Lek. Piotr Tomasz Wysocki

Functional characterization of *Disrupted in Renal Carcinoma 3* (*DIRC3*) long non-coding RNA in differentiated thyroid cancers

Rozprawa na stopień doktora nauk medycznych i nauk o zdrowiu w dyscyplinie nauki medyczne

Promotor: prof. dr hab. n. med. Dominika Nowis

Promotor pomocniczy: dr n. med. Monika Kolanowska

Laboratorium Medycyny Doświadczalnej

Obrona rozprawy doktorskiej przed Radą Dyscypliny Nauk Medycznych Warszawskiego Uniwersytetu Medycznego

Warszawa, 2022 r.

Keywords: *DIRC3*, thyroid cancer*, IGFBP5*, IGF-1, invasiveness.

Słowa kluczowe: *DIRC3*, rak tarczycy, *IGFBP5*, IGF-1, inwazyjność.

Piotr Tomasz Wysocki, M.D.

Functional characterization of *Disrupted in Renal Carcinoma 3* (*DIRC3*) long non-coding RNA in differentiated thyroid cancers

A thesis submitted for the degree of Doctor of Philosophy

Supervisor: Prof. Dominika Nowis, M.D., Ph.D.

Auxiliary supervisor: Dr Monika Kolanowska, Ph.D.

Laboratory of Experimental Medicine

Medical University of Warsaw

Warszawa, 2022

Keywords: *DIRC3*, thyroid cancer*, IGFBP5*, IGF-1, invasiveness.

PROJECT FUNDING

This project was supported by following grants:

“Badanie funkcjonalne genu *DIRC3* w zróżnicowanym raku tarczycy”

National Science Centre, Poland (Narodowe Centrum Nauki):
PRELUDIUM 2017/27/N/NZ2/03116. Principal Investigator: Piotr T. Wysocki

“Charakteryzacja funkcjonalna germinalnego wariantu ryzyka dla raka tarczycy: rs11693806”

Medical University of Warsaw, Poland:
The Young Investigator Grant: MB/M/12(26). Principal Investigator: Piotr T. Wysocki

013/RID/2018/19 (Regional Initiative for Excellence)

Polish Ministry of Education and Science, project budget 12,000,000 PLN

Principal Investigator: Dominika Nowis

ACKNOWLEDGEMENTS / PODZIĘKOWANIA

Realizacja pracy nie byłaby możliwa bez wsparcia otrzymanego od wielu osób.

Pragnę złożyć podziękowania:

*Dla mojej Promotor Pani **Profesor dr hab. n. med. Dominiki Nowis** za opiekę naukową, cenne rady, otrzymaną pomocy, zaangażowanie i ogromną życzliwość w trakcie powstawania pracy doktorskiej.*

*Dla mojej Promotor pomocniczej Pani **Doktor n. med. Moniki Kolanowskiej** za poświęcony czas, wskazówki oraz wsparcie merytoryczne w trakcie powstawania koncepcji pracy.*

*Dla Pana **Profesora dr hab. n. med. Krystiana Jażdżewskiego** z Warsaw Genomics za opiekę w pierwszych latach studiów doktoranckich, przedstawienie propozycji tematu pracy, umożliwienie realizacji badań oraz okazaną życzliwości.*

*Dla mojego Kolegi, **Doktora Karola Czubaka** z Laboratorium Medycyny Doświadczalnej WUM za otrzymaną pomoc i współpracę. Dla Pani **Doktor Anny Marusiak** z Międzynarodowego Instytutu Mechanizmów i Maszyn Molekularnych Polskiej Akademii Nauk za cenne rady i umożliwienie realizacji niektórych eksperymentów in vitro. Dla Pana **Doktora Pawła Gaja** z Warsaw Genomics za omówienie aspektów pracy dotyczących NGS. Dla moich Koleżanek i Kolegów z Laboratorium, **Doktor Beaty Rak** oraz **Doktora Wojciecha Gierlikowskiego** z Kliniki Chorób Wewnętrznych i Endokrynologii WUM za wspólnie spędzony czas oraz wiele dyskusji naukowych.*

*Dla Pana **Profesora dr hab. n. med. Jarosława Reguły** oraz Koleżanek i Kolegów z **Kliniki Gastroenterologii Onkologicznej** Narodowego Instytutu Onkologii w Warszawie za miłą atmosferę pracy.*

*I also want to express my gratitude to my Dear Friends, **Prof. Evgeny Izumchenko** from The University of Chicago and **Prof. Mariana Brait** from The Johns Hopkins University, who inspired me and shared with me their love of science. I also want to thank my first P.I., **Prof. David Sidransky** from The Johns Hopkins University, who gave me the opportunity to develop my passion for cancer research.*

Rodzinie dziękuję za otrzymane wsparcie.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AKT	Protein kinase B
ASO	Antisense oligonucleotide
BRAF	*v-Raf murine sarcoma viral oncogene type B*
Cas9	CRISPR-associated protein 9
CDKN1A	*Cyclin-dependent kinase inhibitor 1A*
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
DEG	Differentially expressed gene
DFS	Disease-free survival
DIRC3	*Disrupted in renal carcinoma 3*
DTC	Differentiated thyroid cancer
ECM	Extracellular matrix
eQTL	Expression quantitative trait locus
ER	Estrogen receptor
ERK	Extracellular signal-regulated kinase
FBS	Fetal bovine serum
FTC	Follicular thyroid cancer
GH	Growth hormone
GO	Gene Ontology
gRNA	Guide RNA
GTEx	The Genotype-Tissue Expression (project)
GWAS	Genome-wide association study
HDR	Homology-directed repair
HIF-1	Hypoxia-inducible factor 1
HOTAIR	*HOX transcript antisense RNA*
HPRT1	*Hypoxanthine phosphoribosyltransferase 1*
HR	Hazard ratio
HRAS	*Harvey rat sarcoma virus*
HSPBAP1	*Heat shock 27-kDa-associated protein 1*
HTC	Hürthle cell thyroid cancer
IGF-1	Insulin-like growth 1

IGF-1R	Insulin-like growth factor 1 receptor
IGFBP5	Insulin-like growth factor binding protein 5
KRAS	*Kirsten rat sarcoma virus*
lncRNA	Long non-coding RNA
MALAT1	*Metastasis-associated lung adenocarcinoma transcript 1*
MAPK	Mitogen-activated protein kinase
MEK	MAP kinase kinase (MAPKK)
MMP	Matrix metalloproteinase
mTOR	Mechanistic/mammalian target of rapamycin
mTORC1	mTOR complex 1
MTT	4,5-dimethylthiazol-2-yl-2,5-diphenyltetrazolium bromide
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	Non-homologous end joining
NRAS	*Neuroblastoma RAS viral oncogene homolog*
OD	Optical density
OR	Odds ratio
OS	Overall survival
PAPP-A	Pregnancy-associated plasma protein A
PCA	Principal components analysis
PCR	Polymerase chain reaction
PDK1	Phosphatidylinositol-dependent protein kinase 1
PIK3	Phosphatidylinositol 3-kinase
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol (3,4,5)-trisphosphate
PRE	Putative regulatory element
PTC	Papillary thyroid cancer
PTEN	*Phosphatase and tensin homolog*
PTSC	*PTC susceptibility candidate*
PVT1	*Plasmacytoma variant translocation 1*
RAI	Radioactive iodine
RAS	*Rat sarcoma virus*
RET	*Rearranged during transfection*

RIPA	Radioimmunoprecipitation assay
RNAi	RNA interference
RNP	Ribonucleoprotein
RTK	Receptor tyrosine kinase
S6K	S6 kinase
SD	Standard deviation
sgRNA	Single-guide RNA
SNHG5	*Small nucleolar RNA host gene 5*
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
STC1	*Stanniocalcin 1*
T7EI	T7 endonuclease I
TCGA	The Cancer Genome Atlas
TERT	*Telomerase reverse transcriptase*
TF	Transcription Factor
TKI	Tyrosine kinase inhibitor
TNM	Tumor, Node, Metastasis (classification)
TP53	*Tumor protein P53*
TPM	Transcripts per kilobase million
TSH	Thyroid-stimulating hormone
TSS	Transcription start site
WHO	World Health Organization

ABSTRACT

Differentiated thyroid cancers (DTCs) are endocrine malignancies with a strong but ill-defined hereditary predisposition. Genome-wide association studies highlighted several genomic loci associated with increased risk of DTCs. Relatively strong associations were detected for germline variants located in *disrupted in renal carcinoma 3* (*DIRC3*), a poorly characterized long non-coding RNA gene. This PhD thesis is the first to investigate the functional role of *DIRC3* in thyroid carcinogenesis. Using clinical material and bioinformatic data I have established that *DIRC3* is downregulated in DTCs. *DIRC3* expression level in malignant tissue appeared to influence the risk of DTC recurrence. *DIRC3* was found to be strongly co-expressed with *insulin-like growth factor binding protein 5* (*IGFBP5*), a gene known to regulate cellular response to insulin-like growth factor 1 (IGF-1). A set of comprehensive *in vitro* experiments demonstrated that *DIRC3* transcripts are enriched in the nucleus, where they promote expression of *IGFBP5*. Silencing of *DIRC3* in thyroid cancer cell lines produced a phenotypic dichotomy: it boosted migration and invasiveness, decreased the starvation-induced apoptosis, but also abrogated the MTT reduction rate (the indirect indicator of cell viability and proliferation). Gene rescue experiments indicated that this pro-migratory phenotype was related to the alterations in expression of *IGFBP5*. In contrast, the influence of *DIRC3* on the results of MTT assays appeared be at least partially independent from *IGFBP5*. Transcriptomic profiling of thyroid cancer cells experiencing silencing of *DIRC3* or *IGFBP5* showed a significant redundancy in the activities of both genes. Gene ontology analysis indicated that terms significantly enriched in the response to silencing of *DIRC3* were involved in biological processes related to the cellular migratory potential. I also demonstrated that downregulation of *DIRC3* enhanced the susceptibility of cancer cells to the stimulation with IGF-1, what consequently promoted the oncogenic AKT signaling pathway. Overexpression experiments that utilized *CRISPR activation* (CRISPRa) successfully upregulated *DIRC3*. While this did not elicit changes in expression of *IGFBP5*, the MTT reduction rate was consequently augmented in thyroid cancer cells. Finally, I utilized CRISPR/Cas9 to edit one of the top germline DTC susceptibility variants in *DIRC3*, rs11693806. Modification of a heterozygotic rs11693806[C/G] thyroid cancer cell line into monoallelic [G/-] or homozygotic [G/G] derivatives produced a marked downregulation of *DIRC3* and *IGFBP5*. Furthermore, these genomic modifications phenocopied the pro-migratory effects observed after silencing of

DIRC3. I also confirmed that these genomic modifications resulted in global transcriptomic alterations, often affecting genes involved in different aspects of carcinogenesis.

In conclusion, *DIRC3* emerges as a lncRNA gene functionally implicated in DTCs. Its downregulation stimulates cancer invasiveness, but on the other hand it may produce inhibitory effects in MTT assays. Mechanically, *DIRC3* regulates expression of *IGFBP5*, thus contributing to the altered sensitivity of cancer cells to IGF-1. Accordingly, I propose an interplay between the germline cancer risk variants, *DIRC3* expression and IGF-1 signaling as a mechanism that jointly orchestrates thyroid carcinogenesis.

STRESZCZENIE

Badanie funkcjonalne genu *DIRC3* w zróżnicowanym raku tarczycy

Zróżnicowane raki tarczycy (ang. differentiated thyroid cancers, DTC) są nowotworami układu endokrynnego wykazującymi silne, lecz słabo poznane uwarunkowania genetyczne. Całogenomowe badania asocjacyjne wykazały obecność kilku miejsc w genomie związanych ze zwiększoną zachorowalnością na DTC. Względnie silne asocjacje wykazano dla wariantów germinalnych zlokalizowanych w *Disrupted in Renal Carcinoma 3* (*DIRC3*), słabo scharakteryzowanym genie, którego produktem jest długie niekodujące RNA (ang. long non-coding RNA). W niniejszej pracy doktorskiej po raz pierwszy na świecie zweryfikowałem rolę *DIRC3* w patogenezie raka tarczycy. Wykorzystując materiał pozyskany od pacjentów i dane bioinformatyczne wykazałem, że *DIRC3* ulega istotnemu wyciszeniu w DTC. Uzyskane dane wskazują także, że zmiany w ekspresji *DIRC3* w DTC mogą wpływać na ryzyko wznowy choroby. *DIRC3* wykazał silną ko-ekspresję z *insulin-like growth factor binding protein 5* (*IGFBP5*), pobliskim genem regulującym odpowiedź komórki na insulinopodobny czynnik wzrostu 1 (IGF-1). Używając zróżnicowanych technik badawczych *in vitro* pokazałem, że transkrypty *DIRC3* lokalizują się głównie w jądrach komórkowych, gdzie mogą regulować ekspresję *IGFBP5*. Wyciszenie *DIRC3* w liniach komórkowych raka tarczycy spowodowało dwojaki fenotyp: z jednej strony zwiększało zdolność komórek do migracji i inwazyjności, hamowało apoptozę wywołaną głodzeniem, ale z drugiej strony zmniejszało zdolność do redukcji MTT w badanych punktach czasowych (test pośrednio oceniającym wzrost liczby żywotnych komórek). Transfekcja komórek plazmidem powodującym zwiększoną ekspresję *IGFBP5* wywołała ograniczenie pro-migracyjnego działania obserwowanego w warunkach wyciszania *DIRC3*, lecz nie wpłynęła na wyniki testów MTT. Profilowanie transkryptomu komórek raka tarczycy podlegających wyciszaniu *DIRC3* lub *IGFBP5* wskazało istotną współzależność tych dwóch genów. Analiza ontologii genów (GO) ujawniła, że pod wpływem wyciszenia *DIRC3* istotnej zmianie ekspresji podlegają geny zaangażowane w regulację potencjału migracyjnego komórek. Wykazałem również, że zmniejszenie ekspresji *DIRC3* uwrażliwia komórki nowotworowe na działanie IGF-1, co w konsekwencji aktywuje onkogenny szlak przekazywania sygnałów związany z białkiem AKT. Wynik ten jest zgodny z wcześniej poznaną funkcją białka IGFBP5. Doświadczenia wykorzystujące mechanizmy *CRISPR activation* (CRISPRa) pozwoliły na zwiększenie ekspresji *DIRC3* w liniach

komórkowych. Choć nie wpłynęło to na ekspresję *IGFBP5*, CRISPRa wywołał wzrost liczby żywotnych komórek w teście MTT. W ostaniem etapie projektu wykorzystałem CRISPR/Cas9 do edycji rs11693806, jednego z najważniejszych wariantów germinalnych determinujących ryzyko zachorowania na DTC. Wytworzyłem z heterozygotycznej linii komórkowej raka tarczycy (rs11693806[C/G]) izogeniczne klony komórkowe wykazujące homozygotyczność (rs11693806[G/G]) lub delecję allelu rs11693806[C]. Edycja ta w obu przypadkach spowodowała zmniejszenie ekspresji *DIRC3* oraz *IGFBP5*. Co więcej, powyższe modyfikacje genetyczne indukowały fenotyp komórkowy podobny do tego, który był wywoływany przez wyciszanie *DIRC3* (w tym efekt pro-migracyjny). Potwierdziłem również, że modyfikacja wariantu rs11693806 powoduje w komórkach globalne zmiany transktyptomu, w tym zmiany ekspresji wielu genów uczestniczących w procesach kancerogenezy.

Podsumowując, wyniki uzyskane podczas przygotowywania niniejszej rozprawy doktorskiej wskazują na to, że *DIRC3* jest funkcjonalnie zaangażowany w biologię raka tarczycy. Zmniejszenie ekspresji *DIRC3* wzmaga inwazyjność komórek, ale z drugiej strony może ograniczyć wzrost ich aktywności w teście MTT. Poznane mechanizmy wskazują, że *DIRC3* reguluje transkrypcję *IGFBP5*, przyczyniając się w ten sposób do zmiany wrażliwości komórek nowotworowych na działanie IGF-1. W oparciu o powyższe wyniki postawiłem hipotezę, że interakcja pomiędzy wariantami germinalnymi, ekspresją *DIRC3* oraz szlakiem przekazywania sygnałów z receptora dla IGF-1 stanowi mechanizm regulujący kancerogenezę w gruczole tarczowym.

CHAPTER I: INTRODUCTION

1.1. Thyroid cancer - epidemiology and classification

Thyroid cancer is the most common malignancy of endocrine glands and the 9th most frequent cancer type [1]. Annual incidence is estimated at 586,000 cases worldwide with 4,000 new cases being diagnosed in Poland [1]. Epidemiological data indicate a substantial increase in thyroid cancer incidence in the western world over last 4 decades (incidence in the United States of America has risen by average 3.6% annually) [2, 3]. A substantial part of this increase is attributed to the widespread use of imaging diagnostics that detect a large number of small, often indolent, papillary thyroid carcinomas [4]. Mortality from thyroid cancers has remained low, however some reports indicate that the upsurge in incidence comes with a small rise in the disease-associated mortality [2].

About 95% of thyroid cancers originate from follicular epithelial cells, the major thyroid cell type that is responsible for the production of thyroid hormones (thyroxine and triiodothyronine). The current 4th World Health Organization (WHO) Classification of Tumours of Endocrine Organs divides follicular cell-derived thyroid cancers into five entities: papillary, follicular, Hürthle cell, poorly differentiated, and anaplastic (undifferentiated) carcinomas [5]. Papillary, follicular and Hürthle cell carcinomas retain a significant degree of cytological follicular differentiation and associate with favorable patients' prognosis [6, 7]. Accordingly, these three cancer types are brought together under an umbrella category termed "differentiated thyroid cancer" (DTC). Conversely, anaplastic thyroid cancers demonstrate loss of follicular differentiation, limited treatment options and a dismal prognosis. Poorly-differentiated thyroid cancers are histological and clinical intermediates placed between anaplastic thyroid cancers and DTCs [8].

1.2. Etiology of follicular cell-derived thyroid cancers

Etiology of follicular cell-derived thyroid cancers is poorly understood. Among non-modifiable risk factors are patient's sex, ethnicity and family history of thyroid cancers [3]. Women have about threefold higher risk of developing thyroid cancer than men [1, 2]. Incidence of thyroid cancers is higher in individuals of Caucasian or Asian ethnicity than in the people of African

ancestry [9, 10]. Mean age at the time of diagnosis is 48 years (considerably lower than for most other cancer types) [2], however people who are over 60-year-old tend to develop cancers of more aggressive biology [11].

Contribution of hereditary factors to the pathogenesis of thyroid cancers appears to be one of the highest among all cancer types. Pedigree studies demonstrate approximately five- to nine-fold increase in thyroid cancer risk in the first-degree relatives of DTC patients [12-15]. Similarly, a recent pan-cancer genome-wide association study of ca. 500,000 individuals indicated that thyroid cancers have the second highest heritability estimate among 18 common malignancies [16]. Despite this, only 3-9% of DTCs can be classified as familial cancers, with few *bona fide* hereditary cancer syndromes being recognized in DTC patients [17]. These conditions include: familial adenomatous polyposis syndrome (driven by germline mutations in *adenomatous polyposis coli* [*APC*] gene), Cowden syndrome (caused by germline mutations in *phosphatase and tensin homolog* [*PTEN*]), Werner syndrome, Carney complex and McCune–Albright syndrome [18]. Additionally, a small subset of DTCs may be classified as putative familial cancers, with some novel genes being proposed to drive cancer predisposition (e.g., *PTC susceptibility candidate 3* [*PTSC3*]) [18].

The most well-documented modifiable risk factor of DTCs is a past exposure to ionizing radiation (during radiotherapy, diagnostic procedures or nuclear accidents). The risk of developing thyroid cancer is most significant when the radiation exposure occurs in childhood [19]. Radiation-induced papillary thyroid cancers are molecularly distinct from sporadic cancers as they are primarily driven by chromosomal rearrangements and gene fusions involving the *receptor tyrosine kinase* (*RTK*) genes (e.g. *rearranged during transfection* [*RET*], *neurotrophic tyrosine receptor kinase type 1* [*NTRK1*] and *type* 3 [*NTRK3*]) [17]. Another environmental factor implicated into the pathogenesis of DTCs is iodine intake. Iodine deficiency has been associated with an augmented risk of follicular thyroid cancers. Conversely, some evidence suggest that excessive intake of iodine may increase the risk of papillary thyroid carcinomas [3].

Finally, the risk of DTCs is influenced by anthropometric parameters. Multiple studies have demonstrated that obesity is linked to a higher incidence of thyroid cancers [20-22]. Likewise, a positive association between the DTC incidence and tallness has been observed in ethnically diverse populations. This correlation is stronger for DTCs than for any other major cancer type [20, 23-25]. Additionally, a Danish study showed that the life-long thyroid cancer risk associates more strongly with individual's height and weight documented in childhood than

with the parameters observed in adulthood [26]. The pathophysiological basis of these observations is currently unknown. Some researchers propose that these associations are generated by endocrine and/or paracrine mechanisms involving insulin resistance, insulin-like growth factors (IGFs) or adipokines [3].

1.3. Types of follicular cell-derived thyroid cancers

Papillary thyroid cancer

Papillary thyroid cancer (PTC) represents 80-85% of all follicular-cell derived thyroid cancers [4]. The current WHO classification describes PTC as "*a malignant epithelial tumor demonstrating evidence of follicular cell differentiation and a set of distinct nuclear features. (...) Papillae, invasion or cytological features of papillary thyroid carcinoma are required*" [5]. Cytological features include: ground-glass nuclei, nuclear grooves and irregular nuclear contours [27-29]. PTCs typically have a very favorable prognosis with 5-year overall survival reaching 95-97% [30].

PTCs are characterized by a low mutational burden [31]. Mutually exclusive tumor-driving mutations typically implicate:

(i) *v-Raf murine sarcoma viral oncogene type B* (*BRAF*; 40-60% of cases, typical for the conventional and tall cell variant PTCs),

(ii) *Rat sarcoma virus* (*RAS*) genes (*NRAS, HRAS* or rarely *KRAS*; 15-20% cases, almost exclusively observed in the follicular variant PTCs),

(iii) Gene rearrangements involving *RET*, *NTRK1*, *NTRK3*, *anaplastic lymphoma kinase* (*ALK*) or *BRAF* (10-20% in total; typically detected in the radiation-induced PTCs) [4, 31, 32].

Presence of the most common driver mutation, *BRAF V600E*, is an independent negative prognostic factor. Its presence has been associated with an elevated risk of cancer recurrence and increased patients' mortality [30, 33]. More aggressive PTCs also tend to harbor the *telomerase reverse transcriptase* (*TERT*) promoter mutations (incidence of ca. 10%) [4, 8].

The WHO classification recognizes the conventional (classic) PTCs and multiple histological variants of PTCs [5, 27, 34]. These variants are categorized using various

pathological features (tumor size, histological architecture, cytological appearance, presence of tumor capsule, or capsular invasion).

The PTC variants differ in clinical outcomes:

a) Excellent prognosis variants: papillary microcarcinoma (≤1 cm diameter; constitutes 33% of PTCs), follicular variant (5-27% cases; characterized by the absence of papillary structures, but presence of a typical PTC-like nuclear morphology; *RAS* mutations are characteristic), encapsulated variant (1.3-10%), cribriform-morular variant (rare; *APC* mutations are typical).

b) High-risk and less favorable prognosis variants: tall cell variant (3.8-16.1%; harbors *BRAF* mutations), hobnail variant (0.3-2.7%), solid/trabecular variant (1-3%; *RET*/*PTC* gene rearrangements are typical), columnar cell variant (0.2-2.5%), diffuse sclerosing variant (0.3-1.8%; harbors *RET*/*PTC* rearrangements), oncocytic variant (1.9%).

c) Very rare types with limited data regarding their clinical behavior: desmoid-type fibromatosis/nodular fasciitis-like stroma variant, spindle cell variant, clear cell variant, and Warthin-like variant.

It should be also noted that the encapsulated follicular variant of PTC without capsular or vascular invasion (constituting until recently up to 17% of all PTCs) has been reclassified in 2017 to a new non-malignant category termed *non-invasive follicular thyroid neoplasm with papillary-like nuclear features* (NIFTP) [17]. These lesions are characterized by a very indolent nature resembling benign follicular adenomas and present virtually no recurrences after surgeries [29, 32].

Follicular thyroid cancer

Follicular thyroid cancers (FTCs) constitute 10-15% of follicular-cell derived thyroid cancers [4, 35]. FTCs are histologically characterized by the follicular differentiation of cancer cells, presence of vascular or capsular invasion, and a notable absence of nuclear features typical for PTCs. The WHO classification recognizes three types of FTCs: minimally invasive (with limited capsular invasion only), encapsulated angioinvasive, and widely invasive [5]. 40-month disease-specific survival for these three types is estimated at 97%, 81% and 46%,

respectively [27]. Encapsulated angioinvasive FTCs are further subdivided into carcinomas with limited vascular invasion (<4 vessels involved; better prognosis) or extensive vascular invasion (≥4 vessels invaded; worse outcomes) [27, 32]. Prognosis of patients diagnosed with FTCs is usually good (5-year overall survival of 91%) [35]. Nevertheless, FTCs have a higher propensity for hematogenous spread than PTCs, and distant metastases develop predominantly in the lungs or bones [8, 35].

Molecular landscape of FTCs is marked by *neuroblastoma RAS viral oncogene homolog* (*NRAS*)*, Harvey rat sarcoma virus* (*HRAS*) or *Kirsten rat sarcoma virus* (*KRAS*) mutations (30-40% of FTCs), or *paired box 8 (PAX8)– peroxisome proliferator-activated receptor-γ* (*PPARG)* gene fusions (observed in 30-35% of FTCs). Some FTCs also exhibit mutations in the *TERT* promoter (10-20%), *PTEN* (<10%) or *phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α* (*PIK3CA;* <10%) [4].

Hürthle cell thyroid cancer

Hürthle cell thyroid cancers (HTCs) constitute ca. 5% of the follicular-cell derived thyroid cancers [36]. HTCs were long regarded as variants of FTCs [17], however, the revised 4th WHO Classification of Tumours of Endocrine Organs has recognized HTC as a separate entity [37]. Indeed, Hürthle cell thyroid tumors present distinct histological, molecular and clinical features. Morphologically they are marked by the dominance of enlarged oncocytic (oxyphilic) cells that exhibit eosinophilic granular cytoplasm [4]. HTCs are distinguished from benign Hürthle cell adenomas by the presence of capsular or vascular invasion [29]. HTCs typically develop in older patients, are more common in men, tend to grow large, and may implicate lymph node and distant metastases. They are relatively more refractory to the radioactive iodine treatment and have a slightly worse prognosis than FTCs [4, 34]. Disease-specific mortality in HTCs is estimated at 6% [36].

HTCs display mutations in: the RTK-RAS-BRAF and PI3K-AKT signaling pathway genes, mitochondrial genes, genes involved in protein translation (*eukaryotic translation initiation factor 1A X-linked* [*EIF1AX*]*, mucosal vascular addressin cell adhesion molecule 1* [*MADCAM1*]), chromatin modification or DNA repair. *TERT* promoter mutations are detected in 22% of HTCs. Aberrant karyotype is also common with frequent duplications of chromosomes 5 and 7 [37].

1.4. Management and treatment of differentiated thyroid cancers

Suspicion of a thyroid cancer usually necessitates neck ultrasound and a fine-needle aspiration biopsy. Identification of ultrasound features such as a solid and deeply hypoechogenic appearance, irregular margins, irregular shape and microcalcifications (elements included in TIRADS, the Thyroid Image Reporting and Data System), suggests a malignant character of thyroid lesion and indicates the need of biopsy. Interpretation of cytology is, however, challenging and may give inconclusive results (e.g., FTCs are cytologically indistinguishable from adenomas). Differential diagnosis may be thus supplemented with clinically validated molecular tests that analyze the mutational, gene expression and/or microRNA profiles. Other diagnostic tools used to establish preoperative staging may include: chest X-ray, computed tomography, and (rarely) magnetic resonance imaging. Staging of follicular cell-derived thyroid cancers is determined using the Union for International Cancer Control (UICC) tumor, node, metastasis (TNM) system that accounts for the size of primary tumor, range of extrathyroidal extension, lymph node involvement, presence of distant metastasis, and (uniquely) patient's age at the time of diagnosis (patients ≥55 year-old have worse prognosis) [38].

The mainstay treatment modalities utilized in DTCs are surgery (thyroidectomy) and the radioactive iodine (RAI) therapy. Historically, vast majority of patients underwent a total thyroidectomy, resection of central cervical lymph nodes and the RAI therapy. A large influx of patients diagnosed with early carcinomas has recently expedited changes in therapeutic algorithms aiming to reduce overtreatment. These modifications include: less extensive surgeries (e.g. thyroid lobectomies, less extensive cervical nodal dissections), adjustments in the adjuvant RAI therapy, and avoidance of the thyroid-stimulating hormone (TSH)-suppressive thyroid hormone doses [8, 32]. According to the European Society of Medical Oncology (ESMO) guidelines, thyroid lobectomy is favored in patients without family history of thyroid cancers who present with a small, single carcinoma of favorable cytology without extrathyroidal expansion and unthreatened adjacent structures (the recurrent nerve, esophagus and trachea). Radiation exposure in childhood also disqualifies from this conservative surgical approach. Total thyroidectomy with a prophylactic cervical node dissection is preferred in most other cases. The surgery may be supplemented with the strap muscle resection in T3-T4 stages (i.e., tumors larger than 4 cm or presenting a gross extrathyroidal expansion) and/or a bilateral cervical node dissection (if regional lymph nodes

are involved) [38]. The optimal therapeutic management strategy of thyroid nodules with indeterminate cytology may be also established preoperatively using novel next-generation sequencing Genomic Classifier tests (e.g., ThyroSeq).

The American Thyroid Association (ATA) has developed a recurrence risk stratification system (also adopted by ESMO) to assist in the selection of a proper adjuvant treatment strategy [38, 39]. Using this tool, patients are classified to have a low (<5%), intermediate (5–20%) or high (>20%) risk of recurrence. The classification relies on evaluation of multiple parameters: tumor size, presence of extrathyroidal extension, number and size of involved lymph nodes, presence of distant metastasis, postoperative serum thyroglobulin, cancer histology, presence of vascular invasion, the *BRAF* and *TERT* mutational status, and the radicality of surgical resection [32, 39]. According to the system, RAI is optional in the low-risk patients (low activity of 30 mCi, if used). In the intermediate and high-risk groups, the adjuvant RAI should be administered using doses of 30-100 mCi and ≥100 mCi, respectively. Therapeutic RAI activities (doses of 100-200 mCi) are used in the metastatic and unresectable cancers [38]. This rather complex stratification system has successfully reduced the number of overtreated patients and has not compromised their oncological outcomes [40].

The recurrence risk classification is modified dynamically to assess response to treatment. Responses are classified as: excellent, biochemical incomplete, structural incomplete or indeterminate. This classification is determined using imaging techniques (primarily neck ultrasound), concentration of serum thyroglobulin and anti-thyroglobulin antibodies. The individuals' category determines the follow-up strategy, target TSH level, indicates a need of additional tests (e.g., iodine-131 scan, 2-fluoro-2-deoxy-D-glucose-positron emission tomography), or a necessity of further treatment [38].

In rare cases of progressive iodine-refractory metastatic DTCs that are not eligible for localized forms of treatment, patients are offered oral multiple tyrosine kinase inhibitors (mTKIs): sorafenib, lenvatinib or cabozantinib [38, 41]. The mechanism of their therapeutic activity is primarily attributed to the inhibition of vascular endothelial growth factor receptors (VEGFR) [17]. Novel agents such as RET inhibitors (selpercatinib, pralsetinib) and NTRK inhibitors (entrectinib, larotrectinib) emerge as valuable treatment options in DTCs harboring aberrations in the respective genes [42-45]. Additionally, BRAF inhibitors (vemurafenib, dabrafenib) may be useful in the *BRAF V600*-mutated thyroid cancers [46, 47], Still, responses to the BRAF inhibitors are modest, since DTCs may reciprocal activate the human epidermal growth factor receptor (HER) signaling in response to the BRAF blockade [17].

1.5. Major protein signaling pathways in differentiated thyroid cancers

Two main molecular pathways implicated in DTCs are the mitogen-activated protein kinase (MAPK) and protein kinase B (AKT) signaling cascades [35]. These pathways contribute to thyroid carcinogenesis in various degrees since their signaling output is produced by mutually exclusive driver mutations (Figure 1).

In the classical and tall cell variant PTCs, the constitutively active BRAF V600E oncoprotein maintains a tumor promoting signal *via* sequential phosphorylation of MAP kinase kinase (MAPKK, also known as MEK) and extracellular signal-regulated kinase (ERK, also known as MAPK) [33]. ERK has about 200 downstream effectors in the cytoplasm and nucleus, including certain transcription factors (c-FOS, c-JUN, c-MYC). Ultimately multiple proteins are upregulated, e.g., hypoxia-inducible factor 1α (HIF-1α), vascular endothelial growth factor A (VEGFA), transforming growth factor β1 (TGFB1) and hepatocyte growth factor receptor (MET) [37, 50]. These proteins are implicated in a myriad of pro-tumorigenic processes: cell proliferation, migration, survival, invasiveness and angiogenesis [50]. Additionally, BRAF signaling promotes aberrant gene methylation. This epigenetic alteration silences certain tumor suppressors and iodine metabolism genes (e.g., *sodium/iodide symporter* [*NIS*], *thyroglobulin* [*TG*], *thyroperoxidase* [*TPO*]). Consequently, the follicular differentiation and iodine avidity of cancer cells may be reduced [50-52]. Signaling flux induced by the BRAF V600E oncoprotein is very strong, since its activity is not inhibited by the negative feedback from ERK [33].

In contrast, FTCs and the follicular variant PTCs have a high incidence of activating mutations in *NRAS*, *HRAS* or *KRAS* [33]. RAS oncoproteins are constitutively active. Consequently, MAPK signaling is induced, however, not as strongly as in the *BRAF*-mutant tumors since the negative feedback invoked by ERK is not abrogated. Importantly, among other targets of RAS are phosphatidylinositol 3-kinases (PI3Ks). PI3Ks are signal transducers converting the membrane-bound phosphatidylinositol 4,5-bisphosphate (PIP2) into phosphatidylinositol (3,4,5)-trisphosphate (PIP3). PIP3 acts as a second messenger molecule that activates 3-phosphatidylinositol-dependent protein kinase 1 (PDK1) and protein kinase B (AKT). AKT is a serine/threonine protein kinase with over 100 substrates implicated in numerous signaling cascades, e.g., glycogen synthase kinase 3 (GSK3), mechanistic target of rapamycin (mTOR) and forkhead box O (FoxO) [53]. AKT thus promotes multiple cellular events: proliferation, motility, invasiveness, survival, etc. [37].

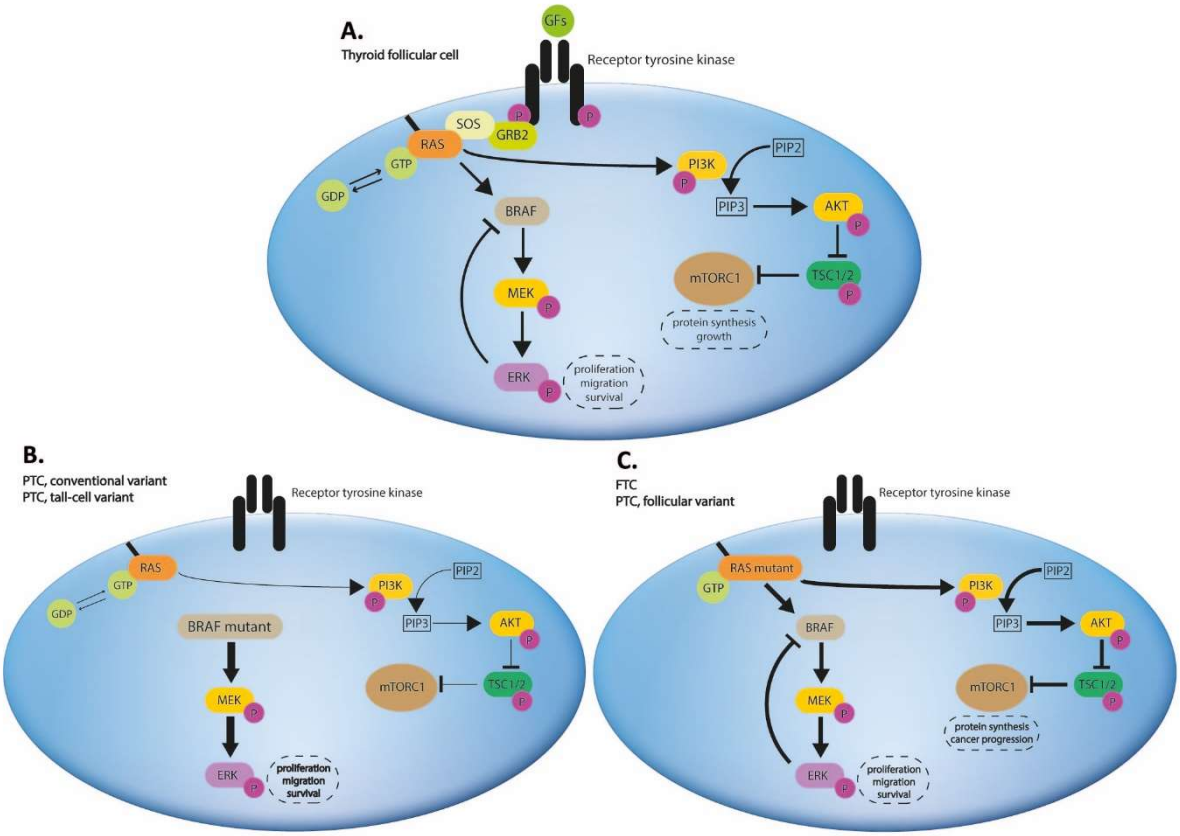


Figure 1. Overview of major signaling pathways involved in thyroid carcinogenesis.

(A). Follicular cells are physiologically stimulated by various growth factors (GFs) that bind to Receptor Tyrosine Kinases (RTKs). Once RTK is stimulated, it becomes internally phosphorylated and may recruit adaptor proteins (e.g., growth factor receptor-bound protein 2, GRB2). GRB2 then binds guanine exchange factors (e.g., Son of Sevenless, SOS). Next, SOS activates RAS proteins (HRAS, NRAS or KRAS) by exchanging guanosine diphosphate (GDP) for guanosine triphosphate (GTP). RAS proteins may then stimulate BRAF (triggering its dimerization with another RAF protein; not shown). Subsequently, BRAF phosphorylates mitogen-activated protein kinase kinase (MEK), which in turn activates extracellular signal-regulated kinase (ERK) to promote the oncogenic phenotype. Additionally, ERK reciprocally inhibits BRAF (a negative feedback mechanism). Importantly, RAS also stimulates phosphoinositide 3-kinases (PI3Ks). PI3Ks catalyze production of phosphatidylinositol-3, 4, 5-triphosphate (PIP3) which acts as a secondary messenger molecule that activates AKT. Consequently, AKT triggers formation of mechanistic target of rapamycin complex 1 (mTORC1) *via* inhibition of the tuberous sclerosis complex proteins (TSC1 and TSC2). mTORC1 upregulates cell proliferation and motility, what is partially related to its ability to enhance protein translation. **(B).** In the *BRAF*-mutant cancers (mostly the conventional and tall cell PTC variants), BRAF is active constitutively. Mutant monomeric BRAF V600E proteins are not inhibited by ERK. Accordingly, a strong pro-mitotic, pro-survival and pro-migratory signal is generated. (C). In the *RAS*-mutant cancers (FTCs and the follicular variant PTCs), RAS oncoproteins simultaneously activate BRAF/MEK/ERK and PI3K/AKT/mTOR pathways. However, the signaling flux produced by wild-type BRAF may be partially inhibited by the negative feedback response from ERK.

Arrow thickness in schemes indicates the magnitude of signaling flow.

PTCs that harbor rearrangements in *RET*, *NRTK1* or *NTRK3* produce ligand-independent receptors that also maintain the dual BRAF-MAPK and PI3K-AKT signal [17]. Moreover, AKT signaling in DTCs may be promoted by disruptive mutations in *PTEN* (the major inhibitor of AKT), or by copy number gains involving the *RTK* or *PI3K* genes [31, 48, 49].

Alterations in the activity of signaling pathways are not exclusively dependent on a few aforementioned driver mutations. Particularly, their output is influenced by epigenetic alterations. Several prominent examples of these events have been described in DTCs. Firstly, aberrant methylation of *PTEN* silences this tumor suppressor in a subset of FTCs and in the follicular variant PTCs [50, 51]. Secondly, thyroid cancers may exhibit altered levels of microRNAs that target members of the signaling cascades. For instance, the *PTEN*-silencing miR-166 and miR-21 are overexpressed in PTCs (high levels of miR-21 are very characteristic for the tall cell variant PTCs) [31, 52], while the *BRAF*-silencing miR-9-5p is frequently downregulated in PTCs [53]. Finally, mounting evidence indicate a similar modulatory role of *long non-coding RNAs* (lncRNAs). For instance, downregulation of *BRAF-activated non-protein coding RNA* (*BANCR*) may stimulate MAPK signaling in PTCs [54]. Some PTCs also overexpress *plasmacytoma variant translocation 1* (*PVT1*), which upregulates *insulin-like growth factor 1 receptor* (*IGF-1R*), a potent activator of PI3Ks and RAS proteins [55].

1.6. Genome-wide association studies and new insights into thyroid carcinogenesis

Family history of thyroid cancers is a recognized risk factor of DTCs. Still, the genetic background of this hereditary predisposition has remained largely elusive. There has been a hope that novel insights into this matter could be provided by *genome-wide association studies* (GWAS). Unfortunately, identification of the biological and clinical relevance of GWAS results remains challenging.

GWAS is a hypothesis-free method for identifying associations between phenotypic traits or diseases, and genomic variants (usually single nucleotide variants, SNVs) [56]. A typical GWAS screens a large number of common SNVs (also known as single nucleotide polymorphisms, SNPs) in a very large population (usually over 10,000 individuals) [56]. The genotyped SNVs act as markers (tags) of particular genomic loci and account for several

neighboring SNVs inherited together across generations (constituting units of linked SNVs organized into haplotype blocks). Since the number of statistical tests in GWAS is huge (at least one per SNV), a very large sample size is required to achieve adequate statistical power. Larger sample size permits dissection of even more association signals (variants with a smaller effect size or with a lower incidence). Still, the methodology of GWAS has inherent limitations. Firstly, GWAS typically identify many weak risk variants of limited prognostic value [57]. Secondly, GWAS mostly evaluate common SNVs and the impact of rare variants may be missed [56]. Thirdly, association signals often locate in noncoding regions with obscure biological relevance and no clear connection to the disease [56]. Accordingly, results of GWAS are almost never straightforward to interpret. Post-GWAS studies should attempt to answer several questions: 1) which SNVs are truly causative? 2) what genes are affected by the causative SNVs? 3) how the causative variants contribute to altered genomic functions? 4) how the implicated genes contribute to the phenotype or disease?

