ\_0003998 is expressed at a high level in the docetaxel-resistant cell lines. Silencing hsa\_circ\_0003998 made LUAD cells more sensitive to docetaxel, indicating a link between the two [181]. In my study, the MTT assay demonstrated that knockdown of circRNA 3\_2\_2a and circRNA 2a\_2 did not affect the proliferation of LUAD cells (PC9 and A549) (data was not shown). Though in the past the function of many circRNAs in cancer was highlighted [182-185], the knowledge about the role for circRNA from LSD1 was still missing.

### 4.3 CircRNAs from LSD1 regulate the alternative splicing of the parental gene LSD1

A novel study revealed SEPALLATA3 (SEP3) circRNAs, which are generated from exon 6, remarkably enhance the abundance of the cognate exon-skipped alternative splicing variation (SEP3.3, lacking exon 6), which in turn drives floral homeotic phenotypes [138]. It was determined that the levels of SEP3.3 variant with exon-skipping was significantly upregulated by overexpression of the circRNA derived from exon 6. Therefore, we speculated that circRNAs from LSD1 locus may also regulate the levels of the LSD1+2a variant, acting as a regulator for alternative splicing of the parental gene. In my study, LSD1+2a was downregulated after the transfection of siRNA for circRNA 2a\_2 in A549 and PC9 cells. Meanwhile, siRNA for circRNA 3\_2\_2a and circRNA 3\_2 did not affect the levels of LSD1+2a. This data shows that circRNA 2a\_2 can regulate the alternative splicing of LSD1 to include more exon 2a into the transcripts. To validate the regulation by circRNAs, we established the transgenic expression approaches. Interestingly, the levels of LSD1+2a were upregulated after circRNA 3\_2\_2a and circRNA 3\_2 overexpression in PC9 cells. The levels of LSD1+2a were elevated after circRNA 3\_2\_2a and circRNA 2a\_2 overexpression in PSAE cells. Thus, we found that the levels of LSD1+2a were increased after the overexpression of circRNA 3\_2\_2a, circRNA 3\_2 and circRNA 2a\_2. Thereby, we found that circRNAs from LSD1 may modulate the alternative splicing of the parental gene.

Evidence from many studies shows that circRNAs modulate host gene expression [186, 187]. CircENO1 knockdown by siRNA resulted in the decrease of the protein and mRNA levels of ENO1. CircENO1 binds to miR-22-3p through a classic ceRNA

## Discussion

mechanism. Thus, in LUAD cells, circENO1/miR-22-3p/ENO1 modulated glycolysis, which in turn influenced proliferation, migration, and EMT [188]. Moreover, COL6A3-derived circRNAs with an ORF (open reading frame) are involved in encoding a unique 198-amino-acid (aa) functional peptide. By increasing COL6A3 mRNA stability, circCOL6A3 (hsa\_circ\_0006401) peptides reduced the host gene's expression at both the protein and mRNA levels. Nevertheless, circURI1 was also shown to have the potential to regulate alternative splicing of genes. By directly interfering with heterogeneous nuclear ribonucleoprotein M (hnRNPM) to control the alternative splicing of genes implicated in the migration of cells, circURI1 suppressed metastasis in gastric cancer [189]. Remarkably, circRNAs linked to R-loop were shown to modulate alternative splicing of their parent genes in stem-differentiating xylem (SDX) by a genome-wide profiling investigation. They confirmed that overexpression of circIRX7 increased the enrichment of the R-loops. Overexpression of circIRX7 was then hypothesized to reduce long transcript levels while simultaneously increasing short transcript levels [137]. Taken together, our study was the first one that detects the regulatory association between circRNAs from LSD1 gene and the regulation of the LSD1 2a variant.

### 4.4 CircRNAs from LSD1 form R-loops with its parental gene

CircRNAs primarily serve as miRNA sponges as part of their functional mechanism, in addition to regulating gene splicing or translation, transcription, and epigenetic regulation via interactions with proteins [190]. CircPRKCI functioned as a sponge for both miR-545 and miR-589, abrogating their suppression of the protumorigenic transcription factor E2F7 in LUAD [191]. Nevertheless, However, circRNA 3\_2\_2a was shown to act as a sponge for a total of five different miRNAs (miR-103a-2-5p, miR-875-3p, miR-92a-5p, miR-608 and miR-518c-5p) [166]. However, further experiments are required to confirm the possible mechanisms of binding to miRNAs. CircXPO1 might interact with IGF2BP1, which would increase the stability of CTNNB1 mRNA, and this would ultimately accelerate the development of LUAD [170]. Additionally, it has been validated that circRNAs may sense translation potential by binding to ribosomes. Also, endogenous circRNAs containing an internal ribosome entry site (IRES) that directly recruits ribosomes [192], can also be

## Discussion

translated [193]. The peptides that circRNAs code for are often shorter than their full-length protein equivalents (circFBXW7-185aa [194, 195]), but they performed similar activities. A new finding showed that a circRNA from SEPALLATA3 controls the splicing of its homologous mRNA by forming R-loops [138].

We speculated whether circRNAs from LSD1 bind to genomic DNA directly and form R-loops to regulate the expression of the host gene. Then we conducted immunoprecipitation (IP) with the antibody of R-loops S9.6 and purified RNA from IP and performed PCR with divergent primers to amplify circRNAs. Interestingly, we found that circRNA 3\_2\_2a and circRNA 2a\_2 bound to genomic DNA. The bound RNA was resistant to RNase H but sensitive to combined treatment of RNase H and RNase R, validating that circRNAs from LSD1 transcript bound to genomic DNA leading to R-loop formation. The functions of circRNA on alternative splicing were reported in plants. By forming R-loops, circSEPALLATA3 controls how its parent mRNA is spliced. SEP3 alternative splicing can be manipulated by circRNA exon 6 via the creation of a stable R-loop [138], which could physically delay transcription elongation and so favor the formation of the exon-skipped alternative splicing subtype [196, 197]. However, in our study, circRNAs from LSD1 were shown to increase the inclusion of exon 2a, thus reducing the exon-skipping. Meanwhile, a slower rate of elongation results in a greater level of exon inclusion. For instance, Godoy Herz and colleagues found that controlling transcriptional elongation is one of the mechanisms via which light affects AS in plants [198]. Here, in dark RNA polymerase II elongation is slower, resulting in exon inclusion [198, 199]. Although in general R-loop formation and a reduced RNA polymerase II kinetics leads to exon skipping, in fact both fast and slow RNA Polymerase II rates can affect both exon inclusion or skipping [197, 200]. Taken together, the findings support our hypothesis that circRNAs from LSD1 bind to the genomic DNA of LSD1 forming R-loops, that in turn lead to slow transcriptional elongation by RNA polymerase II and an increase of exon 2a inclusion.

Furthermore, to investigate the genomic DNA bound to circRNAs from LSD1, we designed the primers for exon 2, 2a, 3 and 5. RT-PCR was performed to amplify the purified DNA from IP and we can find that the genomic DNA to form R-loops with circRNAs from LSD1 are from exon 2a and exon 3. It has been validated that

## Discussion

complexes of ElciRNA–U1 snRNP may also bind to the Pol II transcriptional complex at host gene promoters, resulting in increased expression of the target genes [135]. In this study, one probable mechanism is that circRNAs bind to the genomic DNA of LSD1, which causes Pol II to pause at exon 2a and 3 and increases the levels of the alternative splicing variant LSD1+2a.

