Aus der Klinik für Urologie und Urochirurgie

der Medizinischen Fakultät Mannheim

Direktor: Prof. Dr. med. Maurice Stephan Michel

ANLN, TLE2 and MIR31HG transcripts in muscle invasive bladder cancer: a functional and clinical analysis based on molecular subtypes

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum

der

Medizinischen Fakultät Mannheim

der Ruprecht-Karls-Universität

zu

Heidelberg

vorgelegt von

Sheng Wu

aus

Anhui, China

2020

Dekan: Prof. Dr. med. Sergij Goerdt

Referent: Prof. Dr. med. Philipp Erben

Für meine lieben Eltern

CONTENTS

LIST OF ABRREVIATIONS

ANLN	Anilin actin binding protein
ATCC	American Type Culture Collection
BLCA	Bladder Urothelial Carcinoma
Calm	Calmodulin
cDNA	Complementary Deoxyribonucleic Acid
CI	Confidence Interval
CIS	Carcinoma in situ
DFS	Disease-free Survival
DMEM	Dulbecco's Modified Eagle's Medium
DSS	Disease-specific Survival
ECACC	European Collection of Authenticated Cell Cultures
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial-to-Mesenchymal Transition
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FFPE	Fresh Frozen Paraffin Embedded
GISTIC	Genomic Identification of Significant Targets in Cancer
GO	Gene Ontology
GUS	β-Glucuronidase
HER2	Receptor Tyrosine-protein Kinase erbB-2
HR	Hazard Ratio
IHC	Immunohistochemistry
KEGG	Kyoto Encyclopedia of Genes and Genomes
lncRNA	Long non-coding RNA
LVI	Lymphovascular Invasion
MIBC	Muscle-invasive Bladder Carcinoma

MIR31HG	miR-31 Host Gene
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NCBI	National Center for Biotechnology Information
NMIBC	Non-muscle-invasive Bladder Carcinoma
OS	Overall Survival
PFS	Progression-free Survival
PI3K	Phosphoinositide 3-kinase
qPCR	Quantitative Polymerase Chain Reaction
RC	Radical Cystectomy
ROC	Receiver Operating Characteristic
RPKM	Reads Per Kilobase Million
RPMI	Roswell Park Memorial Institute Medium
siRNA	Small Interfering Ribonucleic Acid
TCGA	The Cancer Genome Atlas
TLE2	Transducing-like enhancer protein 2
TPM	Transcripts Per Million

1 INTRODUCTION

1.1 Bladder cancer

1.1.1 Epidemiology

Bladder cancer (BLCA) is the tenth most common cancer worldwide with estimated 550,000 new cases and 200,000 deaths in 2018 [1]. BLCA is more common in men than in women. The incidence of BLCA rises with age, peaking between age 50 years and 70 years. With mean age for male patients of 69 and female patients of 73.4, men are three to four times suffered as women, while the latter suffered from more aggressive forms of BLCA [1]. Incidence rates in both sexes are highest in Southern Europe (Greece with the highest incidence rate in men globally), Western Europe and Northern America (Figure 1).
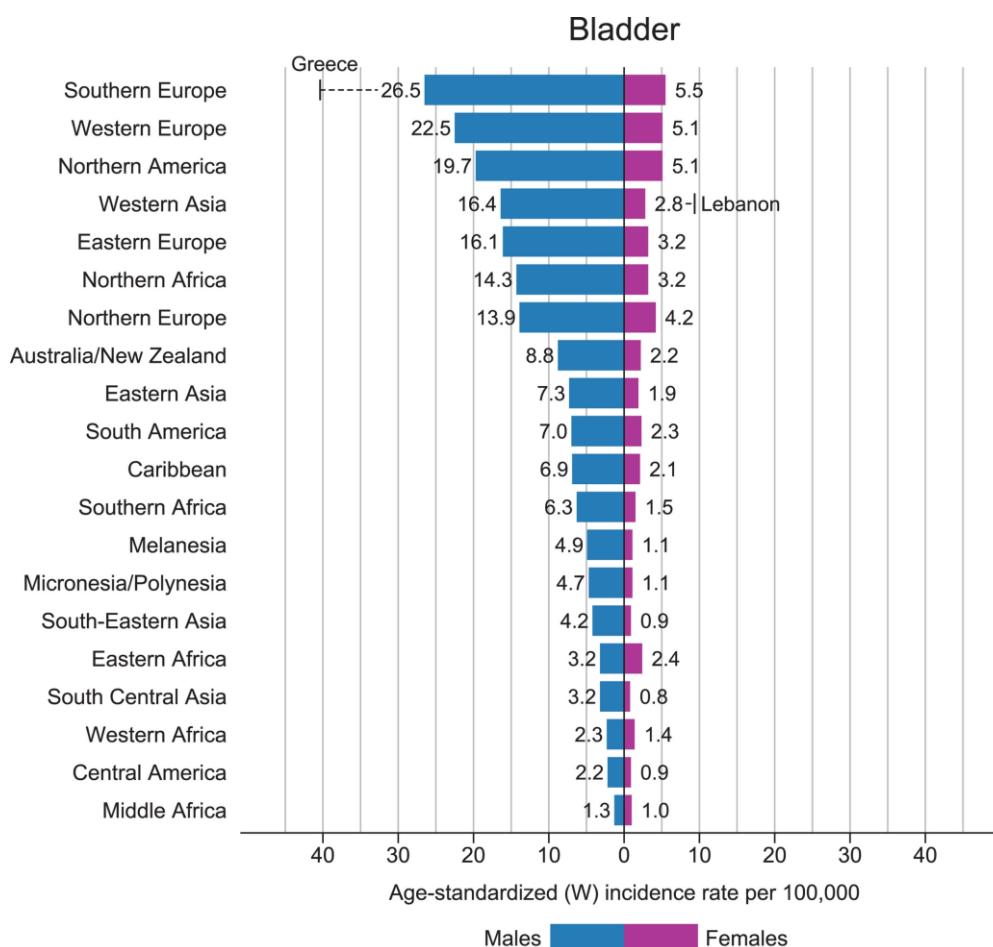
Figure 1. Bar Chart of Region-Specific Incidence Age-Standardized Rates by Sex for Cancers of the Bladder in 2018. Rates are shown in descending order of the world (W) age-standardized rate among men, and the highest national rates among men and women are superimposed. Source: GLOBOCAN 2018.

1.1.2 Risk factors and classification

The main risk factors of the occurrence of BLCA include cigarette consumption and industrially produced chemicals, as well as genetic and molecular abnormalities [2, 3]. Genetic and molecular factors include oncogenes and tumor suppressor genes. It is reported that oncogenes such as tumor protein p63 (TP63), the epidermal growth factor receptors (EGFR), and Ras p21 proteins [4-6] are involved with carcinogenesis and progression of BLCA. Furthermore, tumor suppressor genes probably play a part in the genetic pathogenesis of BLCA, including tumor protein p53 (TP53) and retinoblastoma gene (RB1) [7, 8].

Two main different types of BLCA can be distinguished: non-muscle invasive BLCA with low-grade and muscle invasive cancer with high-grade tumors and high risk of metastasis in regional lymph nodes, lungs, liver and bone [9]. Non-muscle-invasive BLCA (NMIBC; 70%) is not immediately life threatening but often progress, while muscle-invasive BLCA (MIBC; 30%) is responsible for the most cases of metastases and death [10]. The classification of tumors is based on the Tumor-Node-Metastasis (TNM) classification of the Union for International Cancer Control (UICC) of 2010, accordingly include papillary tumors (Ta) and the non-invasive urothelial carcinoma, which infiltrate the subepithelial connective tissue (T1). T2 to T4 are tumors that invade the muscle, the fat tissue and / or surrounding tissue with different depths and counted as the invasive carcinomas [11]. The existing clinical staging and pathological grading of BLCA are depending on observation by naked eye and microscopy. It cannot comprehensively and accurately assess the biological activities of tumor cells, such as proliferation ability, development and invasiveness. Therefore, in-depth research, which clarifies the biological process of BLCA and the molecular regulatory network, can improve early diagnosis, treatment effects, and prognosis.

Currently, several classifications have proposed into sets of molecular classes, including basal and luminal subtype, which partially characterized by KRT5 and KRT20 gene expression [12]. Although a number of groups have reported molecular classifications of BLCA to evaluate the severity and prognosis of this disease, reliable and effective biomarkers for early diagnosis and prognostic prediction are still lacking [13, 14]. Thus, in-depth understanding of molecular events and underlying mechanisms involved in the carcinogenesis of MIBC may provide us effective therapeutic targets and predictive biomarkers, which are urgently needed.

1.1.3 Diagnostic and therapy

Painless macrohematuria is the most common symptom of BLCA. Other typical symptoms are dysuria, increased urinary frequency or frequent urination urgency. However, the positive predictive value of hematuria for BLCA is only 8% [15]. Although cystoscopy is an effective method for detecting BLCA, it also produces many adverse reactions. Common adverse reactions include infection, difficulty urinating, frequent urination, and hematuria [16]. Therefore, it is particularly important to study the mechanism of BLCA occurrence and development and control the recurrence of BLCA. It is necessary to develop non-invasive diagnostic tools for screening and monitoring of BLCA to improve the life quality of patients. For muscle-invasive non-metastatic urothelial carcinoma, radical cystectomy (RC) including bilateral pelvic lymphadenectomy is the curative treatment method of choice [17]. The current therapeutic standard for MIBC is RC with perioperative platinum-based chemotherapy in selected cases [18]. However, the postoperative recurrence rate is very high, and the 5-year recurrence rate can reach 50%-70%. The malignant degree and pathological grade of the tumor are significantly increased when some patients relapse [19].

1.1.4 Novel biomarkers of BLCA

Limits in classification and high recurrence rate lead to the search for novel prognostic biomarkers. In recent years, the increasing number of basic researches on BLCA has shifted the research target from protein molecules to tumor-related genes and their regulatory factors. Different classes of biomarkers were proposed, including coding genes and non-coding genes.
The coding gene anilin actin binding protein (ANLN) is located on chromosome 7q14.2 and encodes for a 1,125 amino acid actin-binding protein that includes a conserved N-terminal actin (F-actin) and myosin binding region and a

conserved C-terminal pH binding domain [20]. Previous studies found that ANLN expression levels were significantly up-regulated in a variety of tumor tissues, including breast, ovarian, colon, lung, and pancreatic cancers [21-26]. Furthermore, ANLN and its encoded protein are highly expressed in BLCA tissues, and their expression levels are positively correlated with the pathological grade and stage of BLCA [27]. However, the involved signaling pathways and the interacting molecular targets in the regulation of BLCA biological function are in discussion.

Transducing-like enhancer protein 2 (TLE2), a member of the TLE gene family, is located on chromosome 19p13.3, and acts as a transcriptional corepressor [28]. Previous studies revealed that the molecular properties of the TLE2 protein point to a function in transcriptional regulation, involved in embryonic neuronal development in conjunction with Hairy/Enhancer of split (HES) proteins [28, 29]. The TLE2 protein inhibits replication-and-transcription-activator-mediated transactivation and lytic reactivation of Kaposi's sarcoma-associated herpesvirus [30]. The function of TLE2 in BLCA has not yet been investigated.

The Wnt/β-catenin pathway was reported to be required for regeneration of the normal urothelium [31]. Also, previous studies have indirectly suggested the involvement of Wnt/β-catenin pathway in malignant behavior of BLCA cells [32, 33]. However, the involved molecules and mechanism of Wnt/β-catenin pathway in MIBC are still unclear. In a previous study, ANLN expression was associated with Wnt/β-catenin signaling in gastric cancer [34]. Also, the Wnt/β-catenin pathway regulates gene expression via T-cell factor/lymphoid enhancer-binding factor 1 (TCF/LEF1) family, which is repressed by transcription factors for TLEs [35, 36]. However, the association between ANLN and TLE2, as well as the association with BLCA was not elucidated yet. Thus, ANLN and TLE2 were chosen as candidates in this study.

1.2 Long noncoding RNA

1.2.1 Characteristics of long non-coding RNA

Long non-coding RNAs (lncRNA) are a class of RNAs that are more than 200 nucleotides in length and lack the ability to encode proteins [37, 38]. In the past, lncRNAs have been considered to be "noise" during transcription process, which do not have any biological function. LncRNA has several structural and expressional characteristics, suggesting that it is not a transcription "noise", but an important functional RNA: 1) lncRNA mostly has an mRNA-like structure with a promoter region and polyA tail and there are different ways of cutting during maturation [39]; 2) most lncRNA expressions are spatiotemporal specific [40]; 3) many lncRNAs display abnormal expression patterns in diseases [41]. Existing research shows that lncRNA, as an important form of epigenetics, plays a regulatory role at multiple levels within the cell (Figure 2).
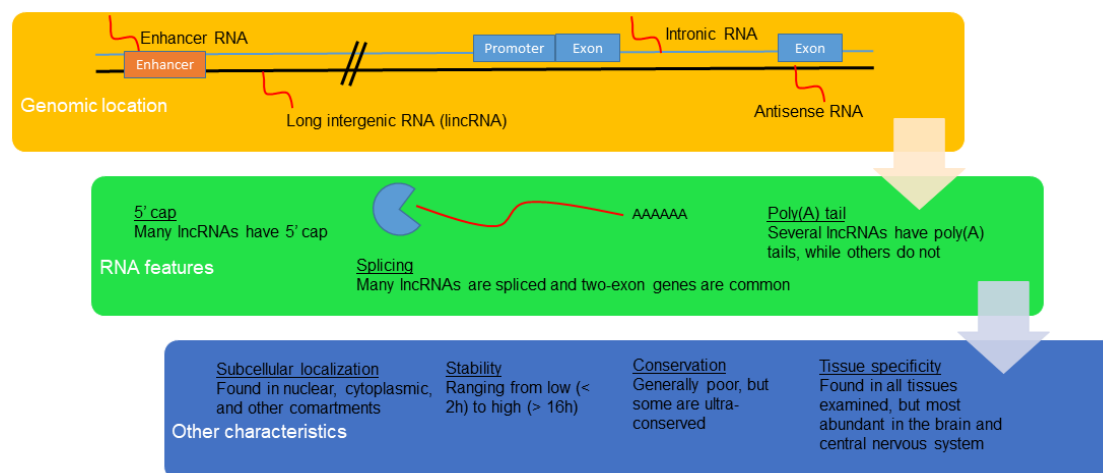
Figure 2. Characteristics of lncRNAs. The genomic location, RNA features and other characteristics of lncRNAs were illustrated. The majority of lncRNA genes are located within 10 kb of protein-coding genes and many lncRNAs are antisense to coding genes or intronic. Many lncRNA genes have two or more exons and about 60% of lncRNAs have polyA+ tails. Other characteristics are involved in the functions of lncRNAs, including subcellular localization, stability, conservation and tissue specificity.

For example, it interferes with the expression of neighboring coding genes; mediates chromatin remodeling and histone modification; forms a

complementary double-strand with the transcript of the coding protein gene, interferes with mRNA shearing; and binds to specific proteins to change the intracellular localization of the protein [42, 43]. Different positions of lncRNAs may exert their gene regulation through the mechanism of reversion. Antisense lncRNA often exhibits positive or negative correlation in expression with its neighboring mRNA, and may achieve multiple transcription and post-transcriptional regulation of coding genes through complementary binding [44]. Intron lncRNAs may alter the transcript cleavage of neighboring coding proteins or enhance their stability [45]. The intergenic lncRNA has been shown to form a complex with the chromatin regulatory protein PRC2 and silence specific gene expression [46]. Although these RNAs are not translated into proteins, in fact, they interact with other molecules such as mRNA to form a complex network, which has many important biological functions and links in the human cells, especially tumors.

1.2.2   LncRNA in cancer

The occurrence and progression of tumors is a complex process involving multiple genes and multiple factors, as well as multiple biological interactions [47, 48]. Most of the basic research in the past decade has focused on detecting BLCA from protein-encoding genes. Related diagnostic markers and in-depth exploration of its role in the pathogenesis of BLCA are urgently needed.

With continuous in-depth research on lncRNA, it is found that lncRNA widely participates in and constitutes the complex regulatory network of gene expression, epigenetic, transcriptional process, post-transcriptional modification, translation process and other levels to regulate gene expression, which ultimately achieve the purpose of playing an important role in tumorigenesis and development [49, 50]. Hepatic cancer, lung cancer, BLCA, breast cancer, pancreatic cancer, glioma, gastric cancer and many other malignancies have been confirmed to have abnormal expression of lncRNA [51-

53]. For example, a study has found that the expression of HOTAIR in primary and metastatic breast cancer is significantly up-regulated. It can enhance the invasion of breast cancer cells by binding to the PRC2 protein, which affects epithelial-mesenchymal transition. And its expression level is related to breast cancer metastasis and prognosis [54]. Lu et al. and Miyoshi et al. reported that lncRNA MEG3 can be used as a tumor suppressor to regulate the occurrence and development of lung cancer and gliomas [55, 56]. It is also found that GAS5 expression was significantly down-regulated in pancreatic cancer tissues, causing the proportion of cells in G0 / G1 phase decreased and the proportion of cells in phase S increased, and GAS5 overexpression could inhibit cell proliferation [57]. Shi et al. confirmed that GAS5 was lowly expressed in non-small cell lung cancer tissues, and regulates the expression of P53 protein and the transcription factor E2F1 by inhibiting cell growth and aggravating apoptosis [58].

1.2.3 LncRNA in BLCA

In recent years, many studies about the expression profile of lncRNAs in BLCA tissues have been emerged, as well as in-depth reports on the regulation of the proliferation and invasion of BLCA cells by specific lncRNA. Currently reported lncRNAs associated with BLCA include H19, UCA1, GAS5, MALAT1, MEG3, TUG1, HOTAIR, ucRAN, etc [59-64]. Luo et al. found that H19 plays a role in BLCA in the form of oncogenes. After the expression of H19 is up-regulated, it is combined with EZH2 to suppress the expression of E-cadherin and eventually promote the distant metastasis of BLCA cells [65]. The results of Zang et al. suggest that the key proteins in the UCA1 and PI3K / AKT signaling pathways have consistent expression trends, suggesting that UCA1 affects the expression level of the transcription factor CREB under the mediation of this pathway, which in turn affects the proliferation and the cell cycle is significantly

regulated [66]. Another study found that down-regulating GAS5 expression in BLCA tissues or cells, and inhibiting GAS5 can promote the proliferation of BLCA cells by regulating cyclin-dependent kinase 6 (CDK6) cyclin [63].

It is known that lncRNA is closely related to the occurrence, development, invasion and metastasis of BLCA [59]. Therefore, by screening differentially expressed lncRNA in BLCA, studying its effect on the biological function of tumor cells, and further revealing its specific mechanism for regulating the occurrence of BLCA, it is expected to open a new world for the diagnosis and treatment of BLCA.

1.2.4   LncRNA MIR31HG in cancer

Recently, miR-31 host gene (MIR31HG, Accession number: NR_027054, also known as LOC554202) has been identified implicated in several cancers, such as breast, colorectal, gastric cancer and pancreatic ductal adenocarcinoma [67-70]. In breast cancer and non-small cell lung cancer (NSCLC) cells, MIR31HG expression was upregulated [67, 71]. While in gastric cancer tissues and cell lines, MIR31HG was poorly expressed [69]. It is also reported that MIR31HG expression was down-regulated in BLCA cell lines and tumor tissues [72]. However, the functional role of MIR31HG in proliferation and migration of BLCA cell lines, and the association with molecular classifications in BLCA is still unknown.

1.3 Aims of this study

Diagnostic and risk stratification of MIBC has barely changed over the last decades. Limitations in histopathological and molecular classification might lead to inadequate treatment decisions and result in high recurrence rates of MIBC. Besides, the expression and function of coding genes ANLN, TLE2 and

non-coding gene MIR31HG in MIBC are remain unknown. Therefore, the aim of the study was the analysis of gene expression and function of ANLN and TLE2, and lncRNA MIR31HG in BLCA cells and tissues.

Therefore, the following questions were addressed in this work:
- How is the gene expression of ANLN and TLE2, and lncRNA MIR31HG in MIBC?
- How are the genes associated with the clinical and histopathological parameters?
- Do the genes have a prognostic role and how is the association with survival of patients after RC?
- How are the genes effects on proliferation and migration of BLCA cells?
- How are the genes involved with signaling pathways and molecular subtypes of MIBC?

To explore the correlation with histopathology and survival data of BLCA patients, expression of ANLN and TLE2 were investigated in our cohort by qRT-PCR and validated in published datasets. For further revealing the potential mechanism, the involved signaling pathways and the impact of ANLN and TLE2 on MIBC molecular subtypes were also evaluated. In addition, expression of MIR31HG and its splicing variants were explored in BLCA cell lines and correlated to survival in MIBC cohort. For functional analysis, knockdown of MIR31HG and a series of *in vitro* experiments were performed to investigate the effects of MIR31HG on proliferation, colony formation and migration of BLCA cells. Transcript-specific knockdown were carried out to reveal if MIR31HG and its splicing variants could associate with molecular subtypes. Further survival analysis in our cohort and validation in published datasets were performed to show if MIR31HG and its splicing variants could act as prognostic biomarkers.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Patients and tissue samples

This study retrospectively enrolled 107 patients who received RC at the Department of Urology and Urosurgery of the University Medical Centre Mannheim, between 2008 and 2012, and who had a histological diagnosis of MIBC (males: n = 79, 74%, median age: 71 years, range: 41–88 years; females: n = 28, 26%, median age: 73 years, range: 47–86 years; Mannheim cohort). All patients were treated with RC and bilateral lymphadenectomy without preoperative or adjuvant chemotherapy or radiotherapy. With the help of the clinic's internal documentation program, the following parameters were collected after examination of the pathology findings: sex, age, T-stage, N-stage, M-stage, grading, lymphovascular invasion (LVI), blood vessel invasion (VI), simultaneous carcinoma in situ (CIS), multifocality, and soft tissue positive surgical margin.

Formalin fixed paraffin embedded (FFPE) tumor tissue samples were evaluated for pathological stage according to the 2017 TNM classification from the Union for International Cancer Control (UICC) [73]. Tumors were graded using the 2017 WHO/ISUP classification (Table 1) [74]. Studies involving human participants were approved by the ethical board of the University of Heidelberg (2015-549N-MA) and performed in accordance with relevant guidelines and regulations. The Cancer Genome Atlas cohort (TCGA, Provisional) contained RNA sequencing data of 407 patients with MIBC and complete clinicopathological and follow-up data (males: n = 300, 73.7%, median age: 68 years, range: 34–90 years; females: n = 107, 26.3%, median age: 72 years, range: 43–90 years; TCGA cohort). In the published TCGA dataset, patients with MIBC were classified into five molecular subtypes according to mRNA

clustering, including basal/squamous, luminal, luminal-infiltrated, luminal-papillary and neuronal subtypes. In a previous study of our group, an alternative classification based on quantification of a 36-gene panel, clustered into 3 subtypes (luminal, basal, infiltrated [p53-like]) on a cohort with MIBC, which also validated by the TCGA cohort [75].

Table 1. Clinicopathological characteristics of patients and specimens of this study.

Clinicopathological Features		n	%
Age	<70	46	43.0
	≥70	61	57.0
Gender	Male	79	73.8
	Female	28	26.2
Grade	Low	5	4.7
	High	102	95.3
Stage	I	5	4.7
	II	12	11.2
	III	69	64.5
	IV	21	19.6
Lymph node metastasis	Negative	75	70.1
	Positive	24	22.4
	Undefined	8	7.5

Note: Stage I tumor was diagnosed as MIBC by transurethral resection of bladder tumor (TURBT).

2.1.2 Database

Expression data in transcripts per million (TPM) of 25 human BLCA cell lines were collected from the Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012) [76], including 20 bladder urothelial cell carcinomas, one bladder squamous cell carcinoma, and four bladder carcinoma cell lines from unknown primaries. Expression of ANLN, TLE2 and MIR31HG was analyzed by Expression Atlas (https://www.ebi.ac.uk/gxa/home) and normalized by TPM. Transcript expression data of 407 BLCA samples and 21 normal samples were

collected from TCGA (https://tcga-data.nci.nih.gov/tcga/) and analyzed by cBioPortal (http://www.cbioportal.org/), which normalized the samples by reads per kilobase million (RPKM). The height of a bar in the resulting bar plots represents the median expression tumor types or normal tissue. Each dot represents the expression of genes in samples in the dot-box plots. The protein-protein interactions were analyzed by STRING (https://string-db.org/). The RNA-protein interactions were analyzed by lncPro (http://bioinfo.bjmu.edu.cn/lncpro/) [77]. GO enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were analyzed by Enrichr (http://amp.pharm.mssm.edu/Enrichr/) [78]. Expression of transcripts, exons and junctions (log2 transformation) was collected and analyzed using the Xena online exploration tool (https://xena.ucsc.edu/) and TSVdb (http://www.tsvdb.com) [79].

2.1.3 Cell lines and cell culture material

In this study six different cell lines were used, including a normal human urothelium cell line (UROtsa), a basal-like urothelial carcinoma cell line (SCaBER), two luminal-like urothelial carcinoma cell lines (RT112 and RT4), and two mixed-type urothelial carcinoma cell lines (UMUC3 and T24). UROtsa cell line was from the left ureter of a 12 year-old girl and immortalized by transfection with a temperature sensitive SV-40 large T antigen gene [80]. RT4 cells established from a recurrent well-differentiated transitional papillary tumor of the urinary bladder of a 63-year-old man in 1968 [81]. RT-112 cells were from an unknown-aged woman with untreated primary urinary bladder carcinoma [82]. UMUC3 cells from an unknown-aged man with untreated primary urinary bladder carcinoma [83]. T24 cells established from the primary tumor of an 81-year-old Caucasian woman with urinary bladder carcinoma in 1970, which described in the literature to produce a variety of cytokines (e.g. G-CSF, IL-6

and SCF) and to carry a p53 mutation [84]. SCaBER cells derived from a 58-year-old male patient, described as squamous cell carcinoma of the human urinary bladder [85].

UROtsa cells were cultured in Roswell Park Memorial Institute medium (RPMI) supplemented with 5% fetal bovine serum (FBS). RT112, RT4, SCaBER, and UMUC3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% FBS. T24 cells were cultured in McCoy's 5A medium containing 10% FBS. UMUC3, SCaBER, RT112, and T24 cells were obtained from the European Collection of Authenticated Cell Cultures (ECACC), RT4 from the American Type Culture Collection (ATCC), and UROtsa cells from a collaborator. Before starting the experiments, all cell lines were authenticated by Multiplexion (Heidelberg, Germany).

2.1.4 Chemicals and reagents

The reagents and kits used in this study as well as their manufacturer and order number are showed in Table 2. The chemicals used in this study are listed in Table 3.

Table 2. Reagents and kits used in this study.

Name	Manufacturer	Order number
Cell Titer 96® Aqueous One Solution Cell Proliferation Assay	Promega (USA)	G3582
Custom siRNA, Standard 0.015 µmol Regular	GE Healthcare (USA)	SO-2799195G
DharmaFECT ™ Transfection 1 Reagent	GE Healthcare (USA)	T-2001-03
Fast SYBR Green Master Mix	Life Technologies (USA)	4385612
Fetal bovine serum (FBS)	Sigma-Aldrich (USA)	F7524
Lincode Set of 4 Upgrade siRNA, MIR31HG	GE Healthcare (USA)	SO-2611084G
McCoy's 5A (Modified) Medium	Invitrogen (USA)	16600-082
M-MLV Reverse Transcriptase	Invitrogen (USA)	28025-013
pdN6 random primer	Roche (Germany)	11034731001
Protease Inhibitor Cocktails	Sigma-Aldrich (USA)	P2714
RNase Away®	Molecular BioProducts (USA)	7003
RNase OUT ™	Invitrogen (USA)	10777-019
RPMI 1640 medium	Invitrogen (USA)	21875-091
Silencer® Select Negative Control siRNA	Ambion (USA)	4390846
Superscript® III Reverse Transcriptase	Invitrogen (USA)	12574035
TaqMan® Fast Advanced PCR Master Mix	Life Technologies (USA)	4444557
RNeasy Mini Kit	QIAGEN (Germany)	74104
XTRAKT FFPE Kit	Stratifyer (Germany)	XTK2.0-96

Table 3. Chemicals used in this study.

Name	Manufacturer	Order number
Agarose	Carl Roth (Germany)	3810.3
Crystal violet	Sigma-Aldrich (USA)	C0775
DMEM, high glucose	Invitrogen (USA)	41965-062
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich (USA)	D2650
DPBS	Invitrogen (USA)	14190
Ethanol 100%	Carl Roth (Germany)	9065.1
Glycine	Sigma-Aldrich (USA)	G8898
Isotonic sodium chloride solution	DeltaSelect (Germany)	7462991
Methanol	Carl Roth (Germany)	8388
Nuclease-free water	Promega (USA)	p1193
Trypan 0.4%	Invitrogen (USA)	15250-061
Trypsin-EDTA 0.25%	Invitrogen (USA)	25200-056

2.1.5 Devices and equipments

The devices and equipments used in this study and their manufacturer are showed in Table 4.