To date at least seven large GWAS have been performed to evaluate the hereditability of DTCs [58-65]. These efforts have identified at least 10 chromosomal loci modulating the thyroid cancer risk in Caucasians. Four of these loci contain germline variants demonstrating particularly robust associations and good cross-study replicability in European, American and Asian populations (Figure 2). These SNVs localize to chromosome 14q13.3 (rs944289 and rs116909374), 9q22.33 (rs965513 and rs1867277), 8q12 (rs2439302) and 2q35 (rs966423 and rs11693806) [58]. Many of these signals have become subjects of post-GWAS studies.

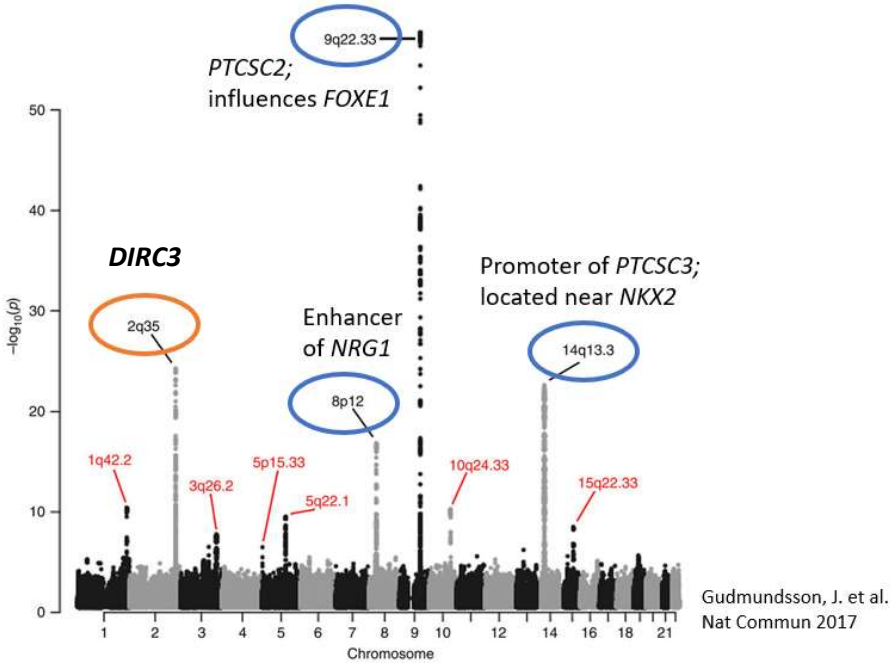


Figure 2. Germline loci contributing to the DTC risk in Caucasian populations.

Variants located in chromosome 2q35 (in *DIRC3*) have the second strongest association with the risk of DTC. In contrast to other major risk loci which have been at least partially characterized (see below), the role of *DIRC3* in thyroid cancers has remained obscure. GWAS Manhattan plot adapted from: Gudmundsson, J. *et al. Nat Commun*. 2017 Feb 14;8:14517 (CC BY license).

DTC risk variants present in chromosome 9q22 (rs965513 and rs1867277) are located in *papillary thyroid carcinoma susceptibility candidate 2* (*PTCSC2*), a long non-coding RNA (lncRNA) gene positioned upstream to *forkhead box E1* (*FOXE1*). Genotypes of rs965513 and rs1867277 were shown to influence expression of *PTCSC2* and *FOXE1*. The latter gene encodes a critical thyroid-specific transcription factor involved in thyroid morphogenesis. It has been demonstrated that *PTCSC2* and *FOXE1* share a bidirectional promoter whose activity is suppressed by myosin-9 (MYH9). This inhibition may be prevented by the interaction of *PTCSC2* transcripts with myosin-9. Reduced expression of *PTCSC2* in PTCs and the resulting downregulation of *FOXE1* upregulates *thrombospondin 1* (*THBS1*) and *insulin-like growth factor binding protein 3* (*IGFBP3*). Consequently, IGFBP3 and THBS1 modulate apoptosis, angiogenesis, metastasis and IGF-1 signaling in PTCs [59, 66, 67].

Cancer risk variants present in chromosome 14q13 (primarily rs944289) are located close to another thyroid-specific lncRNA, *papillary thyroid carcinoma susceptibility candidate 3* (*PTCSC3*). *PTCSC3* is frequently downregulated in PTCs, and its expression is reduced by rs944289[T] allele. This SNV is located in a binding site for CCAAT/enhancer

binding proteins (C/EBP) α and β. Rs944289 may disturb the binding capacity of C/EBPs and thus influence expression of *PTCSC3*. *PTCSC3* was shown to inhibit growth, motility and invasiveness of thyroid cancer cells. This effect was attributed to the suppression of *S100 calcium binding protein A4* (*S100A4*) and downregulation of certain downstream genes (e.g., *vascular endothelial growth factor* [*VEGF*] and *metalloproteinase* [*MMP*]-*9)* [68, 69]. Additionally, *PTCSC3* was shown to inhibit the wingless-related integration site (Wnt) / β-catenin signaling pathway, thus reducing the motility and invasiveness of cancer cells [69-71].

Another risk variant in chromosome 14q13, rs116909374, is located close to *MAP3K12 binding inhibitory protein* (*MBIP1*), a gene proposed to influence production of thyroid hormones. In addition, rs116909374 and the aforementioned rs944289 are situated close to *thyroid transcription factor 1* (*TTF1*). TTF1 is a transcription factor critical for the proper differentiation of thyroid follicles. Nevertheless, mechanistic links between the germline risk variants and *TTF1* have not been established so far [72].

Another risk variant, rs2439302, is positioned in chromosome 8q12 in *neuregulin 1* (*NRG1*). Genomic locus harboring this SNV contains putative enhancers that may influence expression of *NRG1*. In particular, the cancer risk allele rs2439302[G] was found to promote expression of *NRG1* in thyroid cancer cells [73]. *NRG1* is known to stimulate the oncogenic human epidermal growth factor receptor-3 and -4 (HER-3 and HER-4) signaling pathways across many cancer types [74-77].

A number of germline risk variants in chromosome 2q35 (rs966423, rs6759952, rs12990503, rs16857609, rs11693806) cluster in *disrupted in renal carcinoma 3* (*DIRC3*), a poorly characterized lncRNA gene. The biological relevance of these signals has remained obscure.

1.7. *DIRC3* – a gene possibly implicated in thyroid carcinogenesis

DIRC3 was first identified in 2003 as a gene participating in a fusion event in familial renal cell cancers. It was reported that the t(2;3)(q35;q21) translocation observed in these familial tumors generated transcripts that involved *DIRC3* and *heat shock 27-kDa-associated protein 1* (*HSPBAP1*). Still, no experimental inquiry into the function of *DIRC3* was performed [78].

More information on the role of *DIRC3* in cancers has been provided by GWAS. A germline variant in *DIRC3*, rs966423, demonstrated one of the strongest associations with the DTC risk (odds radio, OR = 1.34) in Icelandic, American, Dutch, Spanish, Polish, Kazakh and Chinese populations [60, 79-82]. Additionally, another study reported that rs6759952, a SNV in the linkage disequilibrium (LD) with rs966423, also contributed to the cancer risk [61]. Since rs966423[C] was found to associate with lower TSH levels in the serum, it was proposed that this variant could exert its carcinogenic effect indirectly *via* endocrine mechanisms [60]. Nevertheless, a nominally small hormonal effect induced by rs966423 undermines this hypothesis. Furthermore, a Polish group has associated the rs966423[TT] genotype with increased overall mortality in DTC patients [83]. It has been proposed that while the more common rs966423[C] allele participates in the initiation of thyroid tumorigenesis, the alternative rs966423[T] allele may support cancer progression [83]. Notably, some recent high-resolution GWAS highlighted novel thyroid cancer susceptibility variants in *DIRC3*: rs16857609[T], rs12990503[G] and rs11693806[C] [62, 63, 84]. In particular, rs11693806 presented a significantly stronger association with the thyroid cancer risk than rs966423 (allelic OR = 1.43 and OR = 1.34, respectively) [62]. Conditional analysis of both variants preserved the association signal only for rs11693806. Rs11693806 is a sole known risk variant located in an exon of *DIRC3* (all other risk variants are positioned in introns) [62].

DIRC3 germline variants also influence human height [85-88], the risk of mitral valve prolapse [89], and the risk of breast cancer [90-93]. In particular, chromatin conformation capture studies conducted in breast cancers revealed that *DIRC3* and a nearby gene, *insulin-like growth factor binding protein 5* (*IGFBP5*), are targets of breast cancer predisposing SNVs located in an intragenic region downstream of *DIRC3*. These SNVs were shown to alter expression of *DIRC3* and *IGFBP5* [94]. Their modulatory influence on *IGFBP5*, a putative tumor suppressor in breast cancer [95], was shown to regulate the mammary oncogenesis [96, 97]. *IGFBP5* is also presumed to promote proliferation of thyroid cancer cells [98, 99]. Moreover, certain germline variants located upstream of *IGFBP5* are associated with altered levels of TSH in the blood [100, 101]. IGFBP5 protein is a vital modulator of the IGF-1 signaling cascade, a pathway strongly implicated in many cancer types [102-109].

1.8. Long non-coding RNAs

DIRC3 belongs to a class of genes producing long non-coding RNAs (lncRNAs). LncRNAs are defined as transcripts over 200 nucleotides in length without protein-coding potential. While in the past lncRNAs had been regarded as a transcriptional noise (products of spurious transcription occurring throughout genome), recent years have brought ample of evidence illustrating the importance of lncRNAs in physiology and diseases [110, 111]. In fact, it is currently appreciated that lncRNAs may possess the most diverse and complex functionality of all classes of RNAs [111, 112]. Number of lncRNAs experimentally implicated into carcinogenesis is growing fast. Prominent examples include: m*etastasis associated lung adenocarcinoma transcript 1* (*MALAT1),* which is vital for the distant spread of many cancer types, *survival associated mitochondrial melanoma specific oncogenic non-coding RNA* (*SAMMSON*) essential for the melanoma survival, and *prostate cancer associated 3* (*PCA3*), a clinically relevant biomarker of prostate cancers [113-117]. In humans lncRNAs outnumber protein-coding genes (human genome is estimated to contain ca. 60,000 lncRNA genes), and their expression pattern is more tissue- and cell-specific than that of messenger RNAs (mRNAs) [118, 119]. Mechanistically, lncRNAs interact with proteins, DNA and other RNA species to produce a wide range of cellular events. LncRNAs play roles in the modification of chromatin, assembly of nuclear bodies, regulate gene transcription, influence the mRNA turnover (splicing, degradation, translation), or directly modulate protein signaling pathways [112].

Activities of lncRNAs may be categorized in several ways. Firstly, lncRNAs may exhibit activity in *cis* (i.e., impact elements located close to the proper lncRNA gene), or function in *trans* (i.e., regulate distant genes, RNAs or proteins) [111, 112].

Secondly, lncRNAs may execute different functional archetypes and operate as:

- signals (lncRNAs act as triggers for regulatory processes taking place in specific spatial and temporal contexts),
- scaffolds (lncRNAs act as frames for the assembly of molecular complexes or sub-organelles),
- guides (lncRNAs deliver proteins to specific locations, e.g., certain genomic loci),
- decoys (lncRNAs act as sponges that sequester proteins or microRNAs and thus prevent their binding to proper molecular targets) [120].

Thirdly, in some instances the produced non-coding transcripts may be functionally irrelevant. In these cases, it solely the process of transcription of lncRNAs that induces molecular mechanisms that modulate expression of nearby genes [121]. Moreover, some lncRNAs, contrary to their name, may in reality encode small functional peptides [122].

Since lncRNAs present such a great functional heterogeneity, it is immensely challenging to fully distill their cellular roles. First insights into putative functionalities of lncRNAs have been often provided by the identification of disease-associated germline variants in their genomic loci. It is estimated that at least 7% of lncRNA genes overlap disease-associated SNVs [118]. Several lncRNAs have been discovered to impact thyroid oncogenesis owing to studies that functionally explored germline variants associated with the DTC susceptibility (e.g., *PTCSC2*, *PTCSC3*, *MALAT1* and *HOX transcript antisense RNA* [*HOTAIR*]) [123, 124].

1.9. Potential benefits of the research

LncRNAs are gradually gaining prominence in clinical oncology. Firstly, lncRNAs can be applied as diagnostic marker, e.g., *PCA3* is a highly specific prostate cancer lncRNA that can be detected in urine using the Food and Drug Administration (FDA)-approved tests. Secondly, expression of certain lncRNAs may impact patients' prognosis, e.g., *HOXA transcript at the distal tip* (*HOTTIP*) may predict the risk of disease progression in hepatocellular carcinoma. Thirdly, lncRNAs may act as predictive markers. For example, high levels of *HOTAIR* in ovarian cancers associate with a better treatment response to cisplatin than to carboplatin [125].

Although DTCs can be effectively cured in most patients, some individuals experience recurrences and die from the disease. There is a need to better identify patients who have elevated risk of unfavorable clinical course. The *BRAF* and *TERT* mutational status are established molecular prognostic markers in DTCs [30, 33, 83]. Germline genotype of *DIRC3* has been recently proposed as a novel indicator of the mortality risk in thyroid cancer patients [83]. Nevertheless, the clinical impact of this marker is modest only (mortality hazard ratio, HR ~ 1.5 – 1.6) [62, 83].

Hence, I have hypothesized that elucidation of the function of *DIRC3* in thyroid cancers could expedite development of better risk stratification tools. This could potentially help to identify patients who require more intensive surveillance or different treatment strategies. Furthermore, I expected that this study could highlight unrecognized molecular mechanisms and/or signaling pathways involved in DTCs, some of which could be of therapeutic relevance.

CHAPTER II: STUDY RATIONALE, OBJECTIVES AND AIMS

This PhD thesis has been the first to evaluate the role of *DIRC3* in thyroid carcinogenesis. Given the prognostic information offered by *DIRC3* germline variants, I predicted that dissection of the gene function could provide biological insights of a potential diagnostic or therapeutic value.

It has been shown that lncRNAs contribute to all classical hallmarks of cancer: cell proliferation, viability, growth suppression avoidance, immortality, motility, angiogenesis, altered metabolism and immune evasion [126]. Prominent examples of cancer-modulatory lncRNAs include: *MALAT1* implicated in cancer metastasis and angiogenesis, *HOTAIR* that modulates cell proliferation and invasiveness, and *growth arrest-specific 5* (*GAS5*) which regulates the apoptosis of cancer cells [125]. Functions of lncRNAs are very complex. LncRNAs may interact with various proteins, DNA sequences or other RNA species to control diverse molecular processes (chromatic modifications, DNA methylation, RNA turnover, protein complex assembly, recruitment of transcription factors to specific genomic loci, etc.) [125].

The prime goal of this study was to evaluate how *DIRC3* impacts the biology and phenotype of thyroid cancers. Specific aims were related to the following scientific questions:

1. **Is expression of *DIRC3* altered in DTCs, and does it associate with the pathoclinical characteristics of thyroid cancers?**

2. **Are *DIRC3* transcripts involved in the transcriptomic regulation of protein-coding genes?**

3. **How do alterations in the expression of *DIRC3* influence the phenotype of thyroid cancer cells?**

4. **Does *DIRC3* impact the activity of protein signaling pathways implicated in thyroid carcinogenesis?**

5. **Does rs11693806, the major DTC risk variant in *DIRC3*, influence the phenotypic and transcriptomic profile of thyroid cancer cells?**

6. **Does understanding of the function of *DIRC3* in DTCs provide insights of diagnostic and therapeutic relevance?**

The study has been divided into three parts:

I. Analysis of clinical material: to evaluate expression of *DIRC3* in DTCs, and to correlate the gene expression profile with clinicopathological data.

- Bioinformatic data obtained from large RNA-sequencing projects was used to establish if *DIRC3* is expressed in normal thyroid tissue and in DTCs, and to appraise possible transcriptomic mechanisms invoked by *DIRC3*.
- Analysis of clinical material (DTCs and histologically normal thyroid) was performed to recognize possible alterations in the expression of *DIRC3* in thyroid cancers. Additionally, clinical material was used to validate the correlations in gene expression established in the aforementioned bioinformatic analysis. Relationships between the gene expression profile and patients' clinicopathological characteristics were also tested.

II. *In vitro* experiments: to establish the function of *DIRC3* in thyroid cancer cell lines.

Knock-down and overexpression experiments were performed to evaluate the phenotypic and transcriptomic consequences of perturbations in the expression of *DIRC3* in thyroid cancer cell lines.

III. Genomic editing: to establish if rs11693806, one of the top DTC risk variants in *DIRC3*, contributes to the phenotype of thyroid cancer cell lines.

To evaluate whether rs11693806 could act as a causative SNV, I performed the CRISPR/Cas9-mediated genomic editing of the variant. Phenotypic and transcriptomic consequences of the genomic modification were profiled in a thyroid cancer cell line.

CHAPTER III: MATERIALS AND METHODS

3.1 Clinical material

Clinical material used in this study was collected at the Medical University of Warsaw (Warsaw, Poland). The material consisted of tissue samples obtained from patients who underwent a total or near-total thyroidectomy before 2014 due to the diagnosis of a malignant thyroid tumor. Collection of tissue was approved by the Institutional Review Board at the Medical University of Warsaw. Written informed consents were obtained from all patients.

Tumor and histologically normal thyroid tissue specimens (usually collected from the opposite, unaffected thyroid lobes) were snap-frozen in dry ice and stored at -80 ºC. Tissues were used to isolate total RNA. Archival RNA samples analyzed in this study were obtained from 67 patients.

Diagnosis of a differentiated thyroid cancer (papillary, follicular or Hürthle cell) was established by two expert pathologists using the World Health Organization criteria. Clinical staging was determined using the 2010 American Joint Commission on Cancer staging system (7th edition). Histological evaluation confirmed that cancer samples achieved tumor purity of at least 70%. Histologically normal tissues were free of malignant infiltrations. Patients' clinical and pathological data were anonymized and recorded in secure Excel files. Data included information on the histological type of cancer, clinical stage, regional lymph node metastases, distant metastases, tumor multifocality, angioinvasion, extrathyroidal extension, and patients' sex and age at the time of diagnosis.

Additionally, whole blood samples were collected from all patients. DNA was extracted from 9 mL of peripheral blood using the salting out method [127]. This DNA extraction procedure was completed during a past project [83]. Archival germline DNA was available for all 67 patients who were included in the RNA analysis of tumor samples.

3.2. RNA extraction from tissue and quantification

RNA was extracted previously from thyroid tissue samples. Briefly, 50 to 100 mg of tissue was homogenized and mixed with 1 ml of TRIzol reagent (Thermo Fisher, USA). After 5 minutes of incubation, 0.2 ml of chloroform was added and mixed vigorously. After another 3 minutes,

samples were centrifuged at 12,000 x g for 15 minutes at 4 °C. Afterwards, the upper aqueous layer was transferred to a new tube. RNA was precipitated by adding 0.5 ml of isopropanol to the aqueous layer. Samples were then incubated for 10 minutes. Thereafter, samples were centrifuged at 12,000 x g for 10 minutes at 4 °C. Supernatant was discarded, and white pellets were resuspended in 1 ml of 75% ethanol. After another centrifugation at 7,500 x g for 5 minutes at 4 °C, supernatants were discarded and the remaining pellets were airdried. Finally, the RNA pellets were resuspended in 50 µl of RNase-free water. Yield and purity of RNA were measured using the NanoDrop 2000 instrument (Thermo Fisher, USA). Absorbance A260/A280 ratio of 1.8 to 2.1 was determined as satisfactory for later use. RNA samples were stored at -80 °C.

3.3. Quantitative reverse transcription PCR (qRT-PCR)

Two-step quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to quantify expression of selected genes in thyroid tissue samples and in cancer cell lines.

Firstly, reverse transcription (RT) reactions were used to transcribe RNA into more stable complementary DNA (cDNA). Reactions were performed using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT). Reactions were assembled on ice. Equal amounts of RNA (up to 1000 ng per reaction) were mixed with 200 ng of Random Hexamer Primers (EURx, Poland) and RNase-free water. Samples were incubated in a thermocycler at 70 °C for 5 minutes, and then quickly transferred to ice. Next, following reagents were added to the RNA/hexamer mix: 2 µl of 10 mM deoxyribonucleotide triphosphate mix (EURx, Poland), 4 µL of 5X Reaction Buffer (Promega, USA), 5 U of Ribonuclease Inhibitor (EURx, Poland), 100 U of M-MLV RT (Promega, USA), and RNase-free water (EURx, Poland) up to the total volume of 20 µL. Reactions were mixed and transferred to a thermal cycler (ABI GeneAmp 9700). Incubation was performed as follows: 30 minutes at 16 °C, 90 minutes at 42 °C, 5 minutes at 85 °C, and hold at 4 °C. cDNA samples were stored at -20 °C.

Secondly, quantitative polymerase chain reaction (qPCR) was performed using the LightCycler 480 II system (Roche, Switzerland). SYBR Green assays were used. SYBR Green dye generates fluorescent light once it is intercalated into double-stranded PCR products and excited with a laser beam. Fluorescence light emitted by the dye is directly proportional to the number of copies of PCR products. Relative expression of a target gene is calculated by normalizing its expression level using the expression data obtained for a housekeeping gene

(i.e., a gene demonstrating stable expression level across different physiological and pathological conditions).

qPCR reactions were assembled on ice using 1 µl of diluted cDNA, forward and reverse primers (0.4 µM final, each), 5 µl of LightCycler 480 SYBR Green I Master Mix (Roche, Switzerland), and nuclease-free water (Roche, Switzerland) up to the total volume of 10 µl. Reactions were performed in white 384-well plates in the LightCycler 480 II instrument. qPCR reactions were performed as follows: preincubation at 95 °C for 5 minutes, 45 cycles of amplification (95 °C for 10 seconds, 58 °C for 5 seconds, 72 °C for 7 seconds), a melting curve analysis (95 °C for 5 seconds, 65 °C for 1 minute, ramp to 97 °C with 5 data acquisitions per 1 °C), and a final cooling at 40 °C for 30 seconds. All reactions were performed in technical triplicates. Negative control reactions were performed using multiple water blanks and the negative control RT samples (reverse transcriptase was omitted in these control samples).

Results were analyzed using the LightCycler 480 software (Roche, Switzerland). Melting curve analysis was performed to confirm presence of a single specific product peak for each primer pair. Average cycle threshold (Ct) values were calculated for each triplicate. Relative gene expression was calculated in the reference to the expression of *hypoxanthine phosphoribosyltransferase 1* (*HPRT1*) using the conventional Livak $2^{-\Delta\Delta Ct}$ method [128]. *HPRT1* is a housekeeping gene regarded to be particularly relevant in the context of thyroid studies due to its stable expression in thyroid specimens [129-131]. Additionally, I experimentally validated the stability of *HPRT1* expression in DTC/normal thyroid tissue pairs using other common housekeeping genes: *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) and *glucuronidase beta* (*GUSB*). The sequences of PCR primers are provided in Table 1. Good performance of primer pairs was confirmed using serially diluted RNA samples (i.e., the serial curve analyses).

	Forward primer	Reverse primer
***DIRC3*, total (qPCR)**	CTGGTGTCACATCAGAATCTG	ACAATCAGCATGGGACTCAC
***DIRC3-202* (qPCR)**	GGAATTGTGCGAGTGGAAATG	TGCTAAGATAGAGCCGCAAAG
***DIRC3-203* (qPCR)**	TCATGCTGGCCGATGCTGAG	CCACTGGGACAACCATCTTTTAATCAG
***DIRC3-204* (qPCR)**	CTCATCTGTCCGACGAAGCA	CCCTACTGTCCTGGTGGAGA
***IGFBP5* (qPCR)**	AAGATCGAGAGAGACTCCCGT	CCGACAAACTTGGACTGGGT
***HPRT1* (qPCR)**	CAGAGGGCTACAATGTGATG	TGGCGTCGTGATTAGTGATG
***GAPDH* (qPCR)**	ATCATCTCTGCCCCCTCTG	CCTGCCTTCCTCACCTGAT
***GUSB* (qPCR)**	TTAAGGTGCCAGGTGTCAG	ATGAGGAACTGGCTCTTGG
***MALAT1* (qPCR)**	GACGGAGGTTGAGATGAAGC	ATTCGGGGCTCTGTAGTCCT
***PVT1* (qPCR)**	GCCCCTTCTATGGGAATCACTA	GGGGCAGAGATGAAATCGTAAT
***SNHG5* (qPCR)**	GTGGACGAGTAGCCAGTGAA	GCCTCTATCAATGGGCAGACA
rs11693806 locus, 121 bp (PCR and Sanger sequencing)	GCTTTTGATTGAAATATGGCAGC	CAGGCAACCTCCTTACCTG
rs11693806 locus, 600 bp (PCR in T7EI assay)	GTACCATCTCACCACTCC	CAGAAGGACATTGAACTA

Table 1. PCR primers used in the SYBR Green qPCR assays, PCR of the rs11693806 locus, and in Sanger sequencing.

3.4. DNA amplification and agarose gel electrophoresis

Polymerase chain reactions (PCR) were performed to amplify a genomic fragment harboring rs11693806. Reactions were prepared in a total volume of 50 µl using: 25 ng of genomic DNA, 2X Taq PCR Master Mix (1X final; EURx, Poland), forward and reverse primers (0.4 µM final, each), and nuclease-free water (EURx, Poland). Reaction mixes were assembled on ice and transferred to a thermal cycler (ABI GeneAmp 9700, Applied Biosystems, USA; or Mastercycler nexus gradient, Eppendorf, Germany). PCR was performed using the following program: 95 °C for 3 minutes, 40 cycles of amplification (95 °C for 30 seconds, 54 °C for 30 seconds, 72 °C for 1 minutes), 72 °C for 5 minutes, and cooling to 4 °C. PCR products were stored at -20 °C. The sequences of PCR primers are provided in Table 1.

Amplicons were analyzed using the agarose gel electrophoresis. 1% agarose gels were prepared using Agarose Basica LE (Prona, Belgium), Tris-acetate-EDTA (TAE) buffer and SYBR Safe DNA Gel Stain (Thermo Fisher, USA). PCR products were mixed with 6X DNA Gel Loading Dye (1X final; Thermo Fisher, USA) and loaded on 1 % agarose gels (10 µl per lane). GeneRuler 100 bp Plus DNA Ladder (Thermo Fisher, USA) or Perfect 100 bp DNA Ladder (EURx, Poland) were used to identify sizes of PCR products. Electrophoresis was

carried out at 60V for 45-60 min in an electrophoresis box filled with the TAE buffer. PCR products were visualized under UV light. Amplicon bands corresponding to the PCR products were excised using sterile blades and transferred to new tubes.

DNA fragments were purified from agarose using the GeneMATRIX Agarose-Out DNA Purification Kit (EURx, Poland) according to the manufacturer's instructions. Briefly, 30 µl of Buffer A was applied onto spin-columns. 2.5 volumes of Orange A buffer were added to one volume of the excised gel band. Samples were placed in a thermoblock set to 55 °C for 10 minutes. The dissolved agarose solution was applied to the DNA binding spin-columns and centrifuged at 11,000 x g for 1 minute. Flow-through was discarded. Wash A1 buffer was added to spin-columns. Columns were then spined down at 11,000 x g for 1 minute. Flow-through was discarded. Wash AX2 buffer was then added, the columns were spined down at 11,000 x g for 1 minute, and the flow-through was discarded again. Another centrifugation at 11,000 x g for 1 minute was performed to remove any buffer residues. DNA was eluted from the columns using 50 µl of Elution buffer. Final centrifugation at 11,000 x g for 1 minute was performed to complete the elution. Purified PCR products were stored at -20 °C.

3.5. Sanger sequencing

Sanger sequencing was used to confirm the identify of PCR products and to genotype cancer cell lines. Purified PCR products (isolated from agarose gels) were diluted in nuclease-free water to the concentration of 5 ng/µl. Both forward and reverse PCR primers (Table 1) were used as sequencing primers (at the concentration of 5 µM) in order to obtain sequence reads from each side of the amplicon. Sanger sequencing reactions were performed by Genomed Inc. (Warsaw, Poland) using an automated DNA Sequencer. Sequencing chromatograms were evaluated using the FinchTV software (Geospiza, USA).

3.6. rhAmp SNP genotyping

The genotype of rs11693806 was evaluated using the rhAmp SNP genotyping platform (Integrated DNA Technologies, USA). This novel genotyping method relies on the use of RNase H2 enzyme, which is capable to cleave DNA/RNA hybrids only in the case of a perfect base pair match between strands. Components of the rhAmp SNP genotyping system include a

reverse ("locus") primer and two allele-specific forward ("allele") primers. Both allele-specific forward primers contain: a 5′ tail, a sequence complementary to the target with a single RNA nucleotide (matching only one of the probed alleles), and a 3′ blocking modification (preventing extension of the primer by polymerase). Only if the allele-specific primer is perfectly matched to its DNA target, the 3′ blocking modification is cleaved by RNase H2. The 5′ tail present in both allele-specific primers contains a common sequence necessary for anchoring a universal forward primer, and a dissimilar sequence used for anchoring one of two fluorescent probes. The probes contain either FAM or Yakima Yellow fluorescent dye, and a quencher.

The rhAmp genotyping reactions proceed as follows: (i) the reverse primer and the allele-specific forward primers align to DNA targets; (ii) the 3′ blocking modifications are cleaved by RNase H2 only in the case of a perfect match between the unique RNA base and the genotyped sequence; (iii) Taq Polymerase synthesizes a new DNA strand that additionally integrates sequences present in the 5' tail of the allele-specific forward primer (i.e., sequence complementary to the universal forward primer, and the sequence used to anchor the fluorescent probe). (iv) Subsequent PCR cycles utilize the reverse primer and the universal forward primer. Fluorescent probes hybridize to the amplicons and are degraded during the polymerase extension step (i.e., the FAM/Yakima Yellow dyes are decoupled from quenchers). Accordingly, fluorescent signals are generated and detected by a Real-Time PCR thermocycler.

The rhAmp SNP genotyping system was used to genotype rs11693806 in patients' germline DNA. The assay was also used to genotype cancer cell lines. Reactions were performed according to the manufacturer's guidelines. In short, DNA was diluted to 5 ng/µl using IDTE pH 7.5 buffer (10 mM Tris, 0.1 mM EDTA; Integrated DNA Technologies, USA). The rhAmp Genotyping Master Mix (2X) and rhAmp Reporter Mix (40X) were mixed in 20:1 ratio (2.65 µl of the mix was prepared for each reaction). The mix was then combined with a custom 20X rhAmp SNP Assay (1X final) and PCR-grade water (up to the total volume of 3 µl per reaction). Finally, the reaction mix was combined with 2 µl of DNA (10 ng). Reactions were performed using the LightCycler 480 II instrument (Roche, Switzerland) using the following program: 95 °C for 10 minutes, 40 cycles of amplification (95 °C for 10 seconds, 60 °C for 30 seconds, 68 °C for 20 seconds), 99 °C for 10 minutes, and cooling at 40 °C. Post-amplification endpoint reads were performed using the Dual Color Hydrolysis Probes program (FAM and HEX channels) using the LightCycler 480 software (Roche, Switzerland). The rhAmp assay used in this study was custom designed after a personal communication with

the Integrated DNA Technologies (USA) representatives. The sequences of oligonucleotides are provided in Table 2.

Allele primer 1	/rhAmp-F/ GATAGGATTTCCCAGACATTCGrCGAGG /GT2/
Allele primer 2	/rhAmp-Y/ GATAGGATTTCCCAGACATTCCrCGAGG /GT2/
Locus primer	GCGAACAACACCCACTCTGTTTAAGrCCCTA /GT2/

Table 2. Custom oligonucleotides used in the rhAmp rs11693806 genotyping assay. /rhAmp-F/ and /rhAmp-Y/ denote anchors for the universal forward primer, and the FAM and Yakima Yellow fluorophore probes, respectively. /GT2/ is a proprietary modification blocking primer extension. "r" indicates that the next nucleotide is a ribonucleotide. Reference assay name: CD.GT.NBSR9405.1 (Integrated DNA Technologies, USA).

3.7. Bioinformatic databases

Expression of *DIRC3* was evaluated in public databases of comprehensive RNA-sequencing (RNA-seq) projects. Data for normal tissues was obtained from the Genotype-Tissue Expression (GTEx) project [132]. Data was accessed *via* the GTEx Portal (Broad Institute of MIT and Harvard, USA; https://gtexportal.org/home/; accessed May 2018). Expression of *DIRC3* in papillary thyroid carcinomas was evaluated using The Cancer Genome Atlas (TCGA)-generated data [31, 133]. TCGA data was retrieved from the cBioPortal for Cancer Genomics (Memorial Sloan Kettering Cancer Center, USA; https://www.cbioportal.org/; accessed May 2018) [134]. TCGA data was also applied to evaluate associations between the gene expression profile and patients' clinical outcomes. Additionally, TCGA data was used to examine putative correlations in the expression of *DIRC3* and other genes (i.e., evaluate the gene co-expression).

Genomic and transcriptomic sequences necessary to design custom oligonucleotides (PCR primers, antisense oligonucleotides and the single-guide RNA templates) were obtained from UCSC Genome Browser [135] (https://genome.ucsc.edu/) or Ensembl Genome Browser [136] (https://www.ensembl.org/). The human reference genome assembly GRCh38 was used to design oligonucleotides.

3.8. Culture of cancer cell lines

Human cancer cell lines that were used in *in vitro* experiments include: MDA-T32 (cat. CRL-3351; conventional PTC), MDA-T68 (cat. CRL-3353; follicular variant of PTC), MDA-T120 (cat. CRL-3355; conventional PTC) and K1 (cat. 92030501; conventional PTC). MDA-T32, MDA-T68 and MDA-T120 cell lines were purchased from the American Type Culture Collection (ATCC, USA). K1 cell line was previously obtained from Sigma-Aldrich (The European Collection of Authenticated Cell Cultures, UK). Each of these cell lines represent a unique, validated and comprehensively characterized thyroid cancer model [137-139]. Additionally, the panel was supplemented with MCF-7 (cat. HTB-22, ATCC), a breast cancer cell line. MCF-7 was included due to: 1) previously reported associations of *DIRC3* germline variants with the breast cancer risk [90, 91]; 2) a relatively high expression of *DIRC3* in preliminary experiments; 3) frequent use of this cell line in thyroid cancer studies (e.g., MCF-7 exhibits expression of the iodine symporter and a thyroid hormone-dependent growth) [140-142]. Mutational characteristics of cancer cell lines are provided in Table 3.

Name	**Cancer type**	***Mutated genes***
K1	PTC, conventional	*BRAF, PIK3CA, TERT*
MDA-T32	PTC, conventional	*BRAF, TP53, TERT, CDKN2A*
MDA-T68	PTC, follicular variant	*NRAS*
MDA-T120	PTC, conventional	*BRAF, TP53, TERT, NF2*
MCF-7	breast cancer, estrogen receptor positive	*PIK3CA, TP53, CDKN2A, GATA3*

Table 3. Mutational profile of the utilized cancer cell lines.

Cell cultures were maintained using appropriate media and techniques. MDA-T32, MDA-T68 and MDA-T120 were cultured in RPMI-1640 Medium (ATCC modification; Gibco, UK) supplemented with 10% fetal bovine serum (non-heat inactivated FBS; Euroclone, Italy), 2 mM L-glutamine (Lonza, Switzerland), 1X MEM Non-Essential Amino Acids Solution (NEAA; Gibco, USA) and 100 U/ml of Penicillin-Streptomycin (Sigma-Aldrich, USA). K1 cell line was cultured using a medium mix consisting of: 2 volumes of DMEM (Gibco, UK), 1 volume of Ham′s F12 (Gibco, UK) and 1 volume of MCDB-105 (Cell Applications, USA), supplemented with 2 mM L-glutamine, 10% FBS and 100 U/ml of Penicillin-Streptomycin. MCF-7 was cultured in DMEM high-glucose medium (Gibco, UK) supplemented with 2 mM

L-glutamine, 10% FBS and 100 U/ml of Penicillin-Streptomycin. Culture media were replaced twice a week. Subcultures were obtained by rinsing cells with phosphate-buffered saline (PBS; Gibco, UK) and incubating them with 0.25% trypsin-EDTA (Lonza, Switzerland) for 5-10 minutes. Cells were maintained in sterile plastic flasks and plates. Cultures were kept in dedicated incubators under conventional culture conditions (37 ºC, 5% CO_2, 88% humidity). Cell cultures were regularly inspected and periodically checked for *Mycoplasma spp*. contamination using a PCR assay.

Cell counting was performed using Bürker counting chambers. Briefly, cell suspensions were diluted with 0.4% Trypan Blue solution (1:1 ratio) to assess the viability of counted cells. 10 µl of this mixture was pipetted to the Bürker chamber. Cells present in at least 9 large squares were counted under an optical microscope. In the Transwell experiments, cell counting was performed with the TC20 Automated Cell Counter (Bio-Rad, USA) using dedicated single-use slides. At least 2 measurements were made for each count.

3.9. Nucleic acid extraction from cell lines

Total RNA from cancer cell lines was extracted using the GeneMATRIX Universal RNA Purification Kit (EURx, Poland) according to the manufacturer's manual. Briefly, cell suspensions were centrifuged at 1,000 x g for 2 minutes, cell pellets were lysed with RL buffer, vortexed, and transferred to homogenizing minicolumns. Homogenizing minicolumns were centrifuged at full speed for 2 minutes. Afterwards, wash-throughs were mixed with 99.9% ethanol and transferred to the RNA-binding minicolumns. The columns were washed with DN1 buffer (to remove any remaining traces of DNA) and washed twice with RBW buffer (to eliminate contaminants). Centrifugations at 11,000 x g for 1 minute were performed after each buffer wash. Finally, RNA was eluted from the columns using 50 µl of RNase-free water. Yield and purity of RNA were measured using the NanoDrop 2000 instrument (Thermo Fisher, USA). RNA samples were stored at -80 °C.

DNA from cancer cell lines was isolated using the GeneMATRIX Cell Culture DNA Purification Kit (EURx, Poland) following the producer's instructions. Cell suspensions were centrifuged at 1,000 x g for 2 minutes. Cell pellets were lysed using Lyse C buffer, and then treated with RNase A at room temperature for 5 minutes. Afterwards, the samples were mixed with Sol C and Protein K buffers, and incubated for 10 minutes at 70 °C. Samples were then

mixed with 96% ethanol and transferred to silicone minicolumns. The columns were washed with CX1 and CX2 buffers and centrifuged at 11,000 x g for 1 minute after each buffer wash. Finally, DNA was eluted using 50 µl of Elution buffer. Yield and purity of DNA were measured using the NanoDrop 2000 instrument (Thermo Fisher, USA). DNA was stored at -20 °C.

3.10. Subcellular RNA fractionation

RNA fractionation was performed in cancer cell lines to identify the subcellular localization of *DIRC3* transcripts. The experiment was performed due to two reasons. Firstly, its outcomes could provide early insights into a possible transcriptomic activity of lncRNA transcripts, many of which are functional only in the nucleus. Secondly, the identification of subcellular location of transcripts could indicate the most appropriate gene silencing method. In particular, antisense oligonucleotides have been shown to be superior over the RNA interference (RNAi) mechanisms for the knock-down of nuclear lncRNAs [143].

The RNA fractionation experiment was performed in MDA-T32 and MCF-7, two cell lines that demonstrated the highest expression of *DIRC3*. Experiments were executed using a previously described method with some modifications [144]. Briefly, cells were cultured on 6-well plates until they reached a full confluency. Culture medium was discarded and cells were washed with ice-cold PBS. Next, PBS was removed and the culture plates were placed on ice. Thereafter, 500 µl of cold Nuclei EZ lysis buffer (Sigma-Aldrich, USA) was added to each well. Cell lysates were collected using single-use cell scrapers and transferred to clean 1.7 ml tubes. The lysates were vortexed briefly and kept on ice for 5 minutes. Afterwards, samples were centrifuged at 500 x g for 5 minutes at 4 °C. The upper 300 µl of supernatant was transferred to a new tube and placed aside on ice (this represented the cytoplasmic fraction). The rest of the supernatant was discarded. The remaining pellet was resuspended in 500 µl of Nuclei EZ lysis buffer, mixed by pipetting 7 times, and transferred to a new 5 ml tube. Another 3.5 ml of Nuclei EZ lysis buffer was added and mixed. Samples were placed on ice for 5 minutes. Next, tubes were centrifuged at 500 x g for 5 minutes at 4 ºC. Supernatants were discarded, and the remaining pellets were resuspended in 300 µl of ice-cold radioimmunoprecipitation assay buffer (RIPA; Thermo Fisher, USA). Afterwards, the resuspended pellets (representing the nuclear fraction) were passed through sterile needles (25G) 8 times.

Once the separation of cellular fractions was completed, RNA was extracted using the acid guanidinium thiocyanate-phenol-chloroform method. 2 ml of RNA Extracol (EURx, Poland) was added to 300 µl of each RNA fraction. Samples were then incubated at room temperature for 5 minutes. Afterwards, 400 µl of chloroform was added, the samples were vortexed vigorously for 15 seconds, and left undisturbed for 3 minutes. Afterwards, the tubes were centrifuged at 12,000 x for 15 minutes at 4 ºC. The upper (aqueous) layers were carefully transferred to new tubes. Subsequently, RNA isolations were performed using the GeneMATRIX Universal RNA Purification Kit (EURx, Poland) as described above (chapter 3.9.; starting with the application of aqueous layers to the homogenizing minicolumns).

The nuclear and cytoplasmic RNA fractions were isolated at least three times from both analyzed cell lines. qRT-PCR was performed using 1000 ng of RNA obtained from each fraction. The relative subcellular location of transcripts was estimated by calculating differences in the Ct values obtained for equal mass inputs of nuclear and cytoplasmic RNA (both originating from a single fractionation experiment). Two lncRNAs with known subcellular localizations were used as controls: *MALAT1* (a nuclear lncRNA) and *small nucleolar RNA host gene 5* (*SNHG5*, a cytoplasmic lncRNA) [145-148].

3.11. Gene silencing

Gene silencing experiments were performed using antisense oligonucleotides (ASOs). ASOs typically contain chemical modifications that increase their nuclease-resistance. This study utilized GapmeRs (Qiagen, Germany), proprietary oligonucleotides that contain *locked nucleic acid* (LNA)-modified nucleotides in the ASOs' flanks, and a central "gap" consisting of regular DNA nucleotides. The LNA nucleotides (containing a methylene bridge between the 2′ and 4′ position of ribose) increase the GapmeR's stability and affinity, while the central DNA nucleotides produce DNA/RNA hybrids that elicit degradation of the targeted transcripts by RNase H [149].