To explore the possible correlation between the levels of LSD1+8a and circRNA deriving from this locus, we also studied the expression of this LSD1 transcript isoform. The neuron-specific isoform LSD1+8a variant is crucial in the differentiation in terminal differentiation and maturation of neurons [79, 80, 201, 202]. Additionally, it has been identified in SCLC that LSD1+8a contributes to neural differentiation and chemoresistance [82]. I investigated the levels of LSD1+8a in NSCLC and SCLC. It was found that LSD1+8a was only expressed in some SCLC cell lines, confirming a tissue-specific expression [82]. In addition, a previous finding indicated the LSD1+8a variant is associated with neuroendocrine characteristics for SCLC [82]. To detect the cell viability to LSD1 inhibitors, MTT assay was applied to detect cell viability, but no strong correlation was confirmed between the levels of LSD1+8a and the cell viability to LSD1 inhibitors. Then I conducted siRNA transfection to examine the biological roles performed by LSD1+8a in SCLC. After the downregulation of LSD1+8a, cell viability of SCLC was inhibited. Jotatsu et al. also validated that suppression of LSD1+8a repressed the cell viability in SCLC cell lines, consistent with our finding [82]. Moreover, I also found that the levels of neuroendocrine markers ENO2 were dramatically downregulated by the knockdown of LSD1+8a. Meanwhile, a strong association was determined between the expression of LSD1+8a and neuroendocrine markers (CHGA, SYP, NCAM and ENO2). The difference between our results and their findings might be the cell lines were not totally same. To explore whether circRNAs can regulate the levels of the novel variant LSD1+8a, a novel circRNA with the junction of exon 8a and exon 2 from LSD1 was identified. CircRNA 8a\_2 was not identified in other studies due to the low expression in cells. Knockdown and overexpression of circRNA 8a\_2 upregulated and downregulated the levels of LSD1+8a respectively, revealing that circRNA 8a\_2 regulates the levels of the LSD1+8a variant. However, more subsequent experiments are needed to validate whether circRNA 8a\_2 binds to the genomic DNA of LSD1 acting like circRNA 2a\_2 and circRNA 3\_2\_2a. Meanwhile, whether

## Discussion

certain proteins are interacted with the R-loops composed of circRNA 8a\_2 and DNA is also worthy exploring.

### 6. References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018; 68: 394-424.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin.* 2021; 71: 7-33.
3. Frost N, Griesinger F, Hoffmann H, Langer F, Nestle U, Schutte W, et al. Lung Cancer in Germany. *J Thorac Oncol.* 2022; 17: 742-50.
4. Xia C, Dong X, Li H, Cao M, Sun D, He S, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J (Engl).* 2022; 135: 584-90.
5. Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med.* 2020; 41: 1-24.
6. Harris JE. Cigarette smoking among successive birth cohorts of men and women in the United States during 1900-80. *J Natl Cancer Inst.* 1983; 71: 473-9.
7. Mithoowani H, Febbraro M. Non-Small-Cell Lung Cancer in 2022: A Review for General Practitioners in Oncology. *Curr Oncol.* 2022; 29: 1828-39.
8. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc.* 2008; 83: 584-94.
9. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. *Adv Exp Med Biol.* 2016; 893: 1-19.
10. Kalemkerian GP, Akerley W, Bogner P, Borghaei H, Chow LQ, Downey RJ, et al. Small cell lung cancer. *J Natl Compr Canc Netw.* 2013; 11: 78-98.
11. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature.* 2018; 553: 446-54.
12. Kumar V. Pulmonary Innate Immune Response Determines the Outcome of Inflammation During Pneumonia and Sepsis-Associated Acute Lung Injury. *Front Immunol.* 2020; 11: 1722.
13. Weibel ER. What makes a good lung? *Swiss Med Wkly.* 2009; 139: 375-86.
14. Colby TV, Wistuba II, Gazdar A. Precursors to pulmonary neoplasia. *Adv Anat Pathol.* 1998; 5: 205-15.



## References

15. Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013; 143: e1S-e29S.
16. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. *Nat Rev Cancer*. 2007; 7: 778-90.
17. Subramanian J, Govindan R. Lung cancer in never smokers: a review. *J Clin Oncol*. 2007; 25: 561-70.
18. Zhu Y, Qu N, Sun L, Meng X, Li X, Zhang Y. Solitary pulmonary capillary hemangioma presents as ground glass opacity on computed tomography indicating adenocarcinoma in situ/atypical adenomatous hyperplasia: A case report. *Biomed Rep*. 2017; 7: 515-9.
19. van Meerbeeck JP, Fennell DA, De Ruyscher DK. Small-cell lung cancer. *Lancet*. 2011; 378: 1741-55.
20. Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nat Rev Cancer*. 2017; 17: 725-37.
21. Sinjab A, Rahal Z, Kadara H. Cell-by-Cell: Unlocking Lung Cancer Pathogenesis. *Cancers (Basel)*. 2022; 14.
22. Bareschino MA, Schettino C, Rossi A, Maione P, Sacco PC, Zeppa R, et al. Treatment of advanced non small cell lung cancer. *J Thorac Dis*. 2011; 3: 122-33.
23. Nicholson AG, Chansky K, Crowley J, Beyruti R, Kubota K, Turrisi A, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for the Revision of the Clinical and Pathologic Staging of Small Cell Lung Cancer in the Forthcoming Eighth Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol*. 2016; 11: 300-11.
24. Midha A, Dearden S, McCormack R. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). *Am J Cancer Res*. 2015; 5: 2892-911.
25. Pao W, Miller VA. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. *J Clin Oncol*. 2005; 23: 2556-68.
26. Timar J. The clinical relevance of KRAS gene mutation in non-small-cell lung cancer. *Curr Opin Oncol*. 2014; 26: 138-44.

## References

27. Wistuba, II, Gazdar AF, Minna JD. Molecular genetics of small cell lung carcinoma. *Semin Oncol.* 2001; 28: 3-13.
28. Nau MM, Brooks BJ, Battey J, Sausville E, Gazdar AF, Kirsch IR, et al. L-myc, a new myc-related gene amplified and expressed in human small cell lung cancer. *Nature.* 1985; 318: 69-73.
29. Lazaro S, Perez-Crespo M, Lorz C, Bernardini A, Oteo M, Enguita AB, et al. Differential development of large-cell neuroendocrine or small-cell lung carcinoma upon inactivation of 4 tumor suppressor genes. *Proc Natl Acad Sci U S A.* 2019; 116: 22300-6.
30. McFadden DG, Papagiannakopoulos T, Taylor-Weiner A, Stewart C, Carter SL, Cibulskis K, et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell.* 2014; 156: 1298-311.
31. Augert A, Eastwood E, Ibrahim AH, Wu N, Grunblatt E, Basom R, et al. Targeting NOTCH activation in small cell lung cancer through LSD1 inhibition. *Sci Signal.* 2019; 12.
32. Jia D, Augert A, Kim DW, Eastwood E, Wu N, Ibrahim AH, et al. Crebbp Loss Drives Small Cell Lung Cancer and Increases Sensitivity to HDAC Inhibition. *Cancer Discov.* 2018; 8: 1422-37.
33. Doll R, Hill AB. Smoking and carcinoma of the lung. Preliminary report. 1950. *Bull World Health Organ.* 1999; 77: 84-93.
34. Kim GH, Kim JM, Jee SH, Jung KJ. Associations of biomarkers for exposure to tobacco smoke with lung cancer risk in Korea. *Cancer Biomark.* 2022; 35: 409-17.
35. Cheng ES, Chan KH, Weber M, Steinberg J, Young J, Canfell K, et al. Solid Fuel, Secondhand Smoke, and Lung Cancer Mortality: A Prospective Cohort of 323,794 Chinese Never-Smokers. *Am J Respir Crit Care Med.* 2022; 206: 1153-62.
36. Hematian Larki M, Ashouri E, Barani S, Ghayumi SMA, Ghaderi A, Rajalingam R. KIR-HLA gene diversities and susceptibility to lung cancer. *Sci Rep.* 2022; 12: 17237.
37. Jiang M, Li X, Quan X, Li X, Zhou B. Single Nucleotide Polymorphisms in HMGB1 Correlate with Lung Cancer Risk in the Northeast Chinese Han Population. *Molecules.* 2018; 23.
38. Harbour JW, Lai SL, Whang-Peng J, Gazdar AF, Minna JD, Kaye FJ. Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. *Science.* 1988; 241: 353-7.