Table 4. Devices and equipments used in this study.

Name	Manufacturer
Accurate scales	TSE Systems (Germany)
Culture Insert 2 Well	ibidi (Germany)
Centrifuge 5804	Eppendorf (Germany)
Centrifuge Rotina 430R	Hettich (Germany)
CO_2 incubator	BINDER (Germany)
FlexCycler2	Analytik Jena (Germany)
Fridge freezer Grand Prix	Bosch (Germany)
Fridge freezer No frost	Siemens (Germany)
Herasafe ™ KSP Class II Biological Safety Cabinet	Thermo Fisher Scientific (USA)
Infinite® F200 PRO microplate reader	TECAN (Switzerland)
Inverted microscope Axiovert 10	Zeiss (Germany)
LUNA cell counter	Logos Biosystems (South Korea)
Magnetic stirrer RCT basic	IKA (Germany)
Manual Rotary Microtome	Leica (Germany)
Milli-Q Synthesis Water Purification System	Millipore (USA)
Minishaker MS1	IKA (Germany)
MiniSpin	Eppendorf (Germany)
NanoDrop 1000 Spectrophotometer	Thermo Fisher Scientific (USA)
Pipetus® pipette filler	Hirschmann (Germany)
Reax Top Vortex Mixer	Heidolph (Germany)
Sprout mini centrifuge	Biozym Biozym
StepOnePlus ™ Real-Time PCR System	Applied Biosystems (USA)
Water bath	GFL (Germany)

2.1.6 Small interfering RNAs

Here, a pool of four different small interfering RNAs (siRNAs) against MIR31HG (designated si-MIR31HG, set of 4), and transcript-specific siRNAs against MIR31HGΔE1 and MIR31HGΔE3 were transfected into urothelial cells using

the DharmaFECT1 siRNA transfection reagent (GE Healthcare Dharmacon, Inc., CO, USA). Scrambled siRNA was used as negative control. All sequences of the siRNAs are listed in Table 5.

Table 5. siRNAs used in this study.

Target gene	Sequence (5' → 3')
MIR31HG	GUUGAUGGUUAUUAGUGAA
	GCGCUUUGUGUGAGAAGUU
	AGGUUAUAUCCUAGAGAUC
	CAUAGAACCUUGGAUCCUA
MIR31HGΔE1	CCCAGGAGGAGCUUGGUUUCUGGUU
MIR31HGΔE3	GAGGAUUCAUUCCAAGGUAGAGAUU

2.1.7 Primers and probes

The sequences of primers and probes, which used in this study are listed in Table 6 and Table 7.

Table 6. Primers used in this study.

Gene	Forward primer sequence (5' → 3')	Reverse primer sequence (5' → 3')
ANLN	GCATTAGCAGAAAGCAGCGA	TGGACACTAAACTCTCTGGACT
Calm2	GAGCGAGCTGAGTGGTTGTG	AGTCAGTTGGTCAGCCATGCT
CDK1	AAAACTACAGGTCAAGTGGTAGCCAT	GCATAAGCACATCCTGAAGACTGACT
COBL	TACGCGTGGGACAACAGAAG	GGTCGTTAAACAGCCCTTACT
CTNNB1	GGGTCCTCTGTGAACTTGCTC	TTCTTGTAATCTTGTGGCTTGTCC
EGFR	CGCAAGTGTAAGAAGTGCGAA	CGTAGCATTTATGGAGAGTGAGTCT
FOXM1	GACCACCTGGAGCCCTTTG	GATGTTGGATAGGCTATTGTTGATAGTG
GATA3	GCAATGCCTGTGGGCTCTAC	TTCTGGTCTGGATGCCTTCCT
GUS	GAAAATAYRTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTA
KRT5	CGCCACTTACCGCAAGCT	ACAGAGATGTTGACTGGTCCAACTC
KRT20	GCGACTACAGTGCATATTACAGACAA	CACACCGAGCATTTTGCAGTT
MIR31HG	CTCTGGAGGACAGAGGATTCATTCC	TGGGAGGGTGGTCTGAAACTG
MIR31HGΔE1	GCCTCCCAGGAGGAGC	CCAAACTCTGGAGGACAGAGGATTC
MIR31HGΔE3	CTCTGGAGGACAGAGGATTCATTCC	GGAACACCTGGAGACCTGCT
MKI67	TGCTACTCCAAAGAAGCCTGTG	GTATGAGCTTTCCCTATTATTATGGTAC
RacGap1	GAATGTGCGGAATCTGTTTGAG	TCGCCAACTGGATAAATTGGA
TLE2	AAGCGTCTGAGCGGTATCTG	TGCTGCTGCCCGATGAG

Table 7. Probes used in this study.

Gene	5' Label	3' Label	Sequence (5' → 3')
ANLN	FAM	BHQ1	GAACAGGAAGATGCACTGAATAT CTCCTCAATGTCTTTA
Calm2	VIC	BHQ1	TCGCGTCTCGGAAACCGGTAGC
CDK1	FAM	BHQ1	TCAGACTAGAAAGTGAAGAGGAA GGGGTTCCTAGTACTG
CTNNB1	FAM	BHQ1	ACAAGGAAGCTGCAGAAGCTATT GAAGCTGAGG
FOXM1	FAM	BHQ1	CGAGCAGAAACGGGAGACCTGTG ATGGTGAGG
GATA3	FAM	BBQ	ACAAGCTTCACAATATTAACAGAC CCCTGACTATGAAG
GUS	VIC	BHQ1	CCAGCACTCTCGTCGGTGACTGT TCA
KRT5	FAM	BBQ	TGGAGGGCGAGGAATGCAGACTC A
KRT20	FAM	BBQ	TTGAAGAGCTGCGAAGTCAGATT AAGGATGCT
MIR31HG	FAM	BHQ1	GGGTCTGCTTGTATTCAATGACTG GTCTACGTGGG
MIR31HGΔE1	FAM	BHQ1	CAGGTTTCTGGTCCTCATACCGT GTGGT
MIR31HGΔE3	FAM	BHQ1	GGTAGAGATGGATTCCTGGAAAT ACCTCCTCAAGGCC
MKI67	FAM	TAM	CGAAGTTCACAGTCAATTTAGTAC AGGCCAC
RacGap1	FAM	BHQ2	ACTGAGAATCTCCACCCGGCGCA
TLE2	FAM	BHQ1	GCTCAGATTATCCCCTTCCTGACC CAGGAGCAT

For preparing the working solutions, primers and probes (stock solution 100 pmol/µl) were 1:10 diluted in nuclease-free water to target concentration of 10 pmol/µl. All primers and probes were ordered by Eurofins Genomics (Germany).

2.2 Methods

2.2.1 Cell culture

The cryogenic vials of necessary cell lines were taken from the -80°C freezer and quickly thawed in a water bath at 37°C. The cell suspension was then

transferred into a 15 ml tube with pre-warmed medium and centrifuged at 300 g for 5 min. The supernatant was discarded, the cell pellet was resuspended in an appropriate medium and placed in a 25 cm^2 tissue culture flask with pre-heated medium. The cells were then cultured in an incubator at 37°C in an atmosphere of 5% CO_2 and could grow adherently at the bottom of the flask. Medium was changed every 2 days. For generating a large number of cells, cells were passaged at 70-80% of confluence. Once the cells had dissolved with trypsin from the bottom of the flask, than pre-warmed cell culture medium was added to inactivate the trypsin. After that, the cells were passaged with a ratio of 1: 3 or 1: 4 and transferred into prepared cell culture flasks with fresh medium and incubated.

For determination of cell number, cells were seeded at a defined density, and the number of cells using the Neubauer chamber could be determined. For this purpose, the cells were trypsinized from the flask bottom and then centrifuged in in a 15 ml tube with 5 ml medium at 300 g for 5 min. The supernatant was discarded, the cell pellet was resuspended in 5 ml of fresh medium. 10 μl of this cell suspension were mixed with 90 μl of NaCl and 100 μl of trypan blue. 10 μl of this mixture were then placed in a Neubauer counting chamber and counted four squares under the microscope. The number of total counted cells was multiplied by 50,000 to obtain the number of cells / ml.

2.2.2 Transient transfection

In eukaryotic organisms, the RNA interference (RNAi) is a universal mechanism which is used to prevent viral attacks and gene regulation [86]. In the laboratory, RNA interference (RNAi) is an experimental tool to down-regulate the specific genes and to study the function of particular genes or proteins. In this study, genes using synthetic siRNAs (small interfering or short interfering RNA) were turned off by the siRNA single strand binds to the specific mRNA and induces

their selective degradation.

The siRNA transfection was performed with pre-designed siRNA and DharmaFECT transfection reagent (GE Healthcare, Chicago, IL, USA). For transient transfection of siRNA, cells were treated with indicated culture medium without FBS. First, the cells with the required density and appropriate volume were seeded in complete medium depending on the duplicates or triplicates were made in experiment. The test siRNA and scrambled siRNA were diluted with RNase-free solution to the concentration of 5 μM. Subsequently, the siRNAs and the transfection reagent were diluted to the target concentration with FBS-free medium and incubated for 10 min at room temperature (RT) separately. The mixture with diluted transfection reagent and the per-diluted siRNA was incubated 30 min at RT. Finally, full-medium was added to the siRNA transfection mix. The prepared transfection mixture was pipetted to the previously seeded cells and incubated at 37°C and 5% CO_2 for 24 h, than changing to the medium and further cultured 24-72 h according to the experiment setting or collected directly. An overview of the cell densities and volumes of siRNA, transfection, FBS-free and complete medium for transfection with 10 nM siRNA used can be found in Table 8.

Table 8. Volume of per well for siRNA-transfection with target concentration of 10 nM.

Wells/Platte	Cell density (cells/well)	siRNA solution (μl/well)		DharmaFECT solution (μl/well)		Cell solution (μl/well)
		5μM siRNA (μl)	FBS-free medium (μl)	DharmaFECT (μl)	FBS-free medium (μl)	
96	1×10^4	0.4	9.6	0.2	9.8	180
24	5×10^4	1	49	0.5	49.5	400
12	1×10^5	2	98	1	99	800
6	2×10^5	4	196	2	198	1600

2.2.3 RNA extraction

For cell samples, total RNA was isolated using the RNeasy Mini kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions, based on the buffer with a high salt content and binding of RNA to silica membrane. After washing off steps, contaminants are removed and then eluted RNA in RNase-free water.

First, the cell culture medium was removed individually, and the cells were washed with 5 ml cold PBS and detached with trypsin enzymatically from the bottom of the flask. Cells that could grow in plates were lysed directly after the washing step. The lysis buffer RLT was previously added β-mercaptoethanol in the ratio 1: 100. For the cell counting of less than 1×10^6 cells / ml, the cells were lysed with a volume of 350 µl buffer, for higher number of cells with a volume of 600 µl. Subsequently the lysate was homogenized by repeatedly pulling through an injection needle.

Next, equal volume of 70% ethanol as the lysis buffer (350 µl or 600 µl) was added into the lysate and mixed. 700 µl of this sample were added in an RNeasy column which were in a 2 ml tube, and centrifuged at 10,000 g for 2 min. The flow-through was discarded and the centrifugation step was repeated with the rest of the sample. Then 600 µl wash buffer RW1 and 500 µl of the second wash buffer (RPE) were added to the column consecutively. Between each two steps 1 min centrifugation at 10,000 g was performed and the flow-through was discarded. The RNeasy column was then transferred into a new 2 ml tube and further centrifuged 2 min. Then the column was placed in a nuclease-free 1.5 ml tube and 30 µl of RNase-free water was pipetted directly onto the silica membrane followed by 2 min incubation to elute the RNA. After a final centrifugation (3 min, 10,000 g), the eluate was added in an RNase-free 1.5 ml tube and immediately placed on ice. Finally, the RNA concentration was measured photometrically using the NanoDrop spectrophotometer and the RNA stored at -80°C until further usage.

For FFPE tissues, total RNA were extracted and enriched using the magnetic-bead-based XTRAKT FFPE Kit (Stratifyer, Cologne, Germany), according to the instructions of manufacturer [87]. First, the 10μm FFPE sections were lysed, and the paraffin was melted in 150 μl lyse buffer for 30 min at 80°C with shaking. After cooling, 50 μl Proteinase K was added and incubated for 30 min at 65°C with shaking. 40 μl MagiX-beads were mixed and added to samples for incubating 15 min at RT with shaking. Then the samples were transferred to the special tube rack with magnet. The samples were magnetized by 5 min and washed three times with wash buffer. After discarding the supernatant, 50 μl elution buffer was added and incubated 2 min at RT. The eluate was transferred into a fresh tube after incubation and magnetization, determining concentration with Nanodrop spectrophotometer. The purified RNA could be analyzed by polymerase chain reactions (PCR) directly or stored at -80°C until use.

2.2.4   cDNA synthesis

The reverse transcription converts RNA into complementary DNA (cDNA), which could be used in PCR. For cell samples, it was performed using the M-MLV Reverse Transcriptase kit from Invitrogen (California, USA).

The first step was to dilute RNA samples with nuclease-free water to target concentration of 2 μg, so that a total volume of 10 μl was achieved. For RNA denaturation the diluted RNA samples were incubated for 10 min at 65°C, then placed for 3 min on ice. Next the samples were pipetted with cDNA synthesis Master Mix (Table 9).

Table 9. cDNA synthesis Master Mix for M-MLV reverse transcriptase.

Reagents	Volume (µl)
5x First Strand (M-MLV)-buffer	4
DTT (100 mM)	2
dNTP Mix (10 mM)	1
pdN6 Random Primer (5 mg/ml)	1
RNase OUT (40 U/µl)	1
M-MLV Reverse Transcriptase (200 U/µl)	1
Total Volume	10

The sample was immediately placed in the thermocycler, then incubated 120 min at 37 °C and heated 15 min at 70 °C to inactivate the enzyme. After accomplishment of the cDNA synthesis, the sample was centrifuged briefly and stored at -20 °C.

The reverse transcription for FFPE samples was performed using the Superscript III ® reverse transcriptase kit from Invitrogen. This system is engineered for higher thermostability, with a longer half-life at 50 °C, and reduced RNase H activity, resulting in higher yields of full-length cDNA for more complete gene product representation.

The total volume of each reaction system is 20 µl, including 7 µl primer mix and 8 µl cDNA synthesis Master Mix, as well as 5 µl RNA sample. The primer mix included each 0.5 µl reverse primer, plus 3.5µl nuclease-free water and 1 µl dNTPs. For the negative control, nuclease-free water was used instead of RNA sample. First, 5 µl RNA sample was added in the primer mix, followed by incubating for 10 min at 65 °C, then placed for 3 min on ice. After mixing gently with cDNA synthesis Master Mix (Table 10), the sample was immediately placed in the thermocycler, heated 120 min at 55 °C and then 15 min at 70 °C to inactivate the enzyme. After accomplishment of the cDNA synthesis, the sample was centrifuged briefly and stored at -20 °C.

Table 10. cDNA synthesis reaction system for Superscript III reverse transcriptase.

	Reagents	Volume (µl)
	Reverse Primers (10 pmol/µl)	2.5
Primer mix	dNTPs	1
	nuclease-free water	3.5
	RNA	5
	5x First Strand (Superscript III)-buffer	4
Master Mix	0.1 M DDT	1
	RNaseOUT	1
	Superscript III RT (200U/µl)	2

2.2.5 Quantitative PCR

The PCR method is used for amplification of DNA including DNA polymerase, two sequence-specific oligonucleotides (primers), buffer and nucleotides. First, the double stranded DNAs are separated into single strands, then the primers are bond to the complementary single-stranded DNA (cDNA). In the last step, new DNA strands are synthesized by the DNA polymerase with dNTPs. Since these three steps are repeated for several cycles, the desired DNA sequence is amplified and can be detected.

For cell samples, SYBR Green-based quantitative PCR (qPCR) was performed to determine the gene expression. SYBR Green is a fluorescent dye, which thus leads to the increase of the fluorescence intensity during amplification of cDNA associated with the PCR product. The more product is formed, the more dye is attached. Therefore the fluorescence is proportional to the amount of PCR product formed. The Fast SYBR Green Master Mix from Life Technologies contains the SYBR Green dye, DNA polymerase and dNTPs with nuclease-free water, and cDNA with primer pair for qPCR are prepared in advance (Table 11). After a centrifugation at 1850 g for 5 min, the prepared 96-well plate was placed in StepOne Plus. The PCR process was initiated at 95 °C for 20 s, then followed by 40 cycles of the following temperatures for the amplification of cDNA: 95 °C for 3 s, 60 °C for 30 s and 95 °C for 15 s, and the melting curve was available

after the process (95 °C for 15 s, 60 °C for 1 min and 95 °C for 15 s).

Table 11. SYBR Green qPCR Master Mix.

Reagents	Volume (µl)
Fast SYBR® Green Master Mix	5
Nuclease-free water	1.25
Forward primer (10 pmol/µl)	0.375
Reverse primer (10 pmol/µl)	0.375
cDNA	3
Total Volume	10

For FFPE samples, TaqMan probes-based qPCR was performed to determine the gene expression. The TaqMan reaction system is a Master Mix, which contains DNA polymerase and dNTPs, together with nuclease-free water, cDNA, two primers and a probe (Table 13). The probe is labeled with a fluorescent reporter dye at the 5 'end and at the 3' end with a quencher. The primers and probe can hybridize with the cDNA target sequences. During the elongation, the probes which hybridized to the target are cleaved by the DNA polymerase between reporter and quencher, exposing reporter dye and resulting in increased fluorescence.

The prepared master mix was centrifuged for 5 min at 1850 g and provided for amplification of the samples in the StepOne Plus. The PCR process was started initially by denaturation at 95 °C for 20 s, and then followed by 40 cycles of amplification (95 °C for 3 s and 60 °C for 30 s). β-Glucuronidase (GUS) and calmodulin2 (Calm2) were measured as reference genes [88, 89]. As previously described, the relative mRNA expression level was normalized to reference genes or control cell line, and determined using the 40-ΔCT or $2^{-\Delta\Delta CT}$ for cell culture samples and FFPE samples, depending on the specific analysis [90].

Table 12. TaqMan qPCR Master Mix.

Reagents	Volume(µl)
TaqMan® Fast Advanced PCR Master Mix	5
Nuclease-free water	1
Forward primer (10 pmol/µl)	0.375
Reverse primer (10 pmol/µl)	0.375
Probe (10 pmol/µl)	0.25
cDNA	3
Total Volume	10

2.2.6 Cell viability assay

To determine the cell viability, the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega, Wis, USA) was performed. Increased amounts of cells are presented by the increased purple color of the formazan product which is soluble in cell culture medium. The amount of formazan can be measured photometrically and is proportional to the number of living cells in culture. Cells were seeded with 100 µl complete medium in a 96-well plate on the first day of the experiment, with the desired density and incubated overnight at 37 °C and 5% CO_2. The samples were also pipetted in triplicate with the empty cell culture medium as control group. On the second day, the experimental group and control group at appropriate concentrations with a final volume of 100 µl serum-free medium were added to the cells and the cells were cultured at 37 °C in CO_2 incubator. 20 µl MTS solution was added after 24 h, 48 h and 72 h respectively and the plate was incubated for 3 h in an incubator. Subsequently, the absorbance was measured by microplate reader at 492 nm. Each experiment was performed in triplicates.

2.2.7 Colony formation assay

For colony formation assays, benigne and malignant urothelial cells were seeded in six-well plates at a concentration of 5000 cells per well in triplicates, and cultured for 7 days before staining viable colonies with crystal violet

(Sigma-Aldrich, Darmstadt, Germany). The crystal violet was removed followed by 20 min incubation and colonies were air-dry. Colonies were counted under the microscope at low magnification. The staining intensity of the colonies was quantitated using ImageJ software [91].

2.2.8 Migration assay

Cell migration was evaluated by a wound-healing assay. First, the 2-well culture inserts (ibidi GmbH, Munich, Germany) were placed in empty 6-well or 12-well plate using a sterile tweezer. Next, 70 - 100µl of the appropriate number of cells were seeded into each well. After 24 h incubation, culture inserts were carefully removed and initial pictures of open wound were taken. Every 6 h, snapshot pictures were taken under the inverted microscope until the wound closure. To analyze the results of snapshot pictures, the proportion of the wound area was quantitated using TScratch software [92].

2.2.9 GO enrichment and pathways analysis

2.2.9.1 GO analysis

Gene Ontology (GO) can be divided into three categories, which covering three aspects of biology, including biological process, molecular function, and cellular component. The results of the GO enrichment analysis may suggest the pathological changes of gene function during the occurrence and development of BLCA.

This method calculates the hypergeometric distribution relationship in the GO classification by screening differentially expressed genes related to BLCA. The calculation result is expressed by false discovery rate (FDR). The smaller the FDR value, the more enriched the differential genes are in the GO classification. Therefore, by the GO analysis, it can be predicted which gene may be related

to the cellular functional changes [93].

2.2.9.2 Pathway analysis

Pathway analysis refers to the methods of identifying related genes or proteins within a pathway or building pathway de novo. It mainly analyzes the functional pathways or metabolic pathways that may be involved in the target gene, which is particularly important in the study of mechanisms. The database in KEGG Pathway integrates the current knowledge related to the molecular interaction network, such as channels and associations [94]. According to the functions and signaling pathways that molecules participate in, the molecular composition of each signaling pathway is displayed in the form of Pathway Map. By calculating the hypergeometric distribution relationship between the target gene and pathway, the corresponding p value is obtained. The smaller the p value, the more enriched the differentially expressed genes appear in the pathway.

2.2.10 Statistical analysis

Statistical analyses were performed with SPSS 20.0 (IBM, Chicago, IL, USA) and GraphPad Prism 6.0 (GraphPad Software, La Jolla, CA, USA). A Kolmogorov–Smirnov (K-S) test was used to determine whether the data were normally distributed. Student's t-tests were used to compare between groups of normally distributed numerical data, while Mann-Whitney U and Kruskal-Wallis tests were used to compare the non-normally distributed numerical data. Linear regression was used to determine the efficiency of amplification. Spearman tests were used to test the correlation between different gene expressions. The cut-off values ($2^{-\Delta\Delta CT}$ value) of the groups with high and low ANLN, TLE2 and MIR31HG expression were determined by receiver operating

characteristic (ROC) curve analysis in the Mannheim cohort [95]. Similarly, in the TCGA cohort, the cut-off value (log2 value) of the groups with high and low ANLN, TLE2 and MIR31HG expression was determined by ROC curve analysis. The Cox regression model was used for univariable and multivariable analysis to calculate hazard ratio (HR). The diagnosis age, gender, smoking status, stage, lymph node statues and gene expression were counted in multivariable analysis. Survival rates of patients were calculated by the Kaplan-Meier method, and comparison was made by the log-rank test. In all cases, $p < 0.05$ was considered statistically significant.

3 RESULTS

3.1 Clinical significance of ANLN, TLE2 and MIR31HG

3.1.1 Patient population and survival analysis

The expression of interested genes and clinical significance of patients with MIBC was analyze in the Mannheim cohort. Demographic and clinical-pathological data of 107 patients with MIBC included in the Mannheim cohort are shown in Table 1. Median follow-up of the entire cohort was 21 months (range 3–121 months) and the median follow-up of surviving patients was 50 months (range 9–121 months). In total, 53 patients (49.53%) suffered a relapse (local relapse n = 6, lymph nodes and/or distant metastases n = 36, unclear metastasis pattern n = 11). Of the 56 (52.34%) patients who died during the follow-up, 38 (35.51%) of them died because of BLCA.

3.1.2 ANLN and TLE2 expression in MIBC tissues

To explore the association of ANLN and TLE2 of patients with MIBC, the expression of the two genes was detected. The gene expression of ANLN and TLE2 was examined in 60 samples of the Mannheim cohort, including 47 males and 13 females. The median expression of ANLN was 34.16 (range from 27.48 to 37.58), with 34.26 (range from 28.15 to 37.58) in male and 33.82 (range from 27.48 to 37.11) in female. The median expression of TLE2 was 28.82 (range from 22.17 to 36.74), with 29.27 (range from 22.30 to 36.74) in male and 27.20 (range from 22.17 to 33.25) in female.

For validation, a published dataset was reanalyzed. Bar plots compared the gene expression profiles across all 31 tumor samples and paired normal tissues from TCGA database. ANLN expression was confirmed to be up-regulated in the majority of tumor samples (26/31, 83.87%), while TLE2 was down-regulated

in the majority of tumor samples (25/31, 80.64%). In BLCA tumor samples, ANLN expression was up-regulated (median expression 16.72) compared to normal tissues (median expression 1.54; Figure 3A), while TLE2 expression was down-regulated (median expression 22.27) compared to normal tissues (median expression 60.99; Figure 3B). The dot-box plots with data normalized for log-scale showed that ANLN expression was significantly higher in BLCA tumor samples than normal tissues (median expression 4.15 vs. 1.34, and p < 0.001, Figure 3C). In contrast, TLE2 expression was significantly lower in BLCA samples than in normal tissues (median expression 4.54 vs. 5.95, and $p < 0.001$, Figure 3D).
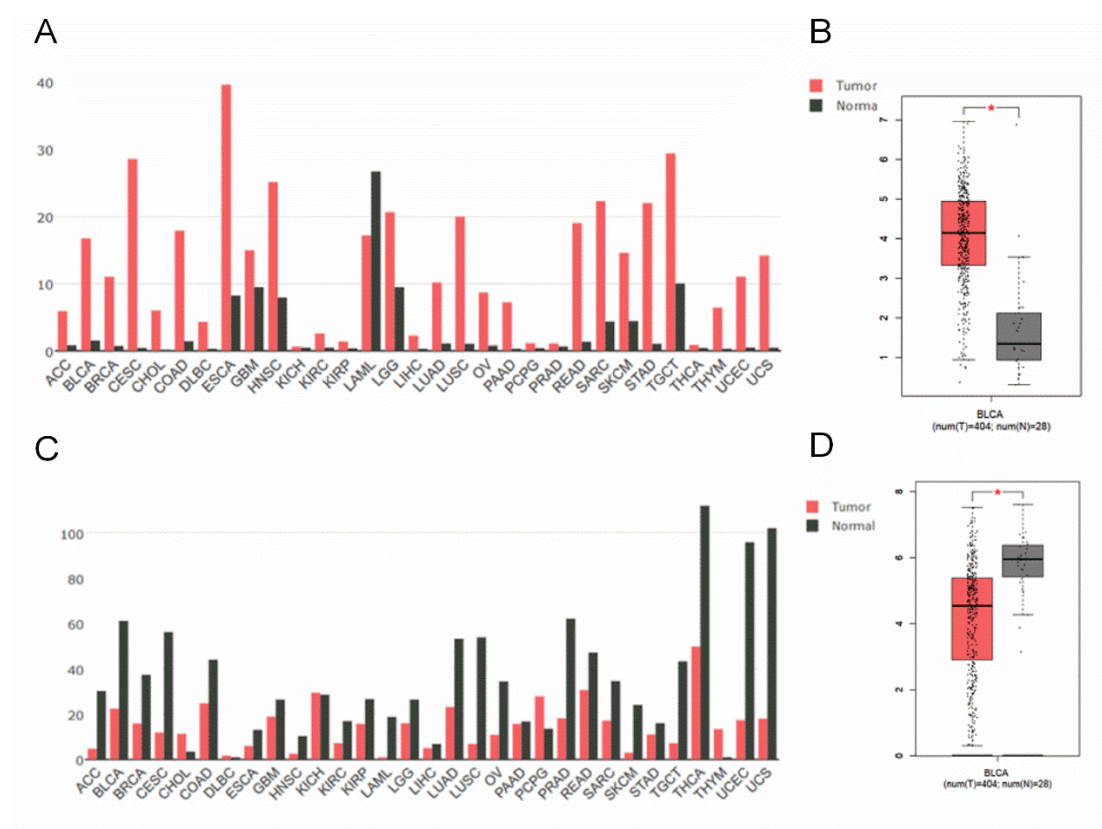
Figure 3. Expression of ANLN and TLE2 in BLCA tumor samples compared with normal tissues. (A) ANLN expression was up-regulated in tumors compared with normal tissues based on TCGA dataset. (B) TLE2 expression was down-regulated in tumors compared with normal tissues based on TCGA data. (C) ANLN expression was significantly higher in BLCA tumor samples than normal tissues based on TCGA and GTEx projects. (D) TLE2 expression was significantly lower in in BLCA tumor samples than normal tissues based on TCGA and GTEx projects.