Gene silencing reactions were performed as follows. Cells were plated in 6-well plates 24 hours before transfections (3 x 10^5 cells per well for MDA-T32, MDA-T68, K1 and MCF-7, or 4 x 10^5 cells per well for MDA-T120). Cells reached confluency of ~70% on the day of transfection. Complete culture medium was replaced with serum- and antibiotic-free medium immediately before transfections. Stock GapmeR solutions were diluted in OptiMEM

(Gibco, UK) in the tube number one, while Lipofectamine 2000 (Thermo Fisher, USA) was diluted in OptiMEM in the tube number two. The solutions were incubated at room temperature for 5 minutes. Afterwards, the contents of both tubes were mixed together (1:1 ratio), and incubated at room temperature for 15 minutes. Thereafter, 250 µl of the mixture was applied drop-wise to each well (attaining a total medium volume of 2 ml in each well). The final amount of Lipofectamine 2000 was 5 µl per well, and the final concentration of GapmeR was 50 nM. The transfection medium was replaced with fresh complete culture medium (containing 10% FBS and antibiotics) five hours after transfections. Cells were collected 72 hours after transfections (for the isolation of total RNA, and in the migration and invasiveness assays), or after 48 hours (in the soft agar assay experiments). In MTT assays transfections were performed directly in 96-well plates. Accordingly, volumes of reagents were scaled down (0.25 µl of Lipofectamine 2000 per well, 0.1 ml of culture media per well).

The GapmeR sequences are provided in Table 4. An early design of a *DIRC3*-targeting GapmeR (termed anti-*DIRC3*-common_1) was found to be non-functional and was not used in the subsequent experimental procedures. GapmeR targeting *MALAT1* was used as a positive control, and a commercial scramble GapmeR was utilized as a negative control. Custom GapmeRs were automatically designed using a dedicated on-line tool offered by Qiagen.

GapmeR	**Sequence**
anti-*DIRC3*-common_2	T*A*A*A*A*T*G*G*C*A*G*G*G*T*G*T
anti-*DIRC3*-common_3	C*T*T*G*A*A*T*A*G*A*G*G*C*T*A
anti-*DIRC3*-202	C*A*C*A*A*A*T*A*A*G*T*A*G*A*C*G
anti-*IGFBP5*	G*G*A*A*T*G*G*T*G*C*G*A*G*T*A*T
anti-*MALAT1*	C*G*T*T*A*A*C*T*A*G*G*C*T*T*T*A
negative control A	A*A*C*A*C*G*T*C*T*A*T*A*C*G*C

Table 4. GapmeR oligonucleotides. * indicates a phosphorothioate backbone modification.

3.12. Migration and invasiveness assays

Migration and invasiveness of cancer cells were tested using Transwell chambers. Firstly, cells were transfected with GapmeRs in 6-well plates. Complete culture medium was replaced with serum-free culture medium 48 hours after transfections. After 24 hours of serum starvation, cells were washed with PBS and detached from wells using the Non-enzymatic Cell Dissociation Solution (Sigma-Aldrich, USA). Afterwards, cells were collected into tubes,

centrifuged (300 x g for 5 minutes), resuspended in serum-free medium, counted, and re-diluted in appropriate serum-free medium. 333 µl of the cell suspension (containing 5 x 10^4 cells) was gently pipetted to the upper compartment of 6.5 mm Transwell with 8.0 µm Pore Polycarbonate Membrane Inserts (Corning, USA) placed in 24-well plates. The lower compartment of these chambers was filled with 750 µl of complete culture medium (containing 10% FBS). Cells were cultured for 22 hours. Afterwards, all liquids were removed. The inserts were immersed in 4% paraformaldehyde (PFA) solution for 5 minutes, and then stained using 0.05% crystal violet solution (Sigma-Aldrich, USA) for 5 minutes. Afterwards, the inserts were rinsed 5 times in water. Any remaining liquids were gently pipetted away. Upper layers of the insert membranes were firmly cleansed with cotton swabs. Finally, the membranes were visualized using an optical microscope (40x magnification). At least 5 random fields of view were photographed for each insert. The number of migrating cells was counted using ImageJ (National Institutes of Health, USA). Results were presented as a mean number of migrating cells per field of view. Each experiment was repeated at least three times.

Cell invasiveness was assessed using the Matrigel-precoated Transwell inserts (BioCoat Matrigel Invasion Chambers with 8.0 µm Pores; Corning, USA). Matrigel is a basement membrane preparation extracted from the Engelbreth-Holm-Swarm murine sarcoma. It consists of extracellular matrix proteins, such as laminin, collagen IV, heparan sulfate proteoglycans, entactin/nidogen and growth factors. Matrigel matrix is commonly used as an *in vitro* model of tissue basement membranes. Frozen inserts were thawed to room temperature immediately before use and rehydrated using warm serum-free culture medium. Subsequently, the invasion assays were executed in the same manner as for the uncoated inserts (described above).

In addition, I tested whether IGF-1 is capable to induce chemotaxis in thyroid cancer cells. This additional experiment was performed due to a hypothetical role of *DIRC3* in the modulation of cellular response to IGF-1. Assays were performed in the same manner as in the regular Transwell migration experiments, except the 24-hour serum-starvation step was omitted (I observed that prior serum starvation strongly abrogated the IGF-1-induced chemotaxis in the tested cell lines). The lower compartment of Transwell chambers was filled with 750 µl of serum-free culture medium containing 100 ng/ml of IGF-1 (Gibco, USA). Each experiment was repeated at least three times.

3.13. MTT assay

MTT assay measures the cellular metabolic activity which indirectly indicates the number of viable cells growing in standard culture conditions. MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, is a soluble tetrazolium dye (yellow color) that is reduced to insoluble formazan crystals (purple color). Reduction of MTT takes place in viable cells. Once the formazan crystals are formed, they may be dissolved using a solution containing sodium dodecyl sulfate (SDS) and diluted hydrochloric acid. Finally, the absorbance of colored solutions is registered at the wavelength of 560-570 nm. The absorbance is directly and nearly linearly proportional to the number of living cells in a sample.

0.1 ml of cell suspensions were plated in 96-well plates (3300 and 5000 cells per well for MDA-T32 and MDA-T120, respectively). Cells were transfected with GapmeRs (in the gene silencing experiments) or sgRNAs (in the CRISPRa overexpression experiments) 24 hours after plating. Culture medium was replaced five hours after transfections (0.2 ml of fresh culture medium was added to each well). Empty wells were filled with PBS to prevent uneven evaporation.

MTT assays were performed in 24-hour intervals (up to 96 hours post-transfection). Briefly, culture medium was replaced with 0.1 ml of OptiMEM. 10 µL of 12 mM MTT solution (Sigma-Aldrich, USA) was added to each well. Negative controls were also included (MTT stock solution was added to 100 µL of OptiMEM in cell-free wells). Plates were incubated for 4 hours at 37 °C. Afterwards, 100 µL of dissolving solution (0.1 g/ml SDS in 0.01 M HCl; Thermo Fisher, USA) was added to each well and gently mixed by pipetting. Plates were incubated overnight at 37 °C. Finally, the well absorbance was measured using the GloMax microplate reader (Promega, USA) using the 560 nm wavelength reads. Absorbance of blank wells was subtracted from the absorbance measured in samples. Results were registered as optical densities (ODs). Four technical replicates were included for each sample. All experiments were repeated at least three times. Mean ODs of all replicates were calculated.

3.14. Apoptosis assay

Apoptosis is a process of programmed cellular death that may be induced by different stimuli (e.g., nutrient starvation, hypoxia, chemicals, immune reactions, pathogens, temperature,

radiation). It is a complex mechanism that requires an intrinsic or extrinsic activation of caspases (proteases), which ultimately activate intracellular proteins that execute the cell death program. Activity of the effector caspases 3 and 7 is frequently used to evaluate the intensity of apoptosis *in vitro*. The magnitude of apoptosis may be modulated by the presence of growth factors. In particular, apoptosis is inhibited by IGF-1, however IGFBP5 or IGFBP-3 may abolish this inhibitory effect [150-153]. Accordingly, I decided to investigate whether silencing of *DIRC3* could influence the apoptosis of thyroid cancer cells that were deprived of serum and L-glutamine.

Apoptosis was measured using the Caspase-Glo 3/7 assay (Promega, USA). This test is based on addition of a substrate (DEVD-aminoluciferin) that is cleaved by active caspase-3/7 released from apoptotic cells. The cleaved substrate stimulates a proprietary luciferase that generates a luminescent signal. The assay was performed according to the manufacturer's instructions. Briefly, MDA-T32 and MDA-T120 cells were plated in 96-well white-walled microplates (Corning, USA). Cells were transfected with GapmeRs 24 hours after plating. Cells were near confluent after 72 hours of incubation. Next, complete culture medium was discarded, and cells were briefly rinsed with PBS. Thereafter, cells were cultured in 0.1 ml of L-glutamine- and serum-free Dulbecco's Modified Eagle Medium (DMEM, Gibco, USA). Blank wells were also filled with DMEM. After 24 hours of starvation, the culture plates were removed from incubators and left to equilibrate to room temperature. The Caspase-Glo 3/7 substrate was reconstituted in the provided buffer, and 0.1 ml of the substrate was added to each well. Luminescence was measured after 1 hour of incubation at room temperature. Measurements were performed using the GloMax microplate reader (Promega, USA). The final luminescence values were expressed as a ratio between the luminescence of samples and blanks. Experiments were repeated three times with technical duplicates in each assay. Mean values for all six replicates were calculated.

3.15. Soft agar assay

Soft agar colony formation assay is a test that measures the anchorage-independent growth, i.e., the hallmark capability of cancer cells to grow without attachment to a solid surface.

A low-melting point agarose (SeaPlaque, Lonza, Switzerland) was added to distilled water to produce 3.2% agarose stock solution. The solution was autoclaved and cooled in a

water bath to 37 °C. The agarose solution was mixed with warm complete RPMI-1640 culture medium (1:3 volume ratio). This mix was immediately poured to 6-well plates (1 ml per well) to produce 0.8% base agarose layer. Plates were swirled briefly to distribute the agarose solution evenly. Next, the plates were cooled at 4 °C for 5 minutes, and transferred to room temperature for immediate use.

MDA-T32 and MDA-T120 cells were plated in 6-well plates, transfected with GapmeRs (as described in chapter 3.11) and cultured for another 48 hours. Afterwards, the cells were rinsed with PBS, trypsinized, pelleted, resuspended in complete culture medium, and counted. Cell suspensions were diluted to 1.15 x 10^4 cells per 1 ml. 5 ml of the diluted cell suspension was swiftly mixed with 0.75 ml of warm 3.2% stock agarose solution (thus producing 0.42% agarose mix). 1 ml of the cell/agarose mixture was applied to the wells previously coated with 0.8% base agarose layer. Plates were gently swirled to distribute solutions evenly and moved to 4 °C for 5 minutes. Finally, wells were covered with 1 ml of complete culture medium. Plates were cultured in a humified incubator for 16 days (MDA-T32) or 21 days (MDA-T120). Culture medium covering the agarose gels was carefully replaced every 5 days. Once the incubation was completed, the covering medium was discarded and gels were covered with 1 ml of staining solution (an aqueous solution of 0.005% crystal violet and 10% ethanol). Staining was performed for at least 1 hour. Thereafter, the staining solution was removed and wells were gently rinsed with water. Finally, plates were placed on a white transilluminator and photographed using a digital camera. Visible colonies were counted using ImageJ (National Institutes of Health, USA). Assays were repeated at least three times with technical triplicates in each experiment.

3.16. RNA sequencing

RNA sequencing (RNA-seq) was performed to assess global transcriptomic alterations triggered by silencing of *DIRC3* or *IGFBP5* in MDA-T32 cell line. Transcriptomic data generated in cells transfected with the gene silencing GapmeRs was compared with the transcriptomic data produced in cells transfected with the negative control GapmeR. Analogous gene expression profiling experiments were executed for MDA-T32 clones that were modified using CRISPR/Cas9. Transcriptomes of the rs11693806-edited clones were compared with the transcriptomic profile obtained for wild-type (unmodified) cells.

Total RNA was isolated from MDA-T32 cells 72 hours after transfections with GapmeRs (anti-*DIRC3*-common_3, anti-*IGFBP5* or negative control A). In experiments utilizing the CRISPR-edited clones, total RNA was isolated 96 hours after cell seeding. All samples were prepared in biological triplicates. Efficient silencing of the GapmeRs' targets was verified using qRT-PCR. Once validated, the samples were used for RNA-sequencing. All steps of Next Generation Sequencing (including the RNA quality control, library preparation, sequencing, bioinformatic analysis) were performed by GENEWIZ (Leipzig, Germany) using the "Standard RNA-seq service".

Briefly, RNA concentration was measured using Qubit (Thermo Fisher, USA) and the sample quality was evaluated using Fragment Analyzer (Agilent, USA). All samples achieved RNA Quality Number (RQN) of 10 (the maximum score). Libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs, USA). The library generation workflow included: enrichment of transcripts (polyA selection utilizing oligo-dT beads), RNA fragmentation, cDNA synthesis with random priming, end repair, 5' phosphorylation and dA-tailing, adapter ligation, and the final PCR enrichment. Sequencing was performed on the Illumina NovaSeq 6000 instrument using paired-end reads (PE 2x150).

Raw reads were quality evaluated using FastQC [154]. The ≥Q30 parameter exceeded 92% in all samples (i.e., 92% of bases were called with an accuracy of ≥99.9%) with the mean quality score of 35.6. Sequence reads were trimmed to remove adapters and poor-quality nucleotides using Trimmomatic v.0.36 [155]. Reads were mapped to the *Homo sapiens* GRCh38 reference genome (available on Ensembl) using STAR aligner v.2.5.2b [156]. BAM files were generated with at least 96% of all reads being successfully mapped in each sample. Unique gene hit counts were calculated using the featureCounts software from Subread package v.1.5.2 [157]. Gene expression levels were expressed as Transcripts Per Kilobase Million (TPM). Differential gene expression was determined using DESeq2 [158]. The Wald test was performed to obtain p-values and log2 fold changes. Genes with the adjusted p-value of < 0.05 (the Benjamini-Hochberg false-discovery rate correction) and the absolute log2 fold change >1 were denoted as differentially expressed genes (DEGs). Gene ontology (GO) analyses were performed using the GeneSCF v.1.1-p2 software [159]. The goa_human GO list was used to obtain clusters of genes related to particular biological processes.

3.17. Plasmids

Plasmids were used to produce cell line derivatives with a stable expression of selected genes. Plasmids were obtained from Addgene (USA) and Invitrogen (USA).

pcDNA3-IGFBP5-V5 plasmid was a gift from Steven Johnson (Addgene plasmid #11608; http://n2t.net/addgene:11608; RRID: Addgene_11608). The plasmid was used to induce overexpression of *IGFBP5* in the gene "rescue" experiments. The pcDNA3 plasmid (Invitrogen, USA) was used as an empty control vector.

SP-dCas9-VPR plasmid was a gift from George Church (Addgene plasmid #63798; http://n2t.net/addgene:63798; RRID: Addgene_63798). This plasmid encodes a nuclease-null Cas9 protein (dead Cas9, dCas9) fused with the VP64-p65-Rta (VPR) domain. VPR is a tripartite transcriptional activator that may potently induce expression of endogenous coding and noncoding genes [160]. dCas9-VPR protein hybridizes with a guide RNA (gRNA), which is delivered separately to the plasmid-transfected cells. Consequently, gRNA anchors the ribonucleoprotein complex to its DNA target. Thereafter, dCas9-VPR is capable to recruit transcriptomic machinery to a specific genomic locus (typically to the gene promoter).

The SP-dCas9-VPR plasmid was used in the *CRISPR activation* (CRISPRa) experiments. The main goal of this procedure was to upregulate *DIRC3* in thyroid cancer cell lines. An important advantage of CRISPRa over the conventional plasmid-mediated gene overexpression is a possibility to elicit transcription of a given gene in its endogenous genomic locus [161]. This is especially important for nuclear lncRNAs whose activity is often nuclear and cis-regulatory. In contrast, expression plasmids produce transcripts in the cytoplasm, and thus cannot recapitulate the nuclear regulatory functions of many lncRNAs. CRISPRa has been successfully applied to functionally characterize several nuclear lncRNAs [161, 162].

The Addgene plasmids were obtained as bacteria in agar stabs (E. coli DH5α strains). Luria broth (LB) agar plates containing 100 μg/mL ampicillin (Gibco, USA) were prepared on 10 cm polystyrene Petri dishes. Bacteria growing in agar stabs were touched with a sterile pipette tip. The tip was gently spread on LB plates. Plates were then incubated overnight at 37 °C. Afterwards, single colonies were collected using sterile tips. The tips were dropped to round-bottom falcon tubes filled with S.O.C. medium (Super Optimal Catabolite medium, Invitrogen, USA). The cultures were incubated in an orbital shaking incubator at 37 °C for 12 hours. Once the S.O.C. medium became hazy, one part of the culture was used to prepare

glycerol stocks (medium was mixed with an equal volume of 50% glycerol and frozen at -80 °C). The remaining medium was used to isolate plasmids using the GeneMATRIX Plasmid Miniprep DNA Purification Kit (EURx, Poland). Briefly, 3 ml of the culture was centrifuged at 12,000 x g for 2 minutes. Bacterial pellets were resuspended in Cell R buffer and lysed using Lysis Blue buffer. The content was mixed gently by inversions. Afterwards, the mix was neutralized using Neutral B buffer and mixed gently by inversions. Next, the tubes were centrifuged at 12,000 x g for 7 minutes. The supernatant was applied to silicone minicolumns, and spined at 11,000 x g for 1 minute. The columns were flushed with Wash PLX1 and Wash PLX2 buffers with centrifugations at 11,000 x g for 1 minute being performed after each buffer wash. After the final centrifugation, the columns were transferred to new test tubes. 50 µl of Elution Buffer was applied to the silicone membranes. After incubation at room temperature for 5 minutes, the columns were centrifuged at 11,000 x g for 1 minute. Plasmid concentrations were measured using the NanoDrop 2000 instrument (Thermo Fisher, USA). Plasmid elutes were frozen at -20 °C. Identities of the isolated plasmids were confirmed with Sanger sequencing (performed by Genomed, Warsaw, Poland). Universal sequencing primers were used for this analysis (CMV-F for SP-dCas9-VPR, and T7 for pcDNA3-IGFBP5-V5).

3.18. Stable cell line generation

Cell lines with stable overexpression of specific proteins were produced by transfecting cell lines with the aforementioned expression plasmids. Selection of transfected clones was performed using geneticin (G418) as all plasmids used in the study contain the neomycin resistance gene (neo) that confers resistance to G418.

Cells were plated in 24-well plates 24 hours before transfections (7.5 x 10^4 cells in 0.5 ml of antibiotic-free culture medium per well). Fugene 6 transfection reagent (Promega, USA) and OptiMEM medium were allowed to equilibrate to room temperature before use. 6 µl of Fugene 6 was added to 94 µl of OptiMEM, mixed gently and incubated at room temperature for 5 minutes. Thereafter, 2 µg of plasmid (SP-dCas9-VPR, pcDNA3-IGFBP5-V5, or empty pcDNA3 plasmid) was added to the Fugene/OptiMEM solution. Plasmids were not added to the control wells (a matching volume of OptiMEM was added instead). Tubes were gently mixed and incubated at room temperature for 15 minutes. Afterwards, 50 µl of the transfection mix was added dropwise to the cells. Culture plates were swirled and placed in a humified incubator for 48 hours. Afterwards, the transfection medium was replaced with

complete culture medium supplemented with penicillin, streptomycin and G418 (Clontech, USA). Cells were cultured for at least three weeks. Culture medium containing G418 was replaced every 4 days. This treatment resulted in a complete elimination of cells in the control wells, and a gradual expansion of the G418-resistant clones in the transfected wells.

The appropriate, cell-selecting concentration of G418 was determined experimentally in each cell line using the concentration-killing curve approach. Briefly, cells were plated in 96-well plates and treated with various concentrations of G418 for two weeks. Afterwards, cell viability was tested using MTT assay. The lowest concentration of G418 that resulted in a complete elimination of cells was selected for later use. The concentration of G418 suitable for the clone selection was determined as: 600 μg/ml for MDA-T32, MDA-T120 and MDA-T68, and 300 μg/ml for K1 cell line. The selected clones were maintained in 500 μg/ml and 200 μg/ml of G418, respectively.

3.19. Western blot

Western blotting was done to evaluate alterations in the activity of IGF1 signaling pathway in thyroid cancer cells undergoing silencing of *DIRC3* or *IGFBP5*.

MDA-T32 cells were plated in 6-well plates (3.0 x 10^5 cells per well) one day before transfections. Transfections were performed as described above (chapter 3.11) using Lipofectamine 2000 and GapmeRs (anti-*DIRC3*-common_3, anti-*IGFBP5*, or negative control A). Serum-free RPMI-1640 medium was removed 5 hours after transfections, and the cells were cultured in complete medium (containing 10% FBS) for another 67 hours. Afterwards, the medium was removed, cells were rinsed with PBS, and cultured in 2 ml of serum-free RPMI-1640 medium containing IGF-1 (20 ng/ml; Gibco, USA). Cells were incubated for another 24 hours. Next, the conditioned medium was collected into marked tubes and stored at 4 ºC. Meanwhile, the cells were rinsed with PBS and cultured in 2 ml of serum-free RPMI medium. Cells were serum-starved for 12 hours. Subsequently, the starvation medium was discarded, and the cells were re-stimulated with 1 ml of the conditioned medium that was collected 12 hours prior (the conditioned medium was returned to the original wells). The conditioned media were prepared in three version: a) unmodified, b) supplemented with additional IGF-1 (extra 30 ng/ml), or c) supplemented with the recombinant human IGFBP5 protein (500 ng/ml; Peprotech, USA).

Cells were stimulated with the respective media for 10 minutes. Afterwards, the medium was removed from each well, culture plates were placed on ice, and cells were briefly rinsed with ice-cold PBS. 90 µl of ice-cold RIPA buffer supplemented with the Halt Protease and Phosphatase Inhibitor Cocktail (both Thermo Fisher, USA) was added to each well. Next, cells were scraped using chilled single-use scrapers. The resulting lysates were collected to 1.7 ml tubes, placed on ice, and agitated on an orbital shaker for 30 minutes. Next, the samples were centrifuged at 12,000 x g for 20 minutes at 4 ºC. Supernatants were transferred to new tubes, aliquoted and frozen at -80 ºC.

Thereafter, protein aliquots were thawed and the protein concentrations were measured using the Pierce BCA Protein Assay Kit (Thermo Fisher, USA) according to the manufacturer's manual. This bicinchoninic acid (BCA) assay was performed in 96-well microplates. Briefly, 10 µl of protein samples and 10 µl of serially diluted protein standards (bovine serum albumin, BSA) were pipetted in triplicates. 200 µl of the BCA working mix (consisting of 50 parts of BCA Reagent A mixed with 1 part of BCA Reagent B) was added to the protein samples and standards. Plates were incubated at 37 °C for 30 minutes. Next, the plates were moved to room temperature. Absorbance in wells was measured using the GloMax microplate reader (Promega, USA) utilizing the 560 nm wavelength reads. Absorbance of blank wells was subtracted from the absorbance obtained for each protein sample. Standard curve was generated for the BSA standards. Protein concentrations in unknown samples were determined using the standard curve method.

Finally, protein samples were analyzed using the polyacrylamide gel electrophoresis (PAGE) and Western blotting. Briefly, protein samples were diluted in RIPA buffer to obtain equal protein concentrations in all aliquots. Thereafter, samples were mixed with Laemmli loading buffer (Bio-Rad, USA) and boiled at 95 °C for 5 minutes. 10 µg of each protein sample was loaded on 10% Mini-PROTEAN TGX polyacrylamide gel (Bio-Rad, USA). Additionally, PageRuler Prestained Protein Ladder (Bio-Rad, USA) was loaded on the gel. Electrophoresis was performed using the Mini-PROTEAN Tetra cells (Bio-Rad, USA) filled with 1x Tris/Glycine/SDS buffer (Bio-Rad, USA). Gels were run at 50 V for 5 minutes, and at 100 V for another 60 minutes. Afterwards, gels were removed from cassettes, immersed into 1x Novex Tris-Glycine Transfer Buffer (Thermo Fisher, USA) and assembled into a "transfer sandwich" containing a polyvinylidene fluoride (PVDF) membrane (Bio-Rad, USA). The "transfer sandwich" and icepacks were placed into a transfer tank (Bio-Rad, USA) filled with the Transfer Buffer. Protein transfer was performed overnight in a cold room at a current of

10 mA. Afterward, PVDF membranes were rinsed with distilled water and blocked with 5% (w/v) non-fat dry milk in Tris-buffered saline/1% Tween-20 (TBS-T) for 1 hour. Thereafter, membranes were incubated with primary antibody solutions overnight at 4 ºC (Table 5). Next, membranes were washed three times with TBS-T, and incubated with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibody at room temperature for 1 hour. Membranes were then washed with TBS-T three times. Finally, protein detection was performed using the SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher USA). Chemidoc Touch System (Bio-Rad, USA) was used for image acquisitions. ImageJ software (NIH, USA) was used for the densitometric analysis of protein blots.

Antibody	Dilution	Catalog number	Manufacturer
Phospho-IGF-I Receptor β (Tyr1135) (DA7A8) Rabbit mAb	1:1000	#3918	Cell Signaling, USA
IGF-I Receptor β (D23H3) XP Rabbit mAb	1:1000	#9750	Cell Signaling, USA
Phospho-AKT (Ser473) (D9E) XP Rabbit	1:2000	#4060	Cell Signaling, USA
AKT (pan) (C67E7) Rabbit mAb	1:1000	#4691	Cell Signaling, USA
Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E) XP Rabbit mAb	1:1000	#4370	Cell Signaling, USA
beta-Actin Antibody, Rabbit mAb	1:1000	#4967	Cell Signaling, USA
Peroxidase AffiniPure Goat Anti-Rabbit IgG [H+L]	1:10,000	111-035-144	Jackson Immuno Research Labs, USA

Table 5. Antibodies used for Western blotting.

3.20. Single-guide RNAs

Guide RNA (gRNA) is a molecule that binds Cas9 nuclease and delivers the newly formed ribonucleoprotein complex to a specific nuclear locus. A common method is to apply gRNAs as single-guide RNAs (sgRNAs). sgRNA integrates two separate components: a custom-designed short crRNA (that hybridizes to the target DNA) and a scaffold tracrRNA (that binds to Cas9). Genomic editing is possible if the targeted genomic locus is located next to the protospacer adjacent motif (PAM) sequence. In the case of SpCas9 (protein originating from

Streptococcus pyogenes), the PAM sequence is 5′-NGG-3′ (where "N" is any nucleotide base). Cas9 performs double-stranded DNA cuts exactly three nucleotides upstream from PAM.

In this study sgRNAs were utilized for the genomic editing of the rs11693806 locus, as well as in the CRISPRa experiments. sgRNAs were designed using two publicly available on-line tools: CHOPCHOP [163] (University of Bergen, Norway; http://chopchop.cbu.uib.no/) and CRISPick [164] (The Broad Institute, USA; https://portals.broadinstitute.org/gppx/crispick/public). These tools permit design of sgRNAs that meet stringent criteria (i.e., presence of PAM in the target region, lack of any significant off-targets, high predicted editing efficiency). The sgRNA that targeted the rs11693806 locus was computationally designed using a sequence input of 100 base pairs upstream and 100 base pairs downstream from rs11693806. Important note: rs11693806 is located directly next to another germline variant, rs4674173. This SNV is annotated as rs4674173[A] in the reference human genomes (GRCh37 and GRCh38). This allele is, however, present in very few populations (principally as a minor allele in the people of African ancestry) and is totally absent in Caucasians (in this population the only allele is rs4674173[G] as per The 1000 Genomes Project). Accordingly, sgRNAs and the rhAmp genotyping assays had to accommodate rs4674173[G].

sgRNAs used in the CRISPRa experiments were primarily designed using an input window of 400 to 50 base pairs upstream from the Transcription Start Site (TSS) of *DIRC3*. This target window usually provides the highest efficiency of the CRISPRa-mediated induction of gene expression [165]. TSS was identified using the Cap Analysis of Gene Expression (CAGE) data generated in FANTOM5 project (Riken, Japan) [166]. FANTOM5/CAGE has been previously shown to be the most reliable method of the TSS annotation for CRISPR experiments [167]. The primary TSS of *DIRC3* was identified using the Zenbu genome browser (https://fantom.gsc.riken.jp/zenbu/).

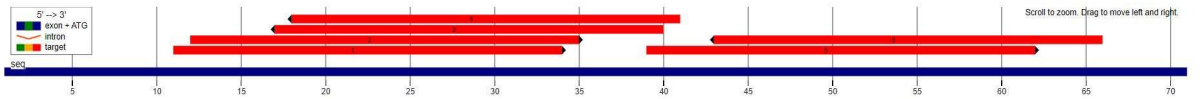
Custom oligonucleotides used for the synthesis of sgRNAs are shown in Table 6. The sgRNA used for the genomic editing of rs11693806 was one of the top picks produced by both CHOPCHOP and CRISPick (Figure 3). This sgRNA had a high on-target efficacy score, no significant off-targets, and its PAM was ideally located exactly three base pairs downstream from rs11693806. This distance has been previously shown to be optimal for efficient editing using the homology-directed repair (HDR) mechanisms [168, 169]. The variant-editing sgRNA was designed to target genomic sequences containing rs11693806[C].

	Target F1 forward primer for template assembly	Target R1 reverse primer for template assembly	Final target sgRNA sequence
sgRNA-*DIRC3*-1	TAATAC GACTCA CTATAG CTTTGA GGGATT ACGGG	TTCTAG CTCTAA AACGCG CCCGTA ATCCCT CAAA	CTTTGAGGGATTACGGGCGC
sgRNA-*DIRC3*-2	TAATAC GACTCA CTATAG TCACGC ACGTAT GCCCC	TTCTAG CTCTAA AACTGC GGGGCA TACGTG CGTG	TCACGCACGTATGCCCCGCA
sgRNA-*DIRC3*-3	TAATAC GACTCA CTATAG GCGCAG GGATAG CTCGC	TTCTAG CTCTAA AACCCA GCGAGC TATCCC TGCG	GCGCAGGGATAGCTCGCTGG
sgRNA-*DIRC3*-4	TAATAC GACTCA CTATAG CACGCA CGTATG CCCCG	TTCTAG CTCTAA AACATG CGGGGC ATACGT GCGT	CACGCACGTATGCCCCGCAT
sgRNA-*DIRC3*-5	TAATAC GACTCA CTATAG CTGCGT GTGAGC AGCTC	TTCTAG CTCTAA AACCCG GAGCTG CTCACA CGCA	CTGCGTGTGAGCAGCTCCGG
sgRNA-*PVT1*	TAATAC GACTCA CTATAG GCCGGG ACCGAG GACGCA CG	TTCTAG CTCTAA AACCGT GCGTCC TCGGTC CCGGC	GCCGGGACCGAGGACGCACG
sgRNA-non-target	TAATAC GACTCA CTATAG ACGGAG GCTAAG CGTCGC AA	TTCTAG CTCTAA AACTTG CGACGC TTAGCC TCCGT	ACGGAGGCTAAGCGTCGCAA
sgRNA-rs11693806[C]	TAATAC GACTCA CTATAG ATTTCC CAGACA TTCGCG A	TTCTAG CTCTAA AACTCG CGAATG TCTGGG AAAT	GATTTCCCAGACATTCGCGA

Table 6. Custom oligonucleotides used for the synthesis of sgRNA templates.

Similarly, sgRNAs used in the CRISPRa experiments were chosen from multiple high-ranking designs obtained from the CHOPCHOP and CRISPick on-line tools. Five non-overlapping sgRNA designs were selected to cover the previously specified target window located upstream from the TSS of *DIRC3*. All sgRNAs were tested empirically and the best performing designs were used in subsequent phenotypic experiments. A previously validated sgRNA inducing overexpression of *PVT1* was used as a positive control [170]. Conversely, a validated sgRNA designed not to target any site in the human genome (the non-target sgRNA) was used as a negative control [171].

A.



Download results: Please select one

View in UCSC genome browser

Rank	Target sequence	Genomic location	Strand	GC content (%)	Self-complementarity	MM0	MM1	MM2	MM3	Efficiency
1	GGATTTCCCAGACATTCGCGAGG	seq:12	+	55	0	0	1	0	3	65.42
2	GATTTCCCAGACATTCGCGAGGG	seq:13	+	50	0	0	1	0	4	64.74
3	TTTCTCCCTCGCGAATGTCTGGG	seq:18	-	50	0	0	1	0	6	52.06
4	TTTTCTCCCTCGCGAATGTCTGG	seq:19	-	50	0	0	1	0	8	29.80
5	AATGCCATGTCCTTTCCTTCTGG	seq:40	+	45	1	1	1	4	26	36.24
6	GACACCAGAAGGAAAGGACATGG	seq:44	-	50	1	1	0	14	42	67.35

B.

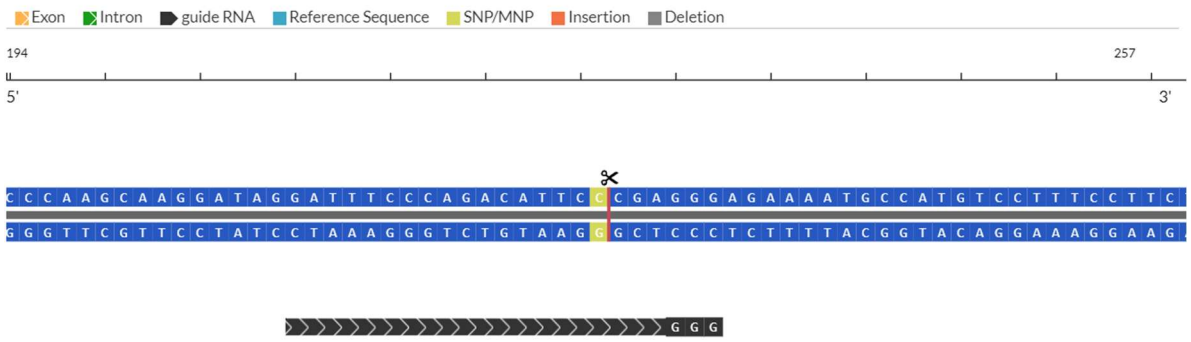


Figure 3. Design of the sgRNA used to edit the rs11693806 locus.

(A). Output from CHOPCHOP, an online sgRNA design tool. The selected design was ranked as number 2. Of note: while the sgRNA is marked to have a single nucleotide mismatch ("MM1"), it is actually on-target (the edited rs11693806 is positioned next to rs4674173, a variant annotated as rs4674173[A] in the human reference genome, but present exclusively as rs4674173[G] in Caucasians). **(B).** The selected guide RNA (black) with the "GGG" PAM sequence, and its genomic target in the *DIRC3* exon. Note that the genomic cut is located directly next to rs11693806 (light green). The scheme was generated using the CRISPR-Cas9 guide RNA design checker (Integrated DNA Technologies, USA).

sgRNAs were synthesized using the GeneArt Precision gRNA Synthesis Kit (Thermo Fisher, USA) according to the manufacturer's instructions. Briefly, two DNA oligonucleotides ("F1" and "R1" target primers) were designed for each sgRNA (Table 6). These oligonucleotides were complementary to each other in their 3' ends. "F1 primer" included the T7 promoter sequence in its 5' end, and was partially complementary to a universal forward primer included in the kit. The 5' end of "R1 primer" was complementary to a tracrRNA fragment also included in the kit. Furthermore, the tracrRNA fragment was partially complementary to a universal reverse primer. Target primers were combined and diluted in IDTE buffer to produce a working mix (0.3 µM of the F1 and R1 oligonucleotides, each). PCR assembly was performed by combining Phusion High-Fidelity PCR Master Mix (Thermo

Fisher, USA), the Tracr fragment and T7 primer mix, the F1/R1 primer mix, and nuclease-free water. PCR was performed in a thermal cycler (Mastercycler nexus gradient, Eppendorf, Germany). The following program was used: initial denaturation at 98 °C for 10 seconds, 32 cycles consisting of 98 °C for 5 seconds and 55 °C for 15 seconds, final extension at 72 °C for 1 minute, and cooling at 4 °C. This reaction produced sgRNA DNA templates. Afterwards, the sgRNA DNA templates were combined with dNTPs mix, 5X TranscriptAid Reaction Buffer and TranscriptAid Enzyme Mix. The solution was incubated at 37 °C for 4 hours to ensure a high yield of *in vitro* transcription (IVT). Afterwards, DNase I was added to the IVT reactions to remove DNA templates. sgRNAs were purified using the gRNA Clean Up Kit (Thermo Fisher, USA). Products were diluted with nuclease-free water, and mixed with 99% ethanol and Binding Buffer. The mixture was transferred to RNA Purification Micro Columns and centrifuged at 14,000 x g for 1 minute. Flow-through was discarded, and the columns were washed with Wash Buffer 1 and Wash Buffer 2, respectively (centrifugations were performed after each buffer application). Thereafter, the columns were dried down with another centrifugation and transferred to new tubes. 20 µl of nuclease-free water was applied to the columns. Final centrifugation at 14,000 x g for 1 minute was performed to elute sgRNAs. The sgRNA yield was measured using NanoDrop 2000. Ready sgRNAs were diluted with nuclease-free water and stored at -80 °C.

3.21. Genomic editing

CRISPR is a modular platform that enables DNA editing in predesigned genomic loci. Mechanically, nucleases (typically Cas9) are directed by gRNAs to create double-strand DNA breaks in a specific genomic site. Eukaryotic cells utilize two major repair mechanisms to repair this damage. The dominant repair pathway is non-homologous end joining (NHEJ), which connects disruptions owing to the presence of “microhomologies” in overhangs of the broken strands. NHEJ is error-prone as when the overhangs are not perfectly matched some terminal nucleotides may be lost. This consequently leads to a small DNA deletion. The alternative repair pathway is termed homology-directed repair (HDR). HDR is a process dependent on the presence of a homologous fragment of DNA which acts a repair template. The template may be either endogenous (second copy of the gene on the sister chromosome), or exogenous (a donor template that is delivered to the nucleus). HDR is an error-free repair method. Experimental donor DNA templates may be, however, intentionally modified to incorporate additional gene

fragments (i.e., to create insertions) or substitute particular nucleotides (e.g., for the precise variant editing). The efficiency of genomic integration of DNA templates is partially dependent on the design of donor oligonucleotides (e.g., the length of complementary ends, asymmetry of donor arms, delivery of donors as double-stranded or single-stranded oligonucleotides, and their resistance to endogenous nucleases) [168, 169, 172, 173]. Here, I used Alt-R HDR donor oligonucleotides (Integrated DNA Technologies, USA), a proprietary modification of oligonucleotides that increases their stability and increases the HDR rate. Alt-R HDR oligonucleotides were designed using an on-line tool (Integrated DNA Technologies, USA). They harbored rs11693806[G] allele which was intended to substitute rs11693806[C] in the edited thyroid cancer cells. Two Alt-R HDR donor oligonucleotides were tested as recommended by the manufacturer: one corresponding to the target strand, and another matching the non-target stand (Table 7). Additionally, Alt-R HDR Enhancer (Integrated DNA Technologies, USA) was utilized to increase the relative frequency of HDR events in the CRISPR-edited cells. This reagent act as a small-molecule inhibitor of NHEJ.

Alt-R HDR donor Oligo + strand ("non-target")	5'- /Alt-R-HDR1/ A*G*C CCA AGA CCC AAG CAA GGA TAG GAT TTC CCA GAC ATT C**C**C GAG GGA GAA AAT GCC ATG TCC TTT CCT TCT GGT GTC A*C*A /ALT-R-HDR2/ -3'
Alt-R HDR donor oligo - strand ("target")	5'- /Alt-R-HDR1/ T*G*T GAC ACC AGA AGG AAA GGA CAT GGC ATT TTC TCC CTC G**G**G AAT GTC TGG GAA ATC CTA TCC TTG CTT GGG TCT TGG G*C*T /ALT-R-HDR2/ -3'

Table 7. Alt-R HDR donor oligonucleotides used for the precise editing of rs11693806. Donor oligonucleotides harbor rs11693806[G] (underlined). Two designs were tested as recommended by the manufacturer: one oligonucleotide corresponds to the stand targeted by sgRNA ("the target stand"), and another donor corresponds to the complementary DNA strand (the "non-target stand"). "*" indicates a phosphorothioate bond between nucleotides. "Alt-R-HDR1" and "Alt-R-HDR2" indicate proprietary modifications of oligonucleotides (Integrated DNA Technologies, USA).

In this study the Cas9/sgRNA complexes were delivered using the ribonucleoprotein (RNP) transfections. This method induces transient editing and consequently produces less off-target cleavages than the plasmid-mediated delivery of the CRISPR/Cas9 components [174]. Transfections were performed as follows. MDA-T32 cells were plated in 24-well plates 24 hours before transfections (0.5 x 10^5 cells in 0.5 ml of complete antibiotic-free culture medium per well). Transfection mix consisting of 25 µl of OptiMEM, 1 µg of TrueCut Cas9 Protein v2 (Thermo Fisher, USA), 200 ng of IVT sgRNAs (200 ng/µl) and 2.5 µl of

Lipofectamine Cas9 Plus Reagent (Thermo Fisher, USA) was prepared in the tube number 1. The tube was vortexed briefly and incubated at room temperature for 5 minutes. Afterwards, 1 µl of 1.65 µM Alt-R donor oligonucleotide (corresponding to either target or non-target strand) was added to the tube number 1 (the final concentration of donor oligonucleotides was 3 nM in each well). 1.5 µl of Lipofectamine CRISPRMAX was added to 25 µl of OptiMEM in the tube number 2. The mixture was incubated at room temperature for 1 minute. Thereafter, the content of tube no. 2 was added to the tube no. 1 and mixed by pipetting. Mixed reagents were incubated for another 15 minutes, and thereafter added to the plated cells (50 µl per well). Finally, 5.5 µl of Alt-R HDR Enhancer (Integrated DNA Technologies, USA) was added to the wells (final concentration of 30 µM). Culture plates were swirled and moved to a humified incubator for 24 hours. Thereafter, the transfection medium was discarded and cells were rinsed with PBS. Fresh complete culture medium was added to each well. Cells were cultured for another 5-7 days to reach confluency. Afterwards, single cell clones were isolated using the dilution cloning method.

Dilution cloning was performed as described before (www.corning.com/catalog/cls/documents/protocols/Single_cell_cloning_protocol.pdf). Briefly, the CRISPR-edited cells were trypsinized and resuspended in fresh culture medium. 4,000 cells (2 x 10^4 cells/ml in 0.2 ml) were added to the upper-left well (A1) of a 96-well plate. The cells were serially diluted in the column A (by a factor of 2; down to the last well, A8). Thereafter, cells in the column A were serially diluted (by a factor of 2) up to the last column (column H). This was done using an 8-channel 200 µl micropipettor. Finally, the medium volumes in each well were brought up to 0.2 ml. Several such plates were prepared in each experiment. Plates were placed in a humified CO_2 incubator at 37 °C. After 4 days of incubation, plates were inspected under an optical microscope to identify wells that were likely to contain single-cell colonies. These wells were marked and observed for a few more weeks. Once the supervised clones reached confluency, they were trypsinized, split, and transferred to two new 96-well replication plates.