## References

39. Little CD, Nau MM, Carney DN, Gazdar AF, Minna JD. Amplification and expression of the c-myc oncogene in human lung cancer cell lines. *Nature*. 1983; 306: 194-6.
40. Li L, Shao M, He X, Ren S, Tian T. Risk of lung cancer due to external environmental factor and epidemiological data analysis. *Math Biosci Eng*. 2021; 18: 6079-94.
41. Markowitz SB. Lung Cancer Screening in Asbestos-Exposed Populations. *Int J Environ Res Public Health*. 2022; 19.
42. Soza-Ried C, Bustamante E, Caglevic C, Rolfo C, Sirera R, Marsiglia H. Oncogenic role of arsenic exposure in lung cancer: A forgotten risk factor. *Crit Rev Oncol Hematol*. 2019; 139: 128-33.
43. Shahadin MS, Ab Mutalib NS, Latif MT, Greene CM, Hassan T. Challenges and future direction of molecular research in air pollution-related lung cancers. *Lung Cancer*. 2018; 118: 69-75.
44. Sundahl N, Lievens Y. Radiotherapy for oligometastatic non-small cell lung cancer: a narrative review. *Transl Lung Cancer Res*. 2021; 10: 3420-31.
45. Li H, Shen Y, Wu Y, Cai S, Zhu Y, Chen S, et al. Stereotactic Body Radiotherapy Versus Surgery for Early-Stage Non-Small-Cell Lung Cancer. *J Surg Res*. 2019; 243: 346-53.
46. Qiu B, Guo W, Zhang F, Lv F, Ji Y, Peng Y, et al. Dynamic recurrence risk and adjuvant chemotherapy benefit prediction by ctDNA in resected NSCLC. *Nat Commun*. 2021; 12: 6770.
47. Reck M, Remon J, Hellmann MD. First-Line Immunotherapy for Non-Small-Cell Lung Cancer. *J Clin Oncol*. 2022; 40: 586-97.
48. Slotman BJ, van Tinteren H, Praag JO, Knegjens JL, El Sharouni SY, Hatton M, et al. Use of thoracic radiotherapy for extensive stage small-cell lung cancer: a phase 3 randomised controlled trial. *Lancet*. 2015; 385: 36-42.
49. Mohammad HP, Smitheman KN, Kamat CD, Soong D, Federowicz KE, Van Aller GS, et al. A DNA Hypomethylation Signature Predicts Antitumor Activity of LSD1 Inhibitors in SCLC. *Cancer Cell*. 2015; 28: 57-69.
50. Hollebecque A, Salvagni S, Plummer R, Niccoli P, Capdevila J, Curigliano G, et al. Clinical activity of CC-90011, an oral, potent, and reversible LSD1 inhibitor, in advanced malignancies. *Cancer*. 2022; 128: 3185-95.

## References

51. Duruisseaux M, Esteller M. Lung cancer epigenetics: From knowledge to applications. *Semin Cancer Biol.* 2018; 51: 116-28.
52. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell.* 2004; 119: 941-53.
53. Chen Y, Yang Y, Wang F, Wan K, Yamane K, Zhang Y, et al. Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc Natl Acad Sci U S A.* 2006; 103: 13956-61.
54. Amente S, Lania L, Majello B. The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochim Biophys Acta.* 2013; 1829: 981-6.
55. Zhang S, Liu M, Yao Y, Yu B, Liu H. Targeting LSD1 for acute myeloid leukemia (AML) treatment. *Pharmacol Res.* 2021; 164: 105335.
56. Martinez-Gamero C, Malla S, Aguilo F. LSD1: Expanding Functions in Stem Cells and Differentiation. *Cells.* 2021; 10.
57. Gu F, Lin Y, Wang Z, Wu X, Ye Z, Wang Y, et al. Biological roles of LSD1 beyond its demethylase activity. *Cell Mol Life Sci.* 2020; 77: 3341-50.
58. Perillo B, Tramontano A, Pezone A, Migliaccio A. LSD1: more than demethylation of histone lysine residues. *Exp Mol Med.* 2020; 52: 1936-47.
59. Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, et al. Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature.* 2007; 446: 882-7.
60. Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, et al. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature.* 2012; 482: 221-5.
61. Adamo A, Sese B, Boue S, Castano J, Paramonov I, Barrero MJ, et al. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat Cell Biol.* 2011; 13: 652-9.
62. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* 2014; 15: 178-96.
63. Boulding T, McCuaig RD, Tan A, Hardy K, Wu F, Dunn J, et al. LSD1 activation promotes inducible EMT programs and modulates the tumour microenvironment in breast cancer. *Sci Rep.* 2018; 8: 73.

## References

64. Ferrari-Amorotti G, Fragliasso V, Esteki R, Prudente Z, Soliera AR, Cattelani S, et al. Inhibiting interactions of lysine demethylase LSD1 with snail/slug blocks cancer cell invasion. *Cancer Res.* 2013; 73: 235-45.
65. Garcia-Bassets I, Kwon YS, Telese F, Prefontaine GG, Hutt KR, Cheng CS, et al. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell.* 2007; 128: 505-18.
66. Nair SS, Nair BC, Cortez V, Chakravarty D, Metzger E, Schule R, et al. PELP1 is a reader of histone H3 methylation that facilitates oestrogen receptor-alpha target gene activation by regulating lysine demethylase 1 specificity. *EMBO Rep.* 2010; 11: 438-44.
67. Metzger E, Wissmann M, Yin N, Muller JM, Schneider R, Peters AH, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature.* 2005; 437: 436-9.
68. Xu S, Wang X, Yang Y, Li Y, Wu S. LSD1 silencing contributes to enhanced efficacy of anti-CD47/PD-L1 immunotherapy in cervical cancer. *Cell Death Dis.* 2021; 12: 282.
69. Xie Q, Tang T, Pang J, Xu J, Yang X, Wang L, et al. LSD1 Promotes Bladder Cancer Progression by Upregulating LEF1 and Enhancing EMT. *Front Oncol.* 2020; 10: 1234.
70. Lim SY, Macheleidt I, Dalvi P, Schafer SC, Kerick M, Ozretic L, et al. LSD1 modulates the non-canonical integrin beta3 signaling pathway in non-small cell lung carcinoma cells. *Sci Rep.* 2017; 7: 10292.
71. Macheleidt IF, Dalvi PS, Lim SY, Meemboor S, Meder L, Kasgen O, et al. Preclinical studies reveal that LSD1 inhibition results in tumor growth arrest in lung adenocarcinoma independently of driver mutations. *Mol Oncol.* 2018; 12: 1965-79.
72. Majello B, Gorini F, Sacca CD, Amente S. Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers (Basel).* 2019; 11.
73. Schulte JH, Lim S, Schramm A, Friedrichs N, Koster J, Versteeg R, et al. Lysine-specific demethylase 1 is strongly expressed in poorly differentiated neuroblastoma: implications for therapy. *Cancer Res.* 2009; 69: 2065-71.
74. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, et al. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell.* 2011; 19: 138-52.