3.1.3 MIR31HG expression was subtype-specific in MIBC tissues

Through the previous part of the experiment, ANLN and TLE2 have opposite expression pattern in BLCA. As a further research object, the expression of lncRNA MIR31HG was detected in this section. The expression of MIR31HG was examined in 107 samples of the Mannheim cohort, including 79 males and 28 females. The median expression of MIR31HG was 26.96 (range from 22.35 to 38.78), with 26.81 (range from 22.35 to 36.95) in male and 27.36 (range from 24.19 to 38.78) in female. Furthermore, *in silico* analyses were performed for validation. In the TCGA dataset, the expression of MIR31HG was down-regulated in MIBC (median expression 0.2435 with range of 0.0351 to 0.6769, n = 407) compared to normal tissues (median expression 0.3549 with range of 0.2243 to 0.4703, n = 23, *p* = 0.0002, Figure 4).
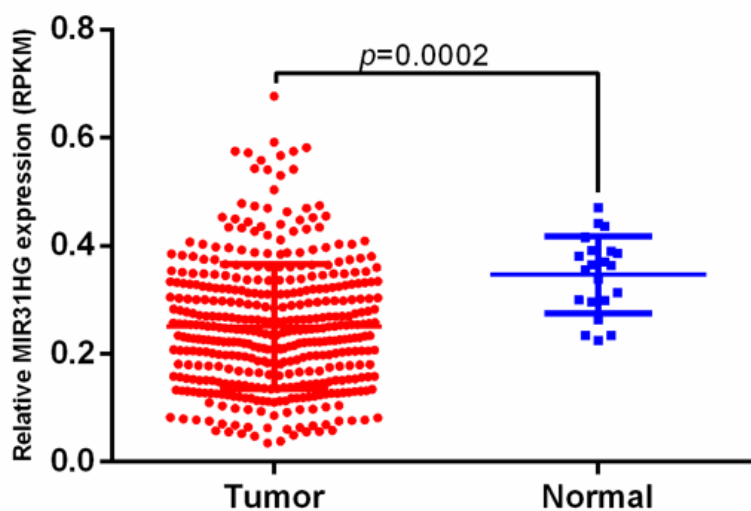

Figure 4. Expression of MIR31HG in BLCA tumor and normal tissue samples. In the TCGA cohort data, MIR31HG was up-regulated in normal tissues compared with tumors.

To better understand the significance of MIR31HG, expression of MIR31HG was analyzed in molecular subtypes of BLCA. Patients with BLCA in the TCGA cohort were classified into basal/squamous, luminal, luminal-infiltrated, luminal-papillary and neuronal subtypes according to mRNA clustering. Among all subtypes, expression of MIR31HG in basal/squamous (median expression 7.10 with range of 0 to 9.96) subtype was the most abundant, which was higher than

in luminal (median expression 3.16 with range of 0.13 to 6.8, *p* < 0.0001), luminal-infiltrated (median expression 4.99 with range of 0 to 9.16, *p* < 0.0001), luminal-papillary (median expression 5.53 with range of 0 to 9.92, *p* = 0.0002), and neuronal (median expression 4.92 with range of 0 to 8.7, *p* = 0.0215) subtypes (Figure 5A). By clustering according to a trichotomous molecular classifications based on an alternative mRNA clustering of the TCGA cohort, it is also determined that the expression of MIR31HG was higher in the basal subtype (median expression 7.10 with range of 0 to 9.96) compared to the luminal (median expression 5.21 with range of 0 to 10.16) and infiltrated (median expression 4.99 with range of 0 to 9.16) subtypes (Figure 5B).
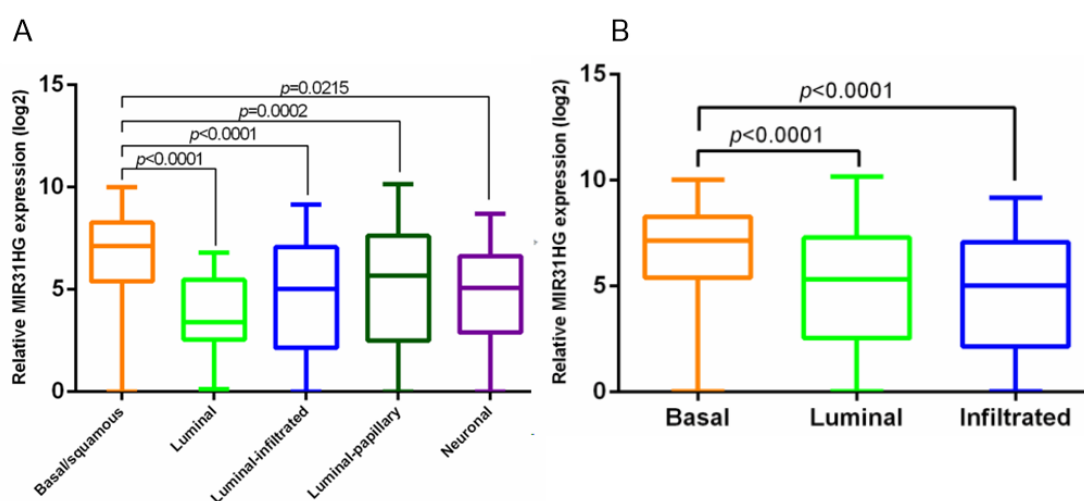

Figure 5. Expression of MIR31HG in molecular subtypes of BLCA. (A)In the TCGA cohort, expression of MIR31HG was higher in basal/squamous subtype compared to luminal-infiltrated, luminal-papillary, and neuronal subtypes. (B) Expression of MIR31HG was higher in basal subtype than in luminal and infiltrated subtypes in patients of the TCGA cohort.

3.1.4 Identification and expression of two splicing variants of MIR31HG

Four transcript variants of MIR31HG were retrieved from the NCBI nucleotide database. Transcript variant 1 (RefSeq ID: NR_027054.2) is the full transcript of MIR31HG, containing four exons and three junctions. For simplification, transcript variant 2 (RefSeq ID: NR_152877.1) lacking exon 1 was named MIR31HGΔE1, and transcript variant 4 (RefSeq ID: NR_152879.1) lacking exon 3 was named MIR31HGΔE3. A model of the gene sequence of the MIR31HG

transcript (ENST00000304425.3) and its two splice variants, MIR31HGΔE1 and MIR31HGΔE3, is shown in Figure 6A.
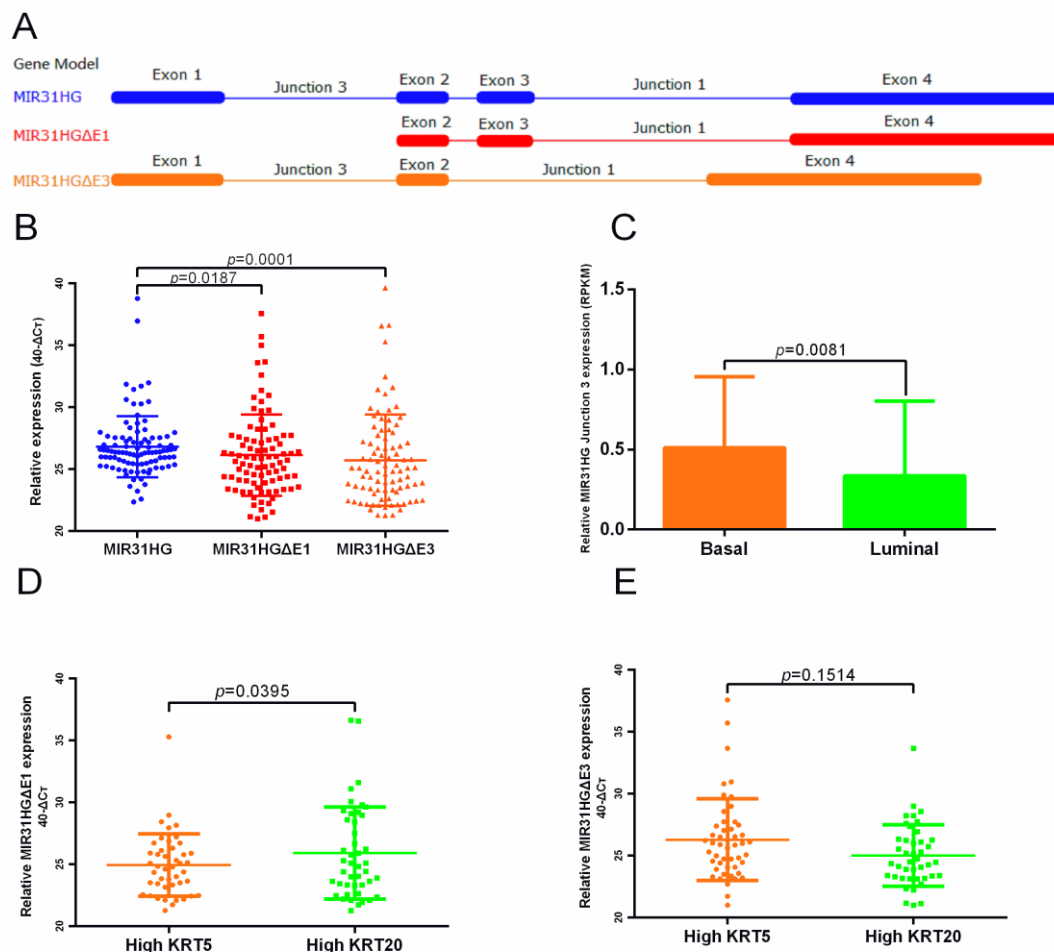

Figure 6. Expression of MIR31HG and two splice variants in the BLCA cohort. (A) Gene model of MIR31HG transcript and splice variants MIR31HGΔE1 and MIR31HGΔE3. (B) In the Mannheim cohort, expression of MIR31HG was significantly higher than in the splice variants, MIR31HGΔE1 and MIR31HGΔE3. (C) In the TCGA cohort, expression of MIR31HGΔE3 was significantly higher in basal compared to luminal subtype. (D) In the Mannheim cohort, expression of MIR31HGΔE1 was higher in the high KRT20 group compared to the high KRT5 group. (E) MIR31HGΔE3 expression was higher in the high KRT5 group compared to the high KRT20 group in the Mannheim cohort.

In the Mannheim cohort, a significantly higher expression of MIR31HG (median expression 34.7 with range of 22.4 to 38.8) was found compared to the transcripts, MIR31HGΔE1 (median expression 35.5 with range of 21.0 to 37.6, p = 0.0187) and MIR31HGΔE3 (median expression 36.7 with range of 21.3 to

39.6, p = 0.0001, Figure 6B). Expression of MIR31HG junction 3, which reflects MIR31HGΔE3 expression, was significant higher in the basal (median expression 0.6267281 with range of 0 to 1.538462) compared to the luminal (median expression 0.3478261 with range of 0 to 1.052632) subtype in the TCGA cohort (p = 0.0081, Figure 6C). KRT5 and KRT20 expression can be used as a marker for basal and luminal phenotypes, respectively. The expression of MIR31HG was analyzed in terms of KRT5 and KRT20 dependent expression. High KRT5 (40-ΔCT value >36.55) and high KRT20 (40-ΔCT value >34.12) were defined as expression above median. MIR31HGΔE1 expression was higher in the high KRT20 group (p = 0.0395, Figure 6D), and MIR31HGΔE3 expression was higher in the high KRT5 group (p = 0.1514, Figure 6E).

3.1.5 ANLN, TLE2 and MIR31HG expressions were associated with copy-number alterations

Analysis of ANLN and TLE2 gene expression in association with copy-number alterations revealed a genetic alteration rate of 10% (41/407) for ANLN and 5% (21/407) for TLE2. Putative copy-number alterations including deep/shallow deletion, diploid, gain, and amplification were acquired from GISTIC. The majority of copy-number alteration signatures for ANLN were gains and amplifications rather than deletions (number 173 vs. 25, 42.51% vs. 6.14%, median expression 10.54 vs. 9.45, and p = 0.0002, Figure 7A). For TLE2, most samples showed deletion variations rather than amplifications (number 143 vs. 54, 35.14% vs. 13.27%, median expression 9.63 vs. 9.77, and p = 0.0394, Figure 7B). Significantly higher expression of ANLN was observed in the subgroup with gain compared to diploid (median expression 10.54 vs. 9.88, and p < 0.0001) and deletion (median expression 10.54 vs. 9.45, and p < 0.0001, Figure 7C). There were no significant differences in TLE2 expression in the subgroups with deletion compared to diploid (median expression 9.63 vs. 9.98,

and p = 0.1458) or gain (median expression 9.63 vs. 9.73, and p = 0.9190, Figure 7D). When gene expression levels were grouped into low and high expression, analysis revealed higher rates of amplification/gains in the subgroup with higher ANLN expression (n = 86, 56.90%) compared to lower ANLN expression (n = 77, 31.95%, Figure 7E). In contrast, amplification was only observed in 12.2% (n = 30) of the TLE2 low expression group and in 14.3% (n = 22) of the TLE2 high expression group (Figure 7F). Furthermore, higher rates of deletion were seen in the TLE2 low expression subgroup than in the TLE2 high expression subgroup (n = 102, 41.46% vs. n = 40, 25.97%).
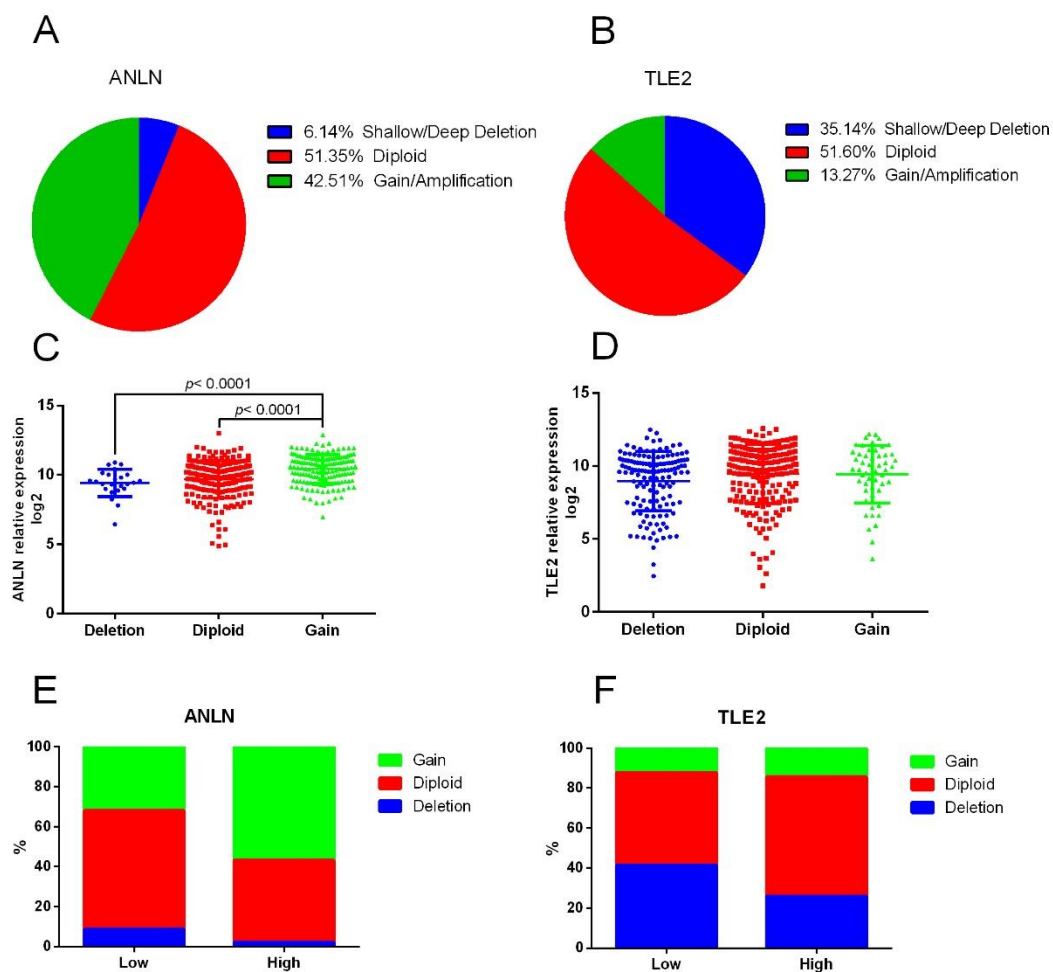
Figure 7. ANLN and TLE2 expression in association with copy-number alterations. (A) ANLN showed 6.14% of shallow/deep deletion, 51.35% of diploid, and 42.51% of gain/amplification. (B) TLE2 showed 35.14% of shallow/deep deletion, 51.60% of diploid and 13.27% of gain/amplification. (C) Significant higher expression of ANLN was observed in the subgroup with gain than diploid (median expression 10.54 vs. 9.88, and $p < 0.0001$) and deletion (median expression 10.54 vs. 9.45, and $p < 0.0001$). (D) No significant differences of TLE2 expression were found in the subgroups with deletion against diploid (median expression 9.63 vs. 9.98, and $p = 0.1458$) or gain (median expression 9.63 vs. 9.73, and $p = 0.9190$). The gene expression levels of ANLN and TLE2 were grouped into high and low expression. (E) In the subgroup with lower ANLN expression, diploid (n = 143, 59.34%) was more frequently observed than gain (n = 77, 31.95%) and deletion (n = 21, 8.71%). The subgroup with higher ANLN expression is associated with a higher percentage of gain (n = 86, 56.90%) than diploid (n = 62, 41.10%) and deletion (n = 3, 2.00%). (F) TLE2 was mainly expressed in the subgroup with deletion and diploid in TLE2 low expression (n = 102, 41.46% in deletion; n = 114, 46.34% in diploid) and TLE2 high expression (n = 40, 25.97% in deletion; n = 92, 59.74% in diploid). Gain was observed in only 12.2% (n = 30) in TLE2 low expression group and in 14.29% (n = 22) in the TLE2 high expression group.

Similar to the analysis for ANLN and TLE2, the association between MIR31HG expression and copy-number alterations were analyzed in the TCGA dataset. The analysis of MIR31HG copy-number alterations revealed a genetic alteration rate of 23% (92/407). Putative copy-number alterations including deep/shallow deletion, diploid, gain, and amplification were also acquired from the GISTIC algorithm. MIR31HG showed 52.97% of shallow / deep deletion (n = 214), 31.44% of diploid (n = 127), and 15.59% of gain/amplification (n = 63, Figure 8A). The expression of MIR31HG was significantly higher in the subgroup with gain (median expression 7.46, $p < 0.0001$) and diploid (median expression 7.12, $p < 0.0001$) compared to those with a deletion alteration (median expression 4.77, Figure 8B). High expression of MIR31HG was found in the subgroup with deletion and diploid alteration than in gain (n = 47, 20.44% in gain; n = 99, 43.04% in diploid; n = 84, 36.52% in deletion), while low expression was observed in the subgroup with deletion alteration, compared to diploid and gain (n = 15, 8.67% with gain; n = 28, 16.18% with diploid; n = 130, 75.15% with deletion, Figure 8C).
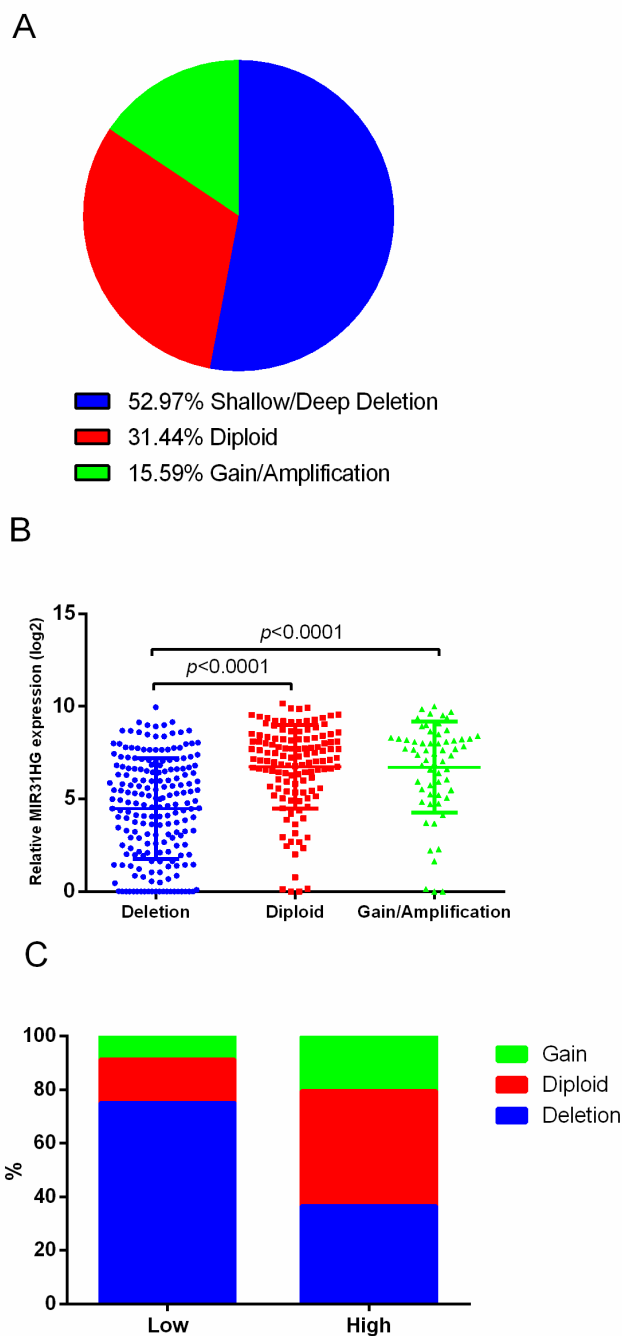
Figure 8. MIR31HG expression in association with copy-number alterations. (A) MIR31HG showed 52.97% of shallow / deep deletion, 31.44% of diploid and 15.59% of gain/amplification. (B) Significantly higher expression of MIR31HG was observed in the subgroup with gain (median expression 7.46) and diploid (median expression 7.12) compared to deletion (median expression 4.77). (C) High expression of MIR31HG was observed in the subgroup with deletion and diploid than in gain (n = 47, 20.44% in gain; n = 99, 43.04% in diploid; n = 84, 36.52% in deletion), while low expression was observed in the subgroup with deletion compared to diploid and gain (n = 15, 8.67% in gain; n = 28, 16.18% in diploid; n= 130, 75.15% in deletion).

3.1.6 ANLN and TLE2 as risk markers for prognostic prediction after RC

After identifying the expression and copy-number alterations of ANLN and TLE2, the association with clinical outcome and parameters was analyzed in the Mannheim and the TCGA cohort. In the Mannheim cohort, patients with high ANLN expression showed worse OS (n = 22 with low expression, n = 38 with high expression; median survival, 21 vs. 10 months, and p = 0.0010) and DSS (n = 20 with low expression, n = 40 with high expression; median survival, 21 vs. 10 months, and p = 0.0060) after RC than patients with low expression (Figure 9A, B). Conversely, patients with high TLE2 expression displayed more favorable OS (n = 29 with high expression, n = 31 with low expression; median survival, 15 vs. 9 months, and p = 0.0236) and DSS (n = 29 with high expression, n = 31 with low expression; median survival, 31 vs. 13 months, and p = 0.2083) than patients with low expression (Figure 9C, D). The gene expression levels were not normally distributed.
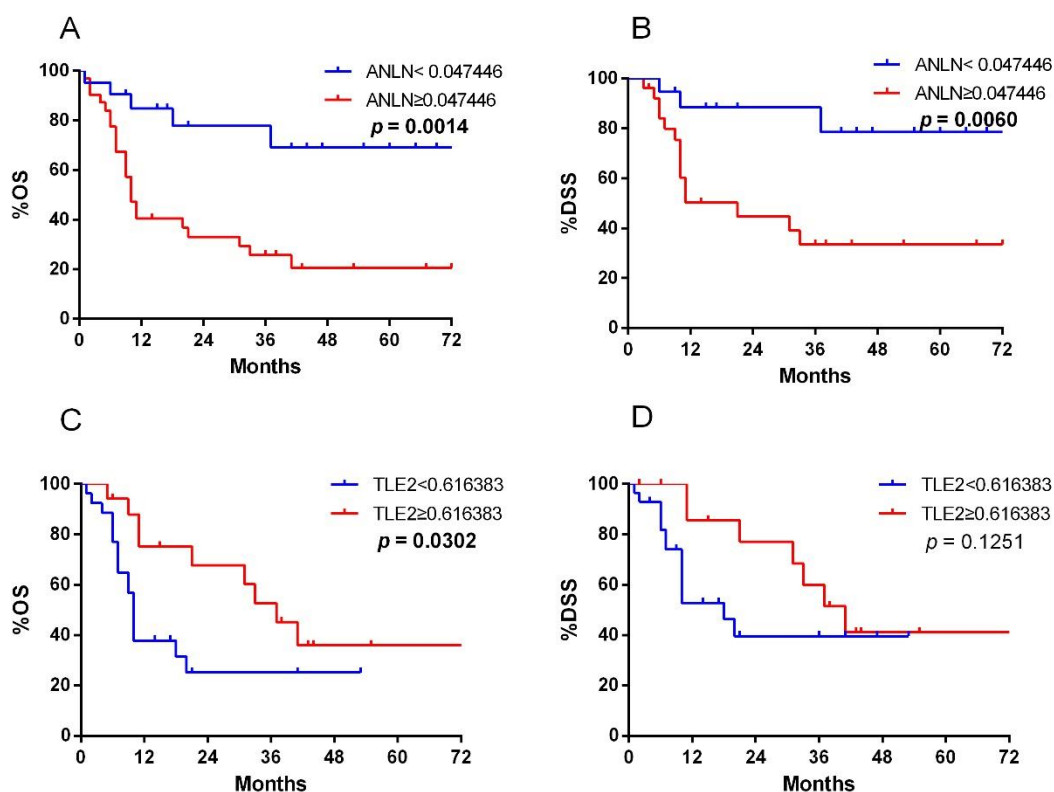
Figure 9. Kaplan–Meier plots of overall survival (OS) and disease-free survival (DSS) survival associated with ANLN and TLE2 risk stratification in the Mannheim cohort. (A, B) The group with high ANLN expression showed worse OS and DSS than the group with low expression. (C, D) The group with high TLE2 expression displayed more favorable OS and DSS than the group with low expression.

To further understand the outcome significance of ANLN and TLE2 in BLCA, multiple regression analysis were performed. In the univariable and multivariable Cox regression analysis, ANLN ($p = 0.0020$ and $p = 0.0390$, respectively) and TLE2 ($p = 0.0120$ and $p = 0.0020$, respectively) expression were independent predictors. Furthermore, lymph node status was identified as an independent prognostic factor by both analyses ($p = 0.0030$ and $p = 0.0240$), LVI ($p = 0.0040$) was only by the univariable Cox regression analysis independent (Table 13). No significant correlation was observed between patient age, patient gender, or stage of tumor and ANLN and TLE2 expression.

Table 13. Uni- and multivariable Cox regression analysis of ANLN and TLE2 with clinicopathological features in the Mannheim cohort.

Factor	Univariable		Multivariable	
	HR (95% CI)	*p* Value	HR (95% CI)	*p* Value
Diagnosis Age				
<70 vs. ≥70	0.584 (0.269–1.269)	0.174	0.874 (0.359–2.127)	0.767
Gender				
Male vs. Female	1.128 (0.405–3.140)	0.957	1.820 (0.587–5.643)	0.300
Stage				
T1/T2 vs. T3/4	0.168 (0.015–1.832)	0.128	0.360 (0.127–1.023)	0.055
LVI				
Negative vs. Positive	0.542 (0.359–0.819)	**0.004**	0.380 (0.048–2.998)	0.358
Lymph node Statues				
Negative vs. Positive	0.549 (0.371–0.813)	**0.003**	0.612 (0.399–0.938)	**0.024**
ANLN				
Low vs. High	0.220 (0.084–0.575)	**0.002**	0.328 (0.114–0.945)	**0.039**
TLE2				
Low vs. High	0.305 (0.121–0.769)	**0.012**	0.172 (0.057–0.519)	**0.002**

Note: HR = hazard ratio, CI = confidence interval, LVI = lymphovascular invasion, significant *p* values are bold.