When cells in the replication plates reached confluency, crude DNA was extracted from one of the replication plates. Culture medium was removed, cells were washed with PBS, and 20 µL of QuickExtract DNA Extraction Solution 1.0 (Lucigen, USA) was added to each well. Cells were scraped with sterile pipette tips and the resulting lysates were transferred to a new PCR microplate. The PCR plate was transferred to a thermocycler. The following incubation program was used: 65 °C for 15 minutes, 68 °C for 15 minutes, 98 °C for 10 minutes, hold at

4 °C. Once the program was completed, the lysates were diluted with IDTE buffer (volume ratio 1:10) and tested using the rhAmp SNP genotyping assay (as described in chapter 3.6). This assay was utilized as a screening test to identify clones that lacked rs11693806[C] (normally present in the unedited heterozygotic MDA-T32 cells). Clones exhibiting this genomic modification were collected from the second replication plate, and further expanded in culture flasks. Once the selected clones achieved significant growth, pure genomic DNA was extracted using the GeneMATRIX Cell Culture DNA Purification Kit. Purified DNA was re-tested using the rhAmp SNP genotyping assay. Finally, DNA was amplified using the conventional PCR and sent for Sanger sequencing of the rs11693806 locus. Validated clones were further expanded and stocked.

3.22. CRISPRa

CRISPR activation (CRISPRa) was used to upregulate *DIRC3* in thyroid cancer cell lines. In these experiments I used cell lines expressing the SP-dCas9-VPR plasmid. Details on the plasmid, production of stable cell lines and sgRNAs are provided in the chapters 3.17, 3.18 and 3.20, respectively. Experiments were performed as follows. Cells expressing SP-dCas9-VPR were plated in 24-well plates 24 hours before transfections (0.5 x 10^5 cells in 0.5 ml of antibiotic-free culture medium). Afterwards, 1000 ng of IVT sgRNAs was added to 25 µl of OptiMEM in the tube number 1. 1.5 µL of Lipofectamine RNAiMAX (Thermo Fisher, USA) was added to 25 µl of OptiMEM in the tube number 2. The tubes were mixed and incubated at room temperature for 1 minute. Thereafter, the content of tube number 2 was added to the tube number 1. Mixed reagents were incubated at room temperature for 15 minutes. Finally, 50 µl of the mixture was applied drop-wise to the plated SP-dCas9-VPR-expressing cells. Cells were then cultured in a humified CO_2 incubator at 37 °C for 72 hours. Afterwards, total RNA was isolated using the GeneMATRIX Universal RNA Purification Kit (EURx, Poland). The sgRNA transfection reactions were scaled-down in MTT assays. MTT assays were performed in 24-hour intervals (up to 96 hours after transfections) as described in chapter 3.13.

3.23. T7 endonuclease assay

T7 endonuclease I (T7EI) assay is a robust method used to evaluate the cleavage efficiency of gRNAs. The technique relies on detection of mismatches between two DNA strands. The mismatches are generated by imprecise (NHEJ-mediated) repairs of DNA breaks. When DNA isolated from the CRISPR-edited cells is PCR-amplified, the denaturation and re-annealing of PCR products forms heteroduplexes between the mutant and wild-type PCR amplicons. Structural deformities in the DNA heteroduplexes are recognized and cut by T7 endonucleases. Consequently, when the PCR products are analyzed using gel electrophoresis, additional DNA bands (corresponding to the cleaved amplicons) are detected.

Cleavage efficiency of sgRNAs was tested using the GeneArt Genomic Cleavage Detection Kit (Thermo Fisher, USA). This test was applied to K1 cells which were transfected with Cas9/sgRNA complexes targeting the rs11693806 locus. Transfection reactions were performed as described in chapter 3.21, except the application of Alt-R donor oligonucleotides and Alt-R HDR Enhancer was omitted. The assay was performed according to the producer's manual. Briefly, cells were harvested 2 days after transfection and spined down at 200 x g for 5 minutes. Cell pellets were resuspended in Cell Lysis Buffer containing Protein Degrader. The resulting lysates were placed in PCR tubes. Afterwards, the following program was run in a thermal cycler: 68 °C for 15 minutes, 95 °C for 10 minutes, hold at 4 °C. Next, the mixture was vortexed and a fraction of the lysate was used in PCR. PCR was assembled using the cell lysate, forward and reverse primers (0.2 μM final, each; see Table 1), AmpliTaq Gold 360 Master Mix and nuclease-free water. PCR was run using the following program: enzyme activation at 95 °C for 10 minutes, 40 cycles of amplification (95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds), final extension at 72 °C for 7 minutes, and a final hold at 4 °C. A fraction of the PCR product was combined with 10X Detection Reaction Buffer and nuclease-free water. Re-annealing of PCR products was performed in a thermal cycler (Mastercycler nexus gradient, Eppendorf, Germany) using the following program: 95 °C for 5 minutes, temperature reduction to 85 °C (ramp -2 °C/second), cooling to 25 °C (ramp -0.1 °C/second), hold at 4 °C. Once the re-annealing procedure was completed, Detection Enzyme (T7EI) was added to the PCR products. The mixture was incubated at 37 °C for 1 hour. The Detection Enzyme was omitted in negative control reactions. Finally, the digested PCR products were run on 1% agarose gel. Products were visualized using a UV transilluminator and photographed.

3.24. Exome sequencing

Whole exome sequencing (WES) was performed to rule out potential off-target alterations in the CRISPR-edited clones. Three lines of MDA-T32 were sequenced: the parental rs11693806 heterozygotic cell line (wild-type), one rs11693806-homozygotic clone (termed here as "D7"), and one clone harboring a small monoallelic deletion in the variant locus (clone named here as "A9"). Genomic DNA was extracted using the GeneMATRIX Cell Culture DNA Purification Kit (EURx, Poland) as described in chapter 3.9.

Whole exome sequencing was performed by GENEWIZ (Leipzig, Germany). The experimental workflow included: RNA quality control, DNA fragmentation, library preparation (exome enrichment), PCR amplification and massive parallel sequencing. Briefly, the concentration and quality of DNA was measured using Qubit (Thermo Fisher, USA) and Fragment Analyzer (Agilent, USA). DNA was ultrasonically sheared, size-selected, and barcoded with adapters. Exome capture (hybridization-based) was performed using Twist Comprehensive Exome kit (Twist Bioscience, USA). This panel covers 36.8 Mb of human protein coding regions (lncRNAs are not included). Libraries were sequenced using NovaSeq 6000 Sequencing System (Illumina, USA) with paired-end reads (PE 2x150).

Raw data was analyzed bioinformatically by GENEWIZ. BCL files generated by the sequencer were converted to FASTQ files using bcl2fastq v2.19 (Illumina, USA). Reads were quality evaluated using FastQC [154]. Sequencing adapters and low-quality bases were trimmed using Trimmomatic 0.39 [155]. Cleaned reads were aligned to the GRCh38 reference genome using Sentieon 202112 (Sentieon, USA). Alignments were sorted and PCR/optical duplicates were marked. Accordingly, binary sequence aligned maps (BAM files) were produced. Single nucleotide variants (SNVs) and small indels were called using Sentieon 202112 (DNAScope algorithm). The Variant Call Format (VCF) files generated by the pipeline were normalized using bcftools 1.13 [175]. Additional VCF files were generated for the CRISPR-edited clones by using the exome of parental MDA-T32 cell line as the reference. Accordingly, variant calls unique to the CRISPR-edited clones were obtained.

Deliverables (principally VCF and BAM files) were manually filtered using VarSifter 1.9 [176] and visually inspected using the Integrative Genomics Viewer (IGV) [177, 178]. Particular attention was paid to the potential off-target sites of the rs11693806-editing sgRNA.

Putative off-targets were identified bioinformatically using Cas-OFFinder [179] and CRISPOR [180] tools.

3.25. Statistical analysis

Categorical variables were presented as frequency and proportions, continuous variables were shown as mean ± standard deviation (SD) unless stated otherwise. Fisher test was used to compare frequencies. Distribution of variables was assessed by plotting histograms. Statistical tests used for the analysis of normally distributed variables were: t-test (for comparison of two groups), one-way ANOVA (Analysis Of Variance) with post-hoc Dunnett's test (for comparison of >2 groups with one dataset serving as a control), or two-way ANOVA with Šidák test (for MTT assays). Analysis of variables with non-normal distribution was performed using non-parametric tests: Wilcoxon signed-rank test (for comparison of the pair-matched specimens), Mann–Whitney U test (for comparison of two independent groups), or Kruskal–Wallis test with Dunn's multiple comparisons test (for comparison of >2 groups). Correlations in gene expression were evaluated using the Spearman's coefficient. Survival data was analyzed using the log-rank test. Differences in the conventional tests were considered significant at p values < 0.05. Statistics were mostly calculated using GraphPad Prism 6 (GraphPad, USA).

Overlaps between gene sets in the RNA-seq profiling experiments were analyzed using the hypergeometric distribution calculator available at http://nemates.org/MA/progs/overlap_stats.html. In this test the total number of genes (coding and non-coding) was 57,500 (i.e., the number of expressed genes detected with RNA-sequencing and analyzed with DESeq2).

CHAPTER IV: RESULTS

4.1. Analysis of clinical material

4.1.1 Examination of public RNA-seq datasets

Expression of *DIRC3* in thyroid tissue was evaluated in publicly available RNA-sequencing datasets. Normal thyroid tissue presented one of the highest levels of *DIRC3* expression among all tissue types analyzed in the GTEx project (Figure 4A). Likewise, transcriptomic data generated in The Cancer Genome Atlas (TCGA) project revealed a relatively high expression of *DIRC3* in PTCs (Figure 4B). TCGA data also indicated a trend for the downregulation of *DIRC3* in PTCs in comparison to unmatched normal thyroid tissues. This data contradicted a recent study that failed to detect *DIRC3* transcripts in PTCs and in cancer-free thyroid tissue [81].

Furthermore, I used TCGA data to evaluate the impact of *DIRC3* expression on clinical outcomes. The user-curated dataset that was created during the analysis can be accessed here: https://www.cbioportal.org/study/summary?id=thca_tcga_pan_can_atlas_2018#sharedGroups=6187b26ef8f71021ce56ebdd,6187b29bf8f71021ce56ebde

Expression heatmaps generated from TCGA data suggested that higher levels of *DIRC3* could associate with a lower incidence of thyroid cancer recurrence (Figure 5A). Using a best fitting threshold of *DIRC3* expression to discriminate the disease-free status, PTCs were classified as either *DIRC3-high* or *DIRC3-low* tumors. Survival plots generated for both groups indicated that elevated expression of *DIRC3* associated with a significantly higher disease free-survival rate (5-year DFS of 100% and 87% in patients classified as having *DIRC3-high* and *DIRC3-low* PTCs, respectively; Figure 5B). No significant interaction was detected for overall survival (Figure 5C). Furthermore, the histological make-up of the *DIRC3-high* and *DIRC3-low* groups was significantly different (Figure 5D). *DIRC3-high* carcinomas were enriched for conventional PTCs (81.9% and 67.3% in the *DIRC3-high* and *DIRC3-low* groups, respectively), but depleted of the follicular (12.6% and 23.1%, respectively) and tall cell variants (2.4% and 8.9%, respectively). Likewise, prevalence of tumor-driving mutations was dissimilar in the expression groups (Figure 5E). These differences appeared to reflect distinct mutational landscapes observed in the histological variants of PTC. Accordingly, *BRAF* and *RET* mutations

were more prevalent in the *DIRC3-high* PTCs, while *NRAS* and *HRAS* mutations were relatively more common in the *DIRC3-low* tumors.

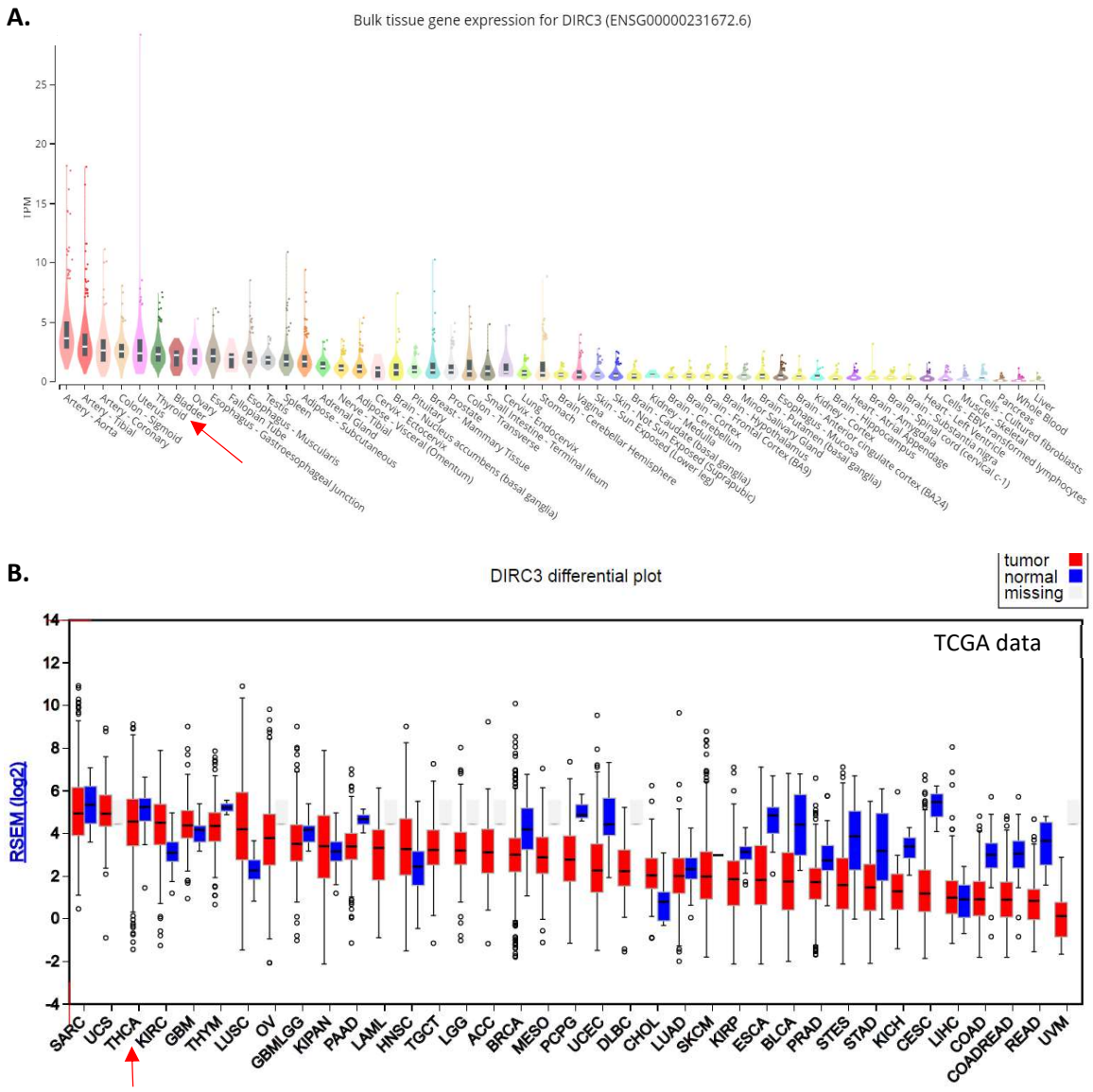
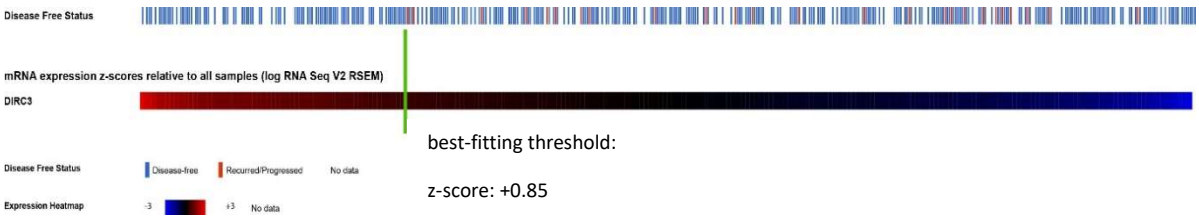


Figure 4. *DIRC3* expression in various normal and malignant tissues.

(A). Expression of *DIRC3* in normal tissues as profiled in the GTEx project (https://gtexportal.org/home/). Results for thyroid are marked with a red arrow. **(B).** Expression of *DIRC3* in malignancies profiled in the TCGA project. Data for PTCs (THCA, thyroid cancers) is marked with a red arrow. Graph was generated in the Firehose portal (http://firebrowse.org/; accessed May 2018). Abbreviations: TPM, transcripts per million; RSEM, RNA-Seq by Expectation-Maximization units.

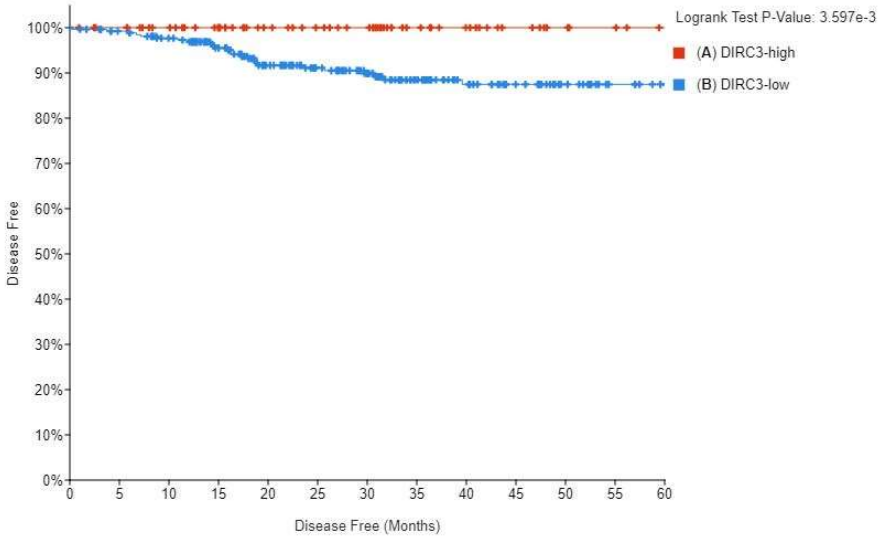
A.



best-fitting threshold:

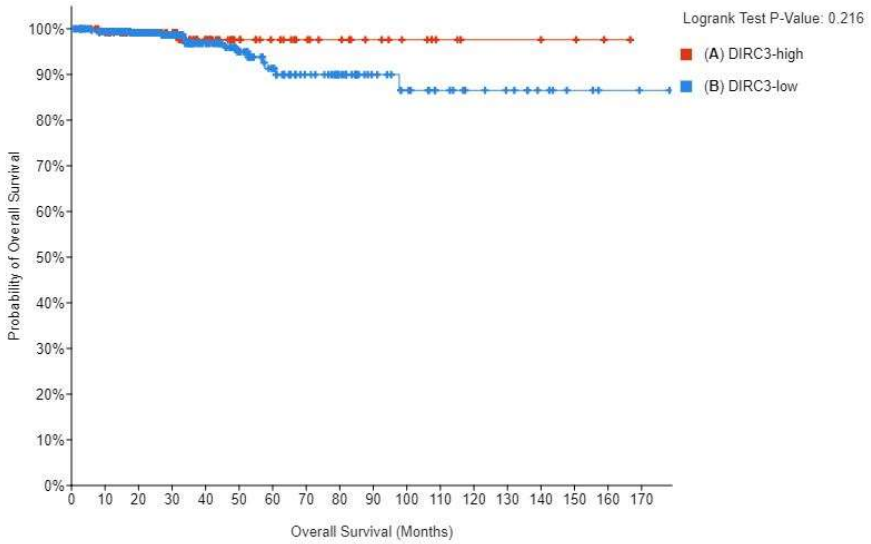
z-score: +0.85

B.



	Number of Cases, Total	Number of Events	Median Months Disease Free (95% CI)
(A) DIRC3-high	**88**	**0**	**NA**
(B) DIRC3-low	**266**	**25**	**NA**

C.



	Number of Cases, Total	Number of Events	Median Months Overall (95% CI)
(A) DIRC3-high	**126**	**2**	**NA**
(B) DIRC3-low	**373**	**14**	**NA**

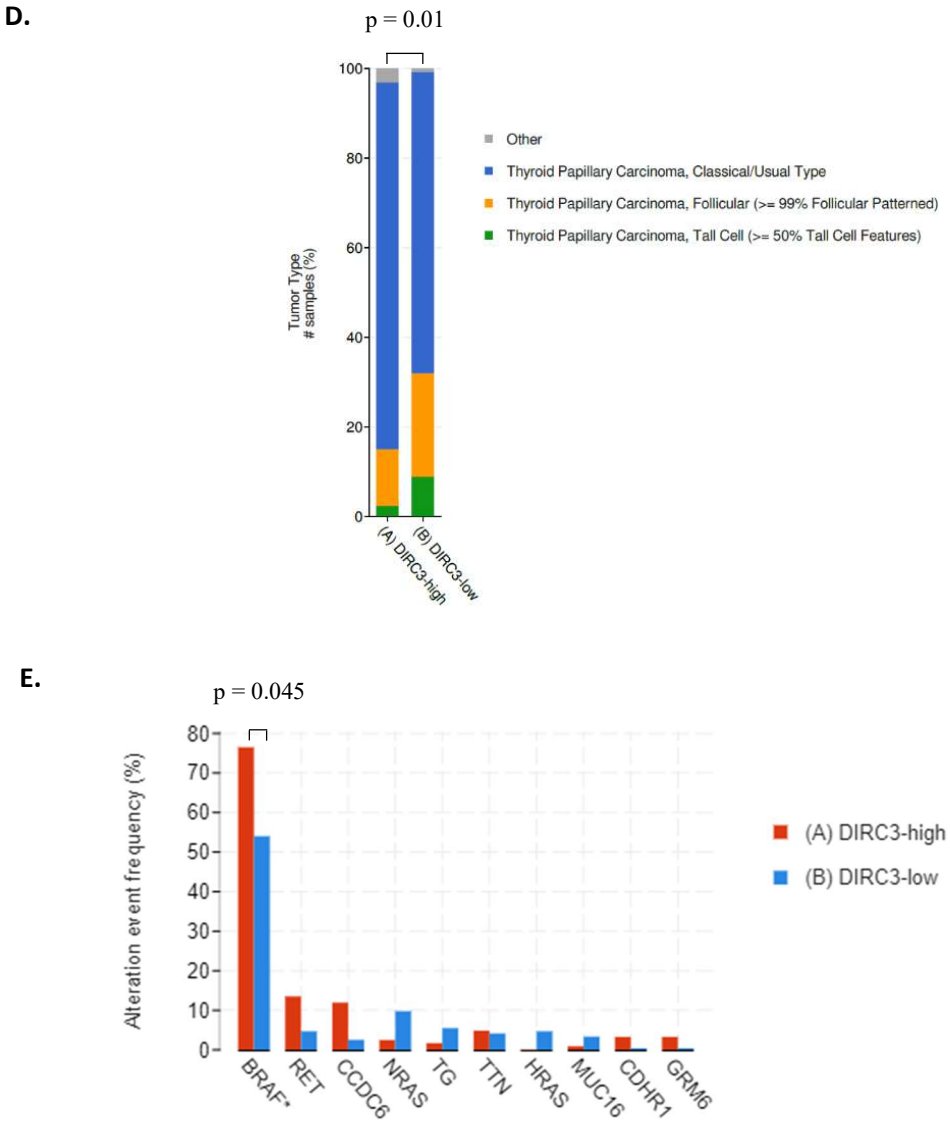


Figure 5. *DIRC3* expression in PTCs in TCGA dataset.

(A). Oncoprint graph illustrating the relationship between expression of *DIRC3* and the disease-free status in PTCs. The best fitting threshold of gene expression used to discriminate the disease-free status is marked with a green line. Accordingly, PTCs were classified as *DIRC3*-high or *DIRC3*-low cancers. **(B).** Disease-free survival of PTC patients according to the *DIRC3* expression status (log-rank test). **(C).** Overall survival of PTC patients according to the *DIRC3* expression status (log-rank test). **(D).** Histological composition of PTCs in the *DIRC3*-high and *DIRC3*-low groups (tested with Kruskal-Wallis test with Benjamini-Hochberg correction) **(E).** Frequency of gene mutations in the *DIRC3*-high and *DIRC3*-low PTCs (one-sided Fisher Exact test with Benjamini-Hochberg correction).

Graphs were prepared using cBioPortal (https://www.cbioportal.org/).

To get insights into a possible transcriptomic function of *DIRC3*, TCGA data was used to identify correlations between expression of *DIRC3* and the expression levels of other genes (Table 8). The analysis revealed that *DIRC3* is strongly co-expressed with *IGFBP5* (Spearman coefficient = 0.79), a gene located approximately 570,000 base pairs downstream from *DIRC3* (Table 8 and Figure 6A). Noteworthy, PTCs also demonstrate the highest level of *IGFBP5* expression among all cancer types analyzed in the TCGA project (Figure 6B).

Gene	Cytoband	Spearman's correlation	p-value
IGFBP5	2q35	0.79	1.61E-103
DMD	Xp21.2-p21.1	0.64	2.21E-56
LCN6	9q34.3	0.59	2.82E-46
QSOX2	9q34.3	0.57	2.70E-43
FIGNL2	12q13.13	0.56	3.86E-41
ESRRA	11q13.1	-0.56	5.49E-41
WSB2	12q24.23	-0.56	1.31E-40
CDK5RAP2	9q33.2	0.55	4.31E-40
ADTRP	6p24.1	0.55	1.04E-39
SCNN1G	16p12.2	0.55	3.39E-39
PTK7	6p21.1	0.55	6.37E-39
PPM1H	12q14.1-q14.2	-0.55	1.41E-38
USP27X	Xp11.23	0.54	2.04E-38
ZFTA	11q13.1	0.54	4.53E-38
INMT	7p14.3	0.54	5.91E-38
LCN8	9q34.3	0.54	6.56E-38
PLAAT1	3q29	-0.54	8.20E-38

Table 8. List of genes demonstrating the strongest expression correlations with *DIRC3* in PTCs (TCGA data).

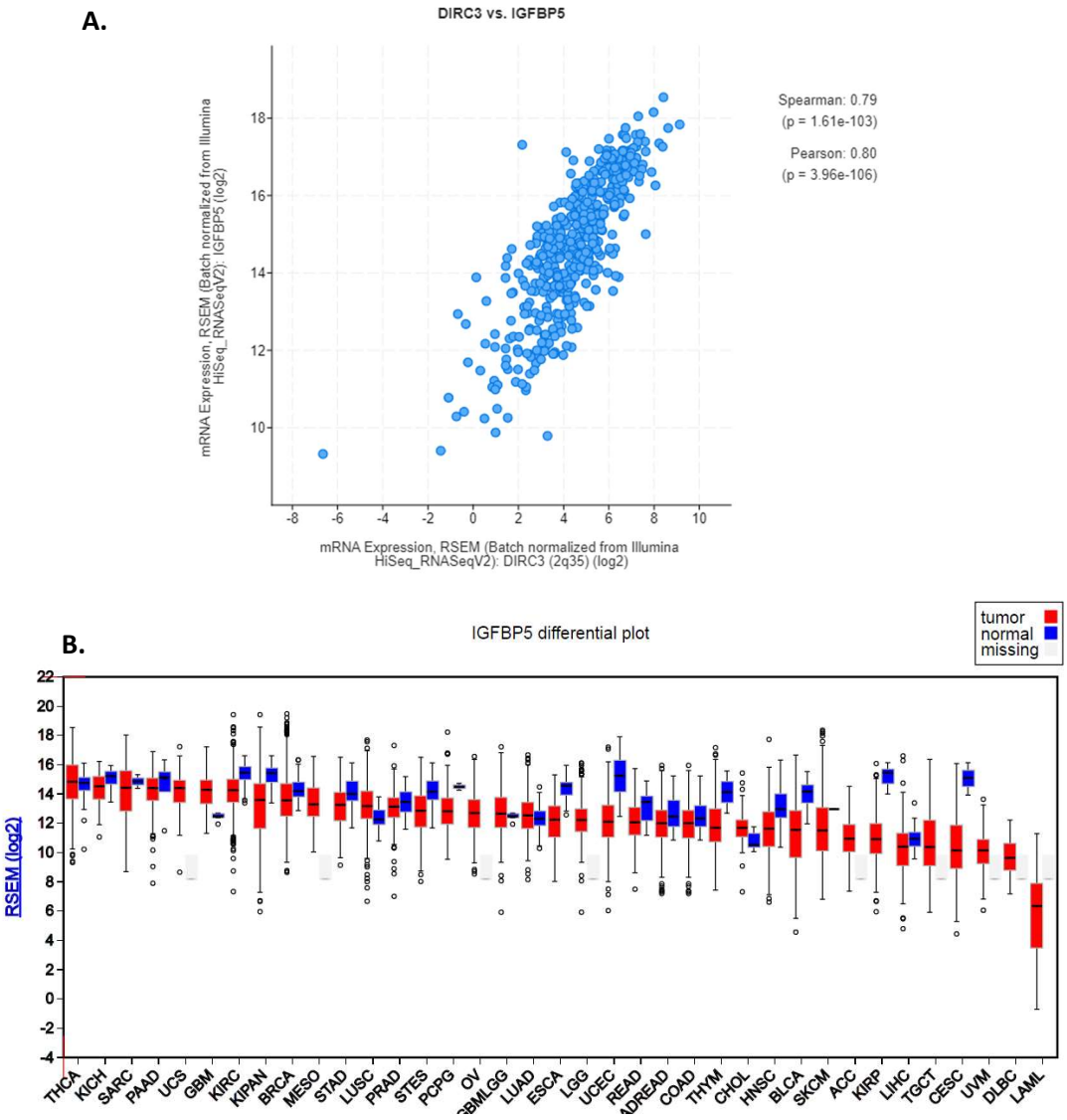


Figure 6. *DIRC3* was co-expressed with *IGFBP5* in PTCs in TCGA data.

(A). Correlation between expression of *DIRC3* and *IGFBP5* in PTCs. Graph was prepared using cBioPortal (https://www.cbioportal.org/). **(B).** Expression of *IGFBP5* in malignancies profiled in the TCGA project. Data for PTCs is shown as THCA. Graph was generated in the Firehose portal (http://firebrowse.org/; accessed May 2018). Abbreviations: TPM, transcripts per million; RSEM, RNA-Seq by Expectation-Maximization.

4.1.2 Analysis of clinical material obtained from DTC patients

Clinical material used in this study consisted of RNA and germline DNA samples obtained from 67 patients who underwent a total or near-total thyroidectomy due to the diagnosis of DTCs. Total RNA was isolated from malignant tumors and histologically normal thyroid tissue

in each patient (normal tissue was usually collected from the opposite, histologically unaffected thyroid lobe). Clinical and pathological data was available for most patients (Table 9).

Sex	Female, n = 55 Male, n = 10 data missing, n = 2
Age at the time of diagnosis	mean 48.1 years (standard deviation: 14.5 y); range 18-80 years
Histological type	PTC, conventional variant, n = 47 PTC, follicular variant, n = 9 FTC, n = 5 FTC + follicular variant of PTC, n = 1 Hürthle cell cancer, n = 3 data missing (DTC), n = 2
Multifocality	Yes, n = 15 No, n = 50 data missing, n = 2
Angioinvasion	Yes, n = 14 No, n = 33 data missing, n = 20
Invasion of thyroid capsule	Yes, n = 23 No, n = 38 data missing, n = 7
T stage **(AJCC/TNM 7th Edition)**	pT1a, n = 25 pT1b, n = 19 pT2, n = 6 pT3, n = 15 data missing, n = 2
N stage **(nodal metastases)**	N0, n = 44 N1, n = 13 data missing, n = 10
M stage **(distant metastases)**	M0, n = 53 M1, n = 4 data missing, n = 10

Table 9. Clinical and pathological information on DTC patients included in the study cohort.

Expression of *DIRC3* and *IGFBP5* was analyzed in the patient-matched DTC and normal thyroid tissue pairs. *DIRC3* was significantly downregulated in DTCs as compared to the normal thyroid specimens (Figure 7A). Moreover, the strong co-expression of *DIRC3* and *IGFBP5* (first identified in the TCGA data) was confirmed in our clinical material (Figure 7B). This suggested that *DIRC3* could play some role in the transcriptomic regulation of *IGFBP5*. No significant alterations in expression of *IGFBP5* were observed in DTCs (Figure 7C).

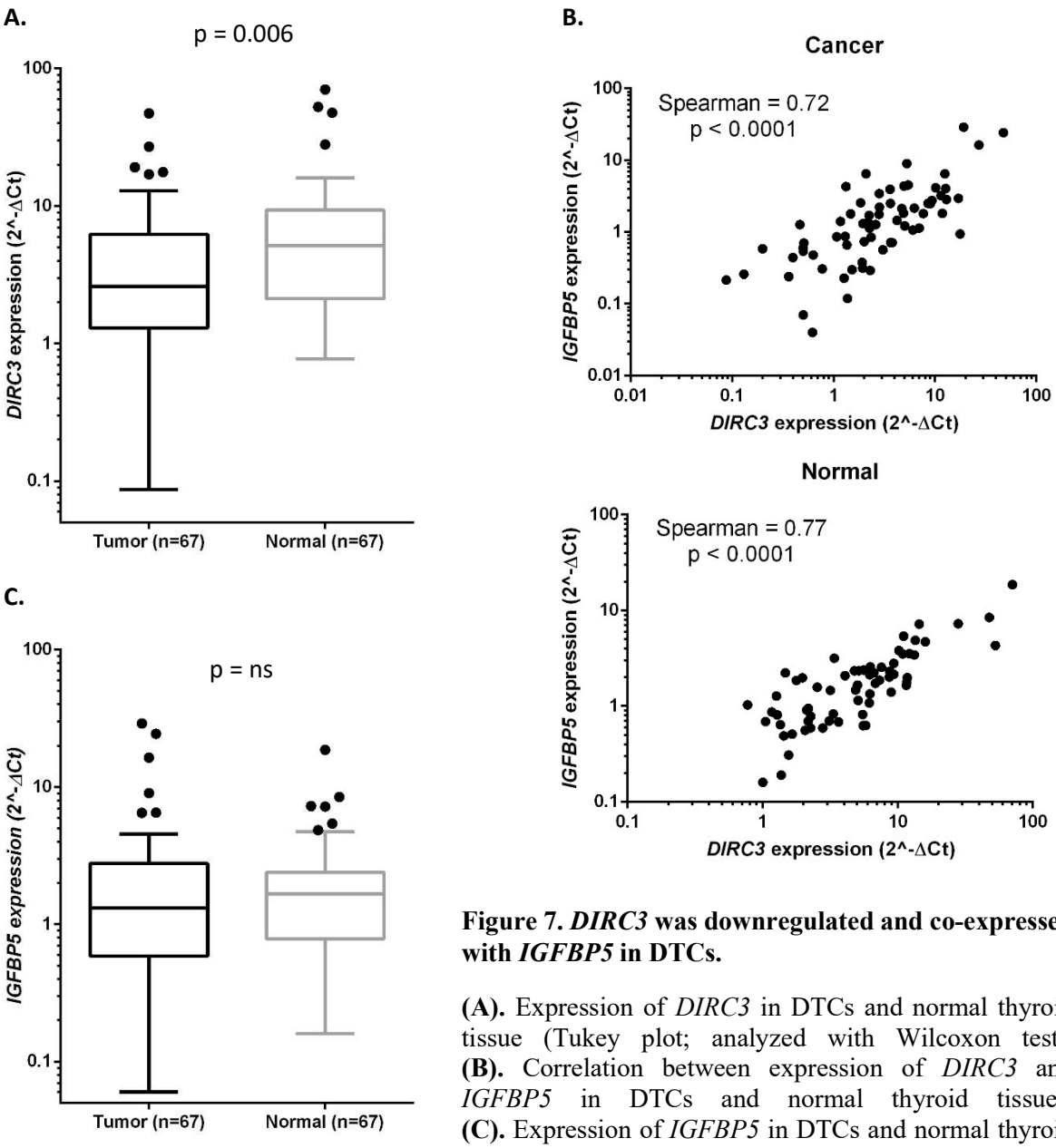


Figure 7. *DIRC3* was downregulated and co-expressed with *IGFBP5* in DTCs.

(A). Expression of *DIRC3* in DTCs and normal thyroid tissue (Tukey plot; analyzed with Wilcoxon test). **(B).** Correlation between expression of *DIRC3* and *IGFBP5* in DTCs and normal thyroid tissues. **(C).** Expression of *IGFBP5* in DTCs and normal thyroid tissue (Tukey plot; analyzed with Wilcoxon test).

Next, the gene expression profile of cancers was analyzed in the context of patients' clinicopathological data. Expression of *DIRC3* was not related to the histological type of DTC, invasion of thyroid capsule, presence of node metastasis, or the incidence of vascular invasion (Figure 8A, 8B, 8C, 8D). On the other hand, cancers that presented distant metastasis were found to express lower levels of *DIRC3* (Figure 8E). Any conclusions must be made with caution as the number of metastatic patients in the cohort was small. Separate analyses were performed exclusively for the conventional PTCs (cPTCs). Observations made in the entire cohort held true for cPTCs (Figure 9). The relationship between gene expression and distant

metastasis was, however, not tested, since the number of cPTCs with distant spread was insufficient. No separate clinicopathological evaluations were attainable for other histological cancer types due to their limited representation in the material.

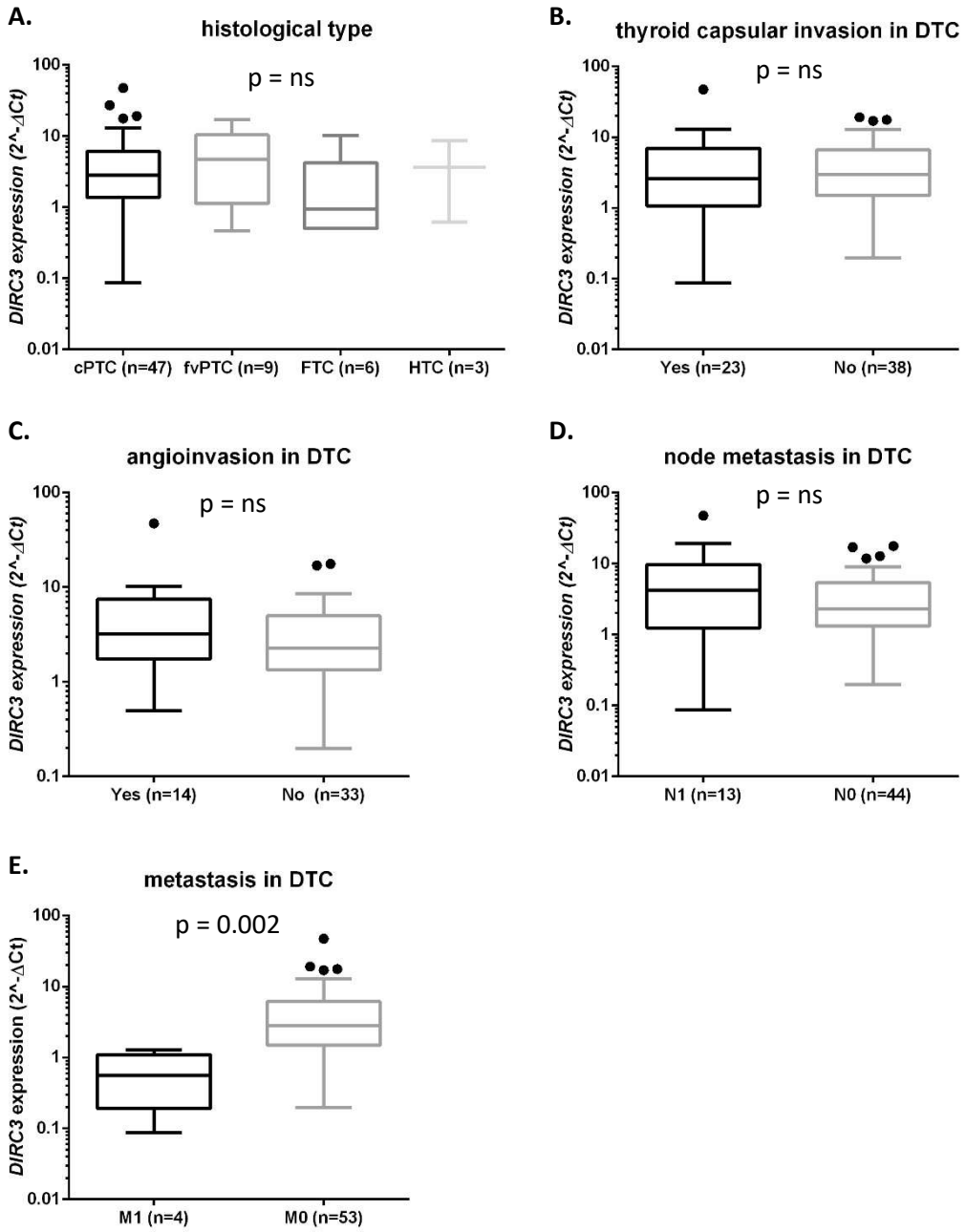


Figure 8. *DIRC3* expression and clinicopathological features of DTCs.

(A). Relationship between *DIRC3* expression and the histological type of DTC (cPTC: conventional PTC; fvPTC follicular variant PTC; FTC: follicular thyroid cancer; HTC: Hürthle cell thyroid carcinoma). Tested with Kruskal-Wallis test. **(B).** Expression of *DIRC3* in DTCs stratified by the presence of capsular invasion, **(C)** angioinvasion, **(D)** lymph node metastasis, and **(E)** distant metastasis. Results were analyzed with Mann–Whitney U test. Graphs are Tukey plots.