## References

75. Harris WJ, Huang X, Lynch JT, Spencer GJ, Hitchin JR, Li Y, et al. The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell*. 2012; 21: 473-87.
76. Maes T, Mascaro C, Tirapu I, Estiarte A, Ciceri F, Lunardi S, et al. ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute Leukemia. *Cancer Cell*. 2018; 33: 495-511 e12.
77. Yatim A, Benne C, Sobhian B, Laurent-Chabalier S, Deas O, Judde JG, et al. NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. *Mol Cell*. 2012; 48: 445-58.
78. Rusconi F, Paganini L, Braida D, Ponzoni L, Toffolo E, Maroli A, et al. LSD1 Neurospecific Alternative Splicing Controls Neuronal Excitability in Mouse Models of Epilepsy. *Cereb Cortex*. 2015; 25: 2729-40.
79. Laurent B, Ruitu L, Murn J, Hempel K, Ferrao R, Xiang Y, et al. A specific LSD1/KDM1A isoform regulates neuronal differentiation through H3K9 demethylation. *Mol Cell*. 2015; 57: 957-70.
80. Hwang I, Cao D, Na Y, Kim DY, Zhang T, Yao J, et al. Far Upstream Element-Binding Protein 1 Regulates LSD1 Alternative Splicing to Promote Terminal Differentiation of Neural Progenitors. *Stem Cell Reports*. 2018; 10: 1208-21.
81. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. *Cancer Cell*. 2020; 38: 60-78 e12.
82. Jotatsu T, Yagishita S, Tajima K, Takahashi F, Mogushi K, Hidayat M, et al. LSD1/KDM1 isoform LSD1+8a contributes to neural differentiation in small cell lung cancer. *Biochem Biophys Rep*. 2017; 9: 86-94.
83. Zhang W, Ruan X, Li Y, Zhi J, Hu L, Hou X, et al. KDM1A promotes thyroid cancer progression and maintains stemness through the Wnt/beta-catenin signaling pathway. *Theranostics*. 2022; 12: 1500-17.
84. Fang Y, Liao G, Yu B. LSD1/KDM1A inhibitors in clinical trials: advances and prospects. *J Hematol Oncol*. 2019; 12: 129.
85. Yang GJ, Lei PM, Wong SY, Ma DL, Leung CH. Pharmacological Inhibition of LSD1 for Cancer Treatment. *Molecules*. 2018; 23.
86. Takagi S, Ishikawa Y, Mizutani A, Iwasaki S, Matsumoto S, Kamada Y, et al. LSD1 Inhibitor T-3775440 Inhibits SCLC Cell Proliferation by Disrupting LSD1

## References

- Interactions with SNAG Domain Proteins INSM1 and GFI1B. *Cancer Res.* 2017; 77: 4652-62.
87. Fiskus W, Sharma S, Shah B, Portier BP, Devaraj SG, Liu K, et al. Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia.* 2014; 28: 2155-64.
88. von Tresckow B, Sayehli C, Aulitzky WE, Goebeler ME, Schwab M, Braz E, et al. Phase I study of domatinostat (4SC-202), a class I histone deacetylase inhibitor in patients with advanced hematological malignancies. *Eur J Haematol.* 2019; 102: 163-73.
89. Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci U S A.* 1976; 73: 3852-6.
90. Hsu MT, Coca-Prados M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature.* 1979; 280: 339-40.
91. Cocquerelle C, Mascrez B, Hetuin D, Bailleul B. Mis-splicing yields circular RNA molecules. *FASEB J.* 1993; 7: 155-60.
92. Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell.* 1993; 73: 1019-30.
93. Ye CY, Chen L, Liu C, Zhu QH, Fan L. Widespread noncoding circular RNAs in plants. *New Phytol.* 2015; 208: 88-95.
94. Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, et al. Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* 2014; 9: 1966-80.
95. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* 2015; 10: 170-7.
96. Dube U, Del-Aguila JL, Li Z, Budde JP, Jiang S, Hsu S, et al. An atlas of cortical circular RNA expression in Alzheimer disease brains demonstrates clinical and pathological associations. *Nat Neurosci.* 2019; 22: 1903-12.
97. Shen S, Wu Y, Chen J, Xie Z, Huang K, Wang G, et al. CircSERPINE2 protects against osteoarthritis by targeting miR-1271 and ETS-related gene. *Ann Rheum Dis.* 2019; 78: 826-36.

## References

98. Liu C, Ge HM, Liu BH, Dong R, Shan K, Chen X, et al. Targeting pericyte-endothelial cell crosstalk by circular RNA-cPWWP2A inhibition aggravates diabetes-induced microvascular dysfunction. *Proc Natl Acad Sci U S A*. 2019; 116: 7455-64.
99. Gupta SK, Garg A, Bar C, Chatterjee S, Foinquinos A, Milting H, et al. Quaking Inhibits Doxorubicin-Mediated Cardiotoxicity Through Regulation of Cardiac Circular RNA Expression. *Circ Res*. 2018; 122: 246-54.
100. Liu J, Song S, Lin S, Zhang M, Du Y, Zhang D, et al. Circ-SERPINE2 promotes the development of gastric carcinoma by sponging miR-375 and modulating YWHAZ. *Cell Prolif*. 2019; 52: e12648.
101. Fu XD, Ares M, Jr. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat Rev Genet*. 2014; 15: 689-701.
102. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*. 2013; 19: 141-57.
103. Zheng X, Chen L, Zhou Y, Wang Q, Zheng Z, Xu B, et al. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Mol Cancer*. 2019; 18: 47.
104. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, et al. The RNA binding protein quaking regulates formation of circRNAs. *Cell*. 2015; 160: 1125-34.
105. Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet*. 2019; 20: 675-91.
106. Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung LH, et al. Exon circularization requires canonical splice signals. *Cell Rep*. 2015; 10: 103-11.
107. Liang D, Tatomer DC, Luo Z, Wu H, Yang L, Chen LL, et al. The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting. *Mol Cell*. 2017; 68: 940-54 e3.
108. Di Timoteo G, Dattilo D, Centron-Broco A, Colantoni A, Guarnacci M, Rossi F, et al. Modulation of circRNA Metabolism by m(6)A Modification. *Cell Rep*. 2020; 31: 107641.
109. Tang C, Xie Y, Yu T, Liu N, Wang Z, Woolsey RJ, et al. m(6)A-dependent biogenesis of circular RNAs in male germ cells. *Cell Res*. 2020; 30: 211-28.



## References

110. Yang Q, Li F, He AT, Yang BB. Circular RNAs: Expression, localization, and therapeutic potentials. *Mol Ther*. 2021; 29: 1683-702.
111. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013; 495: 333-8.
112. Bentley DL. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet*. 2014; 15: 163-75.
113. Huang C, Liang D, Tatomer DC, Wilusz JE. A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs. *Genes Dev*. 2018; 32: 639-44.
114. Chen RX, Chen X, Xia LP, Zhang JX, Pan ZZ, Ma XD, et al. N(6)-methyladenosine modification of circNSUN2 facilitates cytoplasmic export and stabilizes HMGA2 to promote colorectal liver metastasis. *Nat Commun*. 2019; 10: 4695.
115. Panda AC. Circular RNAs Act as miRNA Sponges. *Adv Exp Med Biol*. 2018; 1087: 67-79.
116. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet*. 2016; 17: 272-83.
117. Lu TX, Rothenberg ME. MicroRNA. *J Allergy Clin Immunol*. 2018; 141: 1202-7.
118. Pan H, Li T, Jiang Y, Pan C, Ding Y, Huang Z, et al. Overexpression of Circular RNA ciRS-7 Abrogates the Tumor Suppressive Effect of miR-7 on Gastric Cancer via PTEN/PI3K/AKT Signaling Pathway. *J Cell Biochem*. 2018; 119: 440-6.
119. Liu J, Li H, Wei C, Ding J, Lu J, Pan G, et al. circFAT1(e2) Promotes Papillary Thyroid Cancer Proliferation, Migration, and Invasion via the miRNA-873/ZEB1 Axis. *Comput Math Methods Med*. 2020; 2020: 1459368.
120. Wang L, Tong X, Zhou Z, Wang S, Lei Z, Zhang T, et al. Circular RNA hsa\_circ\_0008305 (circPTK2) inhibits TGF-beta-induced epithelial-mesenchymal transition and metastasis by controlling TIF1gamma in non-small cell lung cancer. *Mol Cancer*. 2018; 17: 140.
121. Zang J, Lu D, Xu A. The interaction of circRNAs and RNA binding proteins: An important part of circRNA maintenance and function. *J Neurosci Res*. 2020; 98: 87-97.
122. Conlon EG, Manley JL. RNA-binding proteins in neurodegeneration: mechanisms in aggregate. *Genes Dev*. 2017; 31: 1509-28.