These results were validated in the TCGA cohort. Higher ANLN expression was associated with worse OS (n = 246 with low expression, n = 161 with high expression; median survival, 18.07 vs. 15.31 months, and *p* = 0.0144) and DFS (n = 246 with low expression, n = 161 with high expression; median survival, 18.82 vs. 13.99 months, and *p* = 0.0045, Figure 10A, B). In contrast, higher TLE2 expression was associated with more favorable OS (n = 250 with low expression, n = 157 with high expression; median survival, 16.69 vs. 18.99 months, and *p* = 0.0054) and DFS (n = 250 with low expression, n = 157 with

high expression; median survival, 15.31 vs. 16.79 months, and $p = 0.0094$, Figure 10C, D). Further analysis of the TCGA BLCA cohort according to clinical stage showed that lower ANLN (n = 77 with low expression, n = 41 with high expression; median survival, 23.43 vs. 19.66 months, and $p = 0.0397$) and TLE2 (n = 67 with low expression, n = 51 with high expression; median survival, 17.94 vs. 20.37 months, and $p = 0.0100$) expression in the pT2 subgroup could be attributed to a good and poor prognosis of OS, respectively (Figure 10E, F). Furthermore, the gene expression of ANLN and TLE2 was analyzed in subtypes of MIBC. Notably, in the TCGA cohort, ANLN expression in patients with MIBC and basal subtype was associated with worse OS (n = 61 with low expression, n = 81 with high expression; median survival, 20.37 vs. 13.96 months, and $p = 0.0467$, Figure 10G) compared with the whole TCGA cohort. Higher TLE2 expression showed better OS in patients with MIBC and luminal subtype (n = 130 with high expression, n = 115 with low expression; median survival, 19.48 vs. 17.87 months, and $p = 0.0181$, Figure 10H) compared with the whole TCGA cohort. The gene expression levels were not normally distributed.
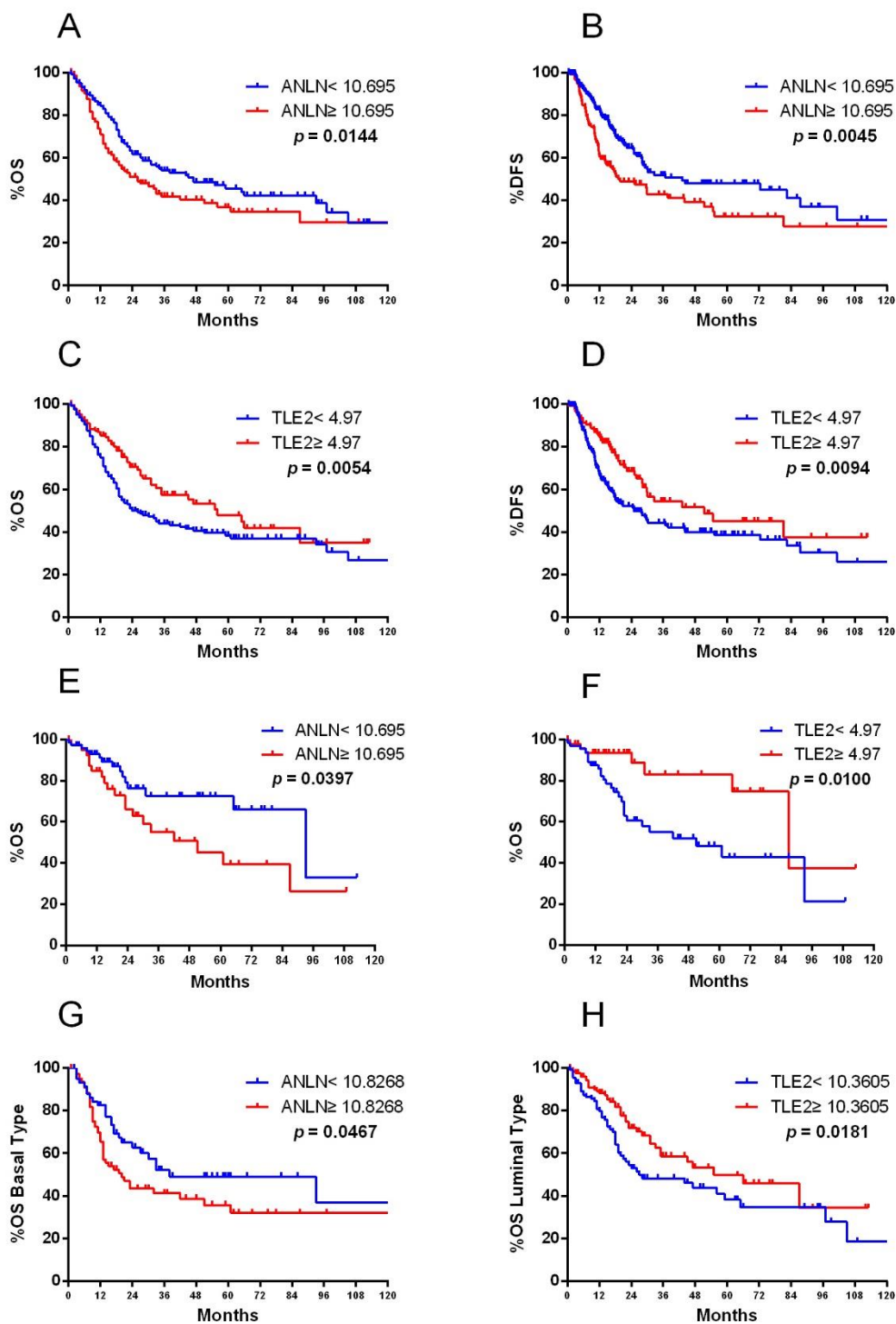
Figure 10. Kaplan–Meier plots of OS and DFS associated with ANLN and TLE2 risk stratification in the TCGA cohort. (A, B) Higher ANLN expression showed worse OS and DFS. (C, D) Higher TLE2 expression showed more favorable OS and DFS. (E, F) In the T2 subgroup, low ANLN and TLE2 expression showed good and poor prognosis of OS, respectively. (G, H) Higher ANLN expression showed worse OS in basal subtype and higher TLE2 showed better OS in luminal subtype.

48

Besides, in the TCGA cohort, correlation between ANLN and TLE2 expression and stage showed that ANLN was significantly expressed in pT3 and pT4 (higher stages, median expression 9.95 vs. 10.16, and $p = 0.0109$, Figure 11A), while TLE2 was dominantly expressed in pT2 (lower stage, median expression 9.88 vs. 9.73, and $p = 0.0228$, Figure 11B).
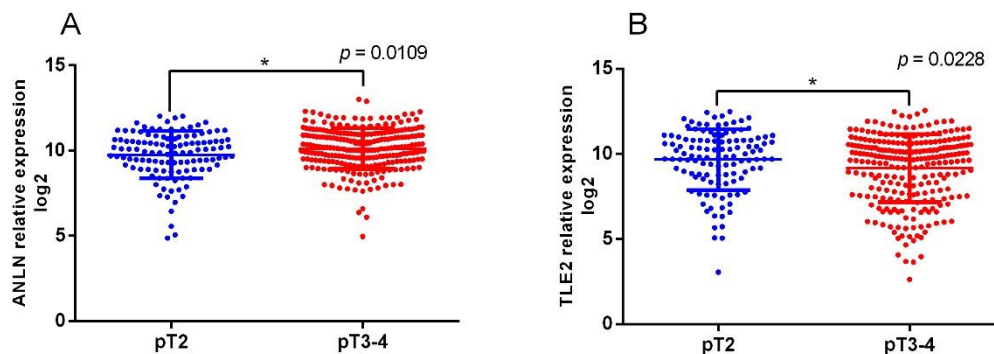

Figure 11. Correlation between ANLN and TLE2 expression and stage from the TCGA cohort. (A) The TCGA cohort showed that ANLN was significantly expressed in higher pT3-4 stages (median expression 9.95 vs. 10.16, $p = 0.0109$). (B) TLE2 was dominantly expressed in lower pT2 stage (median expression 9.88 vs. 9.73, $p = 0.0228$).

In the univariable Cox regression analysis of clinicopathological features, stage ($p < 0.001$), lymph node status ($p < 0.001$), ANLN ($p = 0.0160$), and TLE2 ($p = 0.0060$) expression were independent. Furthermore, stage ($p = 0.0060$), ANLN ($p = 0.0180$), and TLE2 ($p = 0.0400$) were also independent prognostic factors in the multivariable Cox regression analysis (Table 14).

Table 14. Uni- and multivariable Cox regression analysis of ANLN and TLE2 with clinicopathological features in the TCGA cohort.

Factor	Univariable		Multivariable	
	HR (95% CI)	*p* Value	HR (95% CI)	*p* Value
Diagnosis Age				
<70 vs. ≥70	1.260 (0.931–1.705)	0.134	1.195 (0.701–2.038)	0.513
Gender				
Male vs. Female	1.257 (0.902–1.751)	0.177	1.531 (0.724–3.240)	0.265
Smoking Status				
No vs. Yes	1.335 (0.940–1.897)	0.106	1.851 (0.503–6.811)	0.355
Stage				
T2 vs. T3/4	1.950 (1.393–2.731)	**<0.001**	1.646 (1.156–2.342)	**0.006**
Lymph node Statues				
Negative vs. Positive	2.145 (1.596–2.883)	**<0.001**	1.989 (1.461–2.707)	**<0.001**
ANLN				
Low vs. High	1.439 (1.070–1.934)	**0.016**	1.438 (1.064–1.943)	**0.018**
TLE2				
Low vs. High	0.636 (0.460–0.880)	**0.006**	1.415 (1.015–1.973)	**0.040**

Note: HR = hazard ratio, CI = confidence interval, significant *p* values are bold.

3.1.7 MIR31HG as prognostic marker for MIBC patients with basal subtype

Unlike ANLN and TLE2, no significant association between expression of MIR31HG and patient outcome was found in the whole TCGA cohort. Notably, survival analysis showed that patients of the TCGA cohort with basal subtype were significantly associated with OS and DFS based on MIR31HG risk stratification. The group with high MIR31HG expression showed a worse OS compared to the group with low expression (median survival, 28 vs. 16 months, *p* = 0.0081, Fig. 12A). Furthermore, the group with high MIR31HG expression

showed a worse DFS compared to the group with low expression (median survival, 27 vs. 15 months, *p* = 0.1383, Fig. 12B).
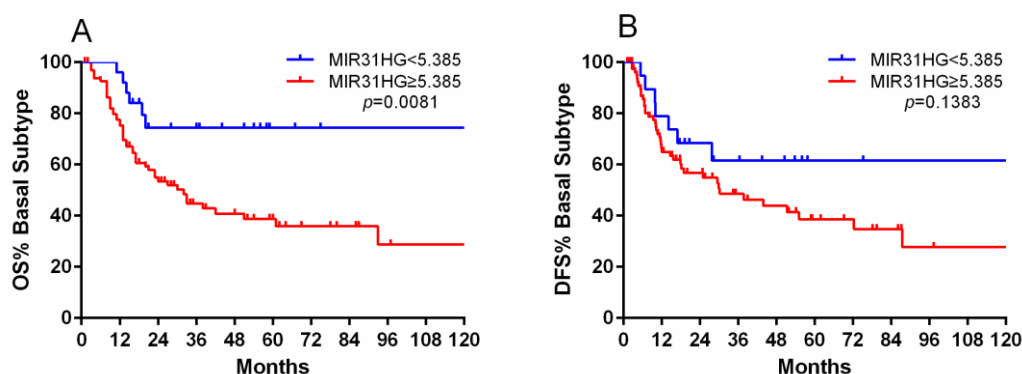

Figure 12. Kaplan-Meier plot of OS and DFS associated with MIR31HG risk stratification. (A) In the TCGA cohort with basal subtype, the group with high MIR31HG expression showed a worse OS compared to the group with low expression (median survival, 28 vs. 16 months, *p* = 0.0081). (B) In the TCGA cohort with basal subtype, the group with high MIR31HG expression showed a worse DFS compared to the group with low expression (median survival, 27 vs. 15 months, *p* = 0.1383).

3.1.8 Splicing variants of MIR31HG as prognostic marker for patients with MIBC

Kaplan-Meier analysis and the log-rank test were used to evaluate the association of the expression of two MIR31HG splice variants with OS and DFS in the Mannheim cohort. For the full-length transcript of MIR31HG, no significant difference in OS (median survival 18 vs. 21 months, *p* = 0.2610, Figure 13A) and DFS (median survival 11 vs. 9 months for MIR31HG, *p* = 0.8978, Figure 13B) was found between high and low expression levels. Tumors with both high MIR31HGΔE1 (median survival 15 vs. 38 months, *p* = 0.0394, Figure 13C) and MIR31HGΔE3 expression (median survival 12 vs. 30 months, *p* = 0.0093, Figure 13E) showed a worse OS. The groups with high MIR31HGΔE1 (median survival 7 vs. 25 months, *p* = 0.0038, Figure 13D) and MIR31HGΔE3 (median survival 9 vs. 15 months, *p* = 0.1252, Figure 13F) expression also showed a worse DFS. Tumors with high expression of

MIR31HG Exon1-2 (Junction 3), which could partially present as MIR31HGΔE3, showed a worse OS than the group with low expression, in the TCGA cohort with basal subtype (median survival 15 vs. 17 months, $p = 0.0298$, Figure 13G). No significant difference was found for DFS (median survival 12 vs. 15 months, $p = 0.5670$, Figure 13H).
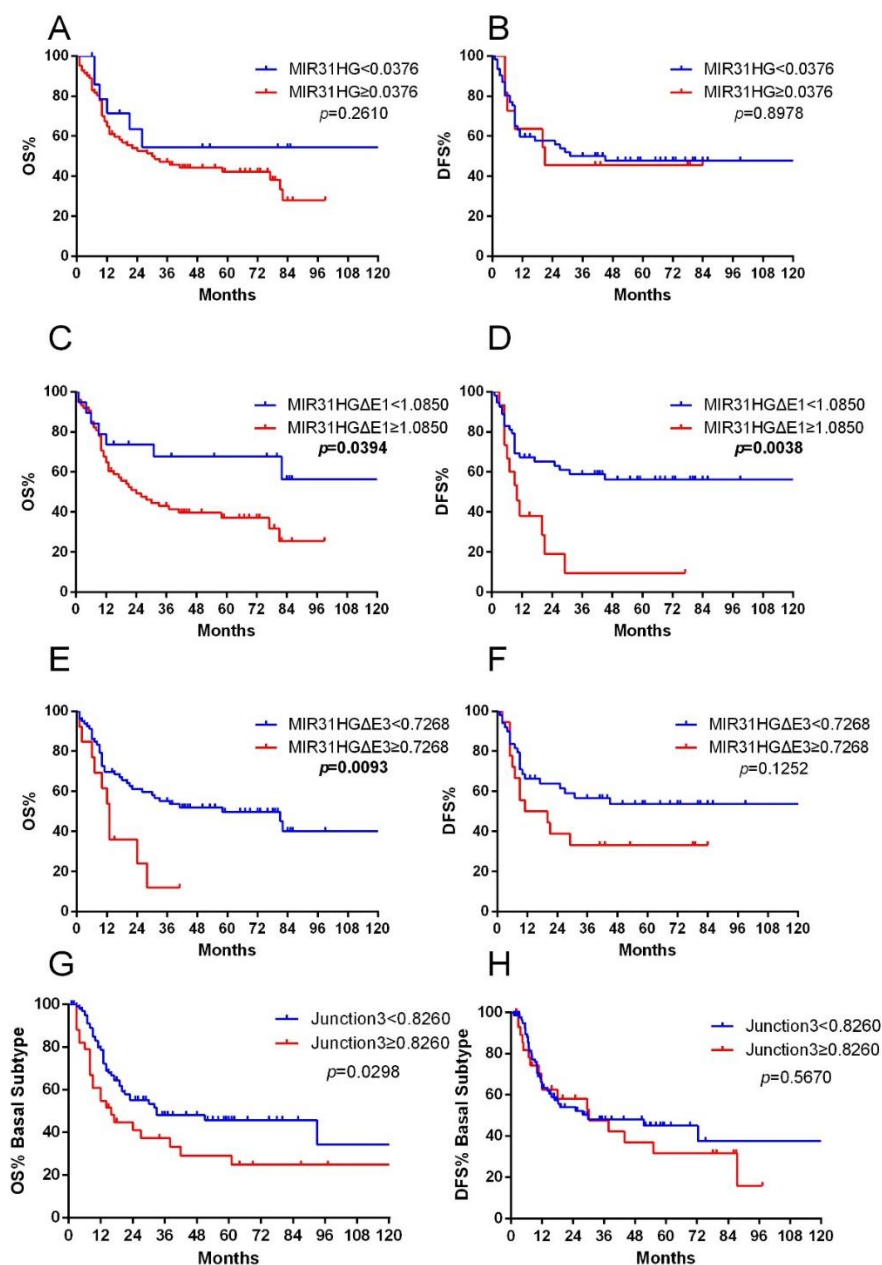
Figure 13. Kaplan-Meier plot of OS and DFS associated with MIR31HG and its splice variants risk stratification. (A, B) In the Mannheim cohort, no significant correlation was found with OS (median survival 18 vs. 21 months) and DFS (median survival 11 vs. 9 months for MIR31HG) in the group with full-length transcript of MIR31HG. (C, D) The group with high MIR31HGΔE1 expression showed a worse OS (median survival 15 vs. 38 months) and DFS (median survival 7 vs. 25 months) compared to the group with low expression. (E, F) The group with high MIR31HGΔE3 expression showed a worse OS (median survival 12 vs. 30 months) and DFS (median survival 9 vs. 15 months) compared to the group with low expression. (G) In the TCGA cohort, the group with high Junction 3 expression showed worse OS than the group with low expression (median survival, 17 vs. 14 months, $p = 0.0298$). (H) No significant difference was observed in DFS between the group with high and low Junction 3 expression (median survival, 17 vs. 15 months, $p = 0.5670$).

In the univariable Cox regression analysis regarding the Mannheim cohort, MIR31HGΔE1 (p = 0.032) and MIR31HGΔE3 (p = 0.007) expression, as well as the stage of cancer, were found to be predictive for patient outcome. In the multivariable Cox regression analysis regarding the Mannheim cohort, only MIR31HGΔE3 was identified as an independent prognostic factor (p = 0.016). No significant correlation was observed between the full-length transcript of MIR31HG and OS of the patients in both analyses, nor with patient age, patient gender, lymph node status, and LVI (Table 15).

Table 15. Uni- and multivariable cox regression analysis of MIR31HG and its splice variants with clinicopathological features in the Mannheim cohort.

Factor	Univariable		Multivariable	
	HR (95% CI)	p	HR (95%CI)	p
Diagnosis Age				
<70 vs. ≥70	0.693(0.401-1.198)	0.189	0.524(0.263-1.043)	0.066
Gender				
Male vs. Female	0.907(0.495-1.662)	0.752	1.077(0.538-2.156)	0.834
Stage				
T2 vs. T3/4	0.401(0.202-0.797)	**0.009**	0.581(0.278-1.216)	0.149
LVI				
Negative vs. Positive	0.608(0.083-4.473)	0.625	0.482(0.061-3.787)	0.487
Lymphnode Statues				
Negative vs. Positive	0.724(0.416-1.260)	0.253	0.936(0.486-1.800)	0.842
MIR31HG				
Low vs. High	1.614(0.688-3.784)	0.271	1.943(0.574-6.577)	0.286
MIR31HGΔE1				
Low vs. High	2.249(0.999-5.067)	**0.032**	1.535(0.642-3.671)	0.335
MIR31HGΔE3				
Low vs. High	2.679(1.312-5.472)	**0.007**	2.507(1.183-5.311)	**0.016**

Note: HR = hazard ratio, CI = confidence interval, LVI = lymphovascular invasion, significant p values are bold.

3.2 Function of ANLN, TLE2 and MIR31HG in BLCA tumorigenesis

3.2.1 ANLN and TLE2 are associated with signaling pathways and therapeutic targets in BLCA

Beside the expression and survival analyses, the functional role was also analyzed. Thus, the protein-protein interactions of ANLN and TLE2 were analyzed by Search Tool for Retrieval of Interacting Genes/Proteins (STRING, Figure 14).
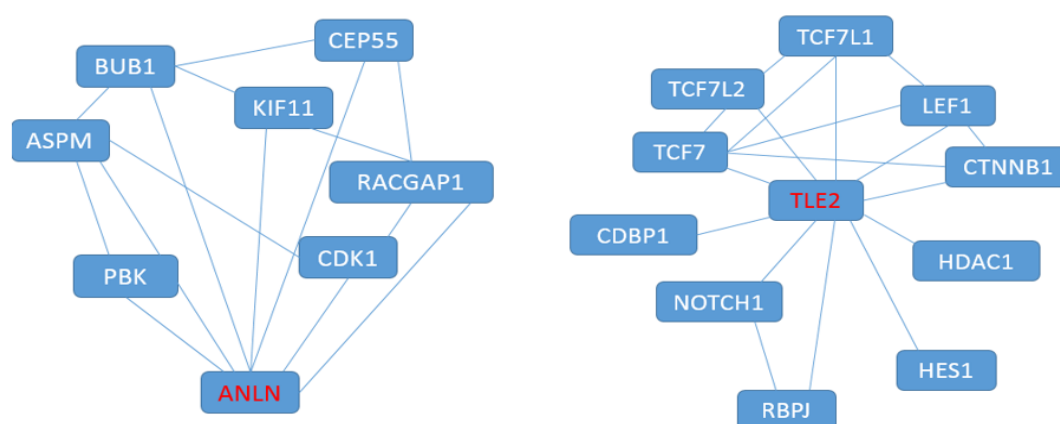

Figure 14. Protein-protein interactions of ANLN and TLE2. Protein-protein interactions predicted by STRING showed the interaction network of ANLN and TLE2 based on curated databases and experimental determination.

ANLN protein was predicted to have interactions with abnormal spindle protein homolog (ASPM), mitotic checkpoint serine/threonine-protein kinase (BUB1), centrosomal protein 55 (CEP55), kinesin family member 11 (KIF11), Rac GTPase activating protein 1 (RACGAP1), cyclin dependent kinase 1 (CDK1), and serine/threonine protein kinase (PBK). TLE2 protein was predicted to have interactions with transcription factor 7 (TCF7), transcription factor 7-like 1 (TCF7L1), transcription factor 7-like 2 (TCF7L2), lymphoid enhancer-binding factor 1 (LEF1), catenin beta-1 (CTNNB1), cyclin D-type binding protein 1 (CDBP1), Notch homolog 1 (NOTCH1), histone deacetylase 1 (HDAC1), hairy and enhancer of split-1 (HES1), and recombination signal binding protein for immunoglobulin kappa J region (RBPJ). The number of counted gene sets and

the false discovery rate (FDR) for each Gene Ontology (GO) term are shown in Table 16.

Table 16. GO enrichment analysis by Enrichr indicated the significant GO terms for ANLN and TLE2 including biological process, molecular function, cellular component, and KEGG pathways.

Gene	Dataset	Pathway ID	Pathway Description	Count in Gene Set	False Discovery Rate
ANLN	Biological Process (GO)	GO:0007067	mitotic nuclear division	10	1.42E-13
		GO:0051301	cell division	9	1.99E-10
		GO:1903047	mitotic cell cycle process	9	7.57E-09
		GO:0000278	mitotic cell cycle	9	1.69E-08
		GO:0051302	regulation of cell division	7	2.14E-08
	Molecular Function (GO)	GO:0005524	ATP binding	6	0.0483
		GO:0043168	anion binding	7	0.0483
	Cellular Component (GO)	GO:0005819	spindle	7	3.16E-08
		GO:0044430	cytoskeletal part	9	4.71E-07
		GO:0015630	microtubule cytoskeleton	8	1.33E-06
		GO:0030496	midbody	5	1.33E-06
		GO:0072686	mitotic spindle	4	1.64E-06
TLE2	Biological Process (GO)	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	9	2.39E-08
		GO:0043588	skin development	7	2.39E-08
		GO:0048864	stem cell development	7	2.93E-08
		GO:2000736	regulation of stem cell differentiation	6	3.28E-08
		GO:0009790	embryo development	9	5.72E-08
	Molecular Function (GO)	GO:0003700	transcription factor activity, sequence-specific DNA binding	9	7.79E-08
		GO:0044212	transcription regulatory region DNA binding	8	7.79E-08
		GO:0003682	chromatin binding	7	5.90E-07
		GO:0001047	core promoter binding	5	2.16E-06
		GO:0070491	repressing transcription factor binding	4	4.10E-06
	Cellular Component (GO)	GO:0005667	transcription factor complex	7	3.16E-08
		GO:0005654	nucleoplasm	11	1.54E-07
		GO:0031981	nuclear lumen	10	3.92E-05
		GO:0070369	beta-catenin-TCF7L2 complex	2	8.60E-05
		GO:0002193	MAML1-RBP-Jkappa- ICN1 complex	2	0.000129
	KEGG Pathways	5216	Thyroid cancer	5	3.57E-10
		4310	Wnt signaling pathway	6	2.38E-09
		4330	Notch signaling pathway	5	2.38E-09
		5213	Endometrial cancer	5	2.38E-09
		5200	Pathways in cancer	7	2.75E-09

The interaction network based on curated databases and experimentally derived results showed that key molecules in signaling pathways were significantly correlated with ANLN and TLE2, including cell proliferation (FDR = 1.99^{-10}), Notch signaling (FDR = 2.38^{-9}), Wnt signaling (FDR = 2.38^{-9}), and hormone receptor (FDR = 3.16^{-8}). The correlation with important therapeutic targets in BLCA, including epidermal growth factor receptor (EGFR) ($p = 7.73^{-31}$ for ANLN and $p = 1.78^{-14}$ for TLE2), Erb-B2 receptor tyrosine kinase 2 (ERBB2) ($p = 3.44^{-13}$ for ANLN and $p = 1.96^{-43}$ for TLE2), fibroblast growth factor receptor 3 (FGFR3) ($p = 6.52^{-9}$ for ANLN and $p = 0.0016$ for TLE2), and programmed death-ligand 1 (PD-L1) ($p = 2.89{-}18$ for ANLN and $p = 1.07^{-18}$ for TLE2) were found to be significantly correlated with ANLN and TLE2 based on TCGA data (Table 17).

Table 17. Correlation of ANLN and TLE2 with key molecules in signaling pathways and therapeutic targets.

Correlated Gene		ANLN		TLE2	
		Correlation Coefficient	*p* Value	Correlation Coefficient	*p* Value
Cell proliferation	CDK1	**0.594**	3.98×10^{-40}	−0.338	2.63×10^{-12}
	RACGAP1	**0.725**	1.41×10^{-67}	**−0.451**	9.46×10^{-22}
	MKI67	**0.711**	7.56×10^{-64}	−0.396	1.07×10^{-16}
	FOXM1	**0.688**	2.98×10^{-58}	**−0.4**	4.46×10^{-17}
Notch signaling	NOTCH1	0.109	0.027294	−0.19	0.000115
	RBPJ	−0.196	6.6×10^{-5}	0.05	0.311468
Wnt signaling	TCF7	0.067	0.175097	0.006	0.89689
	TCF7L1	0.162	0.001066	−0.267	4.75×10^{-8}
	TCF7L2	−0.186	0.000164	0.23	2.84×10^{-6}
	LEF1	−0.101	0.041347	−0.026	0.596286
	CTNNB1	0.237	1.29×10^{-6}	−0.276	1.6×10^{-8}
Hormone receptor signaling	AR	−0.185	0.000171	0.388	4.9×10^{-16}
	ESR1	−0.003	0.944815	−0.048	0.332592
	ESR2	−0.191	0.00011	0.334	4.33×10^{-12}
	FOXA1	−0.38	1.89×10^{-15}	**0.505**	9.11×10^{-28}
	GATA3	**−0.403**	2.33×10^{-17}	**0.65**	2.80×10^{-50}
Therapeutic targets	EGFR	**0.53**	7.73×10^{-31}	−0.368	1.78×10^{-14}
	ERBB2	−0.35	3.44×10^{-13}	**0.613**	1.96×10^{-43}
	FGFR3	−0.283	6.52×10^{-9}	0.156	0.001575
	PIK3CA	**0.426**	2.19×10^{-19}	−0.264	6.34×10^{-8}
	CDK4	0.207	2.68×10^{-5}	−0.267	4.69×10^{-8}
	HRAS	−0.023	0.648147	−0.183	0.000202
	PDCD1	0.0881	0.0754	−0.313	1.07×10^{-10}
	PD-L1	**0.414**	2.89×10^{-18}	**−0.419**	1.07×10^{-18}
	CTLA4	0.0963	0.052	−0.363	3.71×10^{-14}
	EZH2	**0.446**	2.51×10^{-21}	−0.162	0.001009

Note: Correlation coefficient values above 0.4 and below −0.4 are bold.