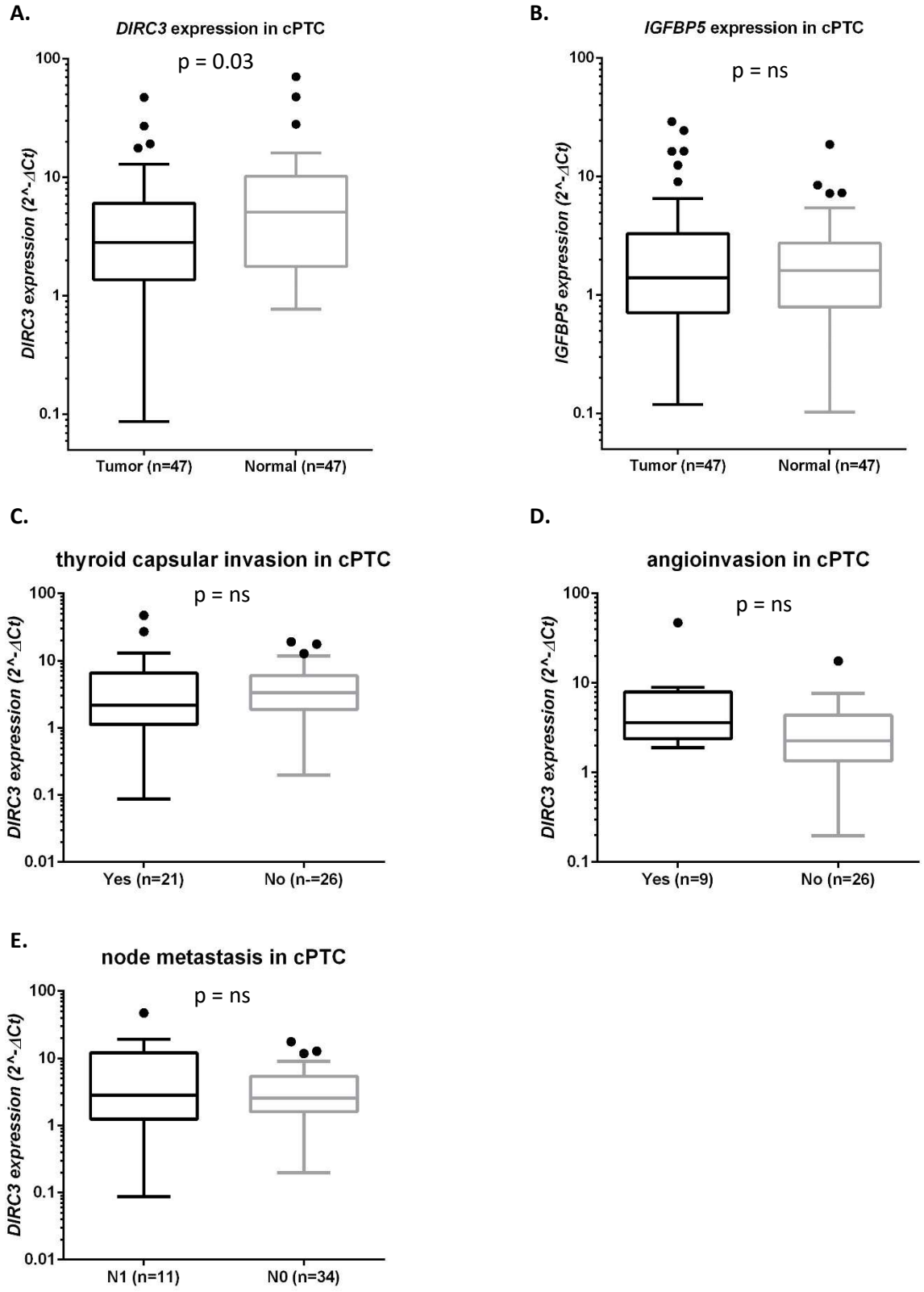


Figure 9. *DIRC3* expression and clinicopathological features of conventional PTCs.

(A). Expression of *DIRC3* in conventional PTCs (cPTCs) and the patient-matched normal thyroid tissue. Tested with Wilcoxon test. **(B).** Expression of *IGFBP5* in cPTCs and patient-matched normal thyroid tissue. Tested with Wilcoxon test. **(C).** Expression of *DIRC3* in cPTCs stratified by the presence of capsular invasion, **(D)** angioinvasion, and **(E)** lymph node metastasis. Analyzed with Mann–Whitney U test. Graphs are Tukey plots.

Additionally, germline DNA was genotyped using a custom PCR assay probing rs11693806 (Figure 10A). Genotypes observed in the study cohort matched very closely the genotypes previously described in European populations (Table 10). Importantly, no associations between the rs11693806 genotype and expression of *DIRC3* or *IGFBP5* in carcinomas were detected (Figure 10B and 10C).

rs11693806 genotype	Study cohort	European population
G/G	35 (52.2%)	54.4%
G/C	27 (40.3%)	38.7%
C/C	5 (7.5%)	6.9%

Table 10. Germline genotypes of rs11693806 in the study cohort and in the European population. Allelic frequencies for the European population were obtained from the Genome Aggregation Database (gnomAD, Broad Institute, USA). Data was retrieved from https://www.ncbi.nlm.nih.gov/snp/rs11693806 (accessed 26/10/2021).

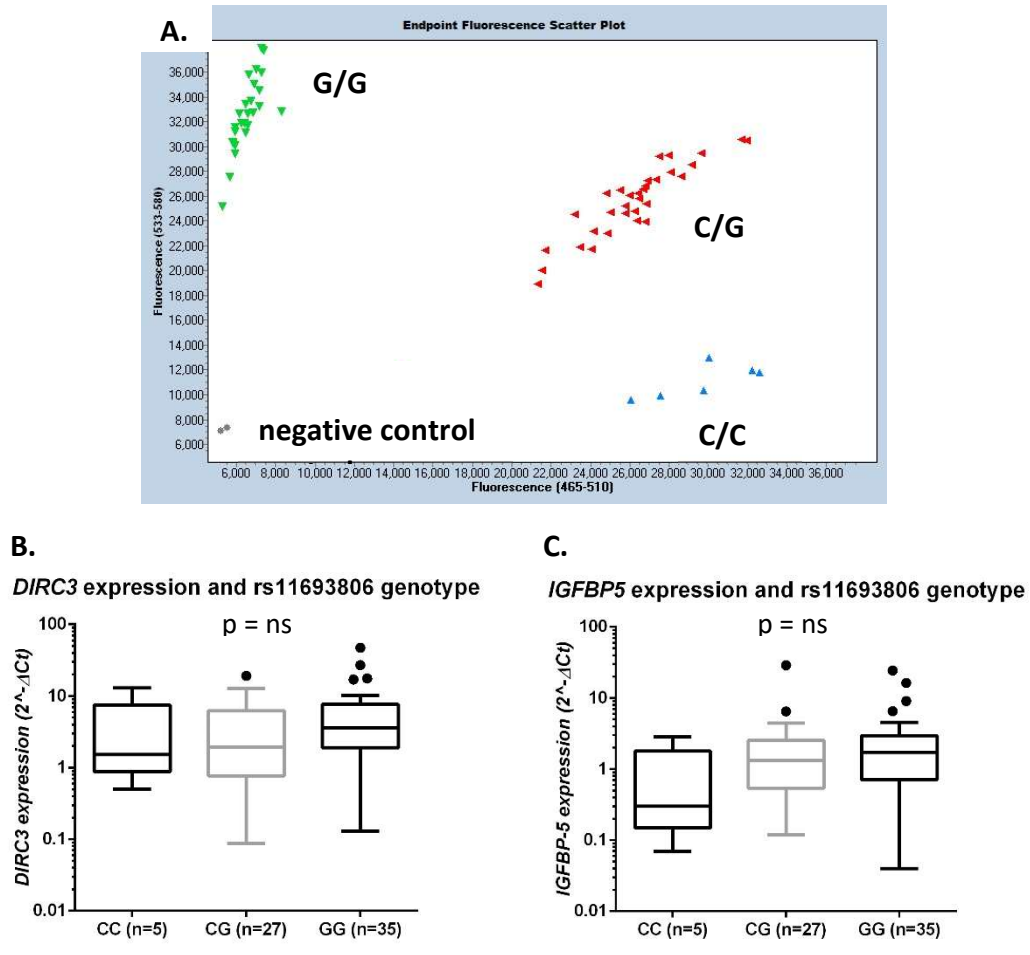


Figure 10. Germline genotype of rs11693806 did not influence expression of *DIRC3* and *IGFBP5* in clinical DTC samples.

(A). Representative results of the rhAmp PCR genotyping assay (germline DNA with appropriate controls). **(B).** Relationships between the rs11693806 germline genotype and expression of *DIRC3*, and **(C)** *IGFBP5* in DTCs. Results were analyzed with Kruskal-Wallis test. Graphs are Tukey plots.

4.2. Functional characterization of *DIRC3* lncRNA *in vitro*

4.2.1. Evaluation of *DIRC3* expression in cancer cell lines

Analysis of the GTEx data revealed that *DIRC3* has four main splice variants. Only two splice variants are prominently expressed across all tissue types evaluated in GTEx. Expression data for thyroid tissue is shown in Figure 11. The longer splice variant is annotated as ENST00000474063.5, and the shorter splice variant is termed ENST00000484635.1. For the sake of simplicity, they are called here *DIRC3-202* and *DIRC3-203*, respectively (following the naming convention used by Ensembl). Two other annotated splice variants do not exhibit expression in the GTEx data for thyroid: ENST00000486365.5 (*DIRC3-204*) and ENST00000423123.1 (*DIRC3-201*). Only *DIRC3-202*, *DIRC3-203* and *DIRC3-204* overlap the previously described germline thyroid cancer risk variants (including rs11693806 that was evaluated in this study).

Accordingly, expression of *DIRC3* in cancer cell lines was profiled for all detectable splice variants (the total expression), or selectively for the particular splice variants (*DIRC3-202*, *DIRC3-203* or *DIRC3-204*). Additionally, ASOs were designed to silence all expressed splice variants simultaneously (using either GapmeR anti-*DIRC3*-common_2 or GapmeR anti-*DIRC3*-common_3), or to silence the longer splice variant selectively (using GapmeR anti-*DIRC3-202*). Locations of qPCR amplicons and the GapmeR target sites are shown in Figure 11.

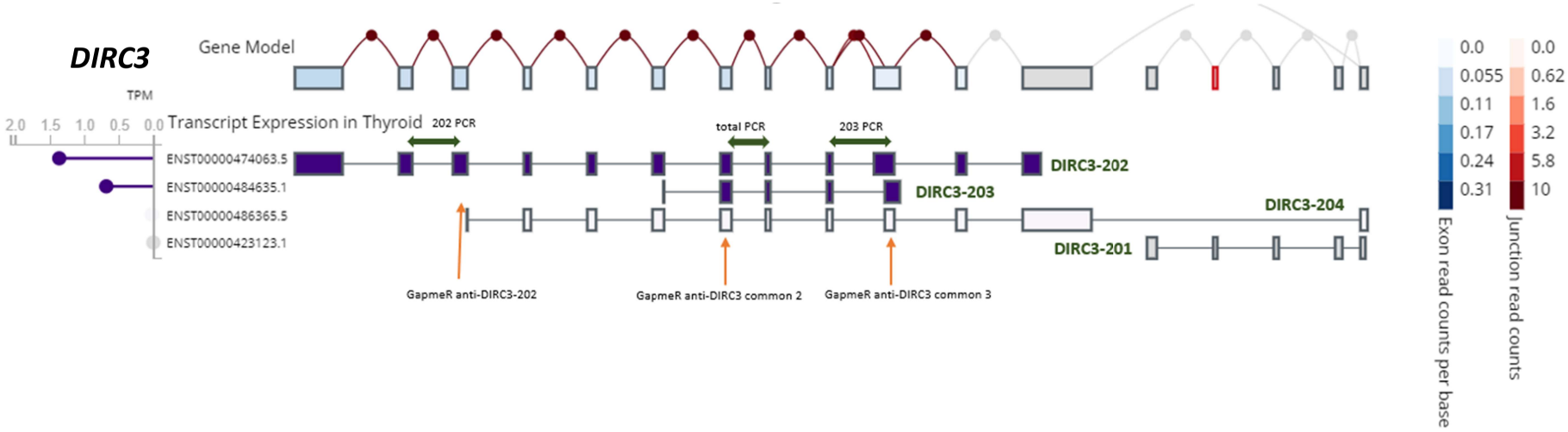


Figure 11. Structure of *DIRC3* and expression of its splice variants in normal thyroid tissue.

Data was retrieved from the GTEx project *via* the GTEx Portal (https://gtexportal.org/home/; accessed May 2018).

DIRC3 is expressed in thyroid in two splice variants: *DIRC3-202* and *DIRC3-203*. qPCR primers were designed to test gene expression (amplicons are shown as green horizontal arrows). Multiple ASOs were designed to silence *DIRC3* (marked as orange vertical arrows).

Expression of *DIRC3* and *IGFBP5* was evaluated in five cancer cell lines: MDA-T32, MDA-T68, MDA-T120, K1 and MCF-7 (Figure 12). The highest expression of *DIRC3* and *IGFBP5* was observed in MDA-T32 and MCF-7. *DIRC3* was not expressed in MDA-T68. Results also confirmed presence of *DIRC3-202* and *DIRC3-203* splice variants, and the absence of *DIRC3-204* in all tested cell lines (in concordance with the GTEx data).

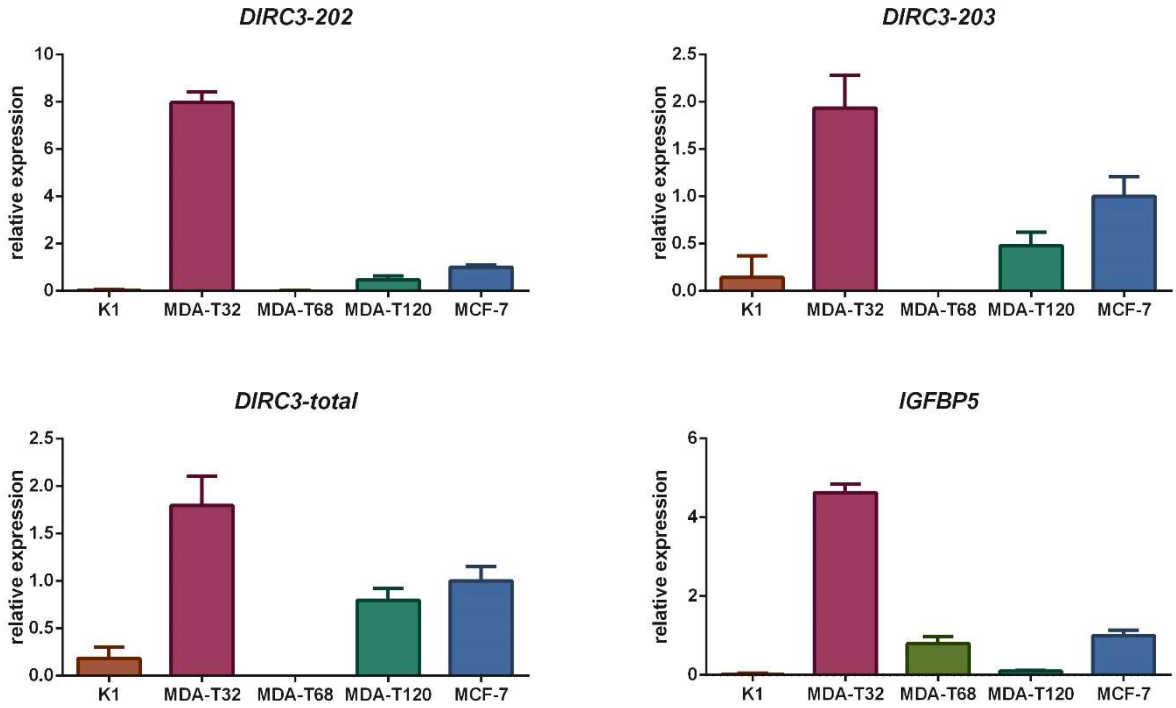
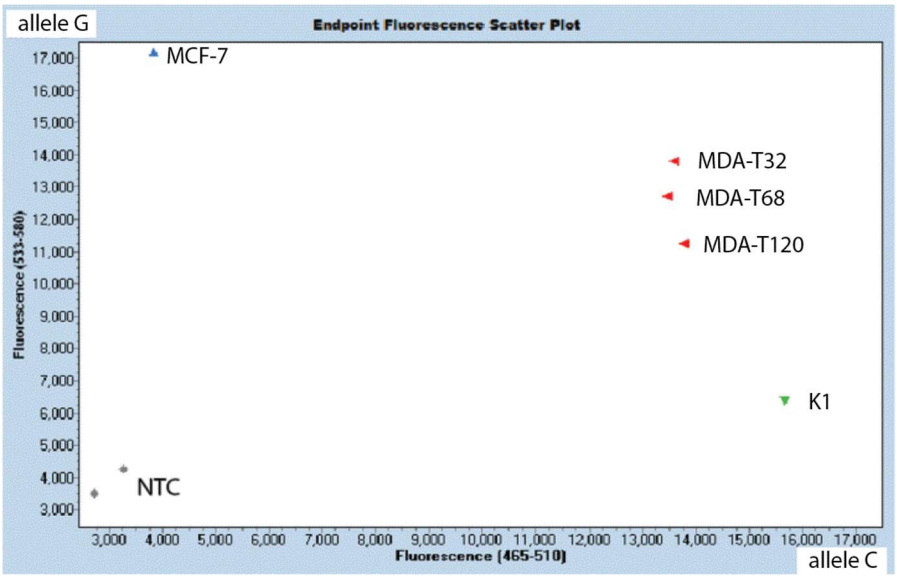


Figure 12. Expression of *DIRC3*, its splice variants and *IGFBP5* in a panel of cancer cell lines.

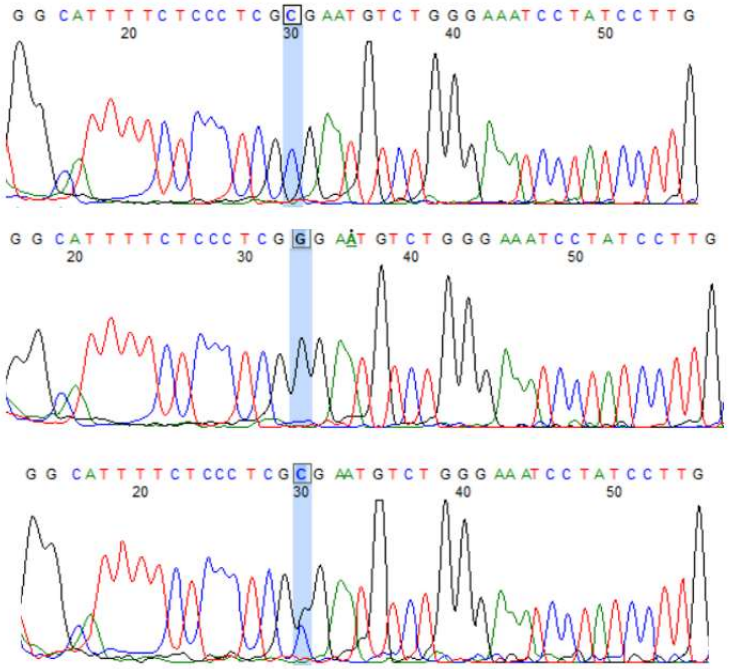
Results are shown as mean ± SD (n = 3). Gene expression in qRT-PCR was normalized using *HPRT1* as a housekeeping gene.

Additionally, rs11693806 was genotyped in cancer cell lines using the rhAmp SNP genotyping assay and Sanger sequencing (Figure 13). MDA-T32, MDA-T68, MDA-T120 were found to be the rs11693806[C/G] heterozygotes, K1 was identified as the rs11693806[C/C] homozygote, and MCF-7 was the rs11693806[G/G] homozygote.

A.



B.



K1: rs11693806[C/C]

MCF-7: rs11693806[G/G]

MDA-T32, MDA-T68, MDA-T120: rs11693806[C/G]

Figure 13. Rs11693806 genotypes of cancer cell lines.

(A). Results of the rhAmp PCR assay in cancer cell lines. **(B).** Sequencing chromatograms of purified PCR products. Rs11693806 is highlighted in blue.

4.2.2. Evaluation of the subcellular localization of *DIRC3* transcripts

Subcellular fractionation of RNA was performed to establish the localization of *DIRC3* transcripts within cells. LncRNAs may be broadly classified as transcripts enriched in the nucleus, cytoplasm, or present in both sites [145]. The subcellular location of *DIRC3* lncRNA was analyzed in two cell lines, MDA-T32 and MCF-7 (cell lines that demonstrated the highest expression of *DIRC3*). Two lncRNAs with well-established subcellular localizations were used as controls: *MALAT1* (a nuclear lncRNA) and *SNHG5* (a cytoplasmic lncRNA) [145-148, 181]. The experiment indicated that *DIRC3* transcripts were enriched in the nuclear fraction of RNA in both analyzed cell lines (Figure 14).

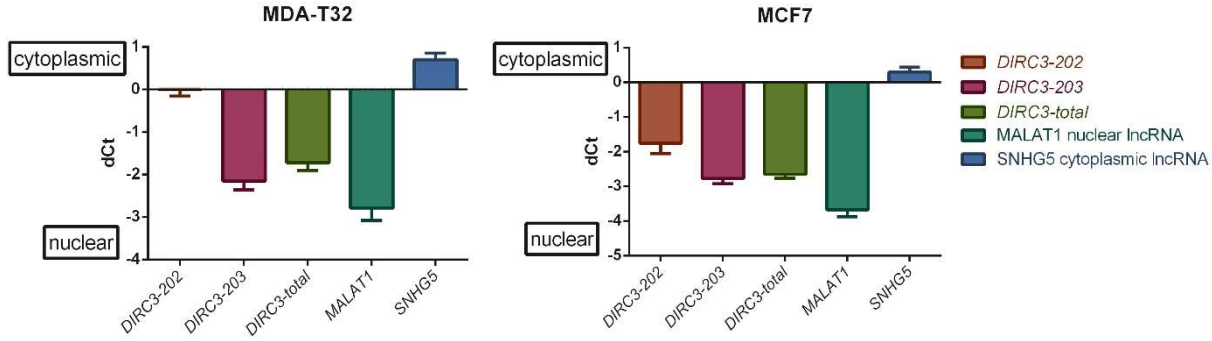


Figure 14. *DIRC3* transcripts were enriched in the nuclear fraction of RNA in MDA-T32 and MCF-7 cell lines.

MALAT1 and *SNHG5* were used as nuclear and cytoplasmic controls, respectively. Relative enrichment of transcripts in each fraction is expressed as ΔCt (the difference in Ct values obtained for a given transcript in the cytoplasmic and nuclear RNA fractions). Values shown are mean ± SD (n = 3).

4.2.3. Evaluation of the efficacy of *DIRC3* silencing using ASOs

ASOs were selected as the most suitable tool for downregulation of *DIRC3* due to the predominant nuclear localization of its products. Three *DIRC3*-targeting GapmeRs were validated and used to transfect cancer cell lines. Two GapmeRs (termed anti-*DIRC3*-common_2 and GapmeR anti-*DIRC3*-common_3) were designed to simultaneously silence both splice variants that were expressed in thyroid cancer cell lines. Moreover, one additional ASO (termed GapmeR anti-*DIRC3-202*) was designed to target only the longer splice variant (*DIRC3-202*). This was done to test whether selective targeting of *DIRC3* splice variants could produce dissimilar phenotypic effects.

Transfection of GapmeRs was performed in four cancer cell lines: MDA-T32 and MCF-7 (cell lines with a relatively high expression of *DIRC3* and *IGFBP5*), MDA-T120 (which had a relatively low expression of *DIRC3* and *IGFBP5*), and MDA-T68 (which had a moderate level of *IGFBP5*, but did not express *DIRC3*). Additionally, cell lines were transfected with a validated *MALAT1*-targeting GapmeR (a positive control). Successful silencing of *MALAT1* proved that transfections were efficient in all tested cell lines (Figure 15A). Next, *DIRC3*-targeting GapmeRs were tested in all cell lines (Figure 15B, 15C, 15D, 15E). qRT-PCR results demonstrated that the GapmeRs successfully downregulated their targets. Importantly, silencing of *DIRC3* was accompanied by the downregulation of *IGFBP5* in the cell lines that expressed *DIRC3* (MDA-T32, MDA-T120 and MCF-7). This was especially evident for the GapmeR anti-*DIRC3*-common_3 which efficiently silenced *IGFBP5* in three cell lines. Accordingly, GapmeR anti-*DIRC3*-common_3 was selected for subsequent phenotypic and transcriptomic studies. Interestingly, GapmeR anti-*DIRC3-202* successfully silenced its target (*DIRC3-202*), however this downregulation did not influence expression of *IGFBP5*. Finally, the *DIRC3*-targeting GapmeRs did not alter *IGFBP5* expression in MDA-T68, the only tested cell line without detectable expression of *DIRC3* (Figure 15E). This result proved that *IGFBP5* downregulation observed in other cell lines was directly related to the *DIRC3*-targeting capabilities of the utilized GapmeRs.

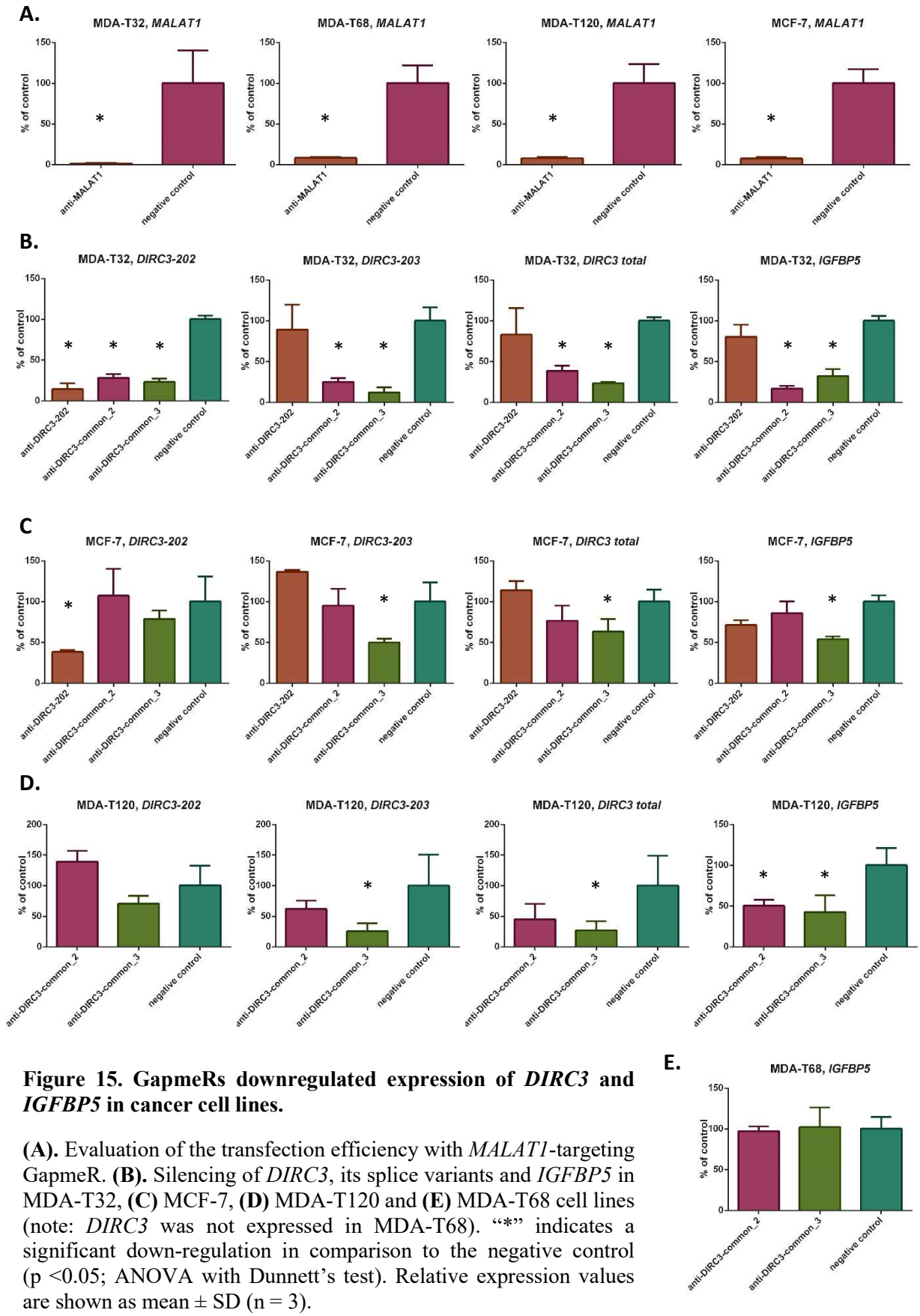


Figure 15. GapmeRs downregulated expression of *DIRC3* and *IGFBP5* in cancer cell lines.

(A). Evaluation of the transfection efficiency with *MALAT1*-targeting GapmeR. **(B).** Silencing of *DIRC3*, its splice variants and *IGFBP5* in MDA-T32, **(C)** MCF-7, **(D)** MDA-T120 and **(E)** MDA-T68 cell lines (note: *DIRC3* was not expressed in MDA-T68). "*" indicates a significant down-regulation in comparison to the negative control ($p < 0.05$; ANOVA with Dunnett's test). Relative expression values are shown as mean ± SD (n = 3).

4.2.4. Evaluation of the phenotypic impact of *DIRC3* silencing

Phenotypic effects of *DIRC3* silencing were evaluated in two PTC cell lines: MDA-T32 and MDA-T120. Five functional assays were applied:

(i) MTT assays to measure the cellular metabolic activity which was utilized to indirectly determine the number of viable cells in samples.
(ii) Transwell assays to quantify cell migration and chemotaxis,
(iii) Matrigel invasion chambers to evaluate cancer cell invasiveness,
(iv) Caspase 3/7 activity to measure the pro-apoptotic effect of serum- and L-glutamine-deprivation,
(v) Soft agar assays to measure the survival and growth capability of cancer cells unable to adhere to a solid surface (the anchorage independent-growth).

4.2.4.1. Effect of DIRC3 silencing in MTT assays

All three *DIRC3*-targeting GapmeRs were used in MTT assays. Cells were monitored for up to 96 hours after transfections. Downregulation of *DIRC3* modestly restrained the MTT reduction rate in MDA-T32 cells. A more prominent inhibitory effect of *DIRC3* silencing was observed in MDA-T120 cells (Figure 16).

Interestingly, no such effect was observed in MTT assays when *DIRC3-202* was silenced selectively. This outcome could indicate possible dissimilar biological and transcriptomic activities of *DIRC3* splice variants.

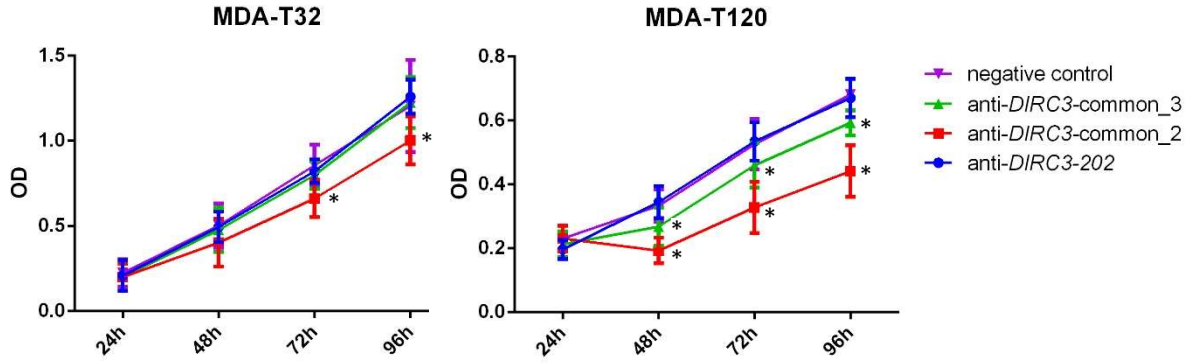
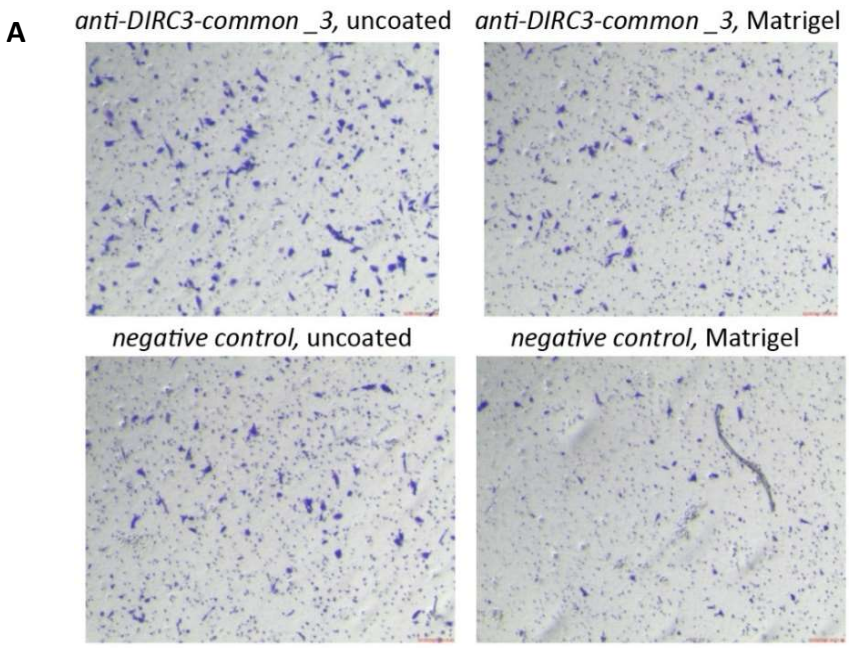


Figure 16. Silencing of *DIRC3* limited the MTT reduction rate in cancer cells.

"*" indicates significant downregulation vs. negative control ($p < 0.05$; two-way ANOVA with Šidák test). Each experiment was repeated three time with technical quadruplicates in each set-up. Results are shown as mean ± SD. Abbreviation: OD, optical density.

4.2.4.2. Effect of DIRC3 silencing on migration and invasiveness of cancer cells.

Transwell assays (utilizing uncoated or Matrigel-coated chambers) were used to measure migration, invasiveness and the IGF-1-induced chemotaxis of MDA-T32 and MDA-T120 cells. Results of these experiments indicated that downregulation of *DIRC3* markedly increased migration and invasiveness in both tested cell lines (Figure 17A and 17B). Similarly, MDA-T32 cells exhibited a potent increase in chemotaxis towards IGF-1 when transfected with the *DIRC3*-targeting GapmeR (Figure 17C and 17D). Use of IGF-1 alone was, however, insufficient to generate a chemotactic response in MDA-T120 cells.



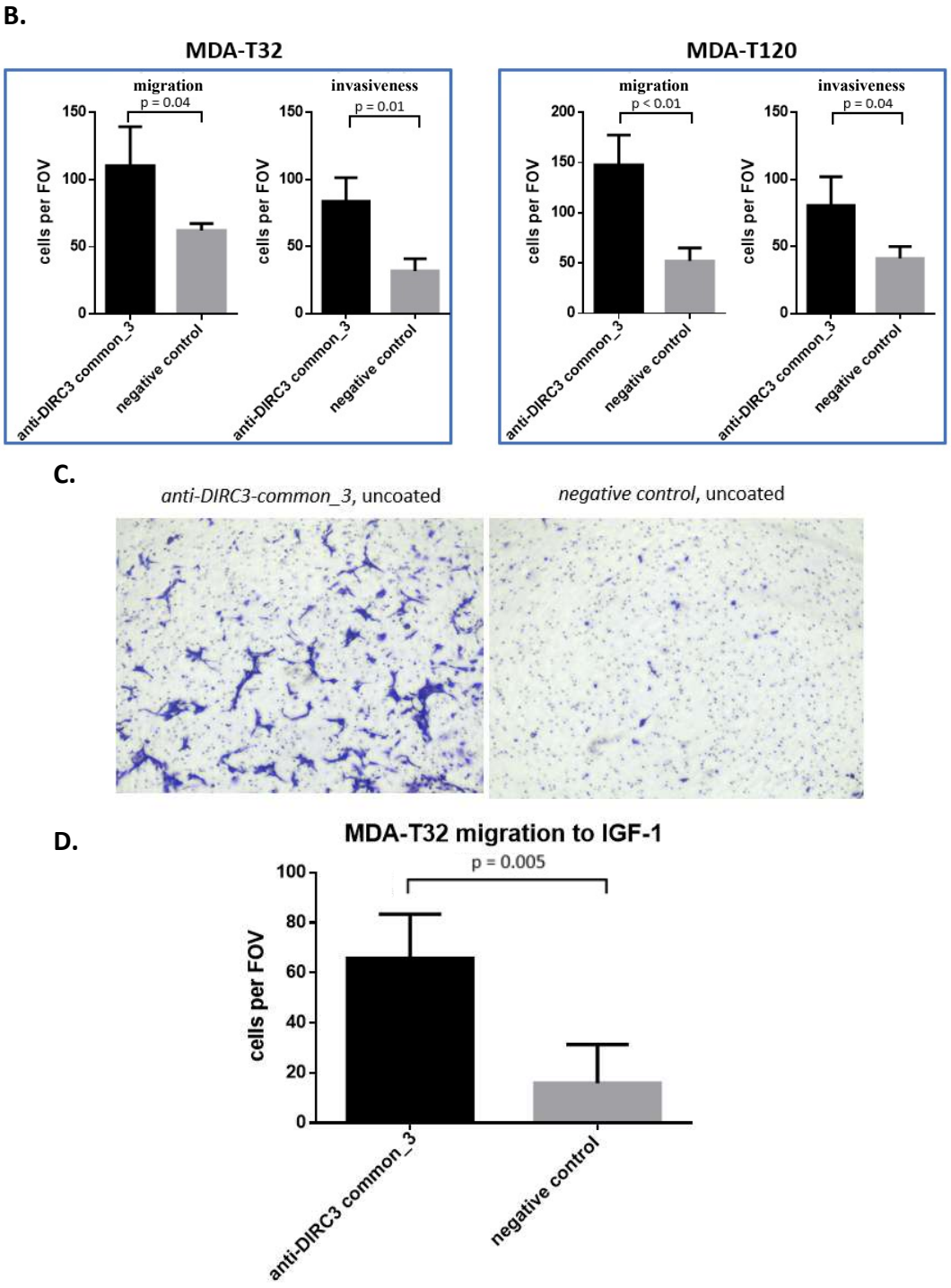


Figure 17. Downregulation of *DIRC3* increased migration and invasiveness of thyroid cancer cells.

(A). Representative images of MDA-T32 cells in Transwell assays (uncoated and Matrigel-coated chambers; chemoattraction with 10% FBS). **(B).** Quantification of migration and invasiveness of MDA-T32 and MDA-T120 cells transfected with GapmeRs (chemoattraction with 10% FBS). **(C).** Chemotactic effect of IGF-1 (100 ng/ml) in MDA-T32 cells transfected with GapmeRs. Representative images of MDA-T32 cells in Transwell assays (uncoated chambers). **(D).** Quantification of migration of MDA-T32 cells transfected with GapmeRs (chemoattraction with IGF-1). All experiments were repeated at least three times. Results are shown as mean ± SD (n = 3). Results were analyzed with t-test. Abbreviation: FOV, field of view.

4.2.4.3. Effect of DIRC3 silencing on the apoptosis of cancer cells

Serum- and L-glutamine-deprivation is capable to induce apoptosis in many cancer cell lines. Hence, I tested if alterations in expression of *DIRC3* could impact the apoptosis susceptibility in cancer cells that were deprived of serum and L-glutamine. The assay showed that silencing of *DIRC3* significantly reduced the caspase 3/7 activity in starved MDA-T120 cells. A similar trend (non-significant) was observed in MDA-T32 cells (Figure 18). Consequently, down-regulation of *DIRC3* appeared to have some anti-apoptotic influence on thyroid cancer cells.

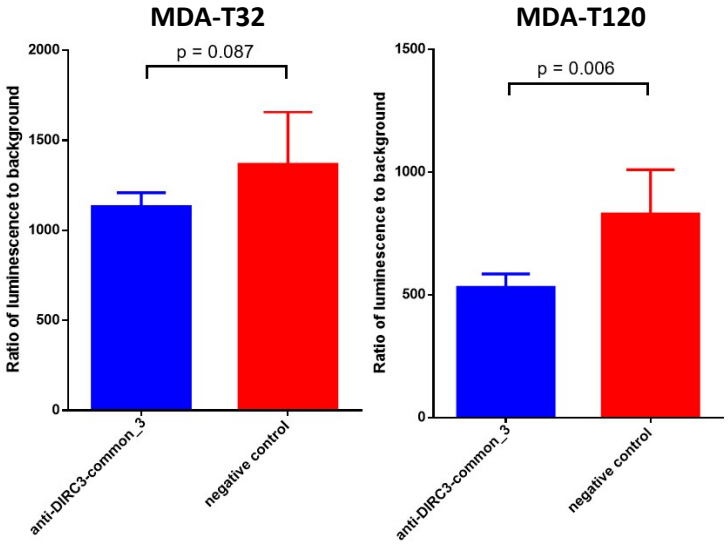


Figure 18. Silencing of *DIRC3* reduced the activity of caspase 3/7 in the serum- and glutamine-starved cancer cells.

Confluent cells were transfected with the *DIRC3*-targeting or negative control GapmeRs, cultured, and then starved for 24 hours. Next, the cells were analyzed with the Caspase Glo 3/7 assay. Experiments were repeated three times with technical duplicates in each assay (total n = 6). Results are shown as mean ± SD. Results were analyzed with t-test.

4.2.4.4. Evaluation of the effect of DIRC3 silencing on the anchorage-independent growth of cancer cells.

Anchorage-independent growth assays were performed in MDA-T32 and MDA-T120 cells that were transfected with GapmeRs. However, since MDA-T120 cell line could not generate a sufficient number of visible colonies, final results were obtained only for MDA-T32. Representative images are shown in Figure 19. The experiment indicated that silencing of *DIRC3* in MDA-T32 cells had no influence on their anchorage-independent growth capability.