## References

123. Yang Q, Du WW, Wu N, Yang W, Awan FM, Fang L, et al. A circular RNA promotes tumorigenesis by inducing c-myc nuclear translocation. *Cell Death Differ.* 2017; 24: 1609-20.
124. Yang ZG, Awan FM, Du WW, Zeng Y, Lyu J, Wu, et al. The Circular RNA Interacts with STAT3, Increasing Its Nuclear Translocation and Wound Repair by Modulating Dnmt3a and miR-17 Function. *Mol Ther.* 2017; 25: 2062-74.
125. Bi W, Huang J, Nie C, Liu B, He G, Han J, et al. CircRNA circRNA\_102171 promotes papillary thyroid cancer progression through modulating CTNNBIP1-dependent activation of beta-catenin pathway. *J Exp Clin Cancer Res.* 2018; 37: 275.
126. Du WW, Yang W, Chen Y, Wu ZK, Foster FS, Yang Z, et al. Foxo3 circular RNA promotes cardiac senescence by modulating multiple factors associated with stress and senescence responses. *Eur Heart J.* 2017; 38: 1402-12.
127. Garikipati VNS, Verma SK, Cheng Z, Liang D, Truongcao MM, Cimini M, et al. Circular RNA CircFndc3b modulates cardiac repair after myocardial infarction via FUS/VEGF-A axis. *Nat Commun.* 2019; 10: 4317.
128. AbouHaidar MG, Venkataraman S, Golshani A, Liu B, Ahmad T. Novel coding, translation, and gene expression of a replicating covalently closed circular RNA of 220 nt. *Proc Natl Acad Sci U S A.* 2014; 111: 14542-7.
129. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, et al. Translation of CircRNAs. *Mol Cell.* 2017; 66: 9-21 e7.
130. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, et al. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell.* 2017; 66: 22-37 e9.
131. Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, et al. Extensive translation of circular RNAs driven by N(6)-methyladenosine. *Cell Res.* 2017; 27: 626-41.
132. Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol Cancer.* 2020; 19: 22.
133. Qin M, Liu G, Huo X, Tao X, Sun X, Ge Z, et al. Hsa\_circ\_0001649: A circular RNA and potential novel biomarker for hepatocellular carcinoma. *Cancer Biomark.* 2016; 16: 161-9.
134. Zhang M, Huang N, Yang X, Luo J, Yan S, Xiao F, et al. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene.* 2018; 37: 1805-14.

## References

135. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol.* 2015; 22: 256-64.
136. Lu T, Cui L, Zhou Y, Zhu C, Fan D, Gong H, et al. Transcriptome-wide investigation of circular RNAs in rice. *RNA.* 2015; 21: 2076-87.
137. Liu X, Gao Y, Liao J, Miao M, Chen K, Xi F, et al. Genome-wide profiling of circular RNAs, alternative splicing, and R-loops in stem-differentiating xylem of *Populus trichocarpa*. *J Integr Plant Biol.* 2021; 63: 1294-308.
138. Conn VM, Hugouvieux V, Nayak A, Conos SA, Capovilla G, Cildir G, et al. A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. *Nat Plants.* 2017; 3: 17053.
139. Zhang HD, Jiang LH, Sun DW, Hou JC, Ji ZL. CircRNA: a novel type of biomarker for cancer. *Breast Cancer.* 2018; 25: 1-7.
140. Lei B, Tian Z, Fan W, Ni B. Circular RNA: a novel biomarker and therapeutic target for human cancers. *Int J Med Sci.* 2019; 16: 292-301.
141. Maass PG, Glazar P, Memczak S, Dittmar G, Hollfinger I, Schreyer L, et al. A map of human circular RNAs in clinically relevant tissues. *J Mol Med (Berl).* 2017; 95: 1179-89.
142. Chen X, Chen RX, Wei WS, Li YH, Feng ZH, Tan L, et al. PRMT5 Circular RNA Promotes Metastasis of Urothelial Carcinoma of the Bladder through Sponging miR-30c to Induce Epithelial-Mesenchymal Transition. *Clin Cancer Res.* 2018; 24: 6319-30.
143. Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, et al. The Landscape of Circular RNA in Cancer. *Cell.* 2019; 176: 869-81 e13.
144. He J, Ren M, Li H, Yang L, Wang X, Yang Q. Exosomal Circular RNA as a Biomarker Platform for the Early Diagnosis of Immune-Mediated Demyelinating Disease. *Front Genet.* 2019; 10: 860.
145. Bahn JH, Zhang Q, Li F, Chan TM, Lin X, Kim Y, et al. The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. *Clin Chem.* 2015; 61: 221-30.
146. Chen L, Shan G. CircRNA in cancer: Fundamental mechanism and clinical potential. *Cancer Lett.* 2021; 505: 49-57.
147. Lei B, Zhou J, Xuan X, Tian Z, Zhang M, Gao W, et al. Circular RNA expression profiles of peripheral blood mononuclear cells in hepatocellular carcinoma patients by sequence analysis. *Cancer Med.* 2019; 8: 1423-33.

## References

148. Yang C, Dong Z, Hong H, Dai B, Song F, Geng L, et al. circFN1 Mediates Sorafenib Resistance of Hepatocellular Carcinoma Cells by Sponging miR-1205 and Regulating E2F1 Expression. *Mol Ther Nucleic Acids*. 2020; 22: 421-33.
149. Li Z, Zhou Y, Yang G, He S, Qiu X, Zhang L, et al. Using circular RNA SMARCA5 as a potential novel biomarker for hepatocellular carcinoma. *Clin Chim Acta*. 2019; 492: 37-44.
150. Jiang Z, Shen L, Wang S, Wu S, Hu Y, Guo J, et al. Hsa\_circ\_0028502 and hsa\_circ\_0076251 are potential novel biomarkers for hepatocellular carcinoma. *Cancer Med*. 2019; 8: 7278-87.
151. Lu J, Zhang PY, Xie JW, Wang JB, Lin JX, Chen QY, et al. Hsa\_circ\_0000467 promotes cancer progression and serves as a diagnostic and prognostic biomarker for gastric cancer. *J Clin Lab Anal*. 2019; 33: e22726.
152. Ye DX, Wang SS, Huang Y, Chi P. A 3-circular RNA signature as a noninvasive biomarker for diagnosis of colorectal cancer. *Cancer Cell Int*. 2019; 19: 276.
153. Tian J, Xi X, Wang J, Yu J, Huang Q, Ma R, et al. CircRNA hsa\_circ\_0004585 as a potential biomarker for colorectal cancer. *Cancer Manag Res*. 2019; 11: 5413-23.
154. Chen L, Nan A, Zhang N, Jia Y, Li X, Ling Y, et al. Circular RNA 100146 functions as an oncogene through direct binding to miR-361-3p and miR-615-5p in non-small cell lung cancer. *Mol Cancer*. 2019; 18: 13.
155. Wei S, Zheng Y, Jiang Y, Li X, Geng J, Shen Y, et al. The circRNA circPTPRA suppresses epithelial-mesenchymal transitioning and metastasis of NSCLC cells by sponging miR-96-5p. *EBioMedicine*. 2019; 44: 182-93.
156. Liu X, Abraham JM, Cheng Y, Wang Z, Wang Z, Zhang G, et al. Synthetic Circular RNA Functions as a miR-21 Sponge to Suppress Gastric Carcinoma Cell Proliferation. *Mol Ther Nucleic Acids*. 2018; 13: 312-21.
157. Gyawali B, Eisenhauer E, Tregear M, Booth CM. Progression-free survival: it is time for a new name. *Lancet Oncol*. 2022; 23: 328-30.
158. Barrett SP, Wang PL, Salzman J. Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife*. 2015; 4: e07540.
159. Hosseini A, Minucci S. A comprehensive review of lysine-specific demethylase 1 and its roles in cancer. *Epigenomics*. 2017; 9: 1123-42.