It is noteworthy that ANLN was positively correlated with cell proliferation markers including CDK1 ($\rho = 0.594$, $p = 3.98^{-40}$), RACGAP1 ($\rho = 0.725$, $p = 1.41^{-67}$), marker of proliferation Ki-67 (MKI67) ($\rho = 0.711$, $p = 7.56^{-64}$), and forkhead box M1 (FOXM1) ($\rho = 0.688$, $p = 2.98^{-58}$). In contrast, TLE2 was negatively correlated with cell proliferation molecules involving CDK1 ($\rho = -0.338$, $p = 2.63^{-12}$), RACGAP1 ($\rho = -0.451$, $p = 9.46^{-22}$), MKI67 ($\rho = -0.396$, $p = 1.07^{-16}$), and FOXM1 ($\rho = -0.400$, $p = 4.46^{-17}$). TLE2 was also correlated with

molecules involved in Wnt signaling including CTNNB1 (ρ = -0.276, p = 1.6^{-8}) and hormone receptors including forkhead box A1 (FOXA1) (ρ = 0.505, p = 9.11^{-28}) and GATA binding protein 3 (GATA3) (ρ = 0.65, p = 2.8^{-50}; Table 18). Further, expression of RACGAP1, MKI67, FOXM1, CDK1, CTNNB1, and GATA3 were examined in the Mannheim cohort. This analysis showed that ANLN was significantly correlated with RACGAP1 (ρ = 0.455, p < 0.0001), FOXM1 (ρ = 0.549, p < 0.0001), MKI67 (ρ = 0.577, p < 0.0001), and CDK1 (ρ = 0.763, p < 0.0001; Figure 15A–D). TLE2 was significantly correlated with GATA3 (ρ = 0.409, p = 0.0012) and CTNNB1 (ρ = 0.363, p = 0.0070) (Figure 15E, F). Similar expression data was also detected in the benigne and malignant urothelial cell lines.
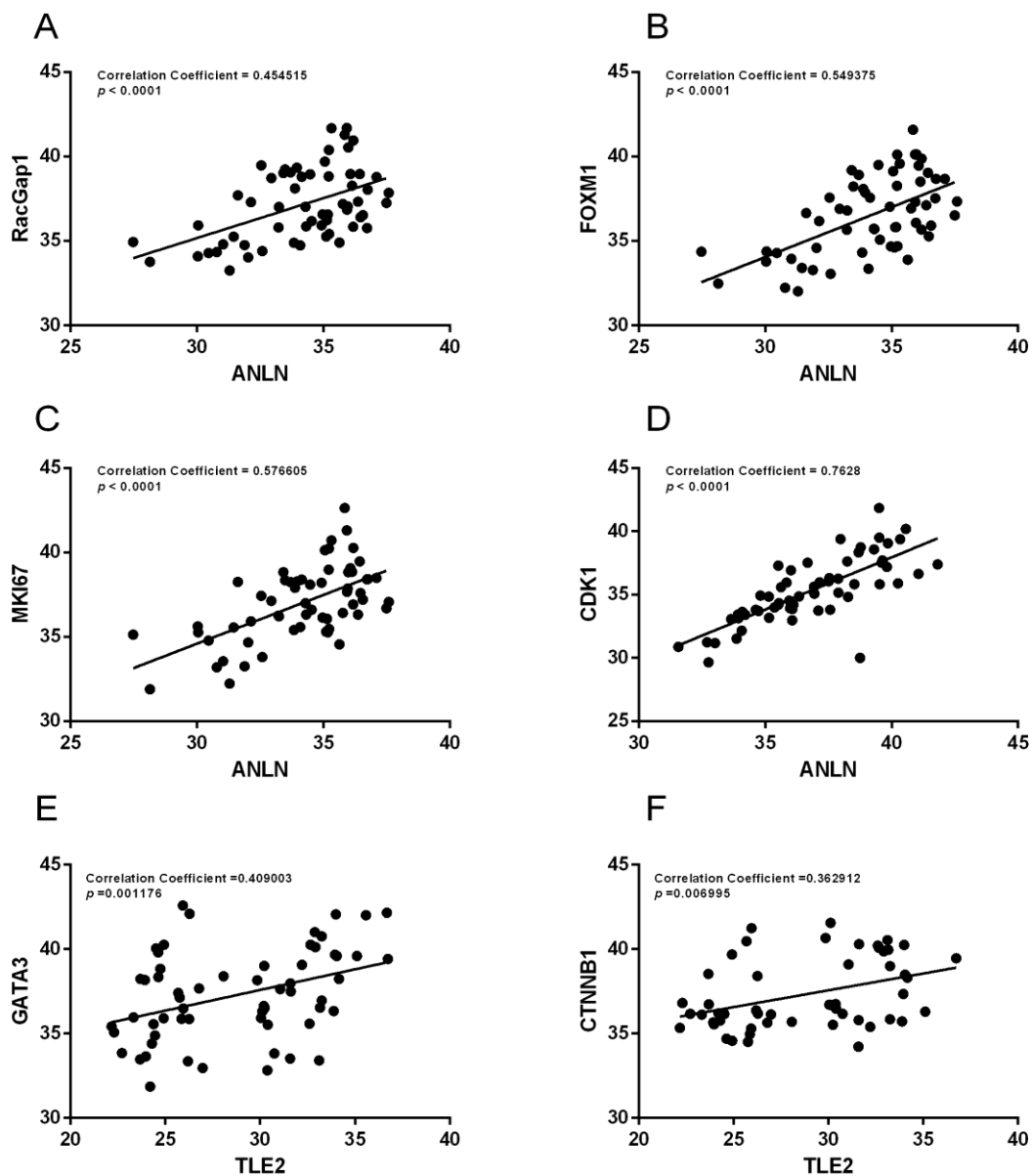
Figure 15. Correlation of selected genes in the Mannheim cohort. ANLN was significantly correlated with RACGAP1 (A), FOXM1 (B), MKI67 (C), and CDK1 (D). TLE2 was significantly correlated with GATA3 (E) and CTNNB1 (F).

3.2.2 Molecular subtype specificity of ANLN and TLE2

To comprehensively elucidate the role of ANLN and TLE2 in BLCA, cell lines and tissues of different molecular subtype were analyzed. Published RNA-seq data from the Cancer Cell Line Encyclopedia showed ANLN and TLE2 expression levels in TPM (transcripts per million) for 25 BLCA cell lines, with

different molecular subtypes of each (basal, luminal, and mixed; Figure 16B). Cell lines that were classified as basal subtypes, including UMUC3 and SCaBER, showed slightly higher expression of ANLN, while TLE2 dominantly expressed in cell lines classified as luminal subtypes, e.g., RT112 and RT4. qPCR based on SYBR Green showed relative expression of ANLN and TLE2 in five malignant urothelial cell lines (RT4, RT112, UMUC3, T24, and ScaBER), which correspond with the RNA-seq data (Figure 16A). Similarly, in patients with BLCA (TCGA, Provisional), expression of ANLN was higher in the basal than in the luminal subtypes (median expression 10.93; range 6.1–13.01 vs. median expression 9.64; range 4.87–12.25, $p < 0.0001$, Figure 16C) according to the mRNA clustering. In contrast, TLE2 showed an opposite trend with a higher expression in luminal than basal subtypes (median expression 10.41; range 5.04–12.55 vs. median expression 7.41, range 1.79–12.25, $p < 0.0001$, Figure 16D).
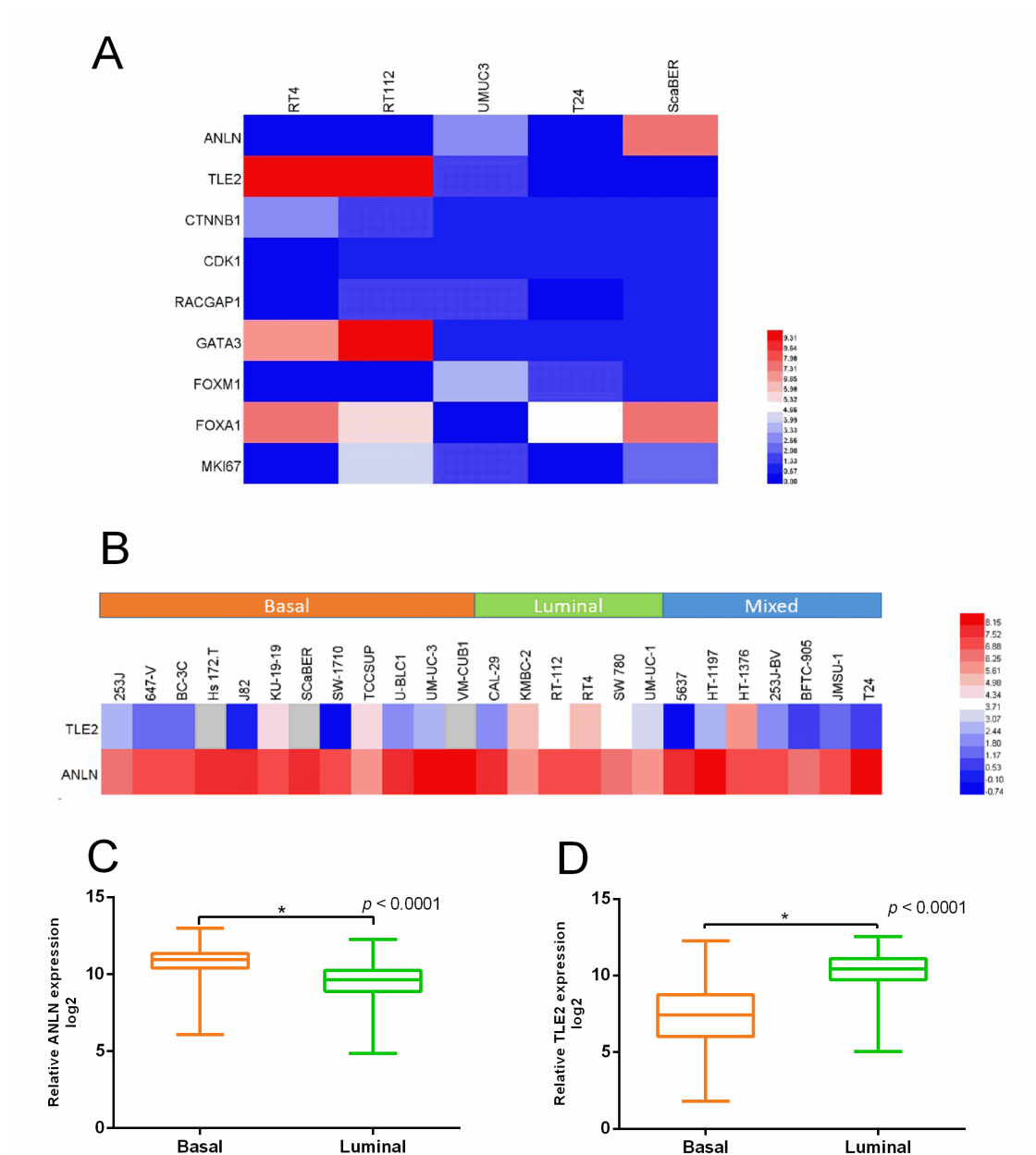
Figure 16. Expression of ANLN and TLE2 in different subtypes of BLCA. (A) qPCR showed relative expression of ANLN and TLE2 in five malignant urothelial cell lines (RT4, RT112, UMUC3, T24, and ScaBER). (B) *In silico* RNA-seq data from the Cancer Cell Line Encyclopedia showed expression level in TPM (transcripts per million) for ANLN and TLE2 in 25 BLCA cell lines with different molecular subtypes of each. (C) Expression of ANLN was higher in basal than luminal subtype (median expression 10.93 with range of 6.1 to 13.01 vs. median expression 9.64 with range of 4.87 to 12.25, $p < 0.0001$) in patients with BLCA. (D) TLE2 expression was higher in luminal than basal subtype (median expression 10.41 with range of 5.04 to 12.55 vs. median expression 7.41 with range of 1.79 to 12.25, $p < 0.0001$).

In addition, the basal subtype marker KRT5 and the luminal subtype marker KRT20 were analyzed in the Mannheim cohort. KRT5 was significantly

correlated with ANLN (ρ = 0.278, p = 0.042; Figure 17A), and KRT20 was significantly correlated with TLE2 (ρ = 0.296, p = 0.026; Figure 17B). The expression of ANLN and TLE2 were also analyzed in the group with high KRT5 and the high KRT20 expression, which were classified based on median expression of KRT5 and KRT20. It was observed that ANLN expression was higher in the high KRT5 group than in the high KRT20 group (p = 0.1119, Figure 17C), and TLE2 expression was higher in the high KRT20 group than in the high KRT5 group (p = 0.1413, Figure 17D).
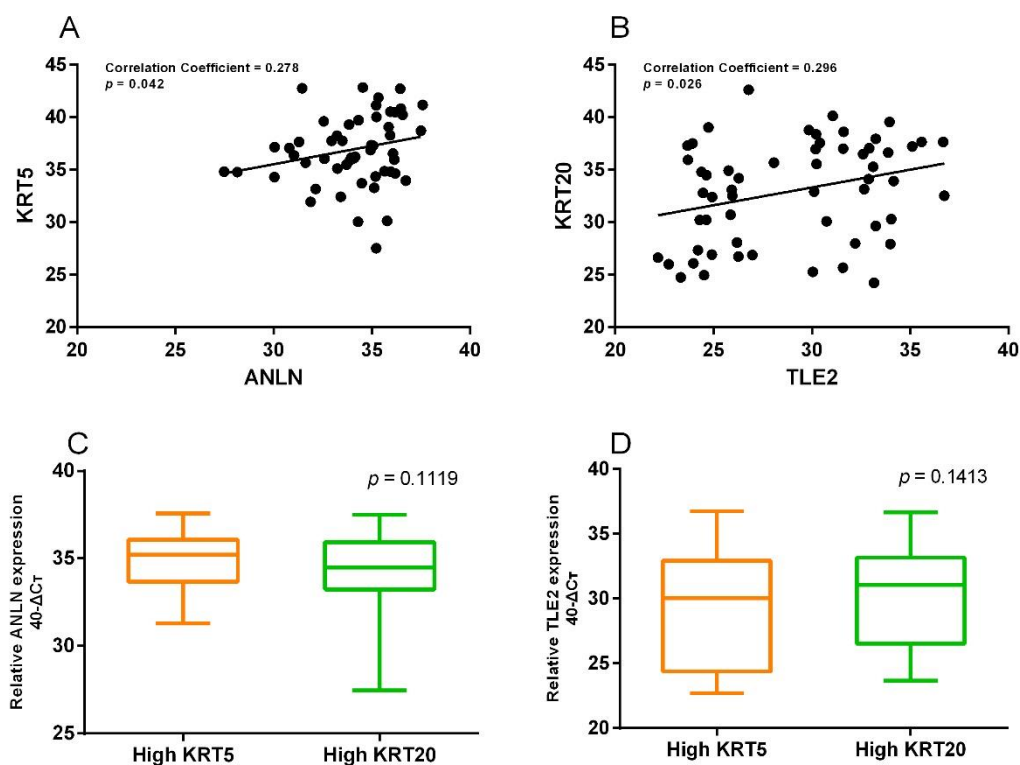
Figure 17. ANLN and TLE2 expression was associated with basal and luminal subtype markers. Correlation of KRT5 and ANLN (A) and KRT20 and TLE2 (B) in the Mannheim cohort. (C) ANLN expression was higher in the high KRT5 group than in the high KRT20 group (p = 0.1119). (D) TLE2 expression was higher in the high KRT20 group than in the high KRT5 group (p = 0.1413).

3.2.3 Expression of MIR31HG in BLCA cell lines with specificity

Similar as ANLN and TLE2, the role of MIR31HG was assessed in BLCA cell lines with different subtypes. Expression of MIR31HG was detected in a normal urothelial cell line (UROtsa) and three BLCA cell lines (SCaBER, UMUC3 and T24) by qRT-PCR. Significantly lower expression of MIR31HG was observed in UMUC3 (mean expression 33.59, $p = 0.0003$) and T24 cells (mean expression 32.34, $p < 0.0001$) compared with UROtsa cells (mean expression 23.64). In contrast, expression of MIR31HG was higher in SCaBER cells (mean expression 35.53, $p = 0.0200$) compared to UROtsa cells. The lowest level of MIR31HG expression was observed in RT112 (mean expression 21.03, $p < 0.0001$) and RT4 cells (mean expression 34.01, $p < 0.0001$, Figure 18).
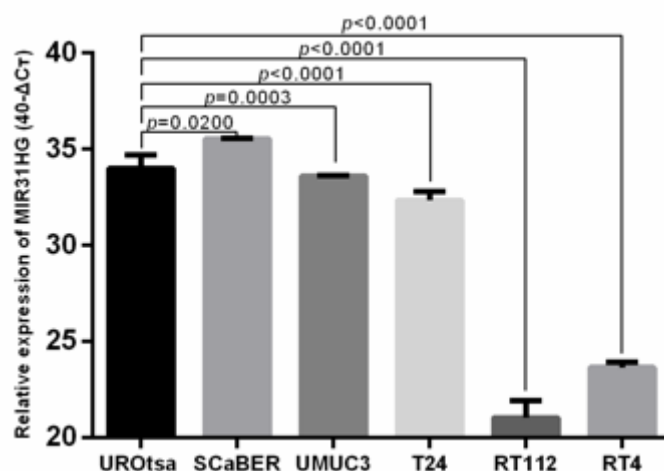
Figure 18. Expression of MIR31HG in BLCA cell lines. In a normal urothelial cell line (UROtsa) and five BLCA cell lines (SCaBER, UMUC3, T24, RT112 and RT4), expression of MIR31HG was detected by qPCR. MIR31HG expression was significantly lower in UMUC3 and T24 cells, and higher in SCaBER cells compared to UROtsa cells. The lowest MIR31HG expression was observed in RT112 and RT4 cells.

MIR31HG expression was also analyzed in RNA-seq data from the Cancer Cell Line Encyclopedia containing 25 BLCA cell lines, including 20 bladder urothelial cell carcinomas, a bladder squamous cell carcinoma, and four bladder carcinoma cell lines from unknown primaries. The data showed that MIR31HG expression levels in SCaBER cells (expression level: 10 TPM) were higher than

in RT112 (expression level: 0 TPM), RT4 (expression level: 0 TPM), and UMUC3 cells (expression level: 3 TPM, Figure 19).
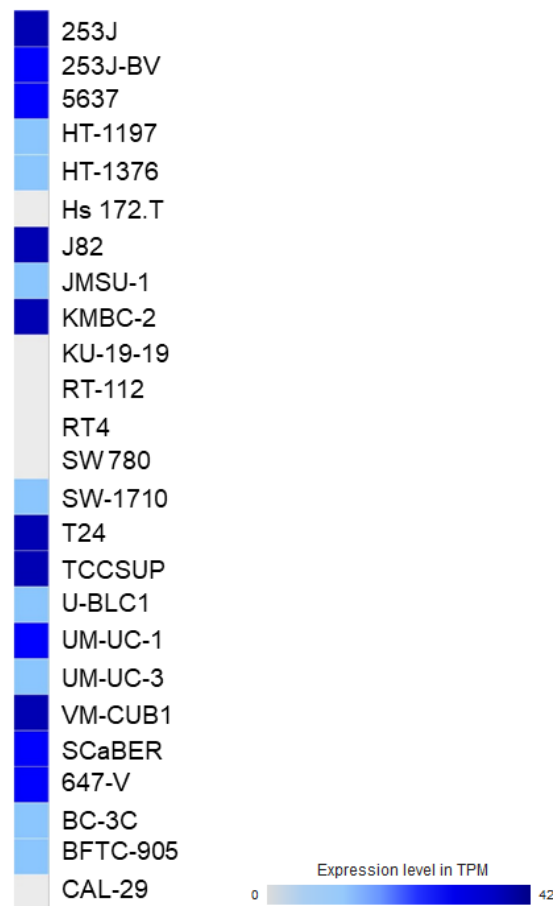

Figure 19. Expression of MIR31HG in BLCA cell lines and tissue samples based on *in silico* data. RNA-seq data from the Cancer Cell Line Encyclopedia showed expression levels in TPM (transcripts per million) for MIR31HG in 25 BLCA cell lines.

3.2.4 MIR31HG is required for BLCA cell proliferation and migration

To evaluate the possible role of MIR31HG in BLCA, a pool of four siRNAs against MIR31HG (designated si-MIR31HG, set of 4) were transfected in T24, UMUC3, and SCaBER cells. T24, UMUC3, and SCaBER cells with MIR31HG knockdown by siRNA showed lower cellular viability at three time points compared to the control group transfected with scramble siRNA (si-NC; T24: p = 0.0068 for 24 h, p = 0.0186 for 48 h, p = 0.0067 for 72 h, Figure 20A; UMUC3: p = 0.0073 for 24 h, p = 0.0096 for 48 h, p = 0.0179 for 72 h, Figure 20B;

SCaBER: p = 0.0047 for 24 h, p = 0.0137 for 48 h, p = 0.0002 for 72 h, Figure 20C).
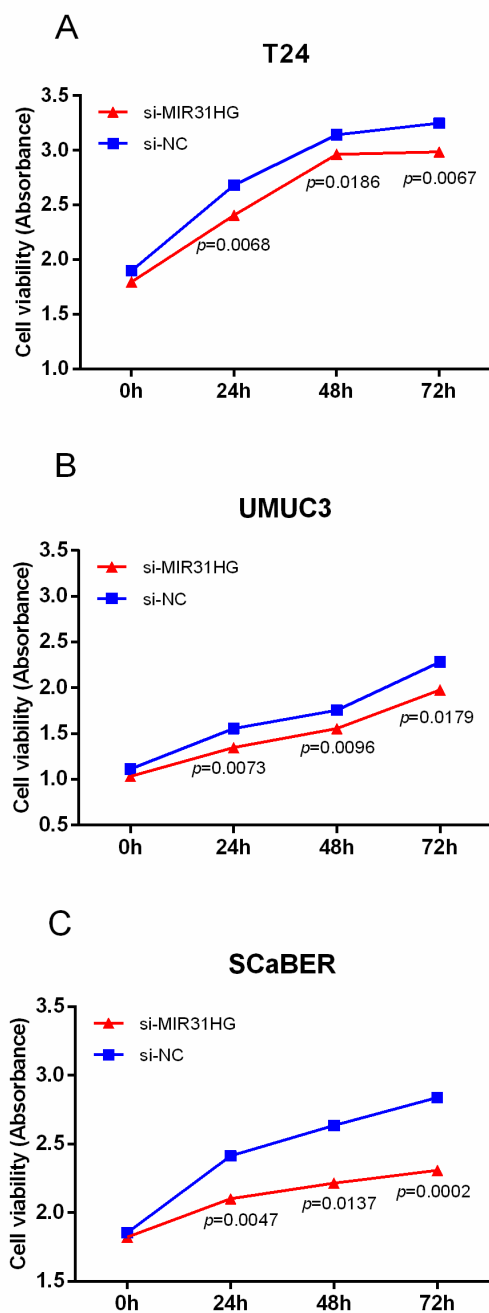
Figure 20. MIR31HG is required for proliferation of BLCA cells. (A-C) Cell viability was quantified using the MTS assay. T24 (A), UMUC3 (B) and SCaBER (C) cells with MIR31HG knockdown by siRNA showed lower cellular viability compared with control group with scramble siRNA (si-NC).

Moreover, colony formation was detected by culturing cells in a 6-well plate and stained with crystal violet after 7 days. T24, UMUC3, and SCaBER cells with

MIR31HG knockdown showed fewer cell colonies compared to the control group (Figure 21A). The staining intensity of the colony formation assay was quantified by calculating the colony intensity percentage. BLCA cells with MIR31HG knockdown showed a reduced formation of colonies compared to the control group ($p = 0.0198$ for T24, $p = 0.03$ for UMUC3, $p = 0.0058$ for SCaBER, Figure 21B).
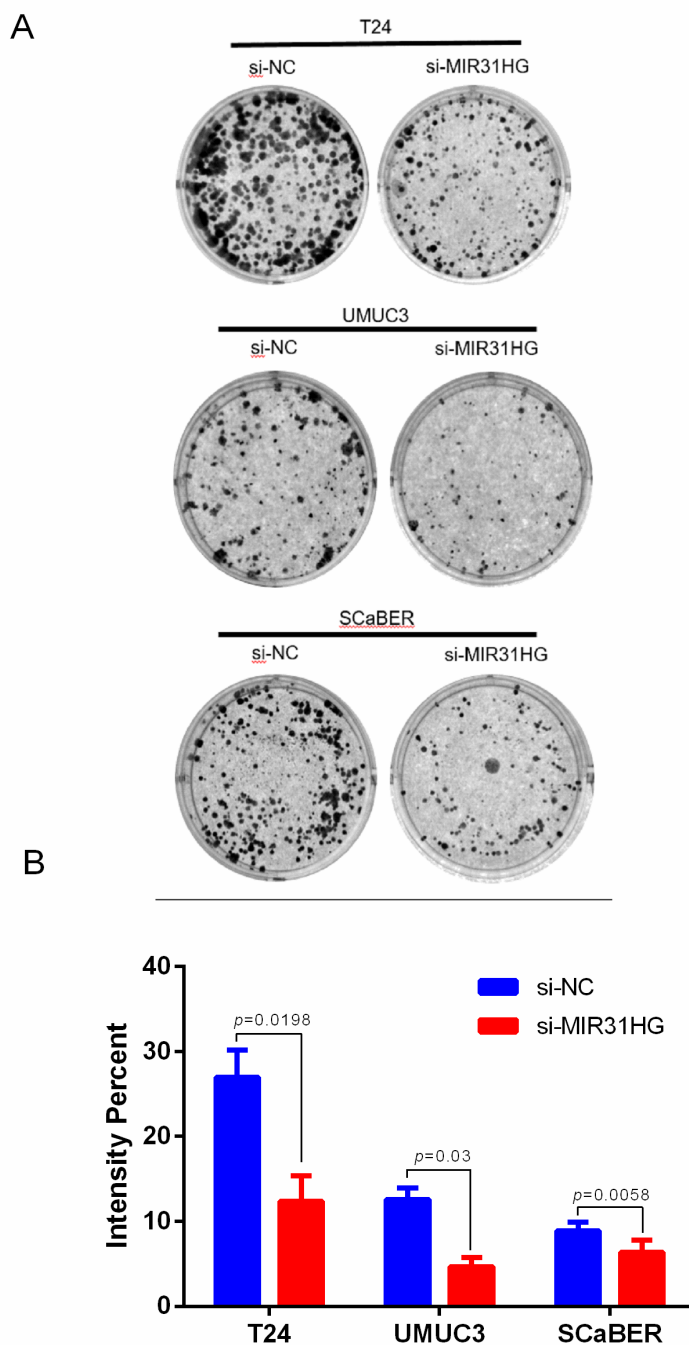
Figure 21. MIR31HG is required for colony formation of BLCA cells. (A) Colony formation was detected by measuring the stained cells after 7 days. T24, UMUC3, and SCaBER cells

with MIR31HG knockdown showed fewer cell colonies compared to control group. (B) Staining intensity percentage of colony formation assay was analyzed by ImageJ software. BLCA cells with MIR31HG knockdown showed decreased staining intensity compared with control group.