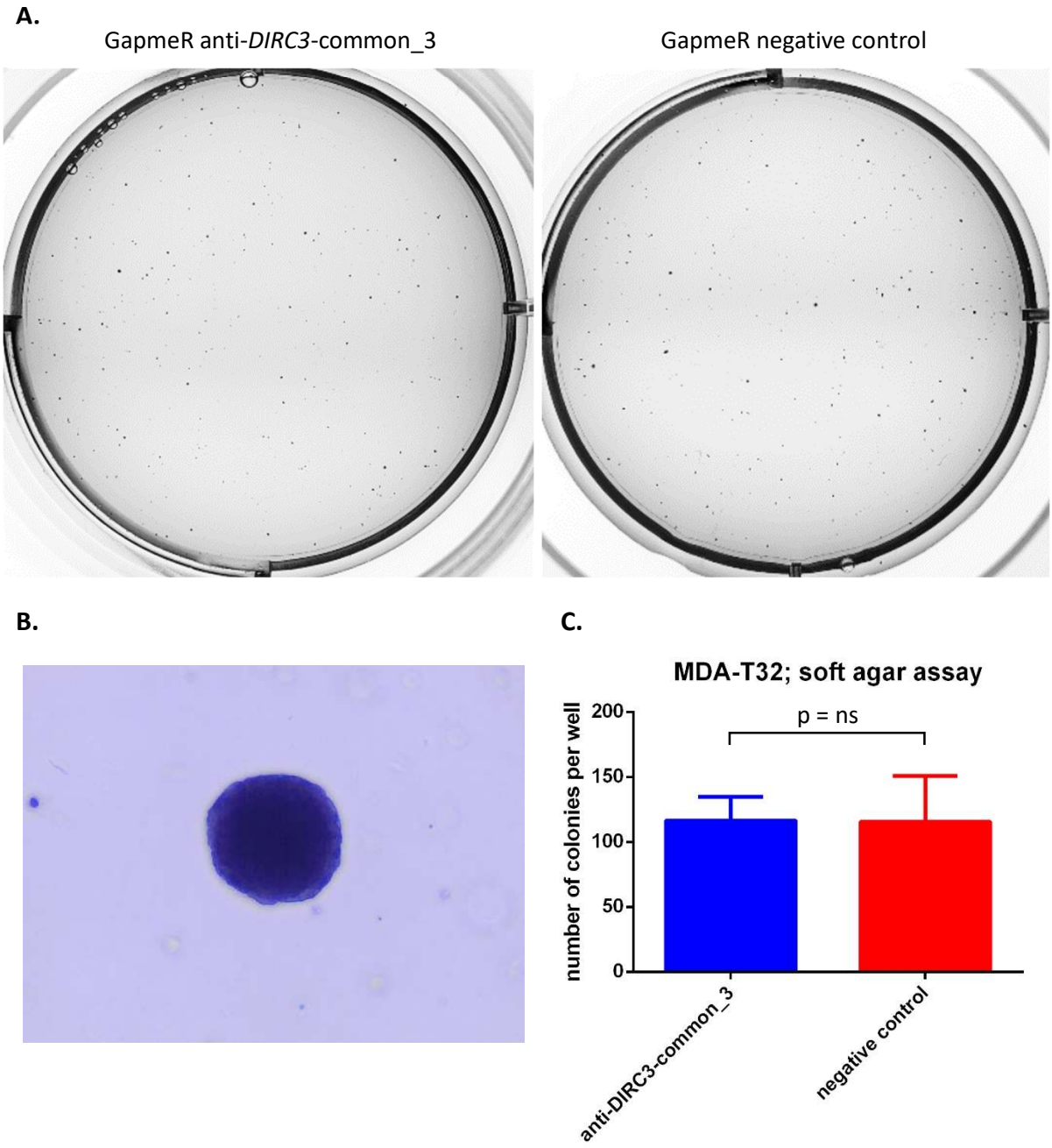


Figure 19. Anchorage-independent growth of MDA-T32 cells was not influenced by silencing of *DIRC3*.

(A). Representative wells showing colonies of MDA-T32 cells in soft agar. **(B).** A typical colony of MDA-T32 cells after 18 days of incubation in soft agar (microscopic image). **(C).** Number of visible colonies per well after transfection with the *DIRC3*-silencing and negative control GapmeRs. Colonies were counted using ImageJ. Results are presented as mean ± SD colonies per one well (three independent experiments with technical triplicates in each set-up, total n = 9). ns: not significant (t-test).

4.2.5. *IGFBP5* rescue experiments

Rescue experiments involving overexpression of *IGFBP5* in thyroid cancer cells were used to establish whether phenotypic effects induced by *DIRC3* silencing were related to the transcriptomic regulation of *IGFBP5*. For this purpose, two cell lines with stable transgene expression were generated: MDA-T32 cell line expressing pcDNA3-IGFBP5-V5 plasmid, and a second MDA-T32 derivative transfected with the empty pcDNA3 vector (a negative control). Positively transfected cells were selected with G418 at 600 μg/ml concentration.

Both cell line derivatives were transfected with the *DIRC3*-targeting GapmeRs. Results of the gene expression analysis (qRT-PCR) are shown in Figure 20. Introduction of pcDNA3-IGFBP5-V5 plasmid to MDA-T32 cells resulted in a strong (~135-fold) upregulation of *IGFBP5*. Transfection with GapmeR anti-*DIRC3*-common_3 significantly downregulated *DIRC3* in both modified cell lines. Yet, the GapmeR significantly downregulated *IGFBP5* only in the cells that expressed the control (empty) plasmid. No such effect was observed in the cells that overexpressed *IGFBP5* from the plasmid. This observation indicated that the *DIRC3*-targeting GapmeR could influence *IGFBP5* expression only when the transcripts were produced from the endogenous nuclear locus. Interestingly, pcDNA3-IGFBP5-V5 also upregulated *DIRC3* (~2.9-fold).

The plasmid-transfected MDA-T32 cell lines were tested in two phenotypic experiments: MTT and Transwell assays. The MTT conversion rate was significantly reduced after silencing of *DIRC3* in both plasmid-expressing cell line derivatives (Figure 21A). In contrast, overexpression of *IGFBP5* successfully negated the pro-migratory phenotype generated by *DIRC3* silencing (Figure 21B and 21C). Collectively, these results suggested that phenotypic effects produced by *DIRC3* could be either *IGFBP5*-dependent (in the case of cell migration), or at least partially independent from *IGFBP5* (in regard to the effects observed in MTT assays).

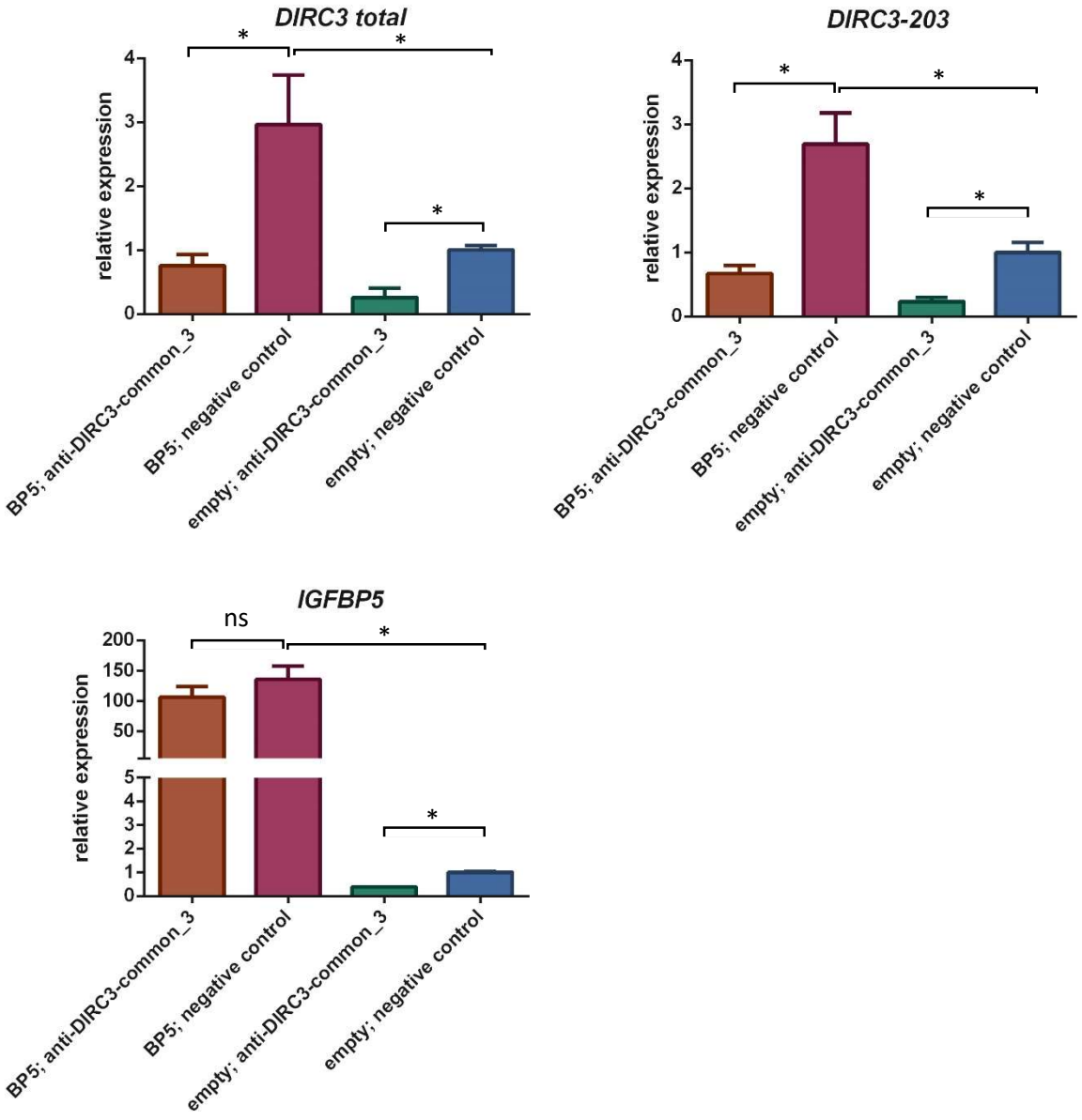


Figure 20. Silencing of *DIRC3* did not influence the level of *IGFBP5* expression from plasmid.

MDA-T32 cells were transfected with pcDNA3-IGFBP5-V5 (abbreviated as "BP5") or control ("empty") pcDNA3 plasmids. Stable cell lines were transfected with GapmeRs (anti-*DIRC3*-common_3 or negative control). Expression of *DIRC3* and *IGFBP5* was analyzed using qRT-PCR. Experiments were performed three times (n = 3). Results are presented as mean ± SD. "*" indicates $p < 0.05$ (t-test). ns: not significant.

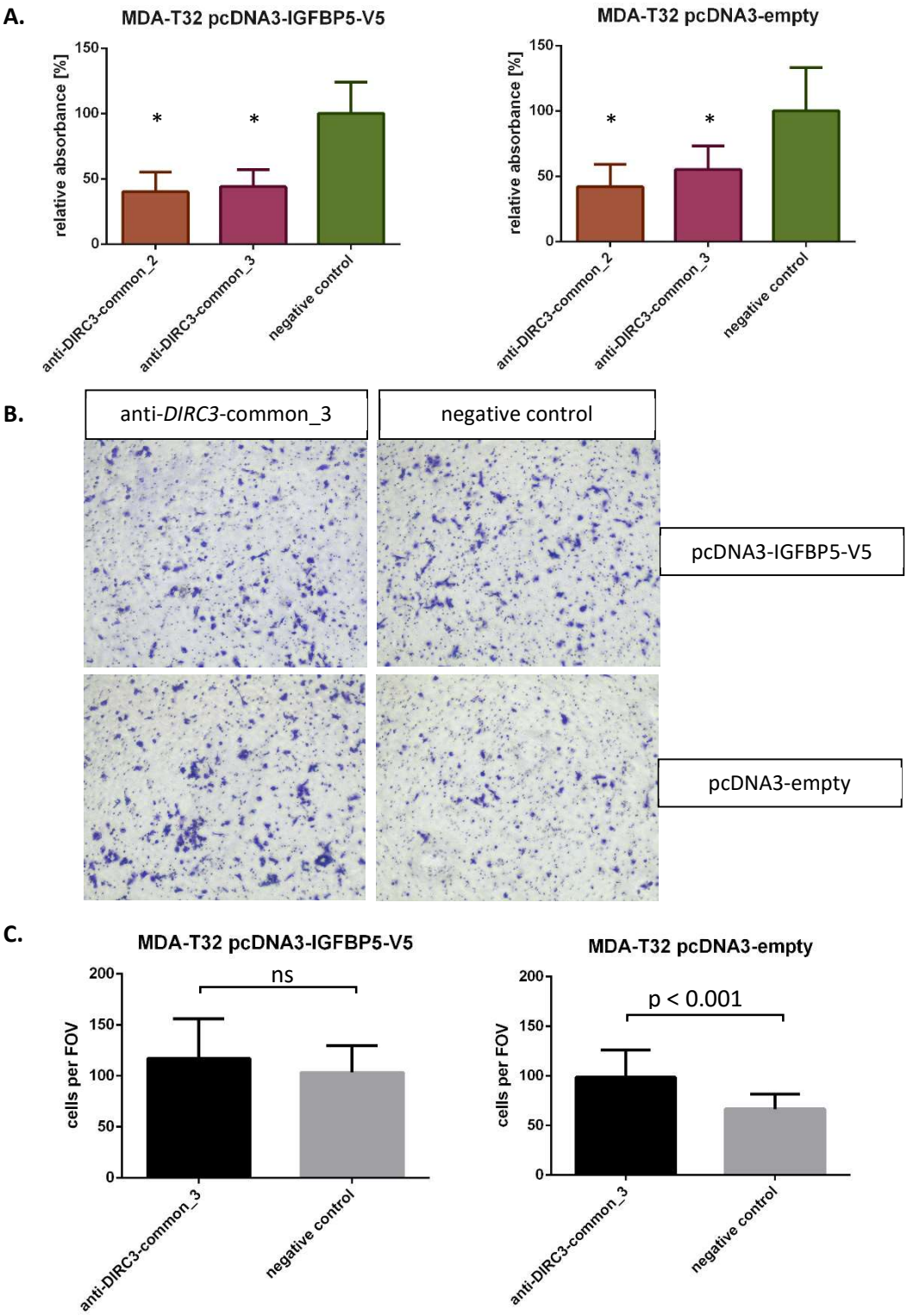


Figure 21. Silencing of *DIRC3* decreased the MTT conversion rate, but did not impact the migratory potential in *IGFBP5*-overexpressing MDA-T32 cells.

(A). Results of MTT assays performed in the plasmid transfected MDA-T32 cells 96 hours after GapmeR transfections. Experiments were performed three times with technical quadruplicates in each assay. "*" indicates $p < 0.05$ (tested with ANOVA in comparison to the negative control). **(B).** Representative fields of view (FOV) in Transwell assays of the GapmeR-transfected and plasmid-overexpressing MDA-T32 cells. **(C).** Migration of cells undergoing silencing of *DIRC3* (chemoattraction with 10% FBS). Experiments were repeated at least three times (n = 3). Results are shown as mean ± SD (analyzed with t-test).

4.2.6. Transcriptomic alterations induced by *DIRC3* silencing

RNA-seq profiling was performed to evaluate changes in cancer transcriptome that were related to the downregulation of *DIRC3* in MDA-T32 cells. An analogous analysis was performed in MDA-T32 cells after silencing of *IGFBP5*. This was done to evaluate the potential overlap in transcriptomic alterations generated by *DIRC3* and *IGFBP5* downregulations.

Total RNA was isolated from cells 72 hours after transfections with GapmeRs (either anti-*DIRC3*-common_3, anti-*IGFBP5* or negative control). Efficient gene silencing was confirmed using qRT-PCR (Figure 22). As expected, silencing of *DIRC3* reduced the abundance of *IGFBP5* transcripts. Interestingly, GapmeR anti-*IGFBP5* also downregulated *DIRC3*. This observation was in line with outcomes of the *IGFBP5* overexpression experiment, where *IGFBP5* upregulation induced expression of *DIRC3* (Figure 20).

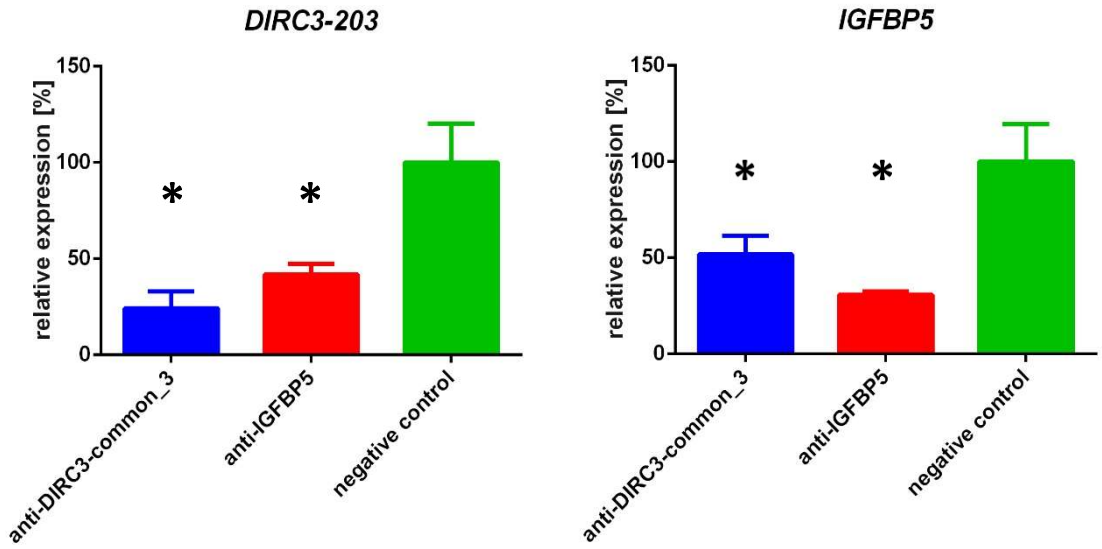


Figure 22. Silencing of *DIRC3* and *IGFBP5* in MDA-T32 samples prepared for RNA-sequencing.

Gene expression was analyzed using qRT-PCR. Transfections were repeated three times (n = 3). Results are shown as mean ± SD. "*" indicates a significant downregulation as compared to the control samples ($p < 0.05$; tested with ANOVA with Dunnett's test).

RNA-seq data was first evaluated using the principal components analysis (PCA). This assessment revealed good clustering of biological replicates, thus indicating that the triplicates had relatively consistent expression profiles (Figure 23).

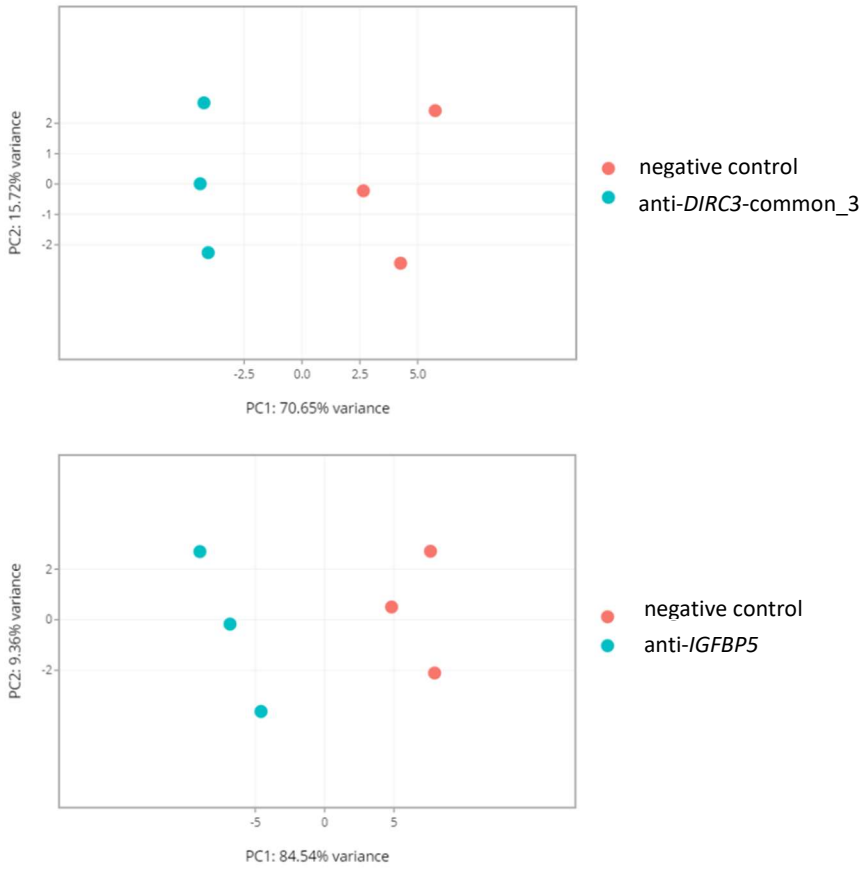


Figure 23. Principal components analysis (PCA) of the GapmeR-transfected MDA-T32 cells.

Cells were transfected with three types of GapmeRs (anti-*DIRC3*-common_3, anti-*IGFBP5* or negative control). Samples were RNA-sequenced and analyzed. PCA is a statistical procedure used for quality assessment and exploratory analysis of high-dimensional data (e.g., gene expression profiles). In the process the number of dimensions present in the original data is reduced into principal components in order to eliminate redundant information. Typically, two first principal components are analyzed since these values define the largest variability observed in a set of samples. Clusters are formed based on the similarity between samples in dimensions that present the largest variances. In this experiment distinct clusters formed by the RNA-seq triplicates confirmed a relative homogeneity of biologically related MDA-T32 samples.

Genes with an adjusted p-value less than 0.05 and absolute log2 fold expression change higher than 1 (as compared to the control samples) were designated as "differentially expressed genes" (DEGs). Silencing of *DIRC3* revealed 198 DEGs (77 genes upregulated, 121 downregulated), while silencing of *IGFBP5* resulted in 631 DEGs (294 genes upregulated, 337 downregulated). Volcano plots presenting prominent DEGs, and heatmaps of top 30 DEGs in the *DIRC3*-silenced and the *IGFBP5*-silenced groups are shown in Figures 24 and 25, respectively.

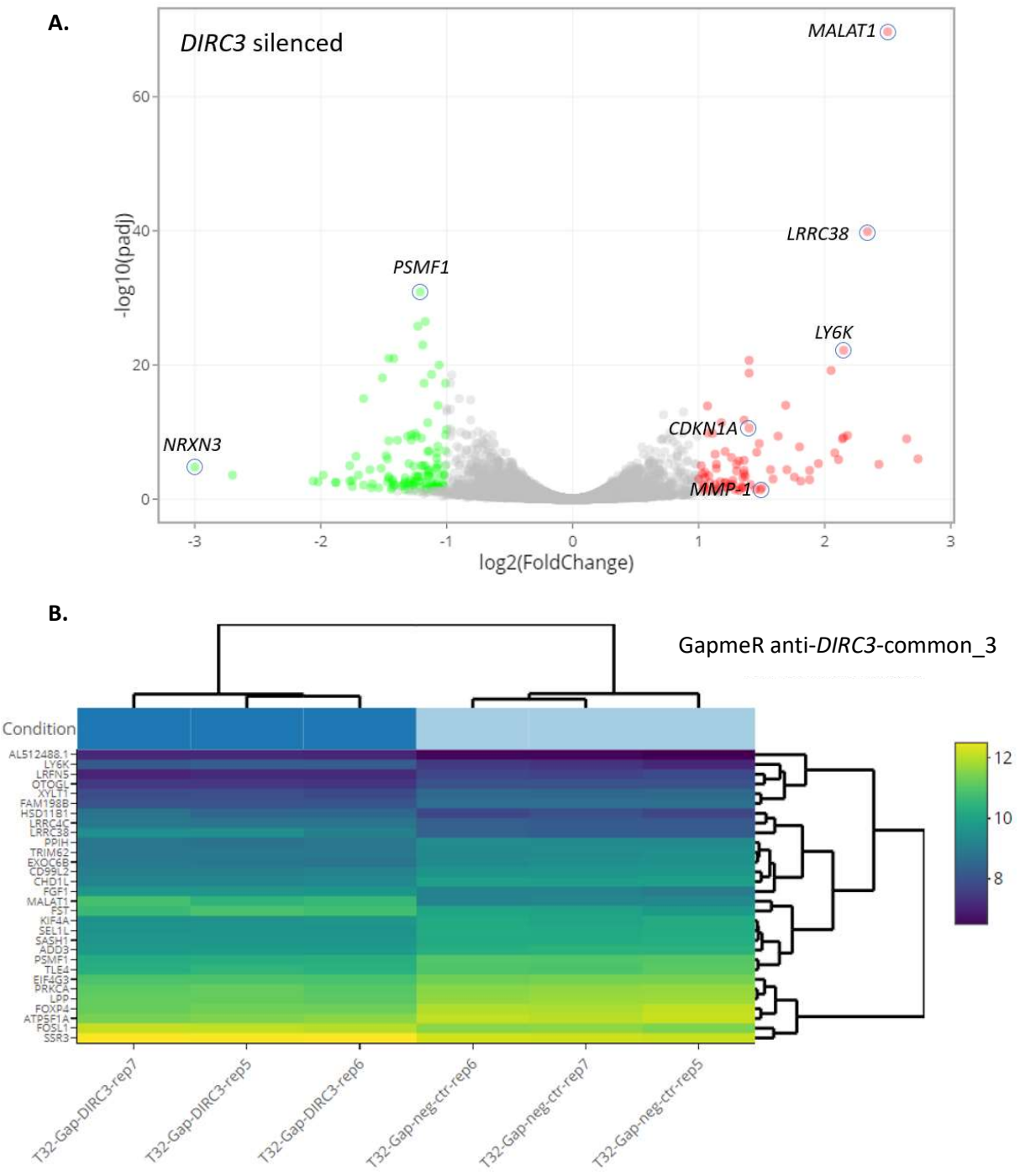


Figure 24. Transcriptome profiling of MDA-T32 cells transfected with the *DIRC3*-targeting GapmeR.

(A). Volcano plot illustrating changes in the gene expression profile after silencing of *DIRC3* as compared to the control samples. DEGs are shown as color dots (red: upregulated, green: downregulated). Certain prominent DEGs are marked. **(B).** A bi-clustering heatmap visualizing the expression profile (log2 transformed expression values) of top 30 *DIRC3*-altered DEGs (sorted by their adjusted p-values). "Gap-DIRC3" and "Gap-neg-ctr" represent samples transfected with anti-*DIRC3*-common_3 and negative control GapmeRs, respectively.

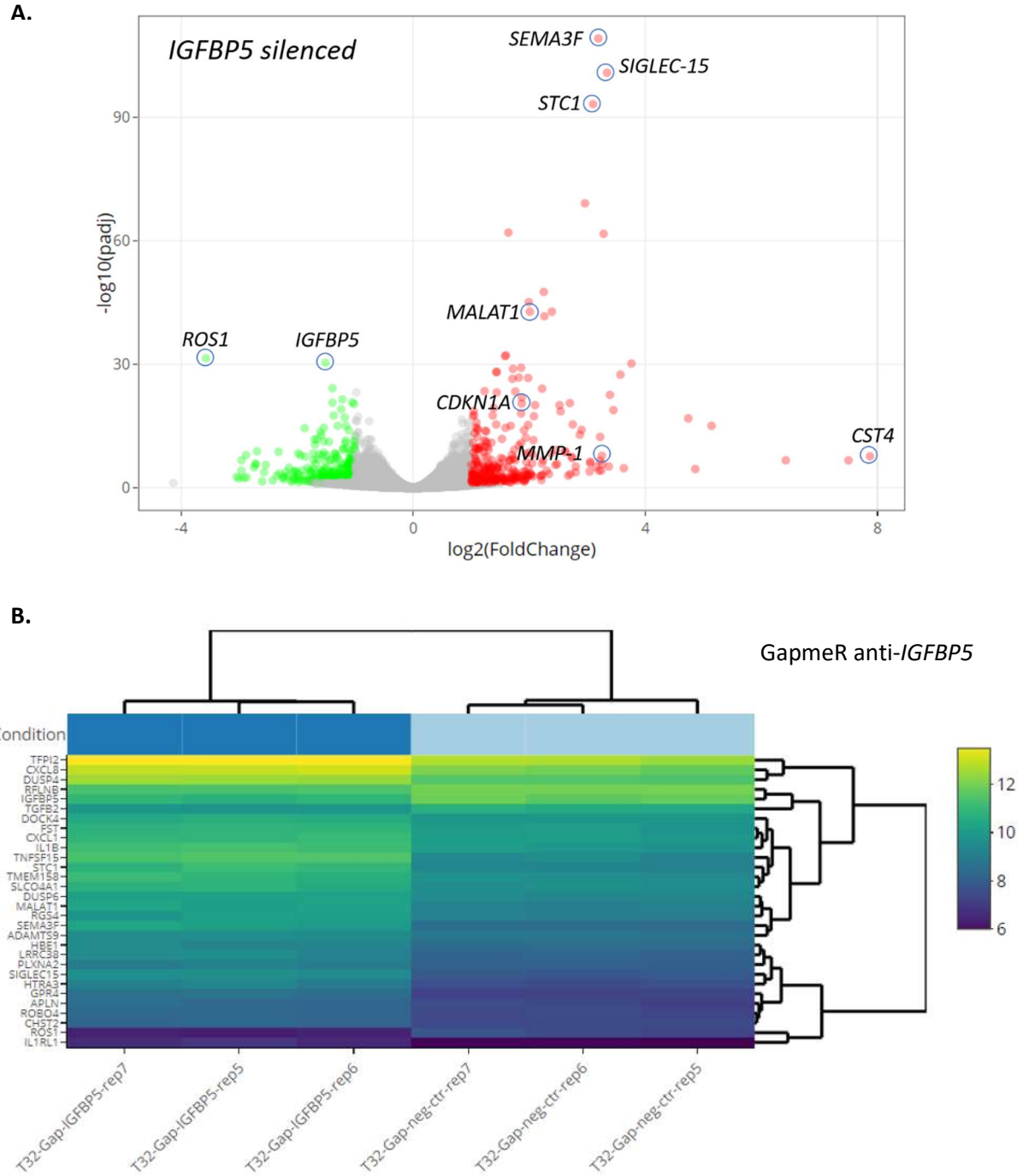


Figure 25. Transcriptome profiling of MDA-T32 cells transfected with the *IGFBP5*-targeting GapmeR.

(A). Volcano plot illustrating changes in the gene expression profile after silencing of *IGFBP5* as compared to the control samples. DEGs are shown as color dots (red: upregulated, green: downregulated). Certain prominent DEGs are marked. **(B).** A bi-clustering heatmap visualizing the expression profile (log2 transformed expression values) of top 30 *IGFBP5*-altered DEGs (sorted by their adjusted p-values). "Gap-IGFBP5" and "Gap-neg-ctr" represent samples transfected with anti-*IGFBP5* and negative control GapmeRs, respectively.

Gene overlap between DEGs (i.e., the *DIRC3*- and *IGFBP5*-regulated DEGs) was significant and comprised of 58 genes (Figure 26). Directions of the expression changes were concordant for all shared DEGs. These common DEGs were manually curated for their putative functional roles in thyroid oncogenesis. Some evidence (experimentally validated functionality, associations with clinical outcomes, or abnormal expression in thyroid carcinomas) was detected in the literature for most shared DEGs (see Appendix Table). Certain genes robustly implicated in thyroid cancers, e.g., *MALAT1*, *matrix metalloproteinase-1* (*MMP-1*), *cyclin dependent kinase inhibitor 1A* (*CDKN1A*) and *stanniocalcin 1* (*STC1*), were among the shared and upregulated DEGs.

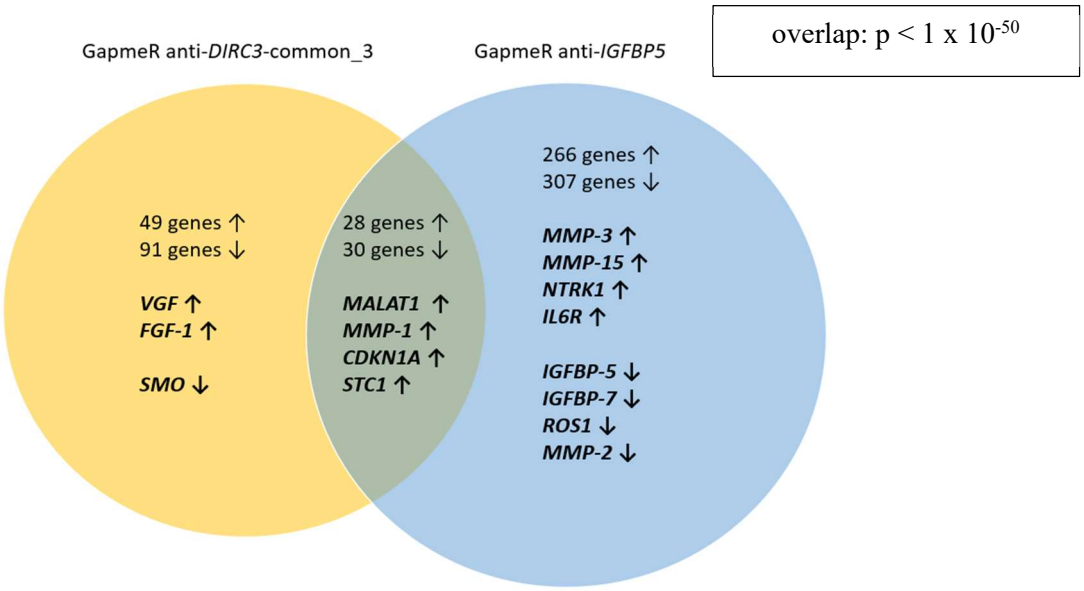


Figure 26. **Overlap between DEGs in the *DIRC3*- and *IGFBP5*-silenced MDA-T32 samples.**

Names of certain genes are provided. Arrows indicate up- or down-regulation. Gene overlap was calculated using the hypergeometric distribution calculator.

Gene ontology (GO) analysis was performed to identify biological processes that could be potentially influenced by *DIRC3* (Figure 27). In agreement with the results of phenotypic experiments, genes involved in the "negative regulation of cell migration" (GO:0030336) were most significantly affected by the *DIRC3* knockdown. Other biological processes that were significantly impacted included: "signal transduction" (GO:0007165), "angiogenesis" (GO:0001525), "cell proliferation" (GO:0008283), "response to growth hormone" (GO:0060416), "positive regulation of protein phosphorylation" (GO:0001934). *IGFBP5* is assigned to 29 GO terms in The Gene Ontology Annotation (GOA) Database. Several of these terms were among the most significantly altered by *DIRC3* silencing (e.g., GO:0030336~negative regulation of cell migration, GO:0007165~signal transduction, GO:0071320~cellular response to cAMP).

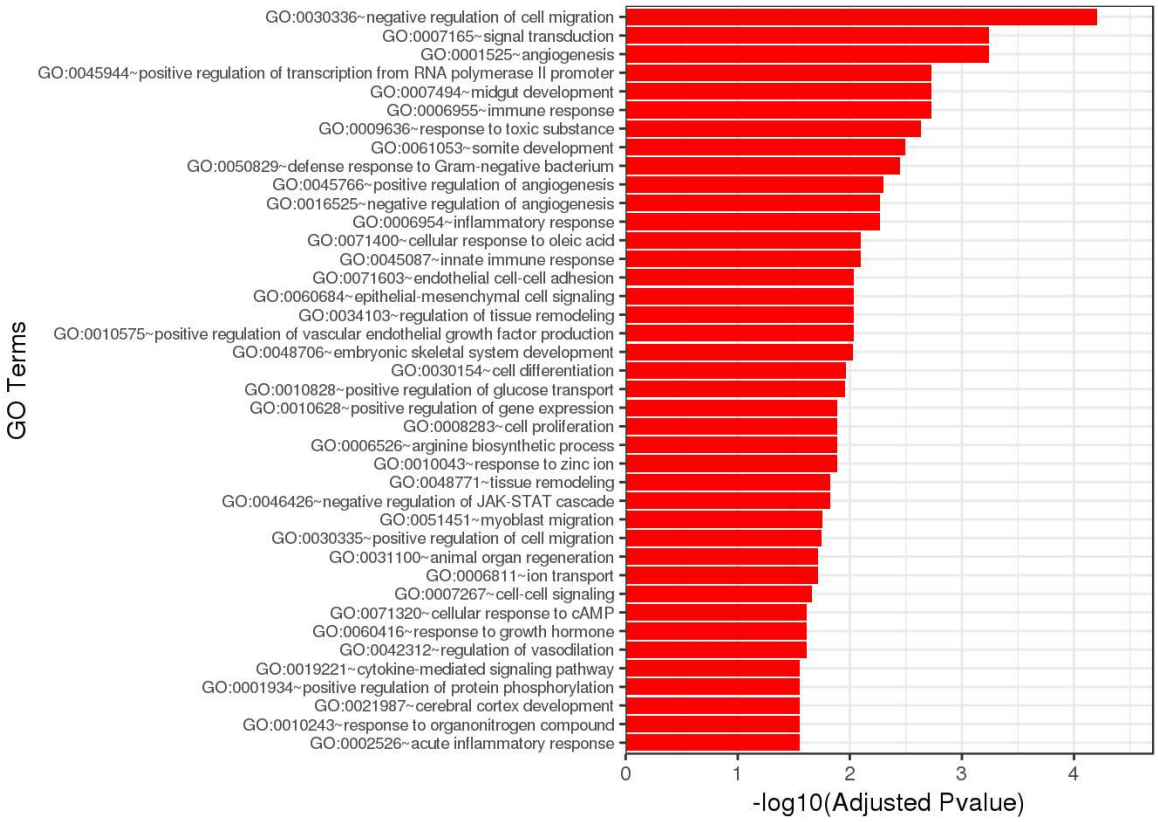


Figure 27. Gene ontology (GO) terms significantly enriched in the *DIRC3*-silenced MDA-T32 cells. Graphs shows 30 of the most significantly enriched GO terms.

4.3. Evaluation of the role of *DIRC3* in IGF-1 signaling

Since *DIRC3* was shown to influence expression of *IGFBP5*, I hypothesized that *DIRC3* could impact IGF-1 signaling in thyroid cancer cells. Remarkably, indirect evidence on the modulatory role of *DIRC3* is provided by previously reported associations between germline variants in *DIRC3* and human height (a physiological trait principally driven by IGF-1) [85-87, 182-186].

MDA-T32 cells were transfected with GapmeRs targeting *DIRC3*, *IGFBP5*, or with the negative control GapmeR. Three days later the complete culture medium was replaced with serum-free medium containing a low concentration of IGF-1. This was done because IGF-1 is known to promote production of IGFBP5 [187-190]. The conditioned medium was collected after 24 hours, while the cells were serum-starved for another 12 hours. Finally, cells were re-stimulated with the previously harvested conditioned medium. The medium was applied either unmodified (i.e., containing only a low amount of IGF-1), enhanced with additional IGF-1, or supplemented with the recombinant human IGFBP5 protein.

Stimulation of cancer cells with the conditioned medium induced phosphorylation of AKT (Figure 28). This effect was, however, much stronger in the cells transfected with the *DIRC3*- or *IGFBP5*- targeting GapmeRs (Figure 28A, lanes 1-2 and 7-8; Figure 28B). Phosphorylation of AKT was further increased when cells were stimulated with additional IGF-1 (lanes 2, 5 and 8). On the other hand, application of the recombinant human IGFBP5 protein prevented AKT phosphorylation (Figure 28; lanes 3 and 6). These findings indicate that upregulation of pAKT in response to the knockdown of *DIRC3* is directly dependent on the stimulatory effect of IGF-1.

In contrast, ERK was strongly phosphorylated across all conditions. This is because the utilized cancer cell line, MDA-T32, harbors *BRAF V600E* mutation (the typical driver mutation in conventional PTCs), which induces a persistent activation of the BRAF/ERK signaling pathway.

In conclusion, silencing of *DIRC3* promoted AKT signaling in thyroid cancer cells. This outcome confirmed that *DIRC3* modulates response to IGF-1 in cancer cells.

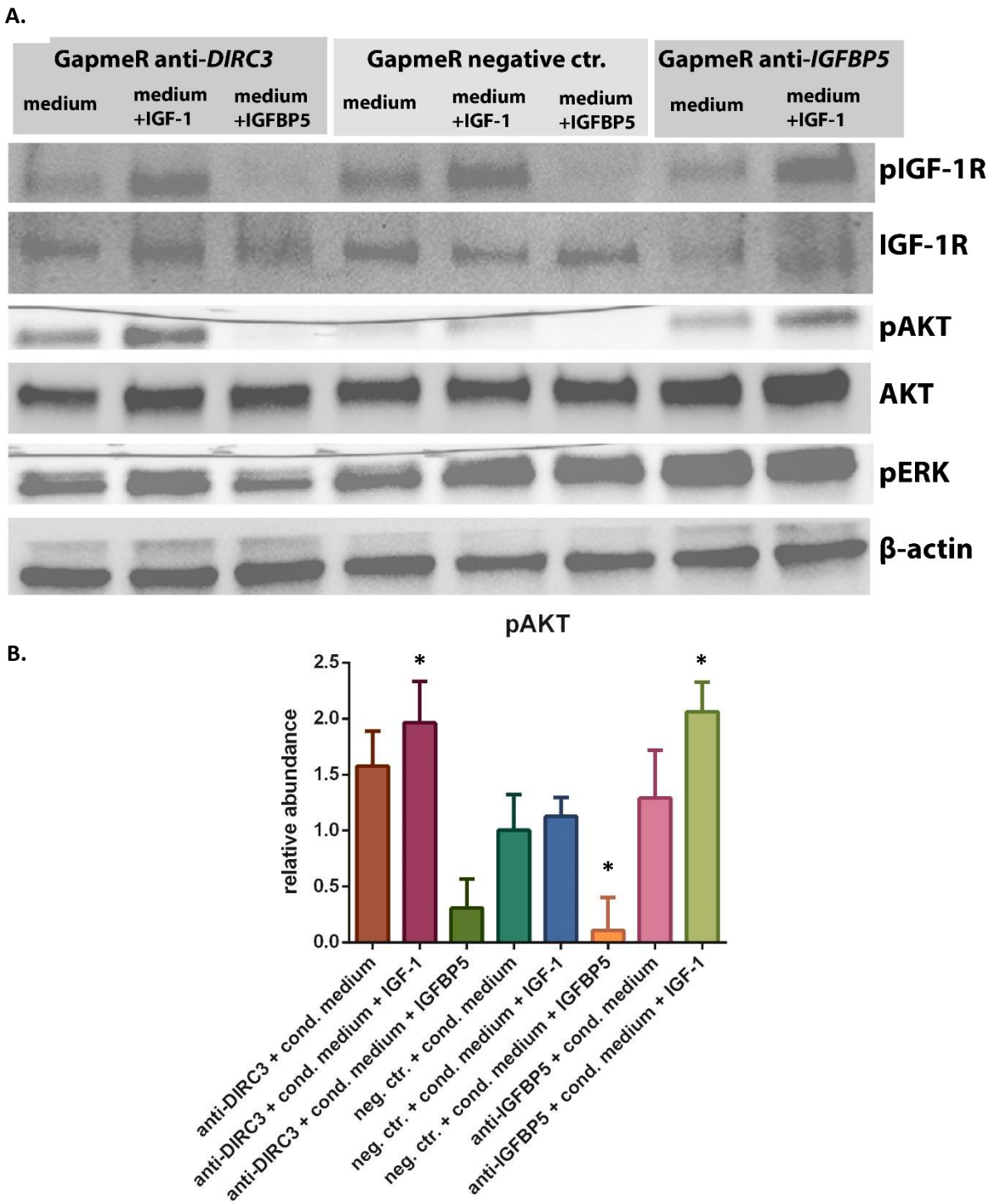


Figure 28. Downregulation of *DIRC3* augmented phosphorylation of AKT in MDA-T32 cells stimulated with IGF-1.