## References

160. Metzger E, Willmann D, McMillan J, Forne I, Metzger P, Gerhardt S, et al. Assembly of methylated KDM1A and CHD1 drives androgen receptor-dependent transcription and translocation. *Nat Struct Mol Biol.* 2016; 23: 132-9.
161. Lv T, Yuan D, Miao X, Lv Y, Zhan P, Shen X, et al. Over-expression of LSD1 promotes proliferation, migration and invasion in non-small cell lung cancer. *PLoS One.* 2012; 7: e35065.
162. Tamaoki J, Takeuchi M, Abe R, Kaneko H, Wada T, Hino S, et al. Splicing- and demethylase-independent functions of LSD1 in zebrafish primitive hematopoiesis. *Sci Rep.* 2020; 10: 8521.
163. Liu YD, Dai M, Yang SS, Xiao M, Meng FL, Chen XW. Overexpression of Lysine-Specific Demethylase 1 Is Associated With Tumor Progression and Unfavorable Prognosis in Chinese Patients With Endometrioid Endometrial Adenocarcinoma. *Int J Gynecol Cancer.* 2015; 25: 1453-60.
164. Nagasawa S, Sedukhina AS, Nakagawa Y, Maeda I, Kubota M, Ohnuma S, et al. LSD1 overexpression is associated with poor prognosis in basal-like breast cancer, and sensitivity to PARP inhibition. *PLoS One.* 2015; 10: e0118002.
165. Wada T, Koyama D, Kikuchi J, Honda H, Furukawa Y. Overexpression of the shortest isoform of histone demethylase LSD1 primes hematopoietic stem cells for malignant transformation. *Blood.* 2015; 125: 3731-46.
166. Song Z, Zhuo Z, Ma Z, Hou C, Chen G, Xu G. Hsa\_Circ\_0001206 is downregulated and inhibits cell proliferation, migration and invasion in prostate cancer. *Artif Cells Nanomed Biotechnol.* 2019; 47: 2449-64.
167. Zhang LX, Gao J, Long X, Zhang PF, Yang X, Zhu SQ, et al. The circular RNA circHMGB2 drives immunosuppression and anti-PD-1 resistance in lung adenocarcinomas and squamous cell carcinomas via the miR-181a-5p/CARM1 axis. *Mol Cancer.* 2022; 21: 110.
168. Wang J, Zhao X, Wang Y, Ren F, Sun D, Yan Y, et al. circRNA-002178 act as a ceRNA to promote PDL1/PD1 expression in lung adenocarcinoma. *Cell Death Dis.* 2020; 11: 32.
169. Peng H, Zhang W, Dong H, Yuan J, Li Y, Li F, et al. CircFAT1 Promotes Lung Adenocarcinoma Progression by Sequestering miR-7 from Repressing IRS2-ERK-mediated CCND1 Expression. *Int J Biol Sci.* 2022; 18: 3944-60.

## References

170. Huang Q, Guo H, Wang S, Ma Y, Chen H, Li H, et al. A novel circular RNA, circXPO1, promotes lung adenocarcinoma progression by interacting with IGF2BP1. *Cell Death Dis.* 2020; 11: 1031.
171. Zhu MC, Zhang YH, Xiong P, Fan XW, Li GL, Zhu M. Circ-GSK3B up-regulates GSK3B to suppress the progression of lung adenocarcinoma. *Cancer Gene Ther.* 2022.
172. Liang Y, Wang H, Chen B, Mao Q, Xia W, Zhang T, et al. circDCUN1D4 suppresses tumor metastasis and glycolysis in lung adenocarcinoma by stabilizing TXNIP expression. *Mol Ther Nucleic Acids.* 2021; 23: 355-68.
173. Wang Y, Ren F, Sun D, Liu J, Liu B, He Y, et al. CircKEAP1 Suppresses the Progression of Lung Adenocarcinoma via the miR-141-3p/KEAP1/NRF2 Axis. *Front Oncol.* 2021; 11: 672586.
174. Dong Y, Qiu T, Xuan Y, Liu A, Sun X, Huang Z, et al. circFBXW7 attenuates malignant progression in lung adenocarcinoma by sponging miR-942-5p. *Transl Lung Cancer Res.* 2021; 10: 1457-73.
175. He F, Zhong X, Lin Z, Lin J, Qiu M, Li X, et al. Plasma exo-hsa\_circRNA\_0056616: A potential biomarker for lymph node metastasis in lung adenocarcinoma. *J Cancer.* 2020; 11: 4037-46.
176. Lu GJ, Cui J, Qian Q, Hou ZB, Xie HY, Hu W, et al. Overexpression of hsa\_circ\_0001715 is a Potential Diagnostic and Prognostic Biomarker in Lung Adenocarcinoma. *Onco Targets Ther.* 2020; 13: 10775-83.
177. Huang Y, Dai Y, Wen C, He S, Shi J, Zhao D, et al. circSETD3 Contributes to Acquired Resistance to Gefitinib in Non-Small-Cell Lung Cancer by Targeting the miR-520h/ABCG2 Pathway. *Mol Ther Nucleic Acids.* 2020; 21: 885-99.
178. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol.* 2014; 32: 453-61.
179. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature.* 2011; 479: 74-9.
180. Xu L, Ma Y, Zhang H, Lu QJ, Yang L, Jiang GN, et al. HMGA2 regulates circular RNA ASPH to promote tumor growth in lung adenocarcinoma. *Cell Death Dis.* 2020; 11: 593.

## References

181. Yu W, Peng W, Sha H, Li J. Hsa\_circ\_0003998 Promotes Chemoresistance via Modulation of miR-326 in Lung Adenocarcinoma Cells. *Oncol Res.* 2019; 27: 623-8.
182. Li B, Zhu L, Lu C, Wang C, Wang H, Jin H, et al. circNDUFB2 inhibits non-small cell lung cancer progression via destabilizing IGF2BPs and activating anti-tumor immunity. *Nat Commun.* 2021; 12: 295.
183. Pan Z, Zhao R, Li B, Qi Y, Qiu W, Guo Q, et al. EWSR1-induced circNEIL3 promotes glioma progression and exosome-mediated macrophage immunosuppressive polarization via stabilizing IGF2BP3. *Mol Cancer.* 2022; 21: 16.
184. Zheng L, Liang H, Zhang Q, Shen Z, Sun Y, Zhao X, et al. circPTEN1, a circular RNA generated from PTEN, suppresses cancer progression through inhibition of TGF-beta/Smad signaling. *Mol Cancer.* 2022; 21: 41.
185. Fan HN, Chen ZY, Chen XY, Chen M, Yi YC, Zhu JS, et al. METTL14-mediated m(6)A modification of circORC5 suppresses gastric cancer progression by regulating miR-30c-2-3p/AKT1S1 axis. *Mol Cancer.* 2022; 21: 51.
186. Liu D, Kang H, Gao M, Jin L, Zhang F, Chen D, et al. Exosome-transmitted circ\_MMP2 promotes hepatocellular carcinoma metastasis by upregulating MMP2. *Mol Oncol.* 2020; 14: 1365-80.
187. Xu X, Zhang J, Tian Y, Gao Y, Dong X, Chen W, et al. CircRNA inhibits DNA damage repair by interacting with host gene. *Mol Cancer.* 2020; 19: 128.
188. Zhou J, Zhang S, Chen Z, He Z, Xu Y, Li Z. CircRNA-ENO1 promoted glycolysis and tumor progression in lung adenocarcinoma through upregulating its host gene ENO1. *Cell Death Dis.* 2019; 10: 885.
189. Wang X, Li J, Bian X, Wu C, Hua J, Chang S, et al. CircURI1 interacts with hnRNPM to inhibit metastasis by modulating alternative splicing in gastric cancer. *Proc Natl Acad Sci U S A.* 2021; 118.
190. Sharma AR, Bhattacharya M, Bhakta S, Saha A, Lee SS, Chakraborty C. Recent research progress on circular RNAs: Biogenesis, properties, functions, and therapeutic potential. *Mol Ther Nucleic Acids.* 2021; 25: 355-71.
191. Qiu M, Xia W, Chen R, Wang S, Xu Y, Ma Z, et al. The Circular RNA circPRKCI Promotes Tumor Growth in Lung Adenocarcinoma. *Cancer Res.* 2018; 78: 2839-51.
192. Macejak DG, Sarnow P. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature.* 1991; 353: 90-4.