Migration was detected by a wound-healing assay. T24, UMUC3, and SCaBER cells with MIR31HG knockdown showed a larger open wound area compared with the control group after 12 h (Figure 22A). The difference in open wound area after 12 h was quantified by calculating the percentage of change in the open wound area (open wound area at 12 h - open wound area at 0 h). BLCA cells with MIR31HG knockdown showed fewer changes in open wound area compared with the control group ($p = 0.0086$ for T24, $p = 0.2206$ for UMUC3, $p = 0.0182$ for SCaBER, Figure 22B).
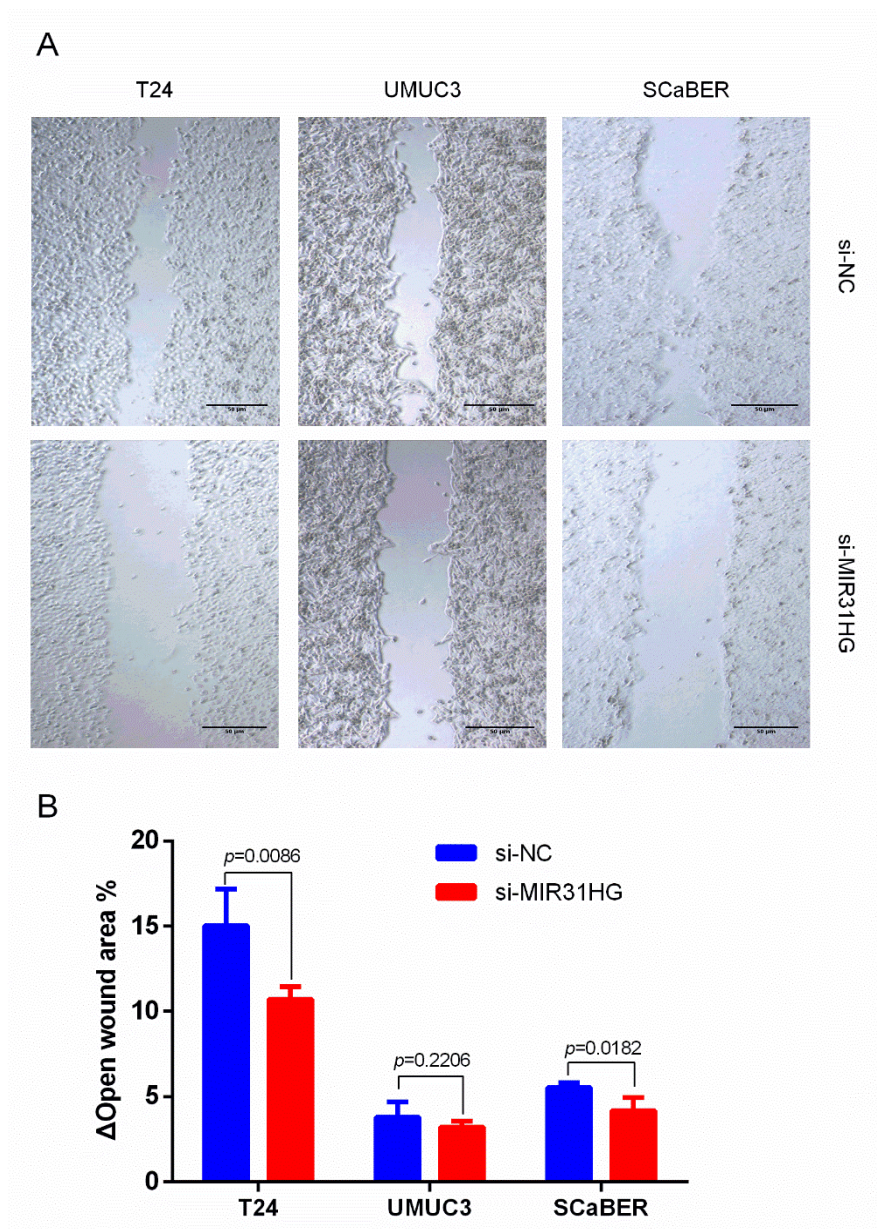
Figure 22. MIR31HG is required for migration of BLCA cells. (A) Wound healing assay was measured using cell culture inserts. T24, UMUC3, and SCaBER cells with MIR31HG knockdown showed more open wound area compared with control group after 12 h. (B) Open wound area was quantified by calculating the percentage of change in open wound area after 12 h. BLCA cells with MIR31HG knockdown showed fewer changes in open wound area compared to the control group.

3.2.5 Splicing variants of MIR31HG regulate BLCA cell proliferation and migration with cell specificity

To further evaluate the possible role of the transcripts MIR31HGΔE1 and MIR31HGΔE3 in BLCA, transcript-specific siRNAs were transfected into BLCA cells. T24, UMUC3, and SCaBER cells with MIR31HGΔE1 and MIR31HGΔE3 knockdown by specific siRNA showed decreased cellular viability with cell specificity (Figure 23A-C). MIR31HGΔE1 knockdown showed significantly decreased absorbance in T24 ($p = 0.0018$, Figure 23A) and UMUC3 ($p = 0.0029$, Figure 23B) cells after 72 h, and MIR31HGΔE3 knockdown resulted in a significant decrease in absorbance in SCaBER ($p = 0.0346$, Figure 23C) cells exclusively after 72 h.
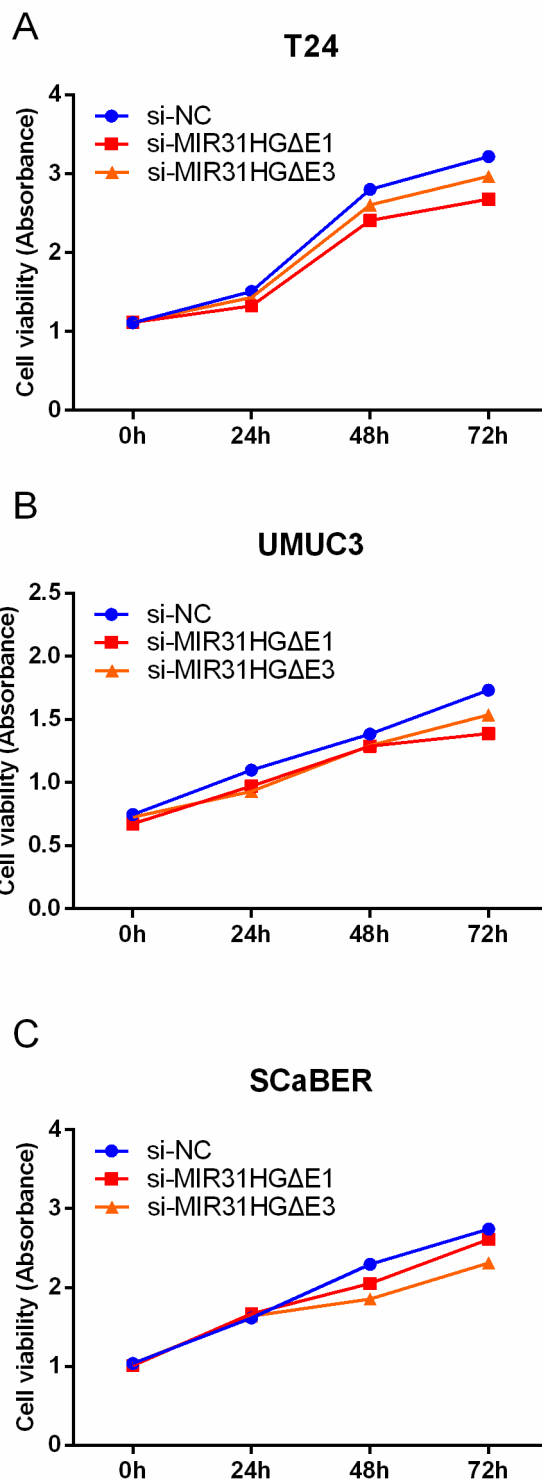
Figure 23. Two splice variants of MIR31HG are required for proliferation of BLCA with cell specificity. (A-C) Cell viability was quantified using the MTS assay. MIR31HGΔE1 knockdown showed a significant decrease in absorbance in T24 ($p = 0.0018$, A) and UMUC3 ($p = 0.0029$, B) cells, and MIR31HGΔE3 knockdown showed a significant decrease in absorbance in SCaBER ($p = 0.0346$, C) cells after 72 h.

Colony formation was detected by culturing cells in 6-well plates and staining with crystal violet after 7 days. T24, UMUC3, and SCaBER cells with MIR31HGΔE1 and MIR31HGΔE3 knockdown by specific siRNA showed fewer cell colonies with cell specificity (Figure 24A). T24 and UMUC3 cells with MIR31HGΔE1 knockdown showed a significantly reduced staining intensity compared with control transfected cells. SCaBER cells with MIR31HGΔE3 knockdown showed significantly decreased staining intensity (T24: $p = 0.0017$ for MIR31HGΔE1, $p = 0.0849$ for MIR31HGΔE3; UMUC3: $p = 0.0001$ for MIR31HGΔE1, $p = 0.7347$ for MIR31HGΔE3; SCaBER: $p = 0.1166$ for MIR31HGΔE1, $p = 0.0022$ for MIR31HGΔE3; Figure 24B).
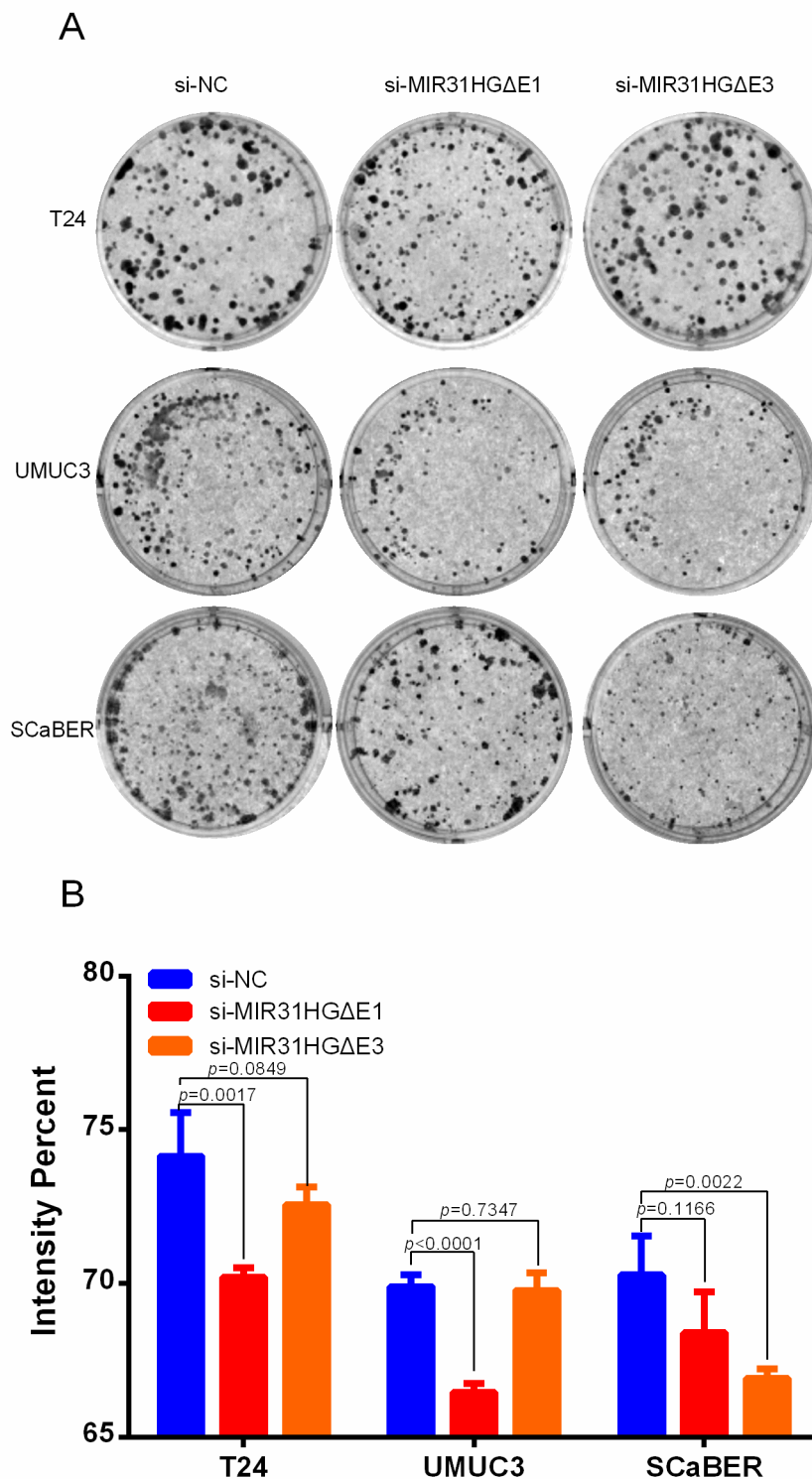
Figure 24. Two splice variants of MIR31HG are required for colony formation of BLCA with cell specificity. (A) Colony formation was detected by measuring cell staining after 7 days. T24, UMUC3, and SCaBER cells with MIR31HGΔE1 and MIR31HGΔE3 knockdown by specific siRNA showed fewer cell colonies with cell specificity. (B) Staining intensity percentage of colony formation assay was analyzed by ImageJ sofware. T24 and UMUC3 cells with MIR31HGΔE1 knockdown and SCaBER cells with MIR31HGΔE3 knockdown showed a significant decrease in staining intensity compared to the control group.

Migration was detected with a wound healing assay using cell culture inserts. T24, UMUC3, and SCaBER cells with MIR31HGΔE1 and MIR31HGΔE3 knockdown by specific siRNA showed a larger open wound area compared with cell specificity after 12 h (Figure 25A). The difference in open wound area after 12 h was quantified by calculating the percentage of change in open wound area (open wound area 12 h - open wound area 0 h). T24 and UMUC3 cells showed fewer changes in open wound area, both in the MIR31HGΔE1 and MIR31HGΔE3 knockdown groups. SCaBER cells showed fewer changes in the open wound area upon MIR31HGΔE3 knockdown (T24: p = 0.0305 for MIR31HGΔE1, p = 0.0105 for MIR31HGΔE3; UMUC3: p = 0.0006 for MIR31HGΔE1, p = 0.0165 for MIR31HGΔE3; SCaBER: p = 0.8160 for MIR31HGΔE1, p = 0.0079 for MIR31HGΔE3; Figure 25B).
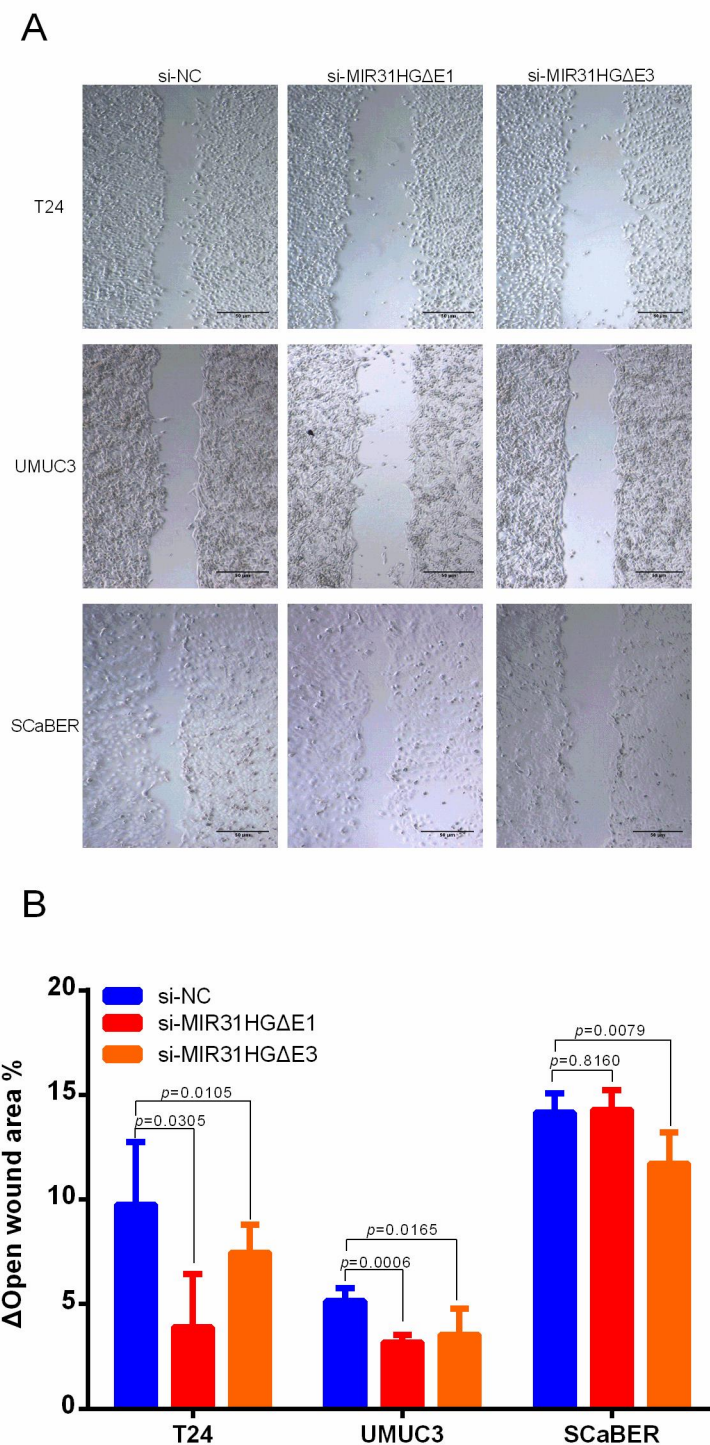
Figure 25. Two splice variants of MIR31HG are required for migration of BLCA with cell specificity. (A) Wound healing assay was measured using cell culture inserts. T24, UMUC3, and SCaBER cells with MIR31HGΔE1 and MIR31HGΔE3 knockdown by specific siRNA showed a larger open wound area compared with cell specificity after 12 h. (B) Difference in open wound area was quantified by calculating the percentage of change in open wound area after 12 h. T24 and UMUC3 cells showed fewer changes in the open wound area, in both MIR31HGΔE1 and MIR31HGΔE3 knockdown groups. SCaBER cells showed fewer changes in the open wound area in the MIR31HGΔE3 knockdown group.

3.2.6 MIR31HG is associated with EGFR pathway

To discover the potential mechanisms underlying the functions of MIR31HG, lncRNA-protein interactions were analyzed by lncPro. lncPro yields a score using amino acid and nucleotide sequences. This score can be used to measure the interaction between a pair of lncRNA and protein. The sequence of MIR31HG and protein sequences were obtained from the NCBI database (http://www.ncbi.nlm.nih.gov/). In total, nine isoforms of EGFR were predicted to interact with MIR31HG. MIR31HG was predicted as interactive with EGFR, phosphoinositide 3-kinase (PI3K) and receptor tyrosine-protein kinase erbB-2 (HER2) protein (score above 50, Table 18).

Table 18. Interaction scores of EGFR, PI3K HER2 protein and MIR31HG.

Protein	Score
EGFR	72.4935
EGFR isoform a	75.0794
EGFR isoform b	93.4501
EGFR isoform c	83.9972
EGFR isoform d	93.8757
EGFR isoform e	72.5006
EGFR isoform f	75.7342
EGFR isoform g	73.3893
EGFR isoform h	82.7141
EGFR isoform i	84.5429
PI3K	75.5458
HER2 isoform a	70.9194
HER2 isoform b	82.7639
HER2 isoform c	79.0902
HER2 isoform d	69.6083
HER2 isoform e	91.4699

To further reveal the interaction with the coded gene, gene expression of EGFR was detected in MIR31HG knockdown cells. To narrow the differences in EGFR expression between cell lines, the $2^{-\Delta\Delta CT}$ values of expression data in T24, UMUC3 and SCaBER cells were normalized by the expression data in UROtsa.

Significantly higher expression of EGFR was observed in SCaBER cells with MIR31HG knockdown compared to the control group (p = 0.0367, Figure 20). No significant difference of expression was observed between T24 (p = 0.4086) and UMUC3 (p = 0.4734) cells with MIR31HG knockdown and control group (Figure 26).
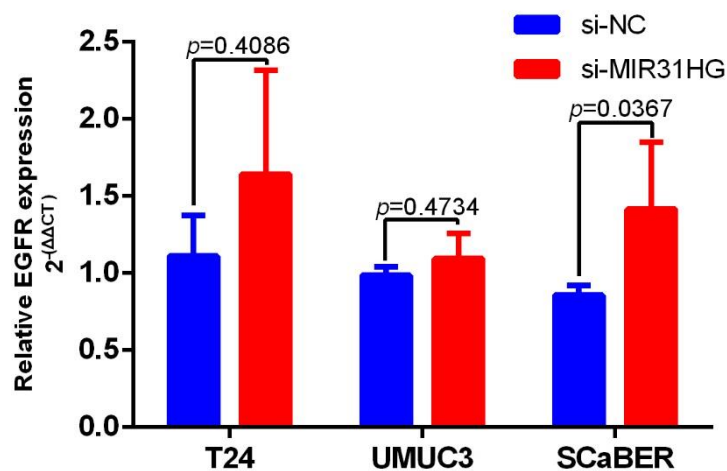

Figure 26. Expression of EGFR in BLCA cell lines with MIR31HG knockdown. EGFR expression was significantly higher in SCaBER cells with MIR31HG knockdown compared to negative siRNA group. Expression of EGFR showed no significant difference between T24 and UMUC3 cells with MIR31HG knockdown and negative siRNA group.

4 DISCUSSION

Despite advances in surgical therapy, patients with MIBC after RC develop into local recurrence or metastases up to 40-50% of cases [96]. The clinical routine primarily uses the TNM stage and grading system to assess the prognosis, due to the tumor stage and lymphogenic metastasis having the greatest influence on the prognosis [97]. Nevertheless, these parameters are limited in predictive ability to the forecast. Therefore, it would be important to identify new biomarkers which enable more accurate prognostic risk and predictive stratification with regard to targeted therapies. In this study, the major aims are evaluating the RNA expression pattern of our interested genes ANLN, TLE2 and MIR31HG with regard to molecular subtype, determining the correlated with the outcome of patients with MIBC and discovering the effects on biological function in BLCA.

4.1 Regulation of ANLN, TLE2 and MIR31HG expression

4.1.1 Expression of ANLN and TLE2

In order to evaluate the translational benefit, gene expression was compared with molecular subtypes, targets, and relevant clinicopathologic parameters in multivariable analyses. As a key regulator of cytokinesis, it is not surprising that ANLN might play a critical role in carcinogenesis [98]. Previous research has established that knockdown of ANLN could significantly inhibit the proliferation of BLCA both *in vitro* and *in vivo*. Furthermore, knockdown of ANLN strongly suppressed the migration and invasion ability of J82 and 5637 bladder cancer cell lines [27]. Additionally, in upper urinary tract urothelial carcinoma, overexpression of ANLN in the nucleus is a poor prognostic factor, which was confirmed by data on protein levels, while low expression of ANLN in the cytoplasm is a poor prognosis marker [99]. In this study, ANLN was extensively

evaluated on mRNA level, showed the similar overexpression in MIBC tissues, which in accord with the protein level in nucleus of upper urinary tract urothelial carcinoma tissues [99]. However, the subcellular expression of ANLN in MIBC still need further investigation.

In this thesis, ANLN expression was strongly correlated with cell proliferation markers including CDK1, RACGAP1, MKI67, and FOXM1. In urothelial carcinoma cell lines, these genes showed a consistently strong positive correlation with the expression of ANLN, which is highly expressed in ScaBER cells and expressed in low amounts in RT4 cells. In patient from the Mannheim cohort, a strong correlation of ANLN with RACGAP1, MKI67, and CDK1 was also found. These results indicate that ANLN could have a crucial role in tumor proliferation of MIBC. It is already known that the proliferation markers MKI67 and RACGAP1 have a significant importance in BLCA [100, 101]. Previous research showed that mRNA expression of MKI67 is significantly correlated and associated with stage and grade in NMIBC [102]. Furthermore, the expression of RACGAP1 correlated significantly with the tumor stage in BLCA after RC, and RACGAP1 was strongly expressed in the early stages of NMIBC patient samples [87]. Additionally, FOXM1 has been shown to be overexpressed on the mRNA and protein levels in BLCA cells, and plays an important role in cisplatin resistance, outcome prediction, and risk stratification of patients with BLCA [103]. These results may indicate that ANLN and its correlated genes may supplement and optimize biomarker panels for molecular characterization of BLCA.

In this study, TLE2 expression correlated with several Wnt pathway components, such as β-catenin (CTNNB1), TCF7, and LEF1 in the TCGA cohort. By analyzing the expression in the Mannheim cohort, it was also found that TLE2 was significantly correlated with CTNNB1. In addition, CTNNB1 is highly expressed in RT4 cells and expressed in lower amounts in ScaBER cells, which is similar to the expression pattern of TLE2. The activation of the Wnt/β-catenin signaling pathway plays an important role in tumorigenesis and

development of various cancers including BLCA [33, 104, 105]. Gain-of-function mutations in CTNNB1 are detected in numerous human cancers [106-108]; therefore, it is necessary to explore the role of Wnt/β-catenin regulated genes in BLCA. Previous research has shown that TLE2 was significantly suppressed and the level of β-catenin protein was increased in esophageal tumor cells, both of which were modulated by NDRG1 overexpression [109]. These results may indicate that the differences of TLE2 expression in different BLCA cell lines may serve as trigger on activating Wnt/β-catenin signaling pathway. Interestingly, ANLN positively correlates with CTNNB1, while TLE2 correlates inversely with CTNNB1, which was also observed in the Mannheim cohort. Although there is no direct correlation between ANLN and TLE2 in both cohorts, the involvement of ANLN and TLE2 underlying the Wnt/β-catenin signaling pathway was preliminarily revealed by this study.

To reveal more details of ANLN and TLE2 expression in BLCA, the RNA-seq data of transcripts based on TCGA cohort were analyzed. Interestingly, the splice variants of ANLN and TLE2 showed different expression levels, with ANLN-201 (ENST00000265748.6), ANLN-202 (ENST00000396068.6), ANLN-210 (ENST00000457743.1), ANLN-212 (ENST00000491782.1), TLE2-201 (ENST00000262953.10), TLE2-202 (ENST00000426948.6), TLE2-204 (ENST00000455444.6), and TLE2-215 (ENST00000590101.5) highly expressed compared with other transcripts. No existing study showed the correlation between clinical or histopathological features of MIBC and these splice variants. These findings indicate that differential expression patterns of ANLN and TLE2 splice variants might potentially have a practical usefulness, which needs further investigations.

This study demonstrates that ANLN and TLE2 show a distinct subtype-specific overexpression in BLCA cell lines. ANLN showed overexpression in basal-like urothelial carcinoma cell lines and patients, while TLE2 showed significantly higher expression in luminal-like urothelial carcinoma cell lines and patients. Interestingly, for patients with MIBC in the whole TCGA cohort, ANLN and TLE2

harbored similar prognostic values for basal and luminal subtypes, respectively. Molecular classification has emerged as a promising research tool beyond histopathology to stratify cancer patients for personalized medicine [110]. Recent years have witnessed increasing interest and research into the molecular basis of BLCA [12]. Several studies subclassified both MIBC and NMIBC through RNA-seq-based data and identified distinct molecular subtypes that correlated with outcome and therapy response [111-113]. MIBC molecular subtypes, which are broadly grouped into basal and luminal subtypes, showed similarities to the molecular phenotypes of breast cancer [114]. It is believed that epithelial-to-mesenchymal transition (EMT) is a critical step in the progression of breast cancer, particularly the basal-like one [115]. The basal and EMT/claudin-low markers were highly expressed in the same subtype of patients with BLCA based on the TCGA database [13]. Due to the significant association with basal subtype of BLCA, it is suggested that ANLN could play a potential role in EMT in BLCA. Also, KRT5 is highly upregulated in basal and KRT20 in luminal subtype in BLCA [12]. qPCR-based molecular subtyping of BLCA by KRT5 and KRT20 mRNA expression is a method associated with the survival of patients with MIBC [116]. It is confirmed that KRT5 and KRT20 showed significant association with ANLN and TLE2 in the Mannheim cohort. The differences of ANLN and TLE2 expression were not significant in the high KRT5 and KRT20 group maybe due the fact that only KRT5 and KRT20 were used for subtype association.

4.1.2 Expression of MIR31HG and its splice variants

In previous studies on multifarious tumors, MIR31HG showed a tissue-specific expression pattern. In breast cancer and non-small cell lung cancer (NSCLC) cells, MIR31HG expression was upregulated [67, 71]. In gastric cancer tissues and cell lines, MIR31HG was poorly expressed [69]. Another study showed that

MIR31HG level is substantially upregulated in oral carcinoma, significantly associated with poor clinical outcomes and representing an independent prognostic predictor [117]. In this study, lower MIR31HG transcript levels were found in luminal-like and mixed-type BLCA cell lines compared with a normal urothelium cell line. Accordingly, down-regulated MIR31HG expression was found in cancer tissues compared to normal tissues, which supports expression results measured by qPCR from a previous study [72]. However, the previous expression results were measured in stage- and type-mixed BLCA tissues, and in this study, MIR31HG was measured in MIBC and associated with multiple molecular subtype respectively. In contrast, MIR31HG was found to be highly expressed in cells lines and clinical tumor samples with the basal subtype compared to luminal and other subtypes, indicating that MIR31HG not only shows tissue specific, but also subtype-specific overexpression in MIBC.

In contrast to a previous study, which reported that miR-31 and MIR31HG are down-regulated in triple-negative breast cancer (TNBC) cell lines of basal subtype [118], the present study shows that MIR31HG is highly expressed in the BLCA cell line of basal subtype and markedly correlates with the survival of patients with MIBC basal subtype. This might be due to different tumor entities and tissue specific expression of MIR31HG. In this study, two MIBC cohorts with multiple molecular subtypes were involved rather than single cell line, which may also lead to the dissimilar results. Further studies of BLCA preclinical models are needed for validation. For basal or squamous-like bladder cancer, molecular signatures were found based on clustering RNA-seq data, including EGFR [119]. EGFR is overexpressed in up to 74% of BLCA tissue specimens, and is amplified in squamous cell carcinomas (SCC) of the bladder [120]. To further investigate the potential relationship between MIR31HG and EGFR, a computational method called lncPro was applied to predict the associations within. All the nine isoforms of EGFR, as well as PI3K and HER2 protein, were predicted as interactive with MIR31HG. The positive interaction score suggested that MIR31HG might be involved in the EGFR/PI3K/AKT signaling

pathway. Furthermore, expression of EGFR was detected in three MIR31HG-knockdown BLCA cell lines. In SCaBER cells, which shows the basal / squamous signature, expression of EGFR was reversely correlated with MIR31HG. According to another study on lung cancer, the protein expression of phosphorylated EGFR and PI3K were repressed in PC9 cells with MIR31HG siRNA transfected. By knocking down MIR31HG, reversal of gefitinib resistance was found by regulation of the EGFR/PI3K/AKT signaling pathway [121]. Taken together, these findings suggested that MIR31HG might potentially correlate with the EGFR pathway.