Western blot of MDA-T32 cells that were transfected with GapmeRs and stimulated with IGF-1. Cells were transfected with GapmeRs (anti-*DIRC3*-common_3, anti-*IGFBP5*, or negative control), cultured in serum-free medium containing IGF-1 (20 ng/ml), starved for 12 hours, and re-stimulated with the conditioned medium for 10 minutes. The conditioned medium was either unmodified, or enhanced with additional 30 ng/ml of IGF-1, or supplemented with IGFBP5 (500 ng/ml). **(A).** Representative blots probed with antibodies detecting pIGF-1R, total IGF-1R, pAKT, pan AKT, pERK and β-actin. **(B).** Relative abundance of pAKT in three independent blots (mean ± SD; n = 3). Semi-quantitative densitometric units (background subtracted) were normalized to the average densitometric value obtained for all eight samples tested in a given blot. Results were analyzed with ANOVA. "*" represent $p < 0.05$ (in comparison to the "neg. ctr + cond. medium: lane 4).

4.4. Phenotypic influence of *DIRC3* overexpression in cancer cell lines

Expression of *DIRC3* was induced using CRISPRa. This technique often permits a faithful recapitulation of physiological functions of nuclear lncRNAs, since these transcripts often exert their effects *in situ* (close to the sites of their transcription).

Stable clones expressing the SP-dCas9-VPR plasmid were successfully generated for MDA-T120 and MDA-T32 cell lines (derivatives were termed here as MDA-T120 dC9-VPR and MDA-T32 dC9-VPR, respectively). These cell lines were functionally validated by transfecting them with sgRNA that targeted the *PVT1* promoter (sgRNA-*PVT1*). As expected, MDA-T120 dC9-VPR and MDA-T32 dC9-VPR cells upregulated expression of *PVT1* in response to the transfection (Figure 29A). Generation of inducible SP-dCas9-VPR derivatives was unsuccessful for MDA-T68 and K1 cell lines.

Next, MDA-T120 dC9-VPR cells were transfected with several sgRNAs designed to target the region directly upstream from the TSS of *DIRC3*. Two out of five utilized sgRNAs successfully upregulated *DIRC3* (Figure 29B). These two sgRNAs (termed sgRNA-*DIRC3*-3 and sgRNA-*DIRC3*-5) were combined. Application of the mixture of two sgRNAs (1:1 ratio) produced a strong synergistic effect on *DIRC3* expression in MDA-T120 dC9-VPR and MDA-T32 dC9-VPR cells (Figure 29B). Accordingly, the combination of sgRNA-*DIRC3*-3 and sgRNA-*DIRC3*-5 (denoted as sgRNA-*DIRC3*-3+5) was used in the subsequent CRISPRa experiments. Surprisingly, while sgRNAs effectively upregulated *DIRC3*, they had no impact on expression of *IGFBP5* (Figure 29B).

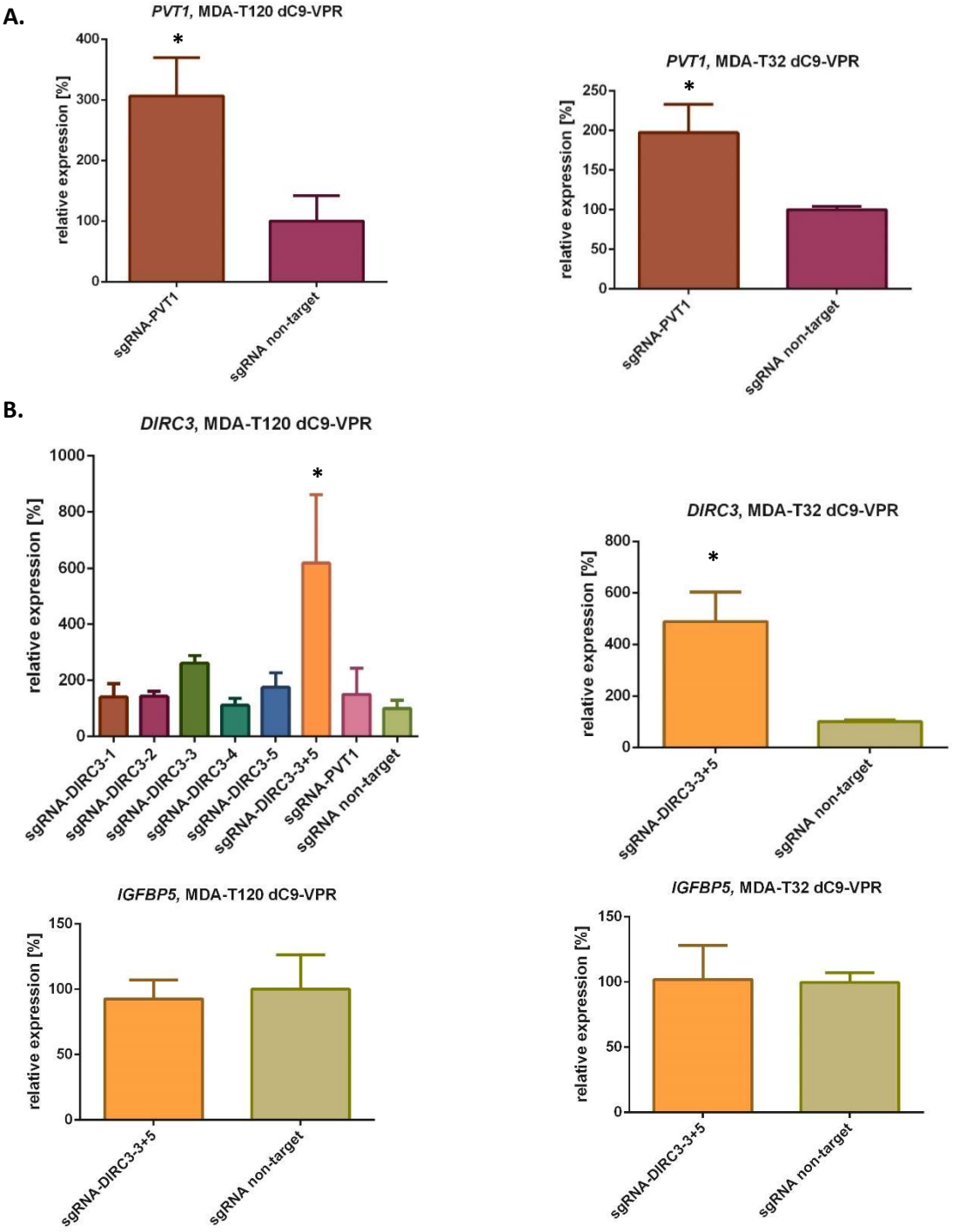


Figure 29. Expression of *DIRC3* was induced in thyroid cancer cell lines with CRISPRa.

(A). Validation of MDA-T32 and MDA-T120 cell lines stably transfected with SP-dCas9-VPR. The positive control sgRNA successfully upregulated *PVT1*. **(B).** Five sgRNAs were designed to target the region directly upstream from the TSS of *DIRC3*. Combination of two sgRNA designs (no. 3 and no. 5) generated the strongest upregulation of *DIRC3*. Results are shown as mean ± SD (n = 3). Results were analyzed with t-test (for two groups) or ANOVA with Dunnett's test (for >2 groups). "*" indicates $p < 0.05$ as compared to the non-target sgRNA.

MDA-T32 dC9-VPR cell line was evaluated using MTT assay. MDA-T120 dC9-VPR cell line could not be reliably evaluated with this assay due to a very slow growth rate.

Transfection with sgRNA-*DIRC3*-3+5 significantly promoted the MTT reduction rate in MDA-T32 dC9-VPR cells (Figure 30). Given that these sgRNAs did not influence expression of *IGFBP5* (as shown in Figure 29), this result could indicate that *DIRC3* regulated the activity observed in MTT assays independently from *IGFBP5*. In fact, this conclusion is in agreement with results of the *IGFBP5* "rescue" experiment (chapter 4.2.5), in which overexpression of *IGFBP5* did not prevent the decrease in MTT conversion rate produced by the *DIRC3*-targeting GapmeRs (see Figure 21A).

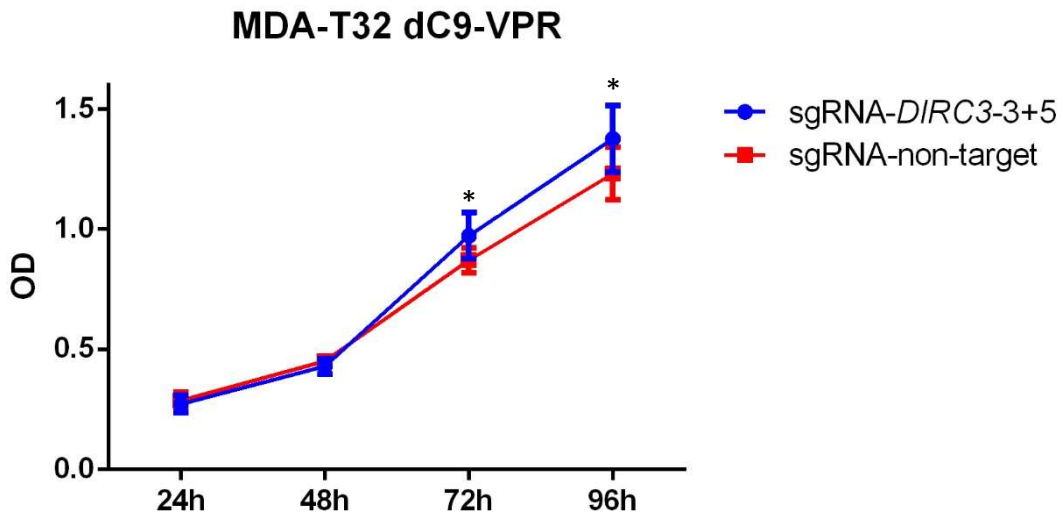


Figure 30. Increase in MTT reduction rate was promoted by *DIRC3* overexpression in MDA-T32 dCas9-VPR cells.

Results analyzed with two-way ANOVA with Šidák test. "*" indicates $p < 0.05$. Each experiment was repeated three time with technical quadruplicates in each set-up. Abbreviation: OD, optical density.

4.5. Genomic editing of rs11693806, the germline DTC risk variant in *DIRC3*

4.5.1. Generation of the rs11693806-edited clones.

Allele rs11693806[C] is associated with increased incidence of DTCs. I decided to evaluate the biological role of this variant using CRISPR. sgRNA was designed to cleave genomic DNA directly in the rs11693806[C] locus. The on-target activity of this sgRNA was evaluated using the T7 endonuclease I (T7EI) assay in K1 cell line (Figure 31). Results of the assay indicated that the desired genomic locus was edited efficiently.

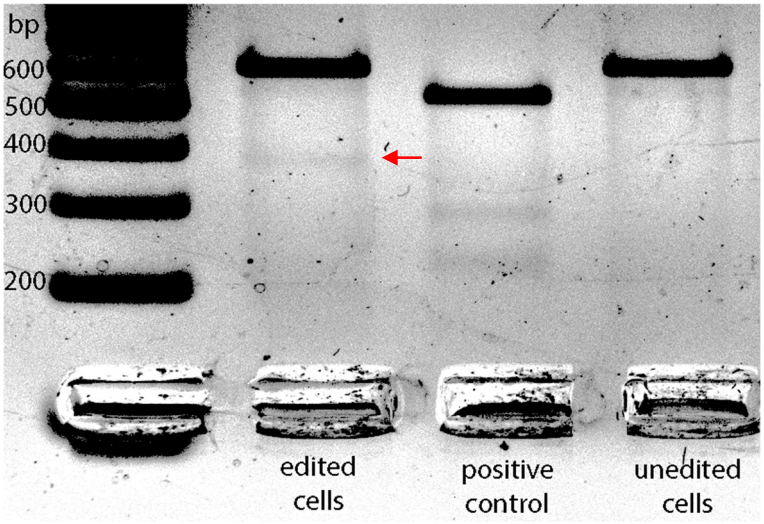


Figure 31. sgRNA successfully targeted the rs11693806[C] allele in T7EI assay.

PCR products encompassing the rs11693806 locus were visualized on 1% agarose gel. Presence of a cleaved PCR product (arrow) in the CRISPR-edited K1 cells indicated a mismatch in the rs11693806 locus. This small band was not observed in the unedited cells. Positive control was a mismatched DNA product included in the GeneArt Genomic Cleavage Detection Kit (Thermo Fisher).

Next, MDA-T32 cells were transfected with Cas9/sgRNA RNP complexes targeting the rs11693806[C] allele. Additionally, donor DNA oligonucleotides (Alt-R oligos) and the NHEJ inhibitor were applied to employ HDR mechanisms into the editing process. Two types of donor oligonucleotides (both containing rs11693806[G]) were tested (one corresponding to the "target" strand, and another to the "non-target" strand). The anticipated outcome was conversion of the heterozygotic MDA-T32 cell line (rs11693806[C/G]) into the homozygotic rs11693806[G/G] derivative.

Once the transfected cells achieved a sufficient growth, single cell clones were isolated using the dilution cloning method. Next, crude (unpurified) DNA was obtained from each clone and tested using the rhAmp SNP genotyping assay. This screening test identified multiple clones that were deprived of rs11693806[C] (Figure 32A). The candidate clones were further expanded, re-tested with PCR and subsequently Sanger sequenced. Among the sequenced clones, one clone had an insertion, several clones harbored small deletions in the rs11693806 locus, and two clones were initially identified as rs11693806[G/G] homozygotes (Figure 32B

and 32C). One of these presumably homozygotic clones was, however, excluded from the analysis due to inconsistent results obtained in Sanger sequencing. The second candidate clone (termed "D7") was consistently shown to be precisely edited into the rs11693806[G/G] homozygote. Clone "D7" was generated utilizing the non-target strand ("+") Alt-R donor oligonucleotide.

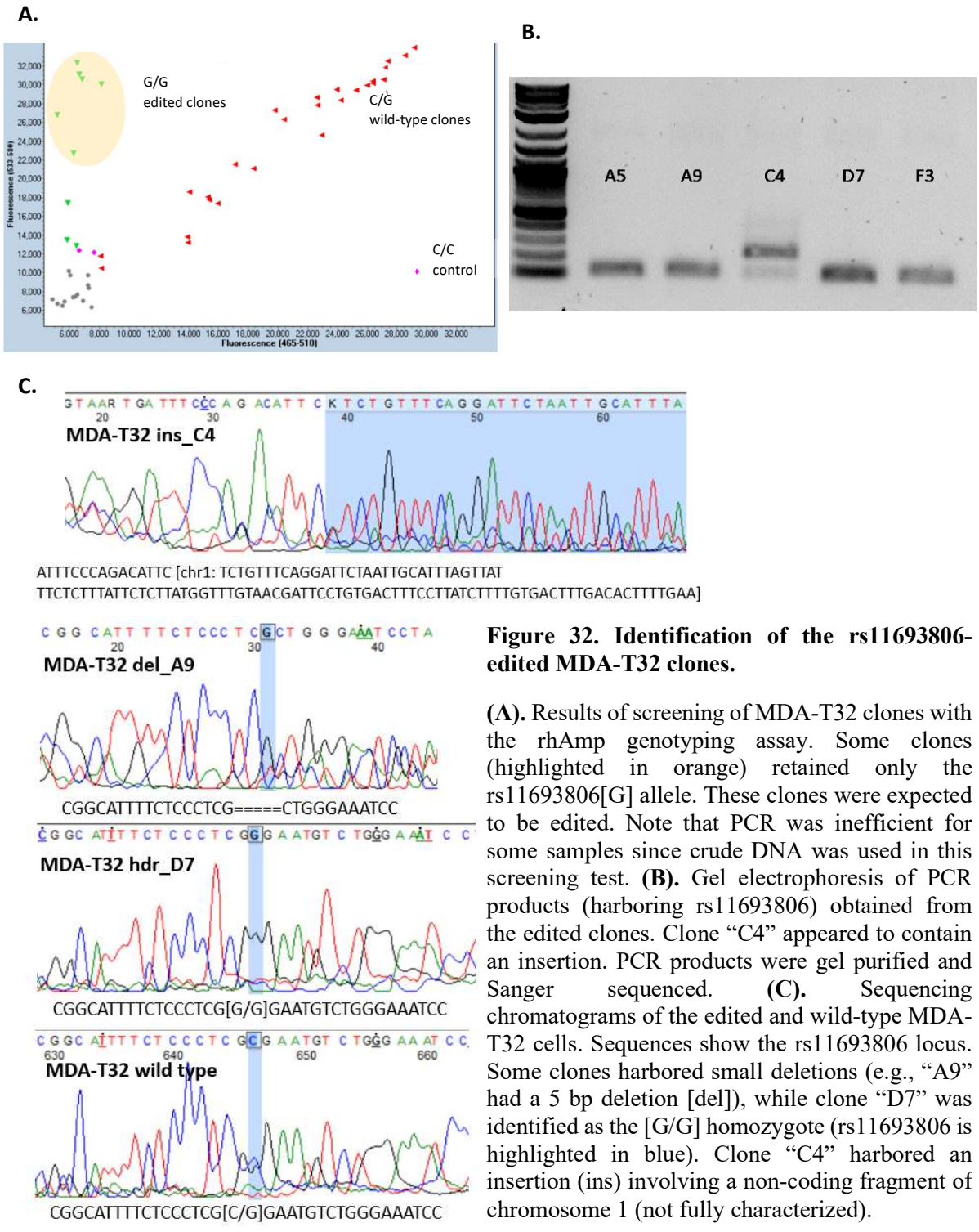


Figure 32. Identification of the rs11693806-edited MDA-T32 clones.

(A). Results of screening of MDA-T32 clones with the rhAmp genotyping assay. Some clones (highlighted in orange) retained only the rs11693806[G] allele. These clones were expected to be edited. Note that PCR was inefficient for some samples since crude DNA was used in this screening test. **(B).** Gel electrophoresis of PCR products (harboring rs11693806) obtained from the edited clones. Clone "C4" appeared to contain an insertion. PCR products were gel purified and Sanger sequenced. **(C).** Sequencing chromatograms of the edited and wild-type MDA-T32 cells. Sequences show the rs11693806 locus. Some clones harbored small deletions (e.g., "A9" had a 5 bp deletion [del]), while clone "D7" was identified as the [G/G] homozygote (rs11693806 is highlighted in blue). Clone "C4" harbored an insertion (ins) involving a non-coding fragment of chromosome 1 (not fully characterized).

4.5.2. *DIRC3* and *IGFBP5* expression in the rs11693806-edited clones.

Gene expression was evaluated in unmodified MDA-T32 cells and in two clones harboring genomic modifications of rs11693806[C] (a monoallelic deletion or precise editing of the SNV). Results of qRT-PCR indicated that both of these clones presented significant downregulations of *DIRC3* and *IGFBP5* as compared to the parental cell line (Figure 33).

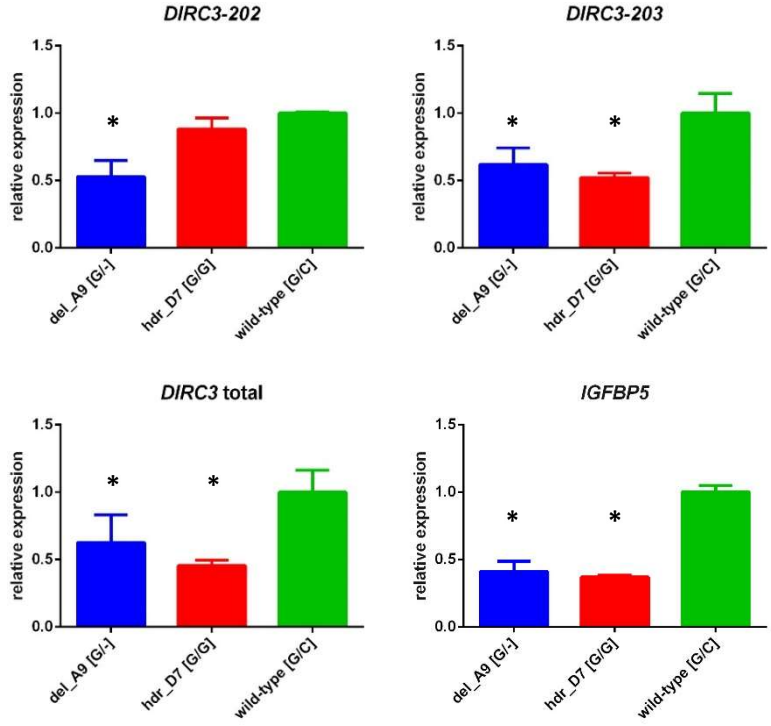


Figure 33. *DIRC3* and *IGFBP5* were downregulated in the rs11693806-edited MDA-T32 clones.

Relative expression of *DIRC3* splice variants (*DIRC3*-202, *DIRC3*-203), total *DIRC3* and *IGFBP5* was analyzed with qRT-PCR (names of genes are shown in captions). Gene expression was tested in the clone "A9" (which harbored a small deletion [del] of the rs11693806[C] allele), in the clone "D7" (which was precisely edited into the rs11693806[G/G] homozygote) and in the wild-type (heterozygotic) MDA-T32 cells. Results are shown as mean ± SD (n = 3) "*" indicates $p < 0.05$ in compassion to wild-type cells (analyzed with ANOVA with Dunnett's test).

4.5.3 Phenotypic evaluation of the rs11693806-edited clones.

Phenotypes of the rs11693806-edited clones were studied using MTT and Transwell migration assays. Clones "A9" and "D7" demonstrated a lower increase in the MTT conversion capability than the parental heterozygotic cells (Figure 34A). On the other hand, the CRISPR-edited clones demonstrated a higher migratory potential in Transwell assays than the wild-type MDA-T32 cells (Figure 34B and 34C). Accordingly, these outcomes phenocopied results of the *DIRC3*-silencing experiments (chapter 4.2.3).

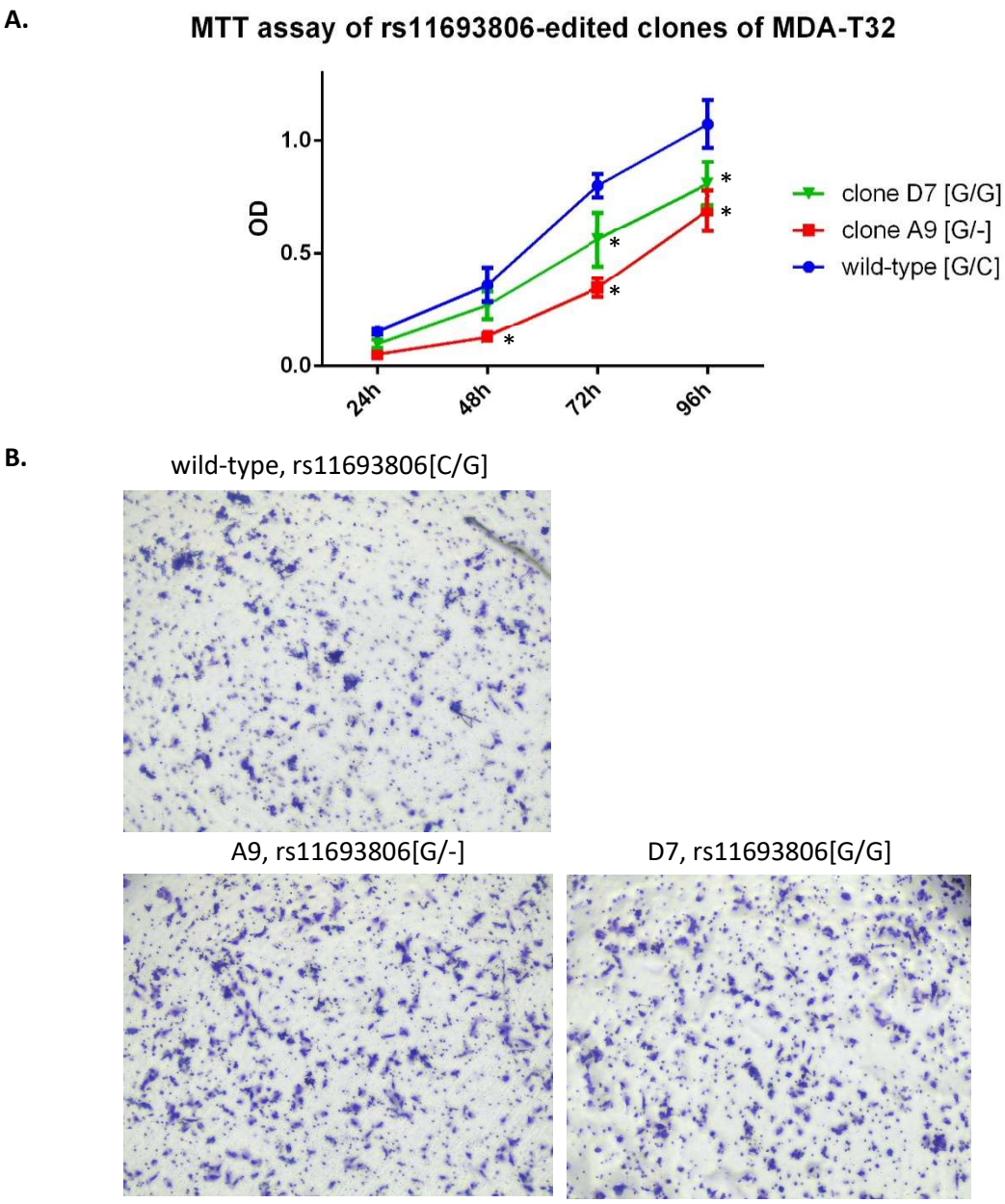


Figure 34. Rs11693806-edited MDA-T32 cells presented reduced MTT conversion rate, but increased migratory potential.

(A). Results of MTT assay of the edited and wild-type MDA-T32 cells. Results are presented as mean ± SD (n = 3) "*" indicates $p < 0.05$ in comparison to wild-type cells; two-way ANOVA with Šidák test). **(B).** Representative fields of view (FOV) in Transwell assays evaluating the edited clones. **(C).** Migration of the CRISPR-edited clones and wild-type MDA-T32 cells. Each experiment was repeated at least three times. Results are presented as mean ± SD (n = 3) "*" indicates $p < 0.05$ in comparison to wild-type cells (ANOVA with Dunnett's test). Abbreviation: OD, optical density.

4.5.4. Whole exome sequencing of the rs11693806-edited clones

I decided to perform whole exome sequencing (WES) to confirm that phenotypic and transcriptomic effects observed in the CRISPR-edited clones were not caused by unintended genomic alterations (i.e., off-target editing effects). Exome profiling of the parental MDA-T32 cell line (wild-type: "wt") and modified daughter lines ("A9" and "D7") produced high quality sequencing results (Figure 35A). In each sample at least 99.94% of unique reads were mapped to the human reference genome (GRCh38). Read depth of ≥20X was achieved in ≥97.39% of target regions.

Mapping of sequence reads revealed a consistent pattern of variant calls (SNVs and indels) across all lines (Figure 35B and 35C). This outcome confirmed that the clones harbored no major deviations in their exomes. Furthermore, protein-coding sequences of the CRISPR-edited clones were computationally compared with the parental exome. The evaluation identified 391 and 408 variant calls (SNVs and indels) that were potentially unique in the exomes of "A9" and "D7" clones, respectively. These putative alterations were filtered and reviewed using VarSifter and IGV software. All indel calls were manually curated (43 and 41 in the "A9" and "D7" clones, respectively). The inspection revealed that none of the variant calls was a credible exome alteration. False-positive signals were caused by various shortcomings in the sequencing and mapping workflow, e.g., inadequate coverage of particular regions, minor imbalances in allelic reads, or sequencing errors in low-complexity genomic regions (Figure 35D).

The edited rs11693806 locus could not be evaluated as non-protein coding regions were not covered by the utilized exome enrichment kit (Figure 36A). Putative off-target sites of sgRNA were predicted *in silico* using CRISPOR and Cas-OFFinder online tools. Potential off-target sites in the exome (sequences with up to five-nucleotide mismatches to the used sgRNA) were reviewed in BAM files (Figure 36B). No alterations were identified in the edited clones in any of these loci (Figure 36C).

In summary, results of WES confirmed that the CRISPR-edited clones were valid daughter cell lines of MDA-T32. The clones harbored no credible off-target alterations in the protein-coding sequences relative to the parental cell line. Hence, any transcriptomic and phenotypic alterations in the clones were very likely to be caused solely by the genomic editing of *DIRC3*.

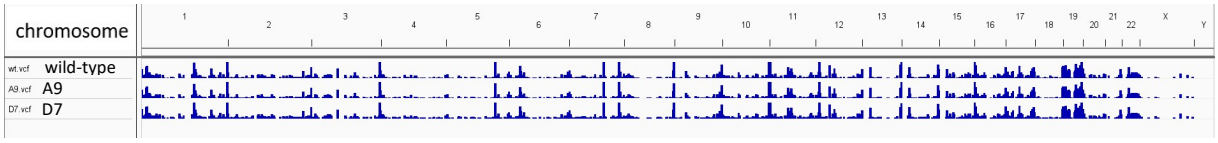
A.

Sample	# Reads	Yield (Mbases)	Mean	% Bases >=30	Mean target coverage	% Target Regions with no coverage	% Target regions >=20X
A9	40,430,510	12,129	36	94.1	86	1.27	97.39
D7	44,730,337	13,419	35.9	93.8	92	1.22	97.46
wt	47,093,355	14,128	35.8	93.3	94	1.18	97.52

B.

Sample	SNV	INDEL	Singleton	Homozygous	Heterozygous	Hemizygous
A9	25,497	926	37	10,481	15,942	0
D7	25,521	916	38	10,465	15,972	0
wt	25,571	911	89	10,414	16,068	0

C.



D.

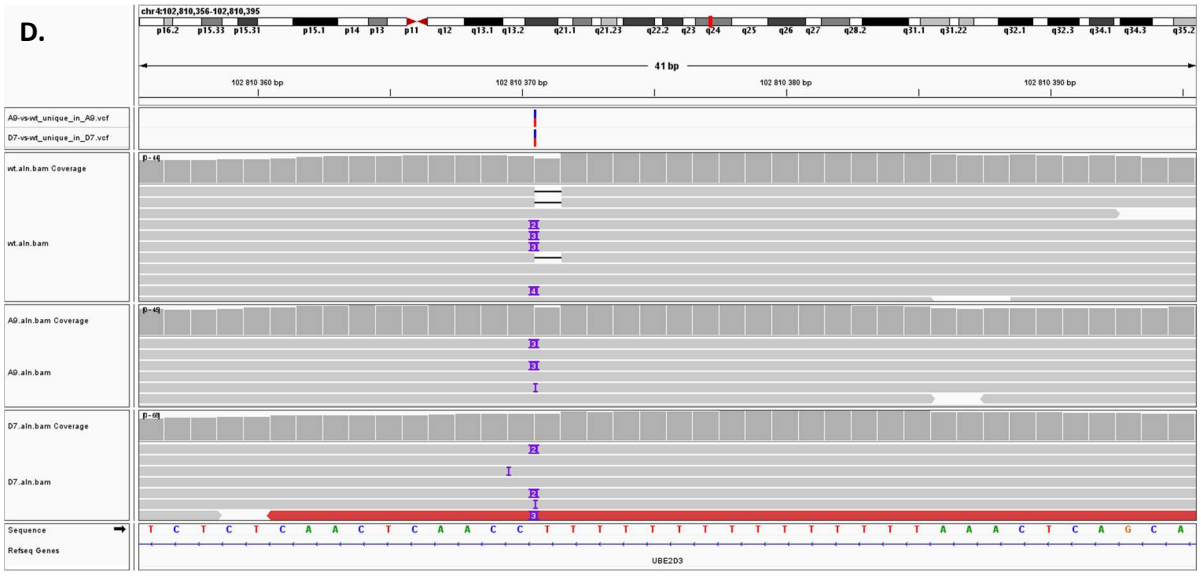
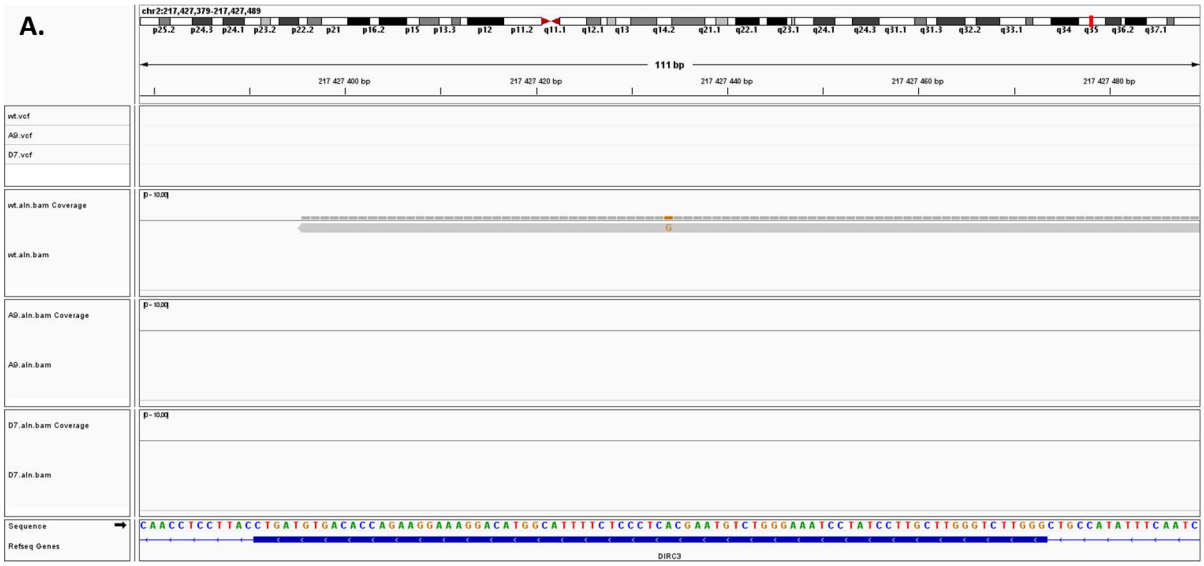


Figure 35. Results of whole exome sequencing of the rs11693806-edited MDA-T32 clones.

(A). Overview of sequencing statistics for the edited clones ("A9" and "D7") and the wild-type cell line ("wt"). **(B).** Variant calls identified in the MDA-T32 clones (relative to the human reference genome GRCh38). **(C).** Histogram representing the pattern of variant calls observed across exome in the edited and wild-type cells (relative to the human reference genome GRCh38). **(D).** Example of a false-positive variant call (marked as a putative deletion unique for the CRISPR-edited clones) in a low complexity poly(T) region.



B.

Off-target Seq; guide: GATTTCCCAGACATTCGCGAGGG	mismatch position	chromosome	start	end	locus
GAAGGCCCAGACATTCGTGAAGG	..***.....*..	chr3	136992015	136992037	exon:IL20RB
AAGCTCCCAGACATTCGAGAAGG	*.**.....*..	chr11	65278902	65278924	exon:POLA2
GACTATCCAGATATTCGCGAGGG	..*.**....*.....	chr6	39725277	39725299	exon:KIF6
GCTTTCCCAGACATTCAGGGAGG	.*.....**.*	chr8	42736844	42736866	exon:CHRNB3
AAATTCCCAGACATGGGCGATGG	*.*.....**....	chr7	143853491	143853513	exon:TCAF1
GATTTCGCAGATAATCGCGACGG	.....*....*.*.....	chr15	74135297	74135319	exon:ISLR2

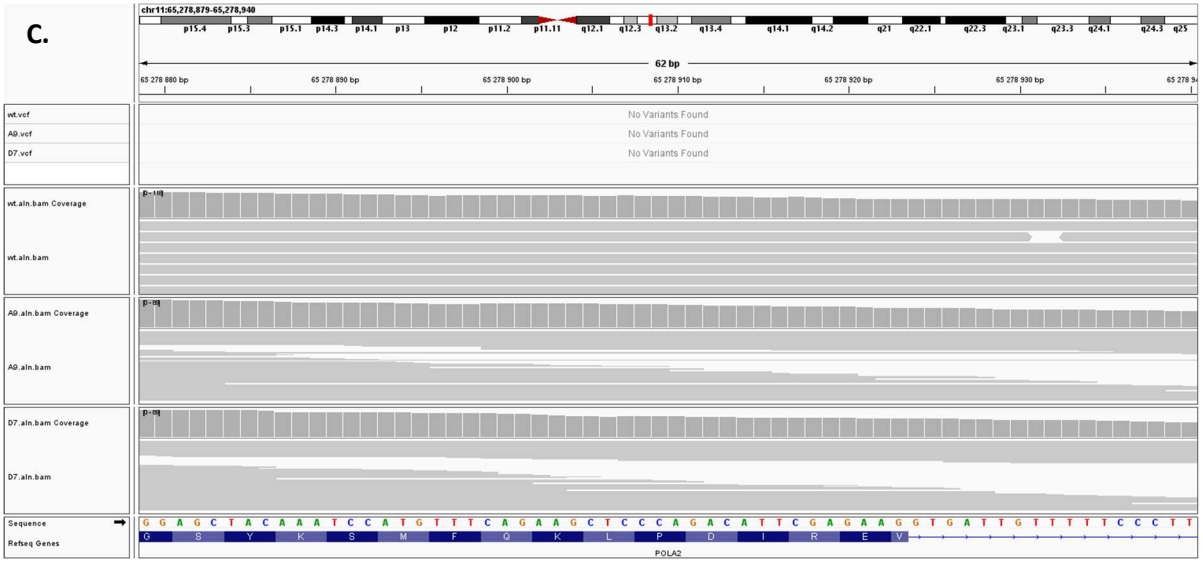


Figure 36. The CRISPR on-target and putative off-target sites in the exome data of MDA-T32 lines (wild-type cells and rs11693806-edited clones).

(A). Reads mapped to the rs11693806 locus. The locus could not be investigated using WES data since the region was not covered by the exome enrichment kit. **(B).** List of top putative exome off-targets for the sgRNA used to edit rs11693806[C]. Off-target sequences were identified using CRISPOR. * represents a mismatch of "Off-target Seq" relative to the "guide". These regions were manually reviewed in the WES data. **(C).** A representative image of one of the putative off-target sites of the rs11693806-editing sgRNA. No variant calls (no off-target editing) were observed in any of the inspected loci. Abbreviation: wt, wild-type.

4.5.5. RNA sequencing of the rs11693806-edited MDA-T32 clones

Transcriptomic landscapes of the CRISPR-edited clones were compared with the transcriptomic profile of the unedited (wild-type) MDA-T32 cells. RNA-seq was performed in the clones that were genomically and functionally validated: "A9" (the deletion clone; rs11693806[G/-]) and "D7" (the homozygotic clone; rs11693806[G/G]). Biological triplicates produced relatively homogenous transcriptomic profiles in each cell line in the PCA analysis (Figure 37A). Analysis of RNA-seq data identified 1511 DEGs (291 upregulated, 1220 downregulated) and 1356 DEGs (245 upregulated, 1111 downregulated) for "A9" and "D7" clones, respectively. The most significant DEGs detected in the CRISPR-edited clones are shown in Figure 37B and 37C. Volcano plots illustrating changes in the gene expression profiles indicated similar transcriptomic landscapes in both rs11693806-edited clones (Figure 38A and 38B). Indeed, the overlap in DEGs was highly significant, and encompassed 44.9% and 55.1% of all DEGs detected in "A9" and "D7" clones, respectively (Figure 38C). Among the shared DEGs was *IGFBP5*, which was downregulated in both edited clones in comparison to the parental wild-type cell line (Figure 39A). *IGFBP5* was among the most significantly affected genes in the "D7" clone (see Heatmap in Figure 37C). This result has confirmed that rs11693806 is critical for the regulation of *IGFBP5* expression. Similarly, *DIRC3* was downregulated in the CRISPR-edited clones (Figure 39B). Nevertheless, expression of the particular exon that harbors rs11693806 appeared to be very low in all samples (Figure 39C). This may suggest that expression of this exon is of limited relevance to the function of *DIRC3* transcripts. Accordingly, it is possible that this locus may impose its transcriptomic and phenotypic effects *via* other mechanisms (e.g., it may harbor PREs).

DEGs common for both rs11693806-edited MDA-T32 clones presented certain overlap with DEGs that were identified in the *DIRC3*-silencing experiments (see chapter 4.2.6). These common DEGs included: *ABCC9, ANGPT2*, AREG*, BEX1*, BGN*, CSF2*, DOK2, EFR3B, ESM1*, FRY, IL23A, INSC*, LINC00520, LRRC38*, MGP*, MMP1*, MYO5B*, NLRP5, RAP1GAP2, RARRES2, RCAN2, SMO, THSD7A, TRIL*, WDR86**. Genes marked with the asterisk were also affected by the *IGFBP5*-targeting GapmeR (i.e., these DEGs were common for all four sample sets). One of the interesting DEGs is *smoothened* (*SMO*), a key component of the hedgehog signaling pathway. *SMO* was downregulated in both CRISPR-edited clones as well as in the *DIRC3*-silencing experiment. On the other hand, its expression was not significantly altered by the *IGFBP5*-targeting GapmeR.

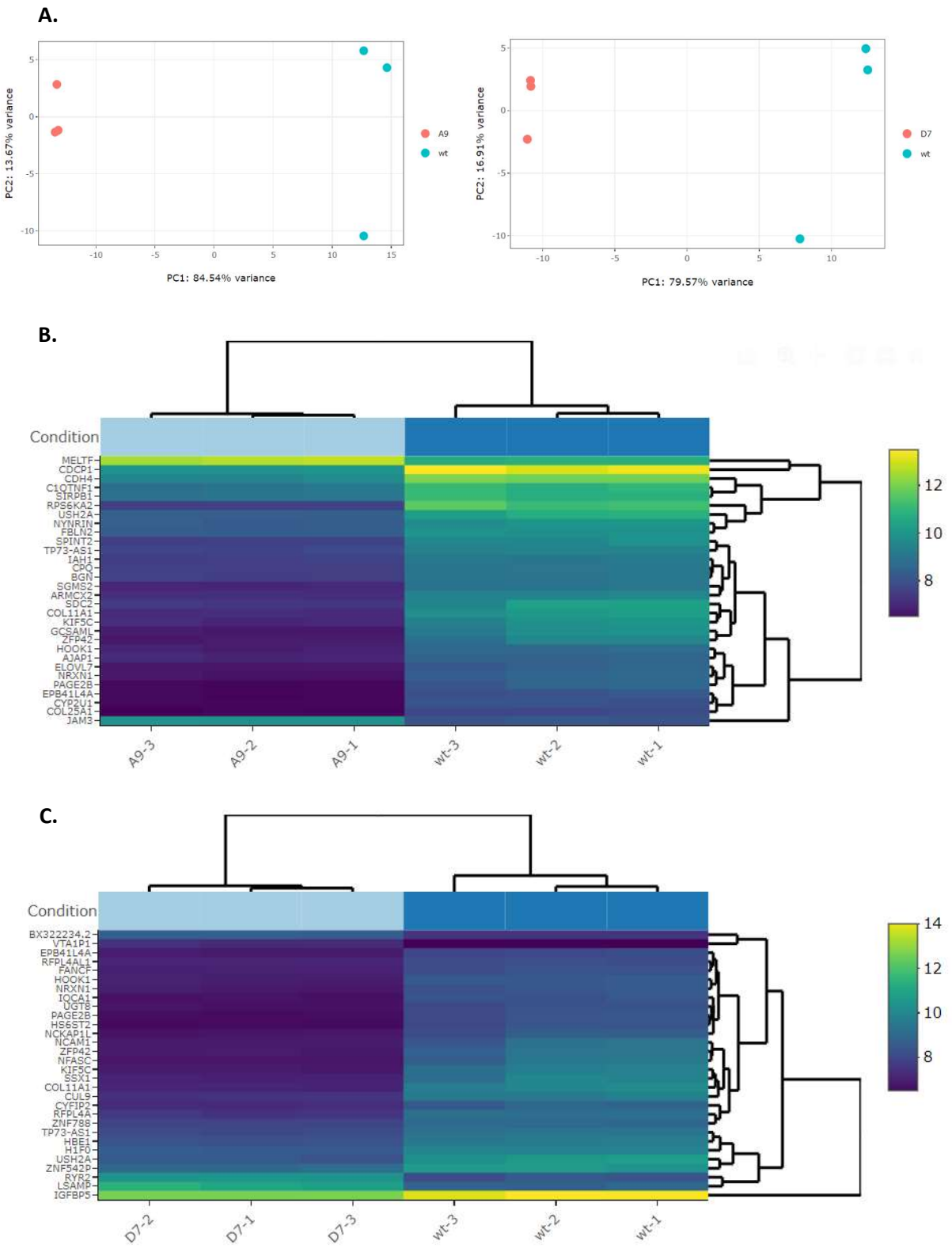


Figure 37. Transcriptomic profile of the CRISPR-edited MDA-T32 cells.