## References

193. Zhang L, Hou C, Chen C, Guo Y, Yuan W, Yin D, et al. The role of N(6)-methyladenosine (m(6)A) modification in the regulation of circRNAs. *Mol Cancer*. 2020; 19: 105.
194. Ye F, Gao G, Zou Y, Zheng S, Zhang L, Ou X, et al. circFBXW7 Inhibits Malignant Progression by Sponging miR-197-3p and Encoding a 185-aa Protein in Triple-Negative Breast Cancer. *Mol Ther Nucleic Acids*. 2019; 18: 88-98.
195. Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J*. 2011; 30: 4414-22.
196. Wongsurawat T, Jenjaroenpun P, Kwoh CK, Kuznetsov V. Quantitative model of R-loop forming structures reveals a novel level of RNA-DNA interactome complexity. *Nucleic Acids Res*. 2012; 40: e16.
197. Dujardin G, Lafaille C, de la Mata M, Marasco LE, Munoz MJ, Le Jossic-Corcus C, et al. How slow RNA polymerase II elongation favors alternative exon skipping. *Mol Cell*. 2014; 54: 683-90.
198. Godoy Herz MA, Kubaczka MG, Brzyzek G, Servi L, Krzyszton M, Simpson C, et al. Light Regulates Plant Alternative Splicing through the Control of Transcriptional Elongation. *Mol Cell*. 2019; 73: 1066-74 e3.
199. Petrillo E, Godoy Herz MA, Fuchs A, Reifer D, Fuller J, Yanovsky MJ, et al. A chloroplast retrograde signal regulates nuclear alternative splicing. *Science*. 2014; 344: 427-30.
200. Fong N, Kim H, Zhou Y, Ji X, Qiu J, Saldi T, et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev*. 2014; 28: 2663-76.
201. Wang J, Telese F, Tan Y, Li W, Jin C, He X, et al. LSD1n is an H4K20 demethylase regulating memory formation via transcriptional elongation control. *Nat Neurosci*. 2015; 18: 1256-64.
202. Coleman DJ, Sampson DA, Sehwat A, Kumaraswamy A, Sun D, Wang Y, et al. Alternative splicing of LSD1+8a in neuroendocrine prostate cancer is mediated by SRRM4. *Neoplasia*. 2020; 22: 253-62.



## Abbreviations

### 7. Abbreviations

LUAD	lung adenocarcinoma
SCLC	small cell lung cancer
NSCLC	non-small cell lung cancer
LSD1	lysine-specific demethylase 1
circRNA	circular RNA
FISH	Fluorescence in situ hybridization
TNM	Tumor-Node-Metastasis
EGFR	epidermal growth factor receptor
KRAS	Kirsten rat sarcoma viral oncogene homologue
PTEN	phosphatase and tensin homolog
NOTCH	neurogenic locus notch homolog protein
CREBBP	CREB binding protein
SNP	single nucleotide polymorphisms
SABR	stereotactic ablative radiotherapy
PD-1	programmed death 1
PD-L1	programmed death-ligand 1
AO	amine oxidase
ESC	Embryonic stem cell
FOXA2	Forkhead Box A2
BMP2	Bone morphogenetic protein
EMT	epithelial–mesenchymal transition
ER	estrogen receptors
PELP1	proline. glutamate and leucine rich protein 1
AR	androgen receptor
AML	acute myeloid leukemia
H3K4me2	histone H3 lysine 4 demethylation
CDDP	cisplatin
TCP	tranlycypromine
HDAC	histone deacetylase
IDO1	indoleamine 2,3-dioxygenase-1
NAE	NEDD8-activating enzyme

## Abbreviations

UTR	untranslated region
tricRNA	tRNA intronic circRNA
exon-intron RNA	EIciRNA
circular intronic RNA	ciRNA
exonic circRNA	ecircRNA
RBP	RNA-binding protein
M6A	N6-methyladenosine
YTHDC1	YTH domain-containing 1
METTL3	methyltransferase-like 3
cdR1	cerebellar degeneration-related protein 1
HCC	hepatocellular carcinoma
PTc	papillary thyroid carcinoma
TCF	T cell factor
IRES	internal ribosome entry site
CSF	cerebrospinal fluid
AUC	area under the curve
HPRT	Hypoxanthine Phosphoribosyltransferase
siRNA	Small interfering RNA
EDTA	Ethylenediaminetetraacetic acid
KCl	Potassium chloride
MgCl <sub>2</sub>	Magnesium chloride
PMSF	Phenylmethanesulfonyl fluoride
NaCl	Sodium chloride
SDS	Sodium dodecyl sulphate
Na <sub>2</sub> HPO <sub>4</sub>	Disodium hydrogenphosphate
KH <sub>2</sub> PO <sub>4</sub>	Potassium dihydrophosphate
TEMED	Tetramethylethylenediamine
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Tetramethylethylenediamine
FBS	Fetal Bovine Serum
PBS	Phosphate buffered saline
K <sub>2</sub> HPO <sub>4</sub>	Dipotassium hydrogenphosphate
TCGA	The Cancer Genome Atlas

## Abbreviations

IP	Immunoprecipitation
β-ME	β-mercaptoethanol
CRC	colorectal cancer
BLCA	bladder carcinoma
BRCA	breast cancer
CEC	cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	cholangiocarcinoma
COAD	colon adenocarcinoma
ESCA	esophageal carcinoma
GBM	glioblastoma
HNSC	head and neck squamous cell cancer
KICH	kidney chromophobe
KIRC	kidney renal clear carcinoma
LIHC	liver hepatocellular carcinoma
LUSC	lung squamous cell carcinoma
PAAD	pancreatic adenocarcinoma
PCPG	pancreatic adenocarcinoma
PRAD	prostate adenocarcinoma
READ	rectum adenocarcinoma
STAD	stomach adenocarcinoma
THCA	thyroid carcinoma
UCEC	uterine corpus endometrial carcinoma
TPM	transcript per million
SEP3	SEPALLATA3
ORF	open-reading frame
COL6A3	collagen type VI alpha 3 chain
hnRNPM	heterogeneous nuclear ribonucleoprotein M
SDX	stem-differentiating xylem