A variety of mature RNAs are formed by different splicing during the transcription of a single gene. These mature RNAs are called the transcripts of the gene, and the process of generating different transcripts is called the alternative splicing process [122]. Alternative splicing affects the amino acid sequence of the translated protein by changing the coding sequence, or affects gene expression by changing the sequence of the untranslated region, thereby affecting biological functions [123, 124]. There are about 30,000 genes in the human genome, and about 95% of the genes can undergo alternative splicing [125]. As a post-transcriptional regulatory mechanism, alternative splicing greatly increases the complexity of genes and the diversity of proteomes [126]. In this study, it is noteworthy that two transcript variants of MIR31HG (MIR31HGΔE1 and MIR31HGΔE3) were identified and their expression was analyzed in BLCA cells and MIBC tissues. Besides the different expression levels in MIBC patient tissues, the two splice variants showed distinguished expression patterns in basal and luminal subtypes, respectively. MIR31HGΔE3 showed high expression in the basal subtype, both in the TCGA cohort and the Mannheim cohort, which is also observed for the group with high KRT5 expression. In contrast, MIR31HGΔE1 showed high expression in luminal subtype tumors in the Mannheim cohort, corresponding to tumors with a high KRT20 expression. To my knowledge, no existing data reported these two splice variants of MIR31HG. A previous study found that PVT1ΔE4, which is a

novel splice variant of lncRNA PVT1, was highly expressed in renal cell carcinoma, and could promote cell proliferation and invasion similar as the full-length transcript [127]. Besides, expression of PVT1ΔE4 was even more abundant than its full-length sibling in renal cell carcinoma. These results may indicate that the full-length transcript and its splice variants could have coordinate or different expression and effect on tumor.

4.2 ANLN, TLE2 and MIR31HG are prognostic markers for MIBC

4.2.1 ANLN and TLE2 as prognostic markers for MIBC

This study retrospectively evaluated the prognostic and clinical impact of ANLN and TLE2 gene expression and to validate these results in published datasets. Higher ANLN transcript levels were found to be associated with worse OS and DSS in the Mannheim cohort, which corresponds with results from previous studies of BLCA and upper urinary tract urothelial carcinoma [27, 99]. Supported results on 40 BLCA patients after RC or transuretheral resection showed that up-regulated expression of ANLN on mRNA level was correlated with poor DSS and DFS [27]. The study of upper urinary tract urothelial carcinoma with surgical intervention showed that, overexpression of ANLN in the nucleus and low expression in the cytoplasm on protein level was significantly associated with poor DSS and metastasis-free survival [99]. The subcellular localization of ANLN and correlation with outcome of MIBC patients still need further research. Together, these results indicate that ANLN, in addition to lymph node status, may be an independent predictor for DFS, and superior to the use of T stage and LVI as predictors.

In contrast to ANLN, TLE2 was found to be dominantly expressed in patients with lower stages of BLCA. Patients with higher TLE2 expression had a more favorable OS and DFS, both in the whole-stage group and in the T2 subgroup. These results, together with univariable and multivariable Cox regression

analysis of the Mannheim and the TCGA cohort, suggest that TLE2 could serve as an independent risk factor for prognostic prediction of patients with BLCA. The most recent research about TLE2 showed an inhibition of replication and transcription in Kaposi's sarcoma-associated herpesvirus [30]. To my knowledge, there is only very limited data about the role of TLE2 in cancer and no data specific to BLCA. Therefore, these findings are the first describing the role of TLE2 in BLCA. In adition, the Kaplan–Meier curves of ANLN and TLE2 with OS and DSS showed separation at earlier time points, but convergence at later time points, between the high and low expression group in the TCGA cohort. It indicated that the predictive capabilities of ANLN and TLE2 expression might be stronger at earlier time points after RC.

Until recently, effective targets following platinum-based chemotherapy were limited for patients with advanced urothelial carcinoma. The most promising option is immunotherapy with programmed cell death 1 (PD-1) / PD-L1 checkpoint inhibitors, coupled with CTLA-4 antibodies [128, 129]. Apart from differences in expression thresholds for defining PD-L1 positivity, the validated biomarkers for optimal patient selection are still unrevealed [130]. In this study, expression of ANLN and TLE2 correlate with the therapeutic targets for BLCA as indicated in existing research and clinical trials. Notably, ANLN was positively and TLE2 was negatively correlated with PD-L1 in the TCGA cohort. TLE2 also showed a negative correlation with PD-1 and the key immunoregulator CTLA-4. These results suggest that ANLN and TLE2 could serve as potential biomarkers for response to immunotherapy and precise therapeutic management of MIBC.

4.2.2 MIR31HG and its splice variants as prognostic markers for MIBC with basal subtype

Similar as ANLN and TLE2, this study evaluated the correlation with survival

and expression of MIR31HG as well as its splice variants in MIBC cohorts. Higher MIR31HG transcript levels were found to be associated with worse OS and DFS in the basal subtype cohort, but not in the whole TCGA cohort. It is the first time to discover the prognostic value of MIR31HG in BLCA, or associated with subtypes of any other tumors. Two groups independently recognized the significance of a distinct basal MIBC subtype [131, 132]. According to classification of the TCGA, basal tumors were divided into two subsets that were largely distinguished by differential expression of biomarkers associated with EMT, which is a reversible developmental process by promoting invasion, metastasis, "stemness", and drug resistance [111, 133]. The potential significance of the mesenchymal basal BLCA was identified using a "claudin-low" gene expression signature in breast cancer [134]. A previous study reported that tumor suppressor microRNA-361 was de-repressed by MIR31HG in osteosarcoma cells, leading to cell growth and mesenchymal phenotype [135]. These results may indicate that high expression MIR31HG could be served as a surrogate marker of poor outcome defined by relative activation of EMT and sponge of tumor suppressor. The discovery that MIR31HG is highly expressed and significantly outcome-correlated in MIBC with basal subtype, has complemented the selection of MIBC markers.

Furthermore, expression of MIR31HGΔE1 and MIR31HGΔE3 was significantly associated with OS and DFS in the Mannheim cohort, rather than the full-length transcript of MIR31HG. In the TCGA cohort, it was demonstrated that MIR31HGΔE3 expression was significantly associated with OS of the basal subtype group. These results, together with univariable and multivariable Cox regression analysis suggested that the alternative splice variants of MIR31HG may serve as potential biomarkers for certain molecular subtypes of MIBC, which could contribute to an individualized BLCA subclassification and therapy decision making.

4.3 MIR31HG and its splicing variants regulate BLCA tumorigenesis

Due to its potential clinical relevance as a marker, the expression of MIR31HG and its underlying biological functions could be vital to the tumor. In this study, by knocking down MIR31HG expression using siRNA, diminished cell proliferation, colony formation, and migration were assessed in BLCA cell lines. This study is the first to highlight the function of MIR31HG in BLCA cells, which indicates that MIR31HG might serve as an oncogene in certain types of BLCA. With the in-depth study of MIR31HG, several downstream targets were found in present researches. For example, it is reported that overexpression of MIR31HG significantly decreased the expression of miR-575, enhanced the suppression of tumorigenicity 7 like (ST7L) in hepatocellular carcinoma (HCC). Thus, MIR31HG regulated ST7L expression through sponging miR-575, and acted as tumor suppressor in HCC [136]. Another published research showed that the level of MIR31HG in esophageal squamous cell carcinoma (ESCC) tissues was positively correlated with the expression of furin and matrix metalloproteinase 1 (MMP1). When MIR31HG was silenced, the expressions of furin and MMP1 in ESCC cells were significantly inhibited. These results suggest that the involvement of MIR31HG in invasion and migration of ESCC cell may be partly achieved through the furin / MMP1 pathway [137]. In a study in oral cancer, MIR31HG was identified as a hypoxia-inducible lncRNA and forms a complex with hypoxia-inducible factor-1 α (HIF-1α), thus as an adverse prognostic predictor for the cancer progression [117]. In addition, it was reported that MIR31HG could function as an oncogene that promotes pancreatic cancer progression, by acting as an endogenous sponge competing for miR-193b [70]. Similarly, it was shown that silencing of MIR31HG significantly inhibited NSCLC cell migration, invasion, and metastasis by attenuated sponging of miR-214 [71]. The various mechanisms of MIR31HG suggest that downstream pathways could be involved with other coding and non-coding RNAs, which requires further verification.

A number of studies have shown that alternative splicing plays an important role in tumors. For genes with multiple transcripts, certain transcripts or splice variants have specific functions in tumorigenesis and development. For example, the tight junction protein ZO-1 (TJP1) has multiple transcripts, among which the transcript retained by exon 23 (E23+) can inhibit the metastasis of colorectal cancer, while the transcript excised by exon 23 does not function in tumor metastasis [138]. As another example, CD44 is an important cell surface marker, and its alternative splicing are related to EMT, as well as its specific splice variants is related to the occurrence of EMT in tumor cells [139, 140]. There is also a report showed that different splice variants of the same gene can play diametrically opposite functions. VEGF165b and VEGF165 are two splice variants of vascular endothelial growth factor (VEGF), differing by only one exon. Yet their functions in tumors are diametrically opposite. VEGF165b, a protein translated by VEGF transcripts with variable exons, can inhibit tumor angiogenesis. While VEGF165, a protein generated by exon skip splice variant, promotes tumor angiogenesis [141]. As described for VEGF165b and VEGF165, which have diametrical functions, MIR31HG variants MIR31HGΔE1 and MIR31HGΔE3 may potentially have different functions and involved in different pathways.

In this study, similar to the full-length transcript, a series of functional experiments validated the corresponding roles of MIR31HG splice variants in certain types of BLCA cells. Knock down MIR31HGΔE3 resulted in reduced colony formation ability and cell viability solely in SCaBER cells, suggesting that exon-specific or transcript-specific mechanisms could be a new direction in studying basal-like BLCA. Recently, emerging data have demonstrated that RNA splice variants are associated with drug resistance in cancer. The expression of androgen receptor (AR) splicing variants in castration-resistant prostate cancer (CRPC) samples increased significantly. The most common AR splicing variants are AR-V7 and ARv567es [142, 143]. These AR splice mutants result from the alternative splicing of the AR precursor mRNA, and most of them

do not contain ligand binding regions, which are the targets of various AR antagonist drugs. Therefore, these AR splice variants are important factors for insensitive to AR antagonists of CRPC patients. Splice variants in V600E BRAF-mutant-positive malignant melanoma patients were shown to be associated with vemurafenib resistance, indicating that aberrant splicing could be a novel mechanism of acquired resistance [144]. Furthermore, interference of the pre-mRNA splicing modulators sufficiently inhibited formation of these splice variants, suggesting that splice variant-specific siRNAs may be proposed as a therapeutic strategy to overcome drug resistance [145].

5 SUMMARY

ANLN and TLE2 are associated with cancer patient survival and progression. The impact of their gene expression on survival of patients with MIBC treated with RC and subtype association has not yet been investigated. This study provides *in silico* and *in vitro* evidence supporting the prognostic potential of ANLN and TLE2 for patients with MIBC. In the Mannheim cohort, tumors with high ANLN expression were associated with lower OS and DSS, while high TLE2 expression was associated with a favorable OS. The TCGA cohort confirmed that high ANLN and low TLE2 expression was associated with shorter OS and DFS. In both cohorts, multivariable analyses showed ANLN and TLE2 expression as independent outcome predictors. Furthermore, ANLN was more highly expressed in cell lines and patients with the basal subtype, while TLE2 expression was higher in cell lines and patients with the luminal subtype. ANLN and TLE2 are promising biomarkers for individualized BLCA therapy including cancer subclassification and informed MIBC prognosis. These results indicate that developing ANLN and TLE2 as new biomarkers will help to further optimize personalized therapy for these patients.

Growing evidence supports the pivotal role of lncRNAs in the regulation of cancer development and progression. Their expression patterns and biological function in MIBC remain elusive. In this study, a decreased expression of lncRNA MIR31HG was found in BLCA cells and tissues, except in the basal subtype. *In vitro* experiments revealed that knockdown of MIR31HG inhibits cell proliferation, colony formation and migration in BLCA. Survival analysis showed that high expression of MIR31HG was associated with poor OS and DFS in patients with MIBC of basal subtype. Two splice variants of MIR31HG lacking exon 1 (MIR31HGΔE1) and exon 3 (MIR31HGΔE3) were identified to have specific expression patterns in different subtypes of both MIBC cohorts. MIR31HGΔE3 was highly expressed in tumors with basal subtype. After

knockdown of splice variants of MIR31HG, cell proliferation and migration assays showed corresponding roles for the full-length transcript. A high expression of MIR31HGΔE1 and MIR31HGΔE3 was associated with worse OS and DFS in the Mannheim cohort. This study demonstrates that MIR31HG and its splice variants could serve as biomarkers for the classification and prognosis prediction of patients with MIBC. Taking together, this thesis provides new insights into studies for the molecular classification and relevance of RNA variants in MIBC.

6 REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: a cancer journal for clinicians, 68 (2018) 394-424.

[2] P. Brennan, O. Bogillot, S. Cordier, E. Greiser, W. Schill, P. Vineis, G. Lopez-Abente, A. Tzonou, J. Chang-Claude, U. Bolm-Audorff, K.H. Jöckel, F. Donato, C. Serra, J. Wahrendorf, M. Hours, A. T'Mannetje, M. Kogevinas, P. Boffetta, Cigarette smoking and bladder cancer in men: a pooled analysis of 11 case-control studies, International journal of cancer, 86 (2000) 289-294.

[3] C. Cordon-Cardo, Molecular alterations associated with bladder cancer initiation and progression, Scandinavian journal of urology and nephrology. Supplementum, (2008) 154-165.

[4] M.J. Urist, C.J. Di Como, M.L. Lu, E. Charytonowicz, D. Verbel, C.P. Crum, T.A. Ince, F.D. McKeon, C. Cordon-Cardo, Loss of p63 expression is associated with tumor progression in bladder cancer, The American journal of pathology, 161 (2002) 1199-1206.

[5] T.T. Junttila, M. Laato, T. Vahlberg, K.O. Söderström, T. Visakorpi, J. Isola, K. Elenius, Identification of patients with transitional cell carcinoma of the bladder overexpressing ErbB2, ErbB3, or specific ErbB4 isoforms: real-time reverse transcription-PCR analysis in estimation of ErbB receptor status from cancer patients, Clinical cancer research : an official journal of the American Association for Cancer Research, 9 (2003) 5346-5357.

[6] N. Shinohara, T. Koyanagi, Ras signal transduction in carcinogenesis and progression of bladder cancer: molecular target for treatment?, Urological research, 30 (2002) 273-281.

[7] H. Primdahl, H. von der Maase, F.B. Sørensen, H. Wolf, T.F. Ørntoft, Immunohistochemical study of the expression of cell cycle regulating proteins at different stages of bladder cancer, Journal of cancer research and clinical oncology, 128 (2002) 295-301.

[8] T. Habuchi, M. Marberger, M.J. Droller, G.P. Hemstreet, 3rd, H.B. Grossman, J.A. Schalken, B.J. Schmitz-Dräger, W.M. Murphy, A.V. Bono, P. Goebell, R.H. Getzenberg, S.H. Hautmann, E. Messing, Y. Fradet, V.B. Lokeshwar, Prognostic markers for bladder cancer: International Consensus Panel on bladder tumor markers, Urology, 66 (2005) 64-74.

[9] H. Miyamoto, J.S. Miller, D.A. Fajardo, T.K. Lee, G.J. Netto, J.I. Epstein, Non-invasive papillary urothelial neoplasms: the 2004 WHO/ISUP classification system, Pathology international, 60 (2010) 1-8.

[10] M. Babjuk, A. Bohle, M. Burger, O. Capoun, D. Cohen, E.M. Comperat, V. Hernandez, E. Kaasinen, J. Palou, M. Roupret, B.W. van Rhijn, S.F. Shariat, V. Soukup, R.J. Sylvester, R. Zigeuner, EAU Guidelines on Non-Muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016, European urology, 71 (2017)

447-461.

[11] L.H. Sobin, C.C. Compton, TNM seventh edition: what's new, what's changed: communication from the International Union Against Cancer and the American Joint Committee on Cancer, Cancer, 116 (2010) 5336-5339.

[12] W. Choi, S. Porten, S. Kim, D. Willis, E.R. Plimack, J. Hoffman-Censits, B. Roth, T. Cheng, M. Tran, I.-L. Lee, Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy, Cancer cell, 25 (2014) 152-165.

[13] Y. Li, K. Yang, K. Li, H. Liu, S. Zhao, M. Jiao, X. Fu, Clinical and molecular characteristics of bladder urothelial carcinoma subtypes, Journal of cellular biochemistry, 120 (2019) 9956-9963.

[14] J.-P. Volkmer, D. Sahoo, R.K. Chin, P.L. Ho, C. Tang, A.V. Kurtova, S.B. Willingham, S.K. Pazhanisamy, H. Contreras-Trujillo, T.A. Storm, Y. Lotan, A.H. Beck, B.I. Chung, A.A. Alizadeh, G. Godoy, S.P. Lerner, M. van de Rijn, L.D. Shortliffe, I.L. Weissman, K.S. Chan, Three differentiation states risk-stratify bladder cancer into distinct subtypes, Proceedings of the National Academy of Sciences, 109 (2012) 2078-2083.

[15] W.S. Tan, A. Feber, R. Sarpong, P. Khetrapal, S. Rodney, R. Jalil, H. Mostafid, J. Cresswell, J. Hicks, A. Rane, A. Henderson, D. Watson, J. Cherian, N. Williams, C. Brew-Graves, J.D. Kelly, Who Should Be Investigated for Haematuria? Results of a Contemporary Prospective Observational Study of 3556 Patients, European urology, 74 (2018) 10-14.

[16] D.M. Burke, D.C. Shackley, P.H. O'Reilly, The community-based morbidity of flexible cystoscopy, BJU Int, 89 (2002) 347-349.

[17] J. Bellmunt, A. Orsola, J.J. Leow, T. Wiegel, M. De Santis, A. Horwich, Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up, Annals of oncology : official journal of the European Society for Medical Oncology, 25 Suppl 3 (2014) iii40-48.

[18] Neoadjuvant chemotherapy in invasive bladder cancer: update of a systematic review and meta-analysis of individual patient data advanced bladder cancer (ABC) meta-analysis collaboration, European urology, 48 (2005) 202-205; discussion 205-206.

[19] S. Cambier, R.J. Sylvester, L. Collette, P. Gontero, M.A. Brausi, G. van Andel, W.J. Kirkels, F.C. Silva, W. Oosterlinck, S. Prescott, Z. Kirkali, P.H. Powell, T.M. de Reijke, L. Turkeri, S. Collette, J. Oddens, EORTC Nomograms and Risk Groups for Predicting Recurrence, Progression, and Disease-specific and Overall Survival in Non-Muscle-invasive Stage Ta-T1 Urothelial Bladder Cancer Patients Treated with 1-3 Years of Maintenance Bacillus Calmette-Guérin, European urology, 69 (2016) 60-69.

[20] K.G. Miller, C.M. Field, B.M. Alberts, Actin-binding proteins from Drosophila embryos: a complex network of interacting proteins detected by F-actin affinity chromatography, The Journal of cell biology, 109 (1989) 2963-2975.

[21] P.A. Hall, C.B. Todd, P.L. Hyland, S.S. McDade, H. Grabsch, M. Dattani, K.J. Hillan, S.E. Russell, The septin-binding protein anillin is overexpressed in

diverse human tumors, Clinical cancer research : an official journal of the American Association for Cancer Research, 11 (2005) 6780-6786.

[22] Z. Wang, J. Chen, M.-Z. Zhong, J. Huang, Y.-P. Hu, D.-Y. Feng, Z.-J. Zhou, X. Luo, Z.-Q. Liu, W.-Z. Jiang, Overexpression of ANLN contributed to poor prognosis of anthracycline-based chemotherapy in breast cancer patients, Cancer chemotherapy and pharmacology, 79 (2017) 535-543.

[23] J. Schiewek, U. Schumacher, T. Lange, S.A. Joosse, H. Wikman, K. Pantel, M. Mikhaylova, M. Kneussel, S. Linder, B. Schmalfeldt, Clinical relevance of cytoskeleton associated proteins for ovarian cancer, Journal of cancer research and clinical oncology, 144 (2018) 2195-2205.

[24] G. Wang, W. Shen, L. Cui, W. Chen, X. Hu, J. Fu, Overexpression of Anillin (ANLN) is correlated with colorectal cancer progression and poor prognosis, Cancer Biomarkers, 16 (2016) 459-465.

[25] C. Suzuki, Y. Daigo, N. Ishikawa, T. Kato, S. Hayama, T. Ito, E. Tsuchiya, Y. Nakamura, ANLN plays a critical role in human lung carcinogenesis through the activation of RHOA and by involvement in the phosphoinositide 3-kinase/AKT pathway, Cancer research, 65 (2005) 11314-11325.

[26] M. Olakowski, T. Tyszkiewicz, M. Jarzab, R. Krol, M. Oczko-Wojciechowska, M. Kowalska, M. Kowal, G.M. Gala, M. Kajor, D. Lange, E. Chmielik, E. Gubala, P. Lampe, B. Jarzab, NBL1 and anillin (ANLN) genes over-expression in pancreatic carcinoma, Folia histochemica et cytobiologica, 47 (2009) 249-255.

[27] S. Zeng, X. Yu, C. Ma, R. Song, Z. Zhang, X. Zi, X. Chen, Y. Wang, Y. Yu, J. Zhao, Transcriptome sequencing identifies ANLN as a promising prognostic biomarker in bladder urothelial carcinoma, Scientific reports, 7 (2017) 3151.

[28] D. Grbavec, R. Lo, Y. Liu, S. Stifani, Transducin-like Enhancer of split 2, a mammalian homologue of Drosophila Groucho, acts as a transcriptional repressor, interacts with Hairy/Enhancer of split proteins, and is expressed during neuronal development, European journal of biochemistry, 258 (1998) 339-349.

[29] B.G. Hoffman, B. Zavaglia, M. Beach, C.D. Helgason, Expression of Groucho/TLE proteins during pancreas development, BMC developmental biology, 8 (2008) 81.

[30] Z. He, Y. Liu, D. Liang, Z. Wang, E.S. Robertson, K. Lan, Cellular corepressor TLE2 inhibits replication-and-transcription-activator-mediated transactivation and lytic reactivation of Kaposi's sarcoma-associated herpesvirus, Journal of virology, 84 (2010) 2047-2062.

[31] K. Shin, J. Lee, N. Guo, J. Kim, A. Lim, L. Qu, I.U. Mysorekar, P.A. Beachy, Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder, Nature, 472 (2011) 110-114.

[32] H.F. Du, L.P. Ou, C.K. Lv, X. Yang, X.D. Song, Y.R. Fan, X.H. Wu, C.L. Luo, Expression of hepaCAM inhibits bladder cancer cell proliferation via a Wnt/β-catenin-dependent pathway in vitro and in vivo, Cancer Biol Ther, 16 (2015) 1502-1513.

[33] Z. Chen, L. Zhou, L. Wang, G. Kazobinka, X. Zhang, X. Han, B. Li, T. Hou, HBO1 promotes cell proliferation in bladder cancer via activation of Wnt/β-catenin signaling, Molecular carcinogenesis, 57 (2018) 12-21.

[34] N.S. Pandi, M. Manimuthu, P. Harunipriya, M. Murugesan, G.V. Asha, S. Rajendran, In silico analysis of expression pattern of a Wnt/beta-catenin responsive gene ANLN in gastric cancer, Gene, 545 (2014) 23-29.

[35] R.A. Cavallo, R.T. Cox, M.M. Moline, J. Roose, G.A. Polevoy, H. Clevers, M. Peifer, A. Bejsovec, Drosophila Tcf and Groucho interact to repress Wingless signalling activity, Nature, 395 (1998) 604-608.

[36] E. Cinnamon, Z. Paroush, Context-dependent regulation of Groucho/TLE-mediated repression, Current opinion in genetics & development, 18 (2008) 435-440.

[37] J.R. Prensner, A.M. Chinnaiyan, The emergence of lncRNAs in cancer biology, Cancer Discov, 1 (2011) 391-407.

[38] J.T. Kung, D. Colognori, J.T. Lee, Long noncoding RNAs: past, present, and future, Genetics, 193 (2013) 651-669.

[39] K.V. Morris, S. Santoso, A.M. Turner, C. Pastori, P.G. Hawkins, Bidirectional transcription directs both transcriptional gene activation and suppression in human cells, PLoS genetics, 4 (2008) e1000258.

[40] W. Yu, D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A.P. Feinberg, H. Cui, Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA, Nature, 451 (2008) 202-206.

[41] K.M. Schmitz, C. Mayer, A. Postepska, I. Grummt, Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes, Genes & development, 24 (2010) 2264-2269.

[42] C. Mayer, K.M. Schmitz, J. Li, I. Grummt, R. Santoro, Intergenic transcripts regulate the epigenetic state of rRNA genes, Mol Cell, 22 (2006) 351-361.

[43] P. Johnsson, A. Ackley, L. Vidarsdottir, W.O. Lui, M. Corcoran, D. Grandér, K.V. Morris, A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells, Nature structural & molecular biology, 20 (2013) 440-446.

[44] P.G. Hawkins, K.V. Morris, Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5, Transcription, 1 (2010) 165-175.

[45] A.M. Khalil, M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B.E. Bernstein, A. van Oudenaarden, A. Regev, E.S. Lander, J.L. Rinn, Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression, Proc. Natl. Acad. Sci. U. S. A., 106 (2009) 11667-11672.

[46] J. Zhao, T.K. Ohsumi, J.T. Kung, Y. Ogawa, D.J. Grau, K. Sarma, J.J. Song, R.E. Kingston, M. Borowsky, J.T. Lee, Genome-wide identification of polycomb-associated RNAs by RIP-seq, Mol Cell, 40 (2010) 939-953.

[47] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell, 144 (2011) 646-674.

[48] P. Hainaut, A. Plymoth, Targeting the hallmarks of cancer: towards a

rational approach to next-generation cancer therapy, Current opinion in oncology, 25 (2013) 50-51.

[49] B.K. Dey, A.C. Mueller, A. Dutta, Long non-coding RNAs as emerging regulators of differentiation, development, and disease, Transcription, 5 (2014) e944014.

[50] M.J. Hangauer, I.W. Vaughn, M.T. McManus, Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs, PLoS genetics, 9 (2013) e1003569.

[51] Y. Wang, D. Xue, Y. Li, X. Pan, X. Zhang, B. Kuang, M. Zhou, X. Li, W. Xiong, G. Li, Z. Zeng, T. Yang, The Long Noncoding RNA MALAT-1 is A Novel Biomarker in Various Cancers: A Meta-analysis Based on the GEO Database and Literature, J Cancer, 7 (2016) 991-1001.

[52] A.C. Tahira, M.S. Kubrusly, M.F. Faria, B. Dazzani, R.S. Fonseca, V. Maracaja-Coutinho, S. Verjovski-Almeida, M.C. Machado, E.M. Reis, Long noncoding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer, Mol Cancer, 10 (2011) 141.

[53] F. Yang, J. Bi, X. Xue, L. Zheng, K. Zhi, J. Hua, G. Fang, Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells, The FEBS journal, 279 (2012) 3159-3165.

[54] R.A. Gupta, N. Shah, K.C. Wang, J. Kim, H.M. Horlings, D.J. Wong, M.C. Tsai, T. Hung, P. Argani, J.L. Rinn, Y. Wang, P. Brzoska, B. Kong, R. Li, R.B. West, M.J. van de Vijver, S. Sukumar, H.Y. Chang, Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis, Nature, 464 (2010) 1071-1076.

[55] K.H. Lu, W. Li, X.H. Liu, M. Sun, M.L. Zhang, W.Q. Wu, W.P. Xie, Y.Y. Hou, Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and induces apoptosis by affecting p53 expression, BMC Cancer, 13 (2013) 461.

[56] N. Miyoshi, H. Wagatsuma, S. Wakana, T. Shiroishi, M. Nomura, K. Aisaka, T. Kohda, M.A. Surani, T. Kaneko-Ishino, F. Ishino, Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q, Genes to cells : devoted to molecular & cellular mechanisms, 5 (2000) 211-220.

[57] X. Lu, Y. Fang, Z. Wang, J. Xie, Q. Zhan, X. Deng, H. Chen, J. Jin, C. Peng, H. Li, B. Shen, Downregulation of gas5 increases pancreatic cancer cell proliferation by regulating CDK6, Cell and tissue research, 354 (2013) 891-896.

[58] X. Shi, M. Sun, H. Liu, Y. Yao, R. Kong, F. Chen, Y. Song, A critical role for the long non-coding RNA GAS5 in proliferation and apoptosis in non-small-cell lung cancer, Mol Carcinog, 54 Suppl 1 (2015) E1-e12.

[59] E.S. Martens-Uzunova, R. Böttcher, C.M. Croce, G. Jenster, T. Visakorpi, G.A. Calin, Long noncoding RNA in prostate, bladder, and kidney cancer, European urology, 65 (2014) 1140-1151.

[60] I. Ariel, M. Sughayer, Y. Fellig, G. Pizov, S. Ayesh, D. Podeh, B.A. Libdeh, C. Levy, T. Birman, M.L. Tykocinski, N. de Groot, A. Hochberg, The imprinted H19 gene is a marker of early recurrence in human bladder carcinoma, Mol

Pathol, 53 (2000) 320-323.

[61] X.S. Wang, Z. Zhang, H.C. Wang, J.L. Cai, Q.W. Xu, M.Q. Li, Y.C. Chen, X.P. Qian, T.J. Lu, L.Z. Yu, Y. Zhang, D.Q. Xin, Y.Q. Na, W.F. Chen, Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma, Clinical cancer research : an official journal of the American Association for Cancer Research, 12 (2006) 4851-4858.

[62] Y. Han, Y. Liu, L. Nie, Y. Gui, Z. Cai, Inducing cell proliferation inhibition, apoptosis, and motility reduction by silencing long noncoding ribonucleic acid metastasis-associated lung adenocarcinoma transcript 1 in urothelial carcinoma of the bladder, Urology, 81 (2013) 209.e201-207.

[63] L. Ying, Y. Huang, H. Chen, Y. Wang, L. Xia, Y. Chen, Y. Liu, F. Qiu, Downregulated MEG3 activates autophagy and increases cell proliferation in bladder cancer, Molecular bioSystems, 9 (2013) 407-411.

[64] Y. Han, Y. Liu, Y. Gui, Z. Cai, Long intergenic non-coding RNA TUG1 is overexpressed in urothelial carcinoma of the bladder, Journal of surgical oncology, 107 (2013) 555-559.

[65] M. Luo, Z. Li, W. Wang, Y. Zeng, Z. Liu, J. Qiu, Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression, Cancer letters, 333 (2013) 213-221.

[66] M. Luo, Z. Li, W. Wang, Y. Zeng, Z. Liu, J. Qiu, Upregulated H19 contributes to bladder cancer cell proliferation by regulating ID2 expression, The FEBS journal, 280 (2013) 1709-1716.

[67] Y. Shi, J. Lu, J. Zhou, X. Tan, Y. He, J. Ding, Y. Tian, L. Wang, K. Wang, Long non-coding RNA Loc554202 regulates proliferation and migration in breast cancer cells, Biochem Biophys Res Commun, 446 (2014) 448-453.

[68] J. Ding, B. Lu, J. Wang, J. Wang, Y. Shi, Y. Lian, Y. Zhu, J. Wang, Y. Fan, Z. Wang, W. De, K. Wang, Long non-coding RNA Loc554202 induces apoptosis in colorectal cancer cells via the caspase cleavage cascades, J Exp Clin Cancer Res, 34 (2015) 100-100.

[69] F.-Q. Nie, S. Ma, M. Xie, Y.-W. Liu, W. De, X.-H. Liu, Decreased long noncoding RNA MIR31HG is correlated with poor prognosis and contributes to cell proliferation in gastric cancer, Tumour Biol, 37 (2016) 7693-7701.

[70] H. Yang, P. Liu, J. Zhang, X. Peng, Z. Lu, S. Yu, Y. Meng, W.M. Tong, J. Chen, Long noncoding RNA MIR31HG exhibits oncogenic property in pancreatic ductal adenocarcinoma and is negatively regulated by miR-193b, Oncogene, 35 (2016) 3647-3657.

[71] W. Dandan, C. Jianliang, H. Haiyan, M. Hang, L. Xuedong, Long noncoding RNA MIR31HG is activated by SP1 and promotes cell migration and invasion by sponging miR-214 in NSCLC, Gene, 692 (2019) 223-230.

[72] A. He, Z. Chen, H. Mei, Y. Liu, Decreased expression of LncRNA MIR31HG in human bladder cancer, Cancer Biomark, 17 (2016) 231-236.

[73] G.P. Paner, W.M. Stadler, D.E. Hansel, R. Montironi, D.W. Lin, M.B. Amin, Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers, European urology, 73 (2018) 560-569.

[74] S.L. Woldu, A. Bagrodia, Y. Lotan, Guideline of guidelines: non-muscle-invasive bladder cancer, BJU Int, 119 (2017) 371-380.

[75] S. Rinaldetti, E. Rempel, T.S. Worst, M. Eckstein, A. Steidler, C.A. Weiss, C. Bolenz, A. Hartmann, P. Erben, Subclassification, survival prediction and drug target analyses of chemotherapy-naïve muscle-invasive bladder cancer with a molecular screening, Oncotarget, 9 (2018) 25935-25945.

[76] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A.A. Margolin, S. Kim, C.J. Wilson, J. Lehár, G.V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M.F. Berger, J.E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F.A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I.H. Engels, J. Cheng, G.K. Yu, J. Yu, P. Aspesi, Jr., M. de Silva, K. Jagtap, M.D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R.C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J.P. Mesirov, S.B. Gabriel, G. Getz, K. Ardlie, V. Chan, V.E. Myer, B.L. Weber, J. Porter, M. Warmuth, P. Finan, J.L. Harris, M. Meyerson, T.R. Golub, M.P. Morrissey, W.R. Sellers, R. Schlegel, L.A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity, Nature, 483 (2012) 603-607.

[77] Q. Lu, S. Ren, M. Lu, Y. Zhang, D. Zhu, X. Zhang, T. Li, Computational prediction of associations between long non-coding RNAs and proteins, BMC genomics, 14 (2013) 651.

[78] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic acids research, 44 (2016) W90-W97.

[79] W. Sun, T. Duan, P. Ye, K. Chen, G. Zhang, M. Lai, H. Zhang, TSVdb: a web-tool for TCGA splicing variants analysis, BMC genomics, 19 (2018) 405.

[80] J.L. Petzoldt, I.M. Leigh, P.G. Duffy, C. Sexton, J.R.W. Masters, Immortalisation of human urothelial cells, Urological research, 23 (1995) 377-380.

[81] C.C. Rigby, L.M. Franks, A Human Tissue Culture Cell Line from a Transitional Cell Tumour of the Urinary Bladder: Growth, Chromosome Pattern and Ultrastructure, Br J Cancer, 24 (1970) 746-754.

[82] J.R. Masters, P.J. Hepburn, L. Walker, W.J. Highman, L.K. Trejdosiewicz, S. Povey, M. Parkar, B.T. Hill, P.R. Riddle, L.M. Franks, Tissue culture model of transitional cell carcinoma: characterization of twenty-two human urothelial cell lines, Cancer research, 46 (1986) 3630-3636.

[83] H.B. Grossman, G. Wedemeyer, L. Ren, G.N. Wilson, B. Cox, Improved growth of human urothelial carcinoma cell cultures, The Journal of urology, 136 (1986) 953-959.

[84] J. Bubenik, M. Baresova, V. Viklicky, J. Jakoubkova, H. Sainerova, J. Donner, Established cell line of urinary bladder carcinoma (T24) containing tumour-specific antigen, International journal of cancer, 11 (1973) 765-773.

[85] C. O'Toole, S. Nayak, Z. Price, W.H. Gilbert, J. Waisman, A cell line

(SCABER) derived from squamous cell carcinoma of the human urinary bladder, International journal of cancer, 17 (1976) 707-714.

[86] N. Agrawal, P.V. Dasaradhi, A. Mohmmed, P. Malhotra, R.K. Bhatnagar, S.K. Mukherjee, RNA interference: biology, mechanism, and applications, Microbiology and molecular biology reviews : MMBR, 67 (2003) 657-685.

[87] T. Martini, J. Heinkele, R. Mayr, C.A. Weis, F. Wezel, S. Wahby, M. Eckstein, T. Schnoller, J. Breyer, R. Wirtz, M. Ritter, C. Bolenz, P. Erben, Predictive value of lymphangiogenesis and proliferation markers on mRNA level in urothelial carcinoma of the bladder after radical cystectomy, Urologic oncology, 36 (2018) 530.e519-530.e527.

[88] T. Tramm, B.S. Sørensen, J. Overgaard, J. Alsner, Optimal reference genes for normalization of qRT-PCR data from archival formalin-fixed, paraffin-embedded breast tumors controlling for tumor cell content and decay of mRNA, Diagnostic Molecular Pathology, 22 (2013) 181-187.

[89] J. Breyer, R.M. Wirtz, W. Otto, P. Erben, M.C. Kriegmair, R. Stoehr, M. Eckstein, S. Eidt, S. Denzinger, M. Burger, In stage pT1 non-muscle-invasive bladder cancer (NMIBC), high KRT20 and low KRT5 mRNA expression identify the luminal subtype and predict recurrence and survival, Virchows Archiv, 470 (2017) 267-274.

[90] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method, methods, 25 (2001) 402-408.

[91] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nature Methods, 9 (2012) 671-675.

[92] T. Gebäck, M.M.P. Schulz, P. Koumoutsakos, M. Detmar, TScratch: a novel and simple software tool for automated analysis of monolayer wound healing assays, Biotechniques, 46 (2009) 265-274.

[93] S. Patel, M.T. Park, M.M. Chakravarty, J. Knight, Gene Prioritization for Imaging Genetics Studies Using Gene Ontology and a Stratified False Discovery Rate Approach, Frontiers in neuroinformatics, 10 (2016) 14.

[94] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs, Nucleic Acids Res, 45 (2017) D353-d361.

[95] M.R.D. Zanetti, C.D. Petricelli, S.M. Alexandre, A. Paschoal, E. Araujo Júnior, M.U. Nakamura, Determination of a cutoff value for pelvic floor distensibility using the Epi-no balloon to predict perineal integrity in vaginal delivery: ROC curve analysis. Prospective observational single cohort study, Sao Paulo Medical Journal, 134 (2016) 97-102.

[96] J.P. Stein, G. Lieskovsky, R. Cote, S. Groshen, A.-C. Feng, S. Boyd, E. Skinner, B. Bochner, D. Thangathurai, M. Mikhail, D. Raghavan, D.G. Skinner, Radical Cystectomy in the Treatment of Invasive Bladder Cancer: Long-Term Results in 1,054 Patients, Journal of Clinical Oncology, 19 (2001) 666-675.

[97] R.E. Hautmann, R.C. de Petriconi, C. Pfeiffer, B.G. Volkmer, Radical cystectomy for urothelial carcinoma of the bladder without neoadjuvant or

adjuvant therapy: long-term results in 1100 patients, European urology, 61 (2012) 1039-1047.

[98] S. Zhang, L.H. Nguyen, K. Zhou, H.-C. Tu, A. Sehgal, I. Nassour, L. Li, P. Gopal, J. Goodman, A.G. Singal, Knockdown of anillin actin binding protein blocks cytokinesis in hepatocytes and reduces liver tumor development in mice without affecting regeneration, Gastroenterology, 154 (2018) 1421-1434.

[99] P.-I. Liang, W.T. Chen, C.-F. Li, C.-C. Li, W.-M. Li, C.-N. Huang, H.-C. Yeh, H.-L. Ke, W.-J. Wu, C.-Y. Chai, Subcellular localisation of anillin is associated with different survival outcomes in upper urinary tract urothelial carcinoma, Journal of clinical pathology, 68 (2015) 1026-1032.

[100] L. Wang, M. Zhou, C. Feng, P. Gao, G. Ding, Z. Zhou, H. Jiang, Z. Wu, Q. Ding, Prognostic value of Ki67 and p63 expressions in bladder cancer patients who underwent radical cystectomy, International urology and nephrology, 48 (2016) 495-501.

[101] R. Stone, A.L. Sabichi, J. Gill, I.-l. Lee, P. Adegboyega, M.S. Dai, R. Loganantharaj, M. Trutschl, U. Cvek, J.L. Clifford, Identification of genes correlated with early-stage bladder cancer progression, Cancer Prevention Research, 3 (2010) 776-786.

[102] J. Breyer, R.M. Wirtz, M. Laible, K. Schlombs, P. Erben, M.C. Kriegmair, R. Stoehr, S. Eidt, S. Denzinger, M. Burger, ESR1, ERBB2, and Ki67 mRNA expression predicts stage and grade of non-muscle-invasive bladder carcinoma (NMIBC), Virchows Archiv, 469 (2016) 547-552.

[103] S. Rinaldetti, R.M. Wirtz, T.S. Worst, M. Eckstein, C.A. Weiss, J. Breyer, W. Otto, C. Bolenz, A. Hartmann, P. Erben, FOXM1 predicts overall and disease specific survival in muscle-invasive urothelial carcinoma and presents a differential expression between bladder cancer subtypes, Oncotarget, 8 (2017) 47595.

[104] Z. Pei, X. Du, Y. Song, L. Fan, F. Li, Y. Gao, R. Wu, Y. Chen, W. Li, H. Zhou, Down-regulation of lncRNA CASC2 promotes cell proliferation and metastasis of bladder cancer by activation of the Wnt/β-catenin signaling pathway, Oncotarget, 8 (2017) 18145.

[105] H.-f. Du, L.-p. Ou, C.-k. Lv, X. Yang, X.-d. Song, Y.-r. Fan, X.-h. Wu, C.-l. Luo, Expression of hepaCAM inhibits bladder cancer cell proliferation via a Wnt/β-catenin-dependent pathway in vitro and in vivo, Cancer biology & therapy, 16 (2015) 1502-1513.

[106] K.C. Kurnit, G.N. Kim, B.M. Fellman, D.L. Urbauer, G.B. Mills, W. Zhang, R.R. Broaddus, CTNNB1 (beta-catenin) mutation identifies low grade, early stage endometrial cancer patients at increased risk of recurrence, Modern Pathology, 30 (2017) 1032.

[107] S. Crippa, P.B. Ancey, J. Vazquez, P. Angelino, A.L. Rougemont, C. Guettier, V. Zoete, M. Delorenzi, O. Michielin, E. Meylan, Mutant CTNNB1 and histological heterogeneity define metabolic subtypes of hepatoblastoma, EMBO molecular medicine, 9 (2017) 1589-1604.

[108] A. Ahadova, M. von Knebel Doeberitz, H. Bläker, M. Kloor, CTNNB1-

mutant colorectal carcinomas with immediate invasive growth: a model of interval cancers in Lynch syndrome, Familial cancer, 15 (2016) 579-586.

[109] R. Ai, Y. Sun, Z. Guo, W. Wei, L. Zhou, F. Liu, D.T. Hendricks, Y. Xu, X. Zhao, NDRG1 overexpression promotes the progression of esophageal squamous cell carcinoma through modulating Wnt signaling pathway, Cancer Biol Ther, 17 (2016) 943-954.

[110] N. Hambly, C. Shimbori, M. Kolb, Molecular classification of idiopathic pulmonary fibrosis: personalized medicine, genetics and biomarkers, Respirology, 20 (2015) 1010-1022.

[111] A.G. Robertson, J. Kim, H. Al-Ahmadie, J. Bellmunt, G. Guo, A.D. Cherniack, T. Hinoue, P.W. Laird, K.A. Hoadley, R. Akbani, Comprehensive molecular characterization of muscle-invasive bladder cancer, Cell, 171 (2017) 540-556. e525.

[112] E.J. Pietzak, A. Bagrodia, E.K. Cha, E.N. Drill, G. Iyer, S. Isharwal, I. Ostrovnaya, P. Baez, Q. Li, M.F. Berger, Next-generation sequencing of nonmuscle invasive bladder cancer reveals potential biomarkers and rational therapeutic targets, European urology, 72 (2017) 952-959.

[113] J. Hedegaard, P. Lamy, I. Nordentoft, F. Algaba, S. Høyer, B.P. Ulhøi, S. Vang, T. Reinert, G.G. Hermann, K. Mogensen, Comprehensive transcriptional analysis of early-stage urothelial carcinoma, Cancer cell, 30 (2016) 27-42.

[114] J.S. Damrauer, K.A. Hoadley, D.D. Chism, C. Fan, C.J. Tiganelli, S.E. Wobker, J.J. Yeh, M.I. Milowsky, G. Iyer, J.S. Parker, Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology, Proceedings of the National Academy of Sciences, 111 (2014) 3110-3115.

[115] M. Fedele, L. Cerchia, G. Chiappetta, The Epithelial-to-Mesenchymal Transition in Breast Cancer: Focus on Basal-Like Carcinomas, Cancers, 9 (2017).

[116] M. Eckstein, R.M. Wirtz, M. Gross-Weege, J. Breyer, W. Otto, R. Stoehr, D. Sikic, B. Keck, S. Eidt, M. Burger, C. Bolenz, K. Nitschke, S. Porubsky, A. Hartmann, P. Erben, mRNA-Expression of KRT5 and KRT20 Defines Distinct Prognostic Subgroups of Muscle-Invasive Urothelial Bladder Cancer Correlating with Histological Variants, Int J Mol Sci, 19 (2018).

[117] J.W. Shih, W.F. Chiang, A.T.H. Wu, M.H. Wu, L.Y. Wang, Y.L. Yu, Y.W. Hung, W.C. Wang, C.Y. Chu, C.L. Hung, C.A. Changou, Y. Yen, H.J. Kung, Long noncoding RNA LncHIFCAR/MIR31HG is a HIF-1α co-activator driving oral cancer progression, Nature communications, 8 (2017) 15874.

[118] K. Augoff, B. McCue, E.F. Plow, K. Sossey-Alaoui, miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer, Molecular cancer, 11 (2012) 5-5.

[119] C.G.A.R. Network, Comprehensive molecular characterization of urothelial bladder carcinoma, Nature, 507 (2014) 315-322.

[120] A. Chaux, J.S. Cohen, L. Schultz, R. Albadine, S. Jadallah, K.M. Murphy, R. Sharma, M.P. Schoenberg, G.J. Netto, High epidermal growth factor receptor immunohistochemical expression in urothelial carcinoma of the bladder is not

associated with EGFR mutations in exons 19 and 21: a study using formalin-fixed, paraffin-embedded archival tissues, Human pathology, 43 (2012) 1590-1595.

[121] B. Wang, H. Jiang, L. Wang, X. Chen, K. Wu, S. Zhang, S. Ma, B. Xia, Increased MIR31HG lncRNA expression increases gefitinib resistance in non-small cell lung cancer cell lines through the EGFR/PI3K/AKT signaling pathway, Oncol Lett, 13 (2017) 3494-3500.

[122] O. Kelemen, P. Convertini, Z. Zhang, Y. Wen, M. Shen, M. Falaleeva, S. Stamm, Function of alternative splicing, Gene, 514 (2013) 1-30.

[123] A.J. Ward, T.A. Cooper, The pathobiology of splicing, The Journal of pathology, 220 (2010) 152-163.

[124] T.A. Hughes, Regulation of gene expression by alternative untranslated regions, Trends in genetics : TIG, 22 (2006) 119-122.

[125] Q. Pan, O. Shai, L.J. Lee, B.J. Frey, B.J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing, Nature genetics, 40 (2008) 1413-1415.

[126] B.R. Graveley, Alternative splicing: increasing diversity in the proteomic world, Trends in genetics : TIG, 17 (2001) 100-107.

[127] T. Yang, H. Zhou, P. Liu, L. Yan, W. Yao, K. Chen, J. Zeng, H. Li, J. Hu, H. Xu, Z. Ye, lncRNA PVT1 and its splicing variant function as competing endogenous RNA to regulate clear cell renal cell carcinoma progression, Oncotarget, 8 (2017) 85353-85367.

[128] Y.W. Moon, J. Hajjar, P. Hwu, A. Naing, Targeting the indoleamine 2, 3-dioxygenase pathway in cancer, Journal for immunotherapy of cancer, 3 (2015) 51.

[129] P.C. Tumeh, C.L. Harview, J.H. Yearley, I.P. Shintaku, E.J. Taylor, L. Robert, B. Chmielowski, M. Spasic, G. Henry, V. Ciobanu, PD-1 blockade induces responses by inhibiting adaptive immune resistance, Nature, 515 (2014) 568.

[130] D.H. Aggen, C.G. Drake, Biomarkers for immunotherapy in bladder cancer: a moving target, Journal for immunotherapy of cancer, 5 (2017) 94.

[131] S. Rebouissou, I. Bernard-Pierrot, A. de Reyniès, M.L. Lepage, C. Krucker, E. Chapeaublanc, A. Hérault, A. Kamoun, A. Caillault, E. Letouzé, N. Elarouci, Y. Neuzillet, Y. Denoux, V. Molinié, D. Vordos, A. Laplanche, P. Maillé, P. Soyeux, K. Ofualuka, F. Reyal, A. Biton, M. Sibony, X. Paoletti, J. Southgate, S. Benhamou, T. Lebret, Y. Allory, F. Radvanyi, EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype, Science translational medicine, 6 (2014) 244ra291.

[132] P.L. Ho, A. Kurtova, K.S. Chan, Normal and neoplastic urothelial stem cells: getting to the root of the problem, Nature reviews. Urology, 9 (2012) 583-594.

[133] D.J. McConkey, W. Choi, C.P. Dinney, New insights into subtypes of invasive bladder cancer: considerations of the clinician, European urology, 66 (2014) 609-610.

[134] J. Kardos, S. Chai, L.E. Mose, S.R. Selitsky, B. Krishnan, R. Saito, M.D. Iglesia, M.I. Milowsky, J.S. Parker, W.Y. Kim, B.G. Vincent, Claudin-low bladder tumors are immune infiltrated and actively immune suppressed, JCI insight, 1 (2016) e85902.

[135] Y. Sun, X. Jia, M. Wang, Y. Deng, Long noncoding RNA MIR31HG abrogates the availability of tumor suppressor microRNA-361 for the growth of osteosarcoma, Cancer management and research, 11 (2019) 8055-8064.

[136] S. Yan, Z. Tang, K. Chen, Y. Liu, G. Yu, Q. Chen, H. Dang, F. Chen, J. Ling, L. Zhu, A. Huang, H. Tang, Long noncoding RNA MIR31HG inhibits hepatocellular carcinoma proliferation and metastasis by sponging microRNA-575 to modulate ST7L expression, J Exp Clin Cancer Res, 37 (2018) 214.

[137] K. Sun, X. Zhao, J. Wan, L. Yang, J. Chu, S. Dong, H. Yin, L. Ming, F. He, The diagnostic value of long non-coding RNA MIR31HG and its role in esophageal squamous cell carcinoma, Life sciences, 202 (2018) 124-130.

[138] L. Wan, W. Yu, E. Shen, W. Sun, Y. Liu, J. Kong, Y. Wu, F. Han, L. Zhang, T. Yu, Y. Zhou, S. Xie, E. Xu, H. Zhang, M. Lai, SRSF6-regulated alternative splicing that promotes tumour progression offers a therapy target for colorectal cancer, Gut, 68 (2019) 118-129.

[139] T. Yae, K. Tsuchihashi, T. Ishimoto, T. Motohara, M. Yoshikawa, G.J. Yoshida, T. Wada, T. Masuko, K. Mogushi, H. Tanaka, T. Osawa, Y. Kanki, T. Minami, H. Aburatani, M. Ohmura, A. Kubo, M. Suematsu, K. Takahashi, H. Saya, O. Nagano, Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell, Nature communications, 3 (2012) 883.

[140] L. Prochazka, R. Tesarik, J. Turanek, Regulation of alternative splicing of CD44 in cancer, Cellular signalling, 26 (2014) 2234-2239.

[141] M.R. Ladomery, S.J. Harper, D.O. Bates, Alternative splicing in angiogenesis: the vascular endothelial growth factor paradigm, Cancer letters, 249 (2007) 133-142.

[142] R. Hu, C. Lu, E.A. Mostaghel, S. Yegnasubramanian, M. Gurel, C. Tannahill, J. Edwards, W.B. Isaacs, P.S. Nelson, E. Bluemn, S.R. Plymate, J. Luo, Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer, Cancer research, 72 (2012) 3457-3462.

[143] S. Sun, C.C. Sprenger, R.L. Vessella, K. Haugk, K. Soriano, E.A. Mostaghel, S.T. Page, I.M. Coleman, H.M. Nguyen, H. Sun, P.S. Nelson, S.R. Plymate, Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant, The Journal of clinical investigation, 120 (2010) 2715-2730.

[144] P.I. Poulikakos, Y. Persaud, M. Janakiraman, X. Kong, C. Ng, G. Moriceau, H. Shi, M. Atefi, B. Titz, M.T. Gabay, M. Salton, K.B. Dahlman, M. Tadi, J.A. Wargo, K.T. Flaherty, M.C. Kelley, T. Misteli, P.B. Chapman, J.A. Sosman, T.G. Graeber, A. Ribas, R.S. Lo, N. Rosen, D.B. Solit, RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E), Nature, 480 (2011) 387-390.

REFERENCES

[145] M. Salton, W.K. Kasprzak, T. Voss, B.A. Shapiro, P.I. Poulikakos, T. Misteli, Inhibition of vemurafenib-resistant melanoma by interference with pre-mRNA splicing, Nature communications, 6 (2015) 7103.

7 APPENDIX

7.1 List of figures

7.2 List of tables

8 CURRICULUM VITAE

Personal information
Name: Sheng Wu
Date of birth: Jan. 24th, 1989
Place of Birth: Anhui, China
Nationality: Chinese
Email: wusheng89@hotmail.com
Address: Alphornstrasse 46, 68169, Mannheim

Education
Oct. 2016-present   Ph.D. student (Dr.sc.hum.)
                               Ruprecht-Karls-University Heidelberg
Ph.D. project: Long non-coding RNAs in bladder cancer: functional roles
and clinical evaluation

Sep. 2015 - Sep. 2016   M.Sc. Translational Medical Research
                               Ruprecht-Karls-University Heidelberg
Thesis: Hormones receptors in upper tract urothelial carcinoma: an
evaluation of potential therapeutic targets

Sep. 2012 - July 2015   Master of Medicine
                               Anhui Medical University
Thesis: Expression and clinical significance of PD-L1 and PD-1 in gastric
carcinoma

Sep. 2007 - June 2012   Bachelor of Medicine
                               Anhui Medical University

                               Ph.D. scholarship from China Scholarship
Scholarship   Council

Publications
Wu S, Nitschke K, Heinkele J, Weis C-A, Worst TS, Eckstein M, Porubsky S, Erben P.
ANLN and TLE2 in Muscle Invasive Bladder Cancer: A Functional and Clinical Evaluation
Based on In Silico and In Vitro Data. Cancers. 2019; 11(12):1840.
Wu S, Nitschke K, Worst TS, Fierek A, Weis C-A, Eckstein M, Porubsky S, Kriegmair M,
Erben P. Long noncoding RNA MIR31HG and its splice variants regulate proliferation and
migration: prognostic implications for muscle invasive bladder cancer. Journal of
Experimental & Clinical Cancer Research. Under review.
Wu S, Shao JY, Wang F, Sun GP. Expression and clinical significance of PD-L1 and PD-1
in gastric carcinoma. Acta Universitatis Medicinalis Anhui, 2015, 50(6):821-5.

9 ACKNOWLEDGEMENT

Firstly, I would like to express my sincere appreciation to my mentor Professor Dr. Philipp Erben. I am extremely grateful that he provided me with the opportunity of my scientific research in bladder cancer. I am deeply grateful of his scientific discussion, support and advice in the completion of this thesis. I benefit a lot from his professional knowledge, supervised approach and friendly guidance during my whole PhD project.

I would like to thank Dr. Katja Nitschke and Dr. Thomas Worst for the continuous support in my research, posters and publications. Many thanks to the whole team of the Urological Research Laboratory, in particular to Lena Hoffmann, Sarah Wahby, Cornelia Oslow and very particularly to Annette Steidler, who are warm-hearted and helpful with regard to my work and life. They always supported me patiently and nicely.

Special thanks to my colleagues Jakob Heinkele and Alexander Fierek for providing the precious clinical data. I also thank other co-authors of my publications.

Especially, the Chinese Scholarship Council is gratefully acknowledged for financing my whole project in Germany.

Last but not the least, I would like to thank my family: my parents, my grandmother and uncles for tirelessly supporting me throughout my study and life. I love them forever.