## Supplementary information

### 8. Supplementary information

Supplementary Table 1: Upregulated and downregulated circRNAs in PC9 cells compared to PSEA cells

circRNAs	Gene	log2 Fold Change	Regulation
hsa_circ_2083	PLOD2	-4.5334291	DOWN
hsa_circ_1027	LPAR3	-7.7279086	DOWN
hsa_circ_7445	THSD1	-5.8641796	DOWN
hsa_circ_5783	PSD3	-5.7414605	DOWN
hsa_circ_9896	KLHDC4	4.48542519	DOWN
hsa_circ_4266	VEGFC	-6.8328792	DOWN
hsa_circ_10206	EHD2	-6.1292727	DOWN
hsa_circ_4647	EPHB4	-2.8921775	DOWN
hsa_circ_5288	SLC22A3	-5.820169	DOWN
hsa_circ_8327	ANXA2	-4.9385923	DOWN
hsa_circ_226	CYP24A1	-5.672415	DOWN
hsa_circ_3354	ADAMTS6	-4.101533	DOWN
hsa_circ_4267	VEGFC	-5.4093809	DOWN
hsa_circ_6440	MICAL2	-4.5647774	DOWN
hsa_circ_10009	ANGPTL4	-5.0223579	DOWN
hsa_circ_9880	CDYL2	-4.9541867	DOWN
hsa_circ_5376	MIR31HG	-5.0874522	DOWN
hsa_circ_4568	SEMA3C	-2.7828992	DOWN
hsa_circ_4936	FKBP5	-3.9772741	DOWN
hsa_circ_6182	ZFY	-4.9999896	DOWN
hsa_circ_3903	RELL1	-2.7062661	DOWN
hsa_circ_5324	UHRF2	-2.9475292	DOWN
hsa_circ_4397	MPP6	-3.1497433	DOWN
hsa_circ_7297	DOCK1	-3.0779989	DOWN
hsa_circ_6181	ZFY	-4.8579701	DOWN
hsa_circ_3850	EVC2	-4.6724156	DOWN
hsa_circ_1424	LAMB3	-4.5849531	DOWN

**Supplementary information**

hsa_circ_3348	ADAMTS6	-4.5849531	DOWN
hsa_circ_4001	MTHFD2L	-4.0588867	DOWN
hsa_circ_9896	KLHDC4	4.48542519	UP
hsa_circ_8758	FUT8	3.13537675	UP
hsa_circ_3034	ZDBF2	4.05889239	UP
hsa_circ_16144	AL117329.1	6.25737881	UP
hsa_circ_18536	GRHL2	6.17990004	UP
hsa_circ_10233	PPP1R12C	4.14974528	UP
hsa_circ_5846	RAB11FIP1	5.06608471	UP
hsa_circ_17344	AGAP1	5.93072843	UP
hsa_circ_668	DHRS3	3.47804636	UP
hsa_circ_2499	SRBD1	3.79701142	UP
hsa_circ_15889	JPH2	5.70043115	UP
hsa_circ_4984	TNFRSF21	2.69276577	UP
hsa_circ_519	CCDC134	2.77610386	UP
hsa_circ_3943	UCHL1	3.33628262	UP
hsa_circ_276	TIAM1	3.74146529	UP
hsa_circ_17731	FSTL5	5.50778605	UP
hsa_circ_2309	HPCAL1	2.39045967	UP
hsa_circ_5782	PSD3	2.41192504	UP
hsa_circ_18633	PIR	5.32191979	UP
hsa_circ_6607	LRP5	3.92056322	UP
hsa_circ_16378	TMEM56	5.24791931	UP
hsa_circ_13643	FIRRE	3.67242352	UP
hsa_circ_18937	ME3	5.22881049	UP
hsa_circ_20474	TMEM38A	5.22881042	UP
hsa_circ_6177	RECQL4	4.60485694	UP
hsa_circ_16143	AL117329.1	5.10851635	UP
hsa_circ_1481	DNAH14	3.03562319	UP
hsa_circ_19431	NTN4	5.06608095	UP
hsa_circ_1564	AKT3	3.52356019	UP
hsa_circ_4454	COA1	2.71484246	UP

### 9. Acknowledgements

I would firstly like to thank my supervisor Prof. Dr. Margarete Odenthal for her immeasurable support and guidance throughout my study. Her professional supervision has always been a driving force, actively and enthusiastically guiding me through the accomplishment of my doctoral thesis. She has helped me develop numerous practical skills and knowledge and an eye and mindset for science. I would also like to thank her for her example as a successful biologist and professor. I am highly grateful of Prof. Dr. Reinhard Büttner for giving me the opportunity to work in the Institute of Pathology, University Hospital of Cologne, Germany.

I am highly obliged of the mentoring provided by my tutors Prof. Dr. Niels Gehring and PD. Dr. Catherin Niemann. I would like to thank them for their immense support and the fruitful advice that have helped me to direct my research.

I am very thankful of my colleagues for all the support that I have received from them over the years of my research. I am especially grateful to the fun group of members who celebrated my birthdays from 2020 to 2022. Thanks to them for helping make such nice memories. Besides, Dr. Maria Anokhina supervised me and worked together for this topic. I give special thanks to my perfect friend Dr. Jie Wang who supported and encouraged me a lot in the past three years. Dr. Priya, Dr. Sonja Meemboor, Dr. Xiaojie Yu, Dr. Priya Dalvi, Dr. Lingyu Wang, Hannah Eischeid, Ulrike Koitzsch, Hardik Makawana also helped me a lot during my study.

Lastly, I am deeply grateful for having a family and a girlfriend that has supported me through thick and thin during this entire time. If it wasn't for them, I would have not been able to achieve this in life. Having been encouraged and helped at every stage of my personal and academic life, I have longed to see this achievement come true.

## Curriculum Vitae

Yefeng Shen

Date of birth: 30.10.1991 in Hubei, China

Nationality: Chinese

Address: Kerpener Str. 62, 50924 Köln

Tel.: 01744784936

Email: yefeng.shen@uk-koeln.de

## Education

2010-2015. Clinical Medicine

Sothorn Medical University

2015-2019 Master of Clinical Medicine

Sothorn Medical University

2019-present PhD candidate in Molecular Medicine

University of Cologne

## Presentations

2021 IPMM Student Days (Graduate School)

2022 IPMM Student Days (Graduate School)

## Awards

2019 Oversea Study Program of Guangzhou Elite Project

2021 Chinese Government Award for Outstanding Self-financed Students

Abroad

## Publications

1. **Shen Y**, Wang S, Wu Y. A Novel m6A-Related LncRNA Signature for Predicting Prognosis, Chemotherapy and Immunotherapy Response in Patients with Lung Adenocarcinoma. *Cells*. 2022; 11.
2. Zhong H, Wang J, Zhu Y, **Shen Y**. Comprehensive Analysis of a Nine-Gene Signature Related to Tumor Microenvironment in Lung Adenocarcinoma. *Front Cell Dev Biol*. 2021; 9: 700607.

## Curriculum Vitae

3. Zhu Y, **Shen Y**, Chen R, Li H, Wu Y, Zhang F, et al. KCNQ1OT1 lncRNA affects the proliferation, apoptosis, and chemoresistance of small cell lung cancer cells via the JAK2/STAT3 axis. *Ann Transl Med.* 2021; 9: 891.
4. Fang S, **Shen Y**, Chen B, Wu Y, Jia L, Li Y, et al. H3K27me3 induces multidrug resistance in small cell lung cancer by affecting HOXA1 DNA methylation via regulation of the lncRNA HOTAIR. *Ann Transl Med.* 2018; 6: 440.
5. Li H, **Shen Y**, Wu Y, Cai S, Zhu Y, Chen S, et al. Stereotactic Body Radiotherapy Versus Surgery for Early-Stage Non-Small-Cell Lung Cancer. *J Surg Res.* 2019; 243: 346-53.



## Erklärung

## Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit -einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Frau Prof. Margarete Odenthal betreut worden.

Ich versichere, dass ich alle Angaben wahrheitsgemäß nach bestem Wissen und Gewissen gemacht habe und verpflichte mich, jedmögliche, die obigen Angaben betreffenden Veränderungen, dem Promotionsausschuss unverzüglich mitzuteilen.

...Cologne, 27.02.2023.....  
Datum