(A). Principal components analysis (PCA) of the RNA-seq triplicates. **(B).** A bi-clustering heatmap visualizing the expression profile (log2 transformed expression values) of top 30 DEGs in the "A9" clone (DEGs sorted by adjusted p-values). **(C).** A bi-clustering heatmap visualizing the expression profile (log2 transformed expression values) of top 30 DEGs in the "D7" clone (DEGs sorted by adjusted p-values). Abbreviation: wt, wild-type.

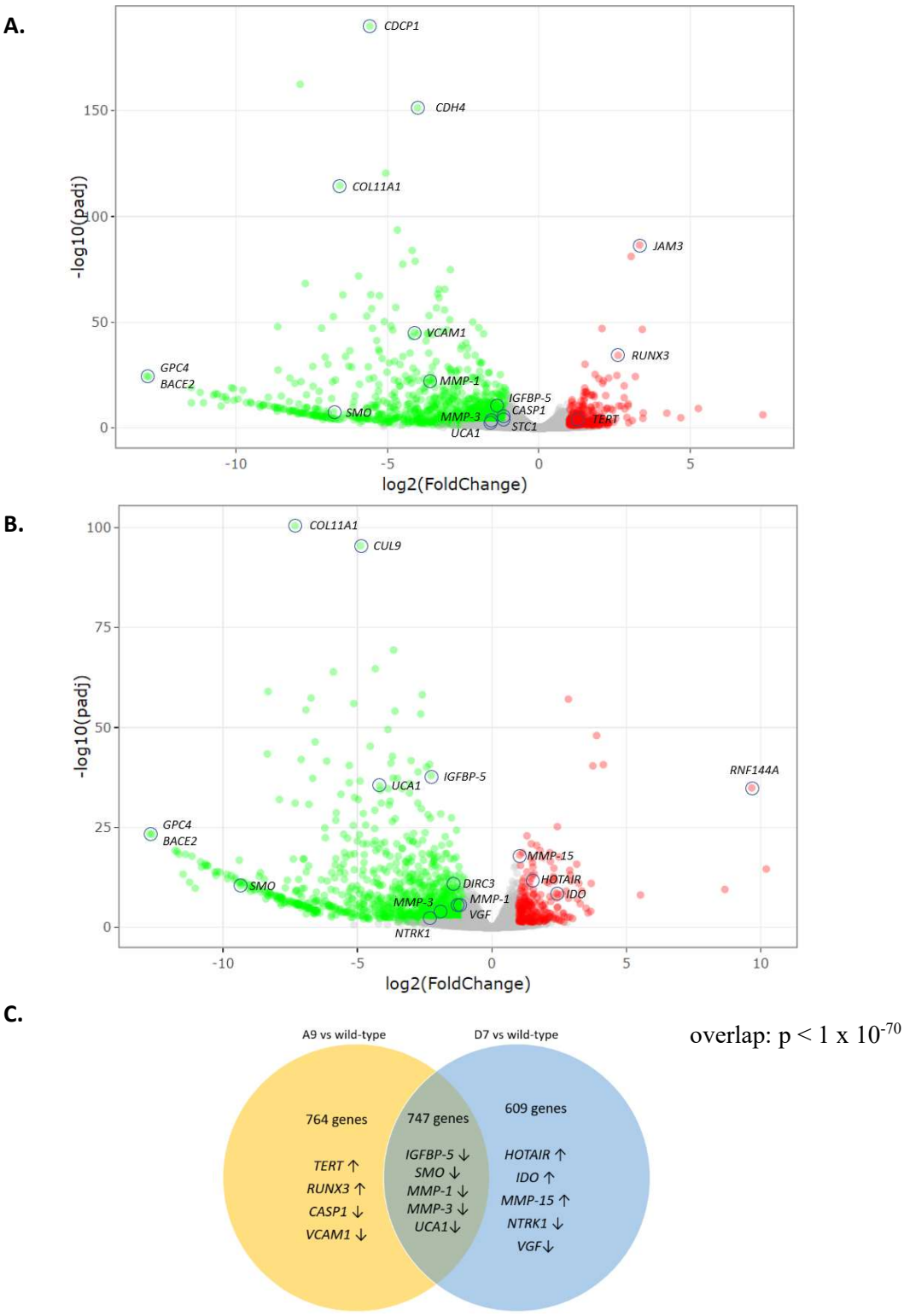


Figure 38. DEGs in the r11693806-edited MDA-T32 clones.

(A). Volcano plot illustrating changes in the gene expression profile in the "A9" clone as compared to the wild-type cells. DEGs are shown as color dots (red: upregulated, green: downregulated). Some prominent DEGs are marked. **(B).** Volcano plot illustrating changes in the gene expression profile in the "D7" clone as compared to the wild-type cells. **(C).** Venn diagram illustrating the overlap between DEGs identified in both CRISPR-edited clones. Names of certain genes are provided. Arrows indicate up- or down-regulation. Gene overlap was calculated using the hypergeometric distribution calculator.

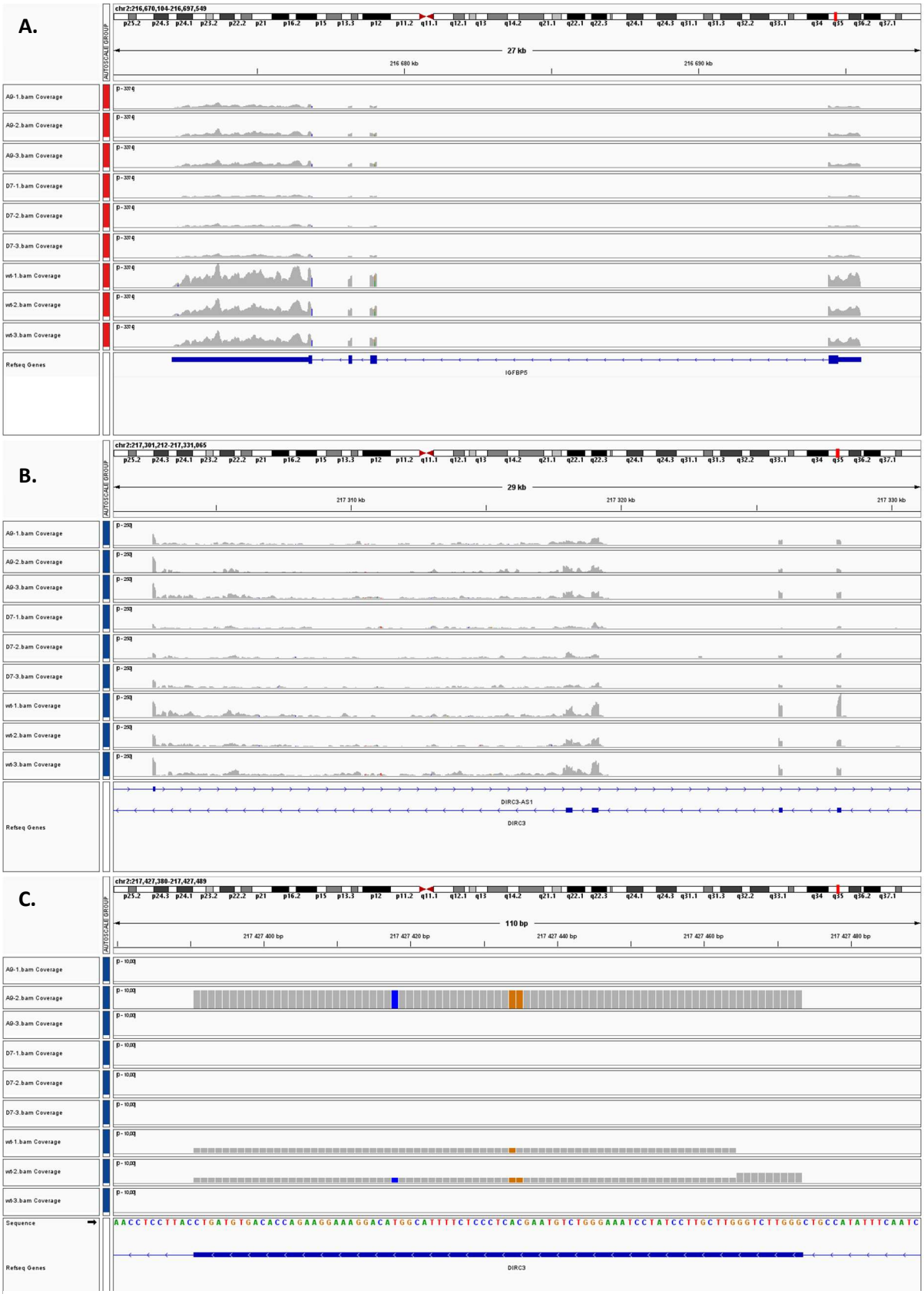


Figure 39. *IGFBP5* and *DIRC3* expression in the CRISPR-edited MDA-T32 clones in RNA-seq.

(A). Expression of *IGFBP5* visualized using the read coverage. **(B).** Expression of *DIRC3* (highly expressed exons) visualized using the read coverage. **(C).** Read coverage of the exon harboring rs11693806 (*DIRC3*). Abbreviation: wt, wild-type.

Gene ontology analysis identified 321 and 284 enriched GO terms enriched in the “A9” and “D7” clones, respectively (9.27% and 8.79% of all analyzed terms). The most significant terms are shown in Figure 40. While terms related to the migratory phenotype were not among these top signals, several GO terms related to cell migration were significantly enriched in “A9” (GO:0030334~regulation of cell migration, adjusted p = 0.016; GO:0016477~cell migration, adjusted p = 0.020) and “D7” clones (GO:0016477~cell migration, adjusted p = 0.012; GO:0030334~regulation of cell migration, adjusted p = 0.029; GO:0030335~positive regulation of cell migration, adjusted p = 0.035; GO:0030336~negative regulation of cell migration, adjusted p = 0.036).

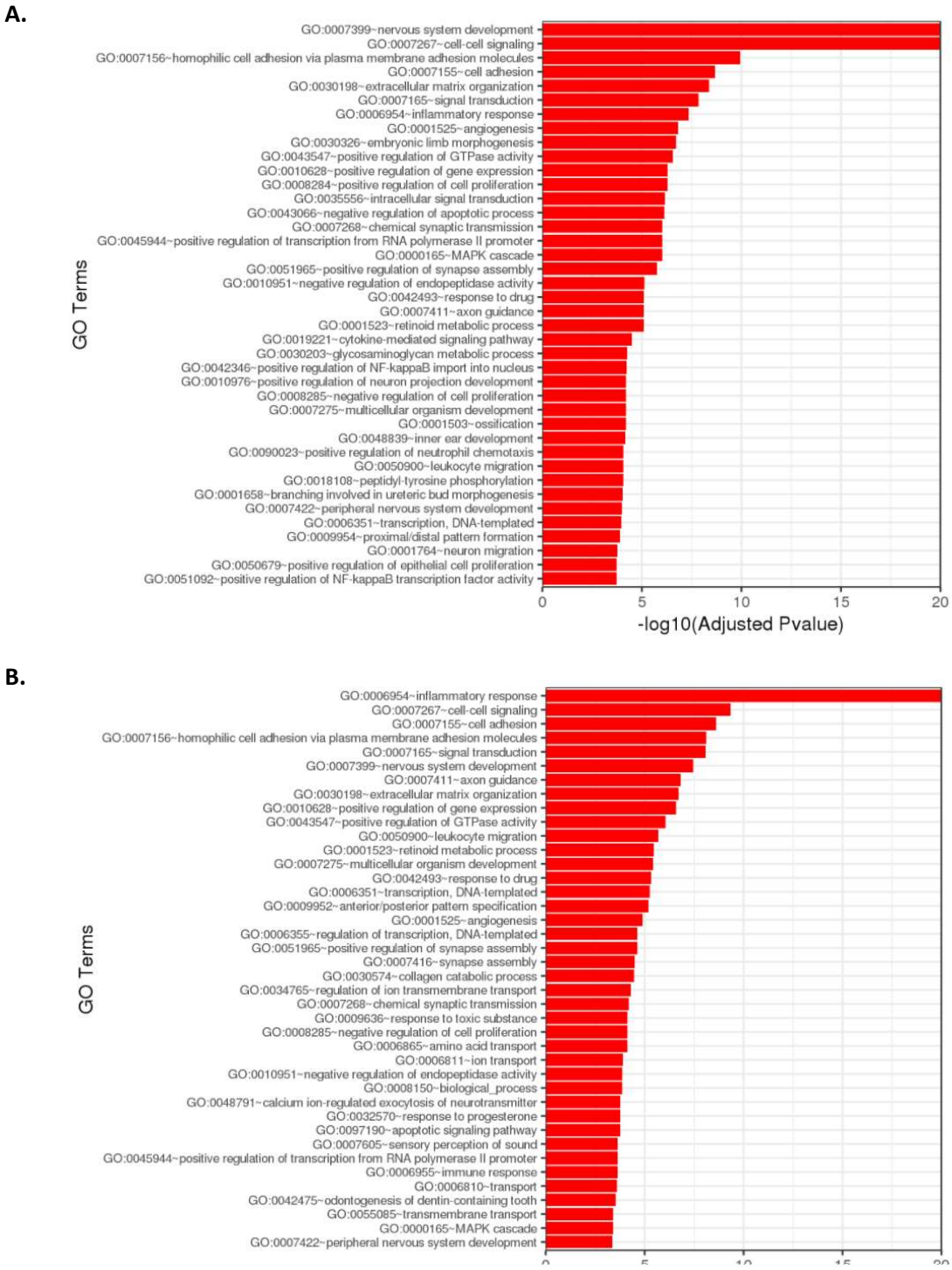


Figure 40. Gene ontology (GO) terms significantly enriched in the rs11693806-edited MDA-T32 clones.

(A). GO of the "A9" (rs116393806[G/-]) clone. **(B).** GO of the "D7" (rs116393806[G/G]) clone.

In summary, results of RNA sequencing indicated that genomic editing of rs11693806 affected expression of numerous genes, in particular *IGFBP5*. Transcriptomic effects observed when the rs11693806[C] allele was lost (either deleted, or substituted with rs11693806[G]) were similar. Genomic alterations of the variant could potentially influence multiple biological processes (e.g., mechanisms related to the regulation of cell signaling, migration, proliferation, apoptosis). Rs11693806 was previously reported to be a germline variant influencing the risk of developing DTCs in humans. Indeed, results of this study confirm that the variant is very likely to be functional. It is reasonable to assume that rs11693806 may influence the IGF-1 signaling pathway in thyroid follicular cells. This hypothesis merits evaluation in a follow-up study.

CHAPTER V: DISCUSSION

5.1. Overview of the study results

The study is the first to report that *DIRC3* is functionally implicated in thyroid carcinogenesis. I demonstrate that *DIRC3* lncRNA is downregulated in DTCs. Analysis of patients' clinicopathological data exposed a nominally significant relationship between the downregulation of *DIRC3* in thyroid cancers and the presence of distant metastases. A limited number of advanced cancers in the study cohort precludes definitive statements, however the putative prognostic value of *DIRC3* expression merits further evaluation. Importantly, this concept is supported by the analysis of TCGA data which indicated a significantly lower risk of disease recurrence in the patients diagnosed with PTCs with high expression of *DIRC3*. Furthermore, I have established that *DIRC3* is strongly co-expressed with *IGFBP5*, a gene located in its indirect proximity in chromosome 2q35.

Clinical findings are supported by the results of *in vitro* experiments. *DIRC3* transcripts localized preferentially in cell nuclei, thus reinforcing the notion that *DIRC3* could participate in the transcriptomic regulation of other genes. Indeed, knockdown of *DIRC3* downregulated expression of *IGFBP5* in multiple cancer cell lines. This phenomenon was, however, not observed when *DIRC3-202* (one of two major splice variants) was silenced selectively. Thus, a dissimilar biological functionality of *DIRC3* splice variants might be proposed. This concept requires elucidation in future studies. In addition, *DIRC3* conveyed phenotypic effects in PTC cell lines. Silencing of *DIRC3* augmented cell migration and invasiveness, diminished the starvation-induced apoptosis, but also reduced cellular activity in MTT assay. The pro-migratory phenotype was, however, reversed by a plasmid-meditated overexpression of *IGFBP5*. This attested that the anti-migratory role of *DIRC3* was chiefly mediated by its ability to regulate expression of *IGFBP5*. On the other hand, this was not the case in MTT assays, as the *IGFBP5* rescue experiment did not prevent the inhibitory effect generated by *DIRC3* silencing. Transcriptomic profiling performed in cancer cells experiencing knock-down of either *DIRC3* or *IGFBP5* indicated a significant redundancy in the activities of these two genes. Differentially expressed genes included multiple entities previously reported to be involved in thyroid tumorigenesis, e.g., *MALAT1*, *CDKN1A*, *MMP-1*. Gene ontology analysis revealed that downregulation of *DIRC3* disturbed most significantly genes involved in the regulation of cell migration. Furthermore, knockdown of *DIRC3* increased the susceptibility of thyroid cancer

cells to stimulation with IGF-1. This enhanced sensitivity to IGF-1 boosted the oncogenic AKT signaling pathway in thyroid cancer cells.

CRISPRa experiments successfully upregulated *DIRC3* in two cancer cell lines. Unexpectedly this was not accompanied by a positive influence on *IGFBP5* expression. Nonetheless, overexpression of *DIRC3* using CRISPRa effectively promoted the increase in number of viable cells in MTT assay. Accordingly, this phenotype appeared to be at least partially *IGFBP*5-independent.

I also demonstrated that rs11693806, a germline variant located in *DIRC3*, imposed transcriptomic alterations in MDA-T32 thyroid cancer cells. CRISPR-mediated modification of the heterozygotic cell line (rs11693806[C/G]) into the homozygotic isogenic derivative (rs11693806[G/G]) induced significant downregulations of *DIRC3* and *IGFBP5*. Similar transcriptomic effects were observed in cells harboring a small monoallelic deletion of the rs11693806[C] allele. Genomic modifications of rs11693806 induced pro-migratory effects that phenocopied outcomes observed in cancer cells in the *DIRC3*-silencing experiments. While the association between the genotype and gene expression was not observed in our clinical material (possibly due to a small sample size), rs11693806 has emerged as a putative cancer causative variant in *in vitro* studies.

5.2. Genes differentially expressed in response to silencing of *DIRC3*

RNA-seq technique was employed to attain transcriptome-based insights into interactions between expression of *DIRC3* and the expression of cancer driver genes. Silencing of *DIRC3* and *IGFBP5* revealed 58 common differentially expressed genes (DEGs) in MDA-T32 thyroid cancer cell line. Manual curation of literature data revealed that many of these DEGs have been convincingly implicated in thyroid carcinogenesis (Appendix Table). Among these DEGs there are several well-known oncogenes and tumor suppressors: *MALAT1*, *MMP-1*, *STC1*, *CDKN1A*. These four genes were upregulated by *DIRC3* silencing in MDA-T32 cell line.

Metastasis associated lung adenocarcinoma transcript 1

MALAT1 was found to be overexpressed in over 20 cancer types. Its upregulation was associated with poor patients' prognosis, drug resistance and metastases in the lung, breast and

liver cancers [191]. Despite strong evidence for its prooncogenic role in many tumor types, a non-canonical cancer suppressor-like role of *MALAT1* has been also suggested in breast cancers [191]. Upregulation of *MALAT1* was observed in PTCs, FTCs and Hürthle cell cancers [192-196]. Expression of *MALAT1* correlated positively with the size of thyroid cancers, presence of lymph node metastases, and advanced disease stage [193]. *MALAT1* promoted the proliferation, migration, invasiveness and angiogenesis of thyroid cancer cells (both *in vitro* and *in vivo*), and its knockdown reduced invasiveness of two PTC cell lines (B-CPAP and K1) [192, 197-199]. *MALAT1* was shown to drive the progression of DTCs by upregulating *IGF2BP2* (*insulin-like growth factor 2 mRNA-binding protein 2*) and *MYC* (a prominent oncogenic transcription factor [TF]) [198].

It may be hypothesized that upregulation of *MALAT1* partially contributed to the pro-invasive phenotype observed in thyroid cancer cells in the *DIRC3*-silencing experiments. Interestingly, *MALAT1* is transcriptionally activated by HIF-1, a transcription factor strongly induced by IGF-1 signaling [200-203]. This fact provides a functional link between *DIRC3*, IGF-1 signaling and *MALAT-1*.

Matrix metalloproteinase-1

MMP-1 (*matrix metalloproteinase-1*) encodes interstitial collagenase, an enzyme that disrupts peptide bonds in the triple-helical collagens. MMP-1 has a wide range of substrates that include collagens of type I, II, III, VII, VIII, X and XI, gelatin, entactin, tenascin, aggrecan, fibronectin and vitronectin [204]. MMP-1 is involved in the degradation of extracellular matrix (ECM). It also drives expression of *vascular endothelial growth factor receptor 2* (*VEGFR2*) and stimulates protease-activated receptor-1 (PAR-1), a protein promoting neo-angiogenesis [205]. The MMP-1/PAR-1 axis increases vascular permeability and enhances hematogenous spread of malignant cells. MMP-1 was reported to promote metastases in chondrosarcomas and carcinomas of the breast, colorectum, head and neck, and bladder [205].

MMP-1 was shown to be significantly upregulated in thyroid cancers as compared to benign thyroid tumors [206-209]. Overexpression of *MMP-1* was especially prominent in cancer cells that infiltrated thyroid capsule or metastasized to lymph nodes [210, 211]. Expression of *MMP-1* corelated positively with the tumor multifocality, presence of laryngotracheal invasion and metastases in thyroid cancers [206, 210, 212]. Accordingly,

upregulation of *MMP-1* could potentially contribute to the pro-invasive phenotype observed in the *DIRC3*-silencing experiments.

Cyclin dependent kinase inhibitor 1A

CDKN1A encodes p21 protein, a potent inhibitor of cell proliferation [213, 214]. *CDKN1A* is a p53-target gene and it act as one of the primary mediators of its antioncogenic activity. Firstly, p21 inhibits cyclin-dependent kinases (CDKs) and thus prevents formation of CDK/cyclin complexes (cyclin D-CDK4/6, cyclin E/A-CDK2 and cyclin B-CDK1). Accordingly, the cell cycle is arrested. Secondly, p21 inhibits proliferative cell nuclear antigen (PCNA), thus abrogating its role in the regulation of DNA synthesis and repair. Thirdly, p21 may bind certain transcription factors (e.g., E2F1, STAT3, MYC) and repress their transactivating functions. Fourthly, p21 interacts with caspases thus modulating the process of apoptosis [213].

Since the loss of *CDKN1A* by itself is insufficient to drive carcinogenesis, p21 is viewed as a tumor suppressor that primarily antagonizes oncogenes. Expression of p21 was shown to be a positive prognostic factor in colorectal, gastric, breast, ovarian, laryngeal and oral cancers [213]. However, the simple view presenting p21 as a tumor suppressor has been perplexed by observations which indicate that p21 may also possess some tumor-promoting properties [213]. Certain studies revealed that overexpression of p21 relates to a dismal prognosis in patients diagnosed with gliomas, sarcomas and carcinomas of the breast, liver, prostate, ovary, uterine cervix and esophagus. These contradictory findings have been at least partially explained by *in vivo* models that showed that the dual function of p21 could be related to its variable subcellular localization, expression of accompanying oncogenes, and the co-existence of *TP53* mutations [213, 214].

Loss of *CDKN1A* is detected in 12.5% of PTCs [215]. Silencing of *CDKN1A* promoted the anchorage-independent growth in K1 cell line [215]. On the other hand, elevated levels of p21 were frequent in undifferentiated, poorly differentiated and differentiated thyroid carcinomas, but rare in normal thyroid follicles [216, 217]. Nuclear staining of p21 decreased progressively with the raising pathological stage of PTCs [217, 218]. In contrast, it was also reported that the level of p21 was higher in large PTCs, and in thyroid cancers that extended beyond the thyroid capsule [219].

Accordingly, the biological role of induction of *CDKN1A* in response to *DIRC3* silencing is uncertain. Further studies are required to evaluate significance of this transcriptomic interaction.

Stanniocalcin 1

Stanniocalcin 1 has been originally discovered as a protein regulating the calcium and phosphate homeostasis in fish. Since stanniocalcin 1 is normally undetectable in the serum of mammals, stanniocalcin 1 is thought to act as a paracrine/autocrine factor in the higher vertebrates [220]. Stanniocalcin 1 is a potent inhibitor of two metalloproteinases, pregnancy-associated plasma protein -A and -A2 (PAPP-A and PAPP-A2) [221]. PAPP-A has a proteolytic activity towards IGFBP-2, IGFBP-4 and IGFBP-5, while PAPP-A2 cleaves IGFBP-3 and IGFBP-5. Degradation of IGFBPs liberates IGFs and increases the local bioavailability of these growth factors. This effect is, however, repressed by stanniocalcin 1, which is capable to inhibit the activity of PAPP-A and PAPP-A2. Accordingly, stanniocalcin 1 prevents the release of IGF-1 and inhibits the IGF-1/IGF-1R signaling axis [221].

STC1 has been also recognized as an oncogene. *STC1* promoted cell proliferation, epithelial-mesenchymal transition (EMT), invasiveness and metastases in cervical, ovarian, breast, laryngeal, esophageal, gastric, colorectal and lung cancers [220]. Mechanistically, these tumor-promoting events were related to the capability of *STC1* to modulate multiple signaling pathways, e.g., NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells), PI3K/AKT, NOTCH1 and JNK (c-Jun N-terminal kinase) [220]. It was also shown that upregulation of *STC1* associated with worse prognosis in the esophageal, gastric and breast cancer patients [220].

Two studies reported that the expression of *STC1* is altered in DTCs and in thyroid cancer cell lines [222, 223]. Silencing of *STC1* reduced proliferation of FTC cells *in vitro*, whereas the recombinant human stanniocalcin 1 produced opposite effects [224]. Moreover, overexpression of *STC1* was described as a negative prognostic marker in thyroid cancers [225, 226].

Accordingly, I hypothesize that overexpression of *STC1* observed in the *DIRC3*-silencing experiments could play a dual role. Firstly, it may act as a tumor-promoting event that enhances the invasiveness of thyroid cancer cells. Secondly, it is plausible that upregulation of *STC1* may constitute a negative-feedback mechanism *via* its inhibitory effect on PAPP-A and PAPP-A2. Accordingly, stanniocalcin 1 may prevent the excessive IGF-1 signaling elicited by the downregulation of *IGFBP5* and/or *DIRC3*. This feedback loop is especially probable since *STC1* is a transcriptional target of HIF-1 [225, 227-229], the transcription factor that is directly promoted by the IGF-1 signaling pathway [152, 188, 230].

Smoothened

Smoothened (*SMO*) was downregulated by silencing of *DIRC3*, as well by the genomic editing of rs11693806[C] allele in MDA-T32 cells. Its expression was, however, unchanged by the knockdown of *IGFBP5*. Downregulation of *SMO* observed in the context of *DIRC3* silencing experiments may thus constitute one of many possible explanations of the *IGFBP5*-indepentent phenotype observed in MTT assays.

SMO is a key component of the Hedgehog signaling pathway. The activity of SMO is normally repressed by Patched (PTCH), a cell membrane-bound receptor. Stimulation of Patched by the Hedgehog ligands leads to the release of SMO. SMO may subsequently relocate to the primary cilia to mobilize glioma-associated oncogene (GLI) transcription factors. GLIs are thus translocated into the cell nucleus where they activate transcription of their target genes (e.g., cyclin D, Snail, N-Myc, IGF-2) [231]. The hedgehog pathway is critical in the regulation of embryonic development, stem cell renewal, proliferation and cellular differentiation. The pathway is frequently upregulated in cancers. Skin basal cell carcinomas and medulloblastomas typically harbor somatic mutations in *SMO* or *PTCH*, while carcinomas of the breast, prostate, esophagus, stomach, biliary tract, pancreas and lung may be stimulated by the Hedgehog ligands [231].

The Hedgehog pathway was reported to be activated in thyroid neoplasms (including PTCs). This activity was shown to promote proliferation of thyroid cancer cells [232]. Interestingly, SMO has been recently linked to the regulation of IGF-1 pathway. SMO was shown to elevate levels of IGF-1R in cancer cells in non-canonical mechanisms that were related to the cellular localization and lysosomal degradation of IGF-1R [233].

To conclude, RNA-seq experiments highlighted several tumor-driving genes that were affected by the transcriptomic and genomic alterations in *DIRC3*. Many of these genes are known to be involved in thyroid carcinogenesis and their functions may potentially explain phenotypes brought by the disturbances in *DIRC3* expression. Still, these conceivable links are purely speculative at this time, and *in vitro* and *in vivo* experiments will be critical to validate these connections. Moreover, DEGs and GO terms illuminate only a part of the global transcriptomic rewiring that takes place in cancer cells. Understanding of this process may be enriched by use of additional bioinformatic tools capable of deconvoluting transcriptomic data into the signaling

pathway activation profiles. Application of this type of computational approach to RNA-seq data may further sanction the role of *DIRC3* in modulation of signaling cascades.

5.3. Role of *DIRC3* in cancer

5.3.1. Discovery of *DIRC3* and insights from GWAS

DIRC3 was originally described as a gene participating in the t(2;3)(q35;q21) translocation in familial renal cell cancers. This chromosomal aberration generated a gene fusion that partnered *DIRC3* with *HSPB1 associated protein 1* (*HSPBAP1*). *HSPBAP1* belongs to the *Jumonji C domain-containing protein* gene family that encompasses several epigenetic regulators [234]. While the enzymatic activity of HSPBAP1 has never been identified, the protein is presumed to influence chromatin remodeling and/or cellular stress response [78, 234]. *HSPBAP1* is upregulated in hepatocellular and prostate cancers, and its high expression level correlated with worse outcomes in prostate cancer patients [234]. The fragment of *HSPBAP1* preserved in the *DIRC3-HSPBAP1* fusion gene was only minimally truncated and thus it retained the heat shock protein family B member 1 (HSPB1)-binding and the Jumonji C domains. Accordingly, it was hypothesized that the retained fragment of *HSPBAP1* could preserve its native functionality. It was also suggested that the relocation of first two exons of *DIRC3* into the chimeric gene could play some oncogenic role [78, 234]. Nevertheless, the hypothesis has never been evaluated.

Novel insights into *DIRC3* have been provided by GWAS projects. Numerous SNVs in *DIRC3* were shown to associate with human height [85-88, 182-186], body weight [87, 235-238] and the risk of mitral valve prolapse [89]. Germline variants in *DIRC3* were also found to impact the risk of two malignancies: breast and thyroid cancers. The risk of estrogen receptor-positive breast cancers was shown to be modulated by the genotypes of rs16857609, rs11693806, rs16857611 and rs4442975 [16, 90-93, 239, 240]. Moreover, an intergenic breast cancer risk variant located downstream of *DIRC3*, rs13387042, was mapped to a putative regulatory element (PRE) that physically interacted with the genomic loci of *IGFBP5* and *DIRC3* [94]. Interestingly, two other studies identified a strong gene enhancer in this intergenic locus (corresponding to the position of a structural variant, esv3594306), which looped to the *IGFBP5* promoter and regulated its activity [96, 97]. This enhancer was bound by estrogen receptor (ERα) and forkhead box A1 (FOXA1), a master regulator of the estrogen receptor activity [96, 97]. A strong candidate for the breast cancer causality in this locus, rs4442975,

was shown to impact binding of FOXA1, regulate chromatin looping, and thus modulate expression of *IGFBP5* [97]. Very recently, another study used fine-mapping analysis from the Breast Cancer Association Consortium Study to refine three loci in chromosome 2q35 [93, 241]. Two of these loci (colocalizing with rs13387042 in the first locus, and rs138522813 and esv3594306 in the second locus) mapped to intergenic regions in the proximity of *DIRC3*, while the third locus (encompassing rs16857609) overlapped ca. 0.5 kb within *DIRC3*. These loci were annotated as PRE1, PRE2 and PRE3, respectively. *IGFBP5* was designated as a putative target gene for all of these PREs [241]. Additional association analysis identified 42 credible causal variants in PRE3 (the element located in *DIRC3*). Two of these variants (rs12694417, rs12988242) were prioritized based on their overlap with the active chromatin markers (indicating the open chromatin modifications, active histone markers [H3K27Ac, H3K4me1], or binding sites for TFs [FOXA1, GATA3, ERα]). Both PRE1 and PRE2 were found to upregulate the activity of *IGFBP5* promoter in a luciferase reporter assay. In contrast, PRE3 modestly enhanced the activity of *IGFBP5* promoter in one breast cancer cell line, however the upregulation did not reach significance in MCF-7 cell line, and was not observed in HepG2 (hepatocyte carcinoma), 293T (embryonic kidney) and HCT116 (colorectal carcinoma) cells. This outcome implied that the activity of PRE3 was cell-type specific. Furthermore, single nucleotide modifications of rs12694417 and rs12988242 did not alter activities of the luciferase reporter constructs, hence dismissing these SNVs located in PRE3 as putative causal variants [241]. In summary, these recent studies have suggested that some breast cancer risk variants located in chromosome 2q35 overlap PREs that modulate expression of *IGFBP5*.

Numerous germline variants in *DIRC3* have been associated with the risk of DTCs: rs966423 [60], rs6759952 [61], rs12990503 [63], rs16857609 [87, 242], rs11693806 [62] and rs772695095 [16]. Associations for rs966423 (OR = 1.34 for allele [C]) and rs6759952 (OR = 1.25 for allele [T]), have been successfully reproduced in diverse European, Asian and American populations [60, 61, 79-82]. Since rs966423 was also shown to influence the concentration of serum TSH, it was proposed that an insufficient stimulation of thyroid follicles by TSH could diminish the follicular differentiation and thus contribute to thyroid tumorigenesis [60]. The rs966423 genotype was also identified as an independent risk factor for the overall mortality in DTC patients (another study, however, failed to reproduce this finding) [83, 243]. Interestingly, while rs966423[C] associated with a higher DTC susceptibility, it was the alternative allele (rs966423[T]) that increased patients' mortality [83].

Subsequent GWAS projects identified even stronger DTC risk variants in *DIRC3*: rs12990503[G] (OR = 1.34-1.38), rs772695095[T] (OR = 1.46), rs16857609[G] (OR = 0.7) and rs11693806[C] (OR = 1.43) [62, 63, 84, 87, 242].

Very recently the EPITHYR consortium conducted GWAS and a fine-mapping analysis of the 2q35 locus in individuals of European and Pacific Islander ancestry. This project replicated the previously reported DTC risk variants (rs6579952, rs11693806, rs1382435, rs966423) and identified two novel susceptibility variants in *DIRC3*: rs57481445 (OR = 1.43) in Europeans, and rs3821098 (OR = 1.44) in a pooled analysis of all ethnicities [242, 244]. Both variants were significantly linked together in the trans-ethnic cohort [244]. Rs57481445 also correlated very strongly with rs11693806 (the variant analyzed in this study) in Europeans ($r^2 = 0.98$). Colocalization analysis that utilized the *expression quantitative trait locus* (eQTL) data from the PancanQTL study [245], highlighted additional intronic variants in *DIRC3*: rs16857611, rs16857609 and rs12990503 [244, 246]. These three SNVs were found to constitute eQTLs associated with the downregulation of *DIRC3* and *IGFBP5* in thyroid tissue. The strongest effect size was observed for rs12990503, the variant previously associated with increased risk of thyroid and breast cancers [244]. In conclusion, the EPITHYR study (published after completion of our project) has been the first to signal a connection between the DTC susceptibility, *DIRC3* genotype and expression of *IGFBP5*.

5.3.2. Novel findings on *DIRC3* in melanoma

Understanding of the function of *DIRC3* has been very recently enriched by a melanoma study. Coe *et al.* investigated the influence of two transcription factors critical in melanomagenesis (melanocyte inducing transcription factor [MITF] and SRY-related HMG-box 10 [SOX10]) on the global expression pattern of lncRNAs [247]. The study described 245 candidate melanoma-associated lncRNAs that were targeted by MITF and SOX10. Among the identified lncRNAs, *DIRC3* was prioritized due to: (i) its impact on the invasive expression signature and patients' clinical outcomes in TCGA data (patients with melanomas expressing low levels of *DIRC3* had shorter survival), (ii) existence of a gene orthologue in mice, and (iii) genomic proximity to *IGFBP5* and *tensin 1* (*TNS1*), two genes previously implicated in melanomagenesis [247]. Coe *et al.* demonstrated that MITF and SOX10 colocalized to two PREs in *DIRC3* and thus suppressed its expression. Accordingly, silencing of *MITF* or *SOX10* upregulated *DIRC3*. Since

DIRC3 and *IGFBP5* were located in the same topologically associated domain (TAD) in normal human epidermal keratinocytes, it was hypothesized that both genes could be transcriptomically related. In line with this notion, silencing of *DIRC3* using ASOs or CRISPR inhibition (CRISPRi) downregulated *IGFBP5* in melanoma cell lines. On the other hand, overexpression of *DIRC3* from a plasmid failed to upregulate *IGFBP5*. This observation indicated that *DIRC3* produced nuclear transcripts that generated regulatory effects only in cell nuclei. Silencing of *DIRC3* or *IGFBP5* promoted the anchorage-independent growth of melanoma cells, but had no effect on cell proliferation or migration [248]. Gene ontology enrichment analysis of RNA-seq data indicated that DEGs shared by *DIRC3* and *IGFBP5* were enriched in the regulators of cell migration, proliferation and differentiation. The study also offered a mechanical model in which recruitment of MITF and SOX10 to regulatory elements colocalizing with *DIRC3* repressed the expression of *DIRC3* and *IGFBP5*. *DIRC3* was thus proposed to be produce melanoma-suppressive non-coding transcripts that prevented binding of the *IGFBP5*-repressive proteins (MITF and SOX10) to PREs located in *DIRC3* gene [247]. While results of the melanoma study are principally in line with some of our findings, it is important to point out several differences and discrepancies.

Firstly, while authors of the melanoma study noted that suppression of *DIRC3* promoted the anchorage-independent growth, no phenotypic effects were observed in regards to cell proliferation and migration [248]. This outcome obviously differs from findings reported in our study (downregulation of *DIRC3* augmented cell migration and invasiveness, repressed MTT reduction and apoptosis, but did not alter the anchorage-independent growth of PTC cells). Conflicting results in soft agar assays might be explained by a more transient silencing effect produced by ASOs (used in our study), while the suppressive effect of CRISPRi (employed in the melanoma study) could have potentially lasted longer. Nevertheless, since initial steps of cellular transformation are critical for colony formation, and because GapmeRs efficiently suppressed *DIRC3* and *IGFBP5* for at least several days, it is also possible that our results faithfully indicated that the anchorage-independent growth was unchanged in thyroid cancer cells. On the other hand, the unaltered proliferation and migration of melanoma cells is puzzling. It was previously reported that overexpression of *IGFBP5* in melanoma inhibited cell proliferation, anchorage-independent growth, migration and invasiveness *in vitro*, and reduced melanoma growth and pulmonary metastasis *in vivo*. Conversely, silencing of *IGFBP5* in melanoma produced opposite, tumor-promoting effects [249, 250]. While these discrepancies may be potentially explained by different melanoma cell lines used in these studies, it is

noteworthy that silencing of *DIRC3* significantly affected the regulators of cell migration and proliferation in the GO enrichment analysis performed in the *DIRC3* melanoma study [247].

Secondly, the model offered by the melanoma study proposes that *DIRC3* lncRNA modulates binding of SOX10 and MITF to PREs located in *DIRC3* gene. Nevertheless, inspection of the proteomic and transcriptomic profiles of thyroid tissues (available from The Protein Atlas [251], GTEx and TCGA projects) indicates that both transcription factors have a very low mRNA expression and are not detectable on the proteomic level in normal thyroid and DTCs. Thus, the mechanistic model proposed for melanoma appears to be insufficient to explain the role of *DIRC3* in DTCs. It is, however, plausible that some thyroid-specific TFs may fit into a similar mechanical scheme. This hypothesis might be verified using the chromatin immunoprecipitation and/or RNA-protein pull-down experiments.

Thirdly, none of the thyroid cancer risk variants overlaps PREs reported in the melanoma study. This suggests that these germline variants are very unlikely to impact functions of SOX10 and MITF. Furthermore, results of our CRISPR-editing experiments indicated that one of major thyroid cancer risk variants, rs11693806, may be directly involved in carcinogenesis. The underlying mechanism of this influence remains unknown. It may be, however, speculated that rs11693806 might influence functions of unrecognized PREs that regulate *IGFBP5*. Given that the exon harboring rs11693806 had a very low expression in our RNA-seq experiments, it is plausible that this SNV imposes its transcriptomic influence by mechanisms that are not related to the function of *DIRC3* transcripts.

Fourthly, the melanoma study did not aim to profile expression of *DIRC3* splice variants. In this study I observed that selective silencing of *DIRC3-202* (one of two major splice variants) did not influence expression of *IGFBP5* (in a marked contrast to the effects produced when all major splice variants were targeted simultaneously). Interestingly, the GapmeR used in our study to silence *DIRC3-202* (anti-*DIRC3-202*) was highly similar to one of the GapmeRs used in the melanoma study. Inspection of designs of both GapmeRs revealed a shift by a single nucleotide in their sequences (relative to the sequence of target lncRNA). Accordingly, the GapmeR used by Coe e*t al.* is extremely likely to silence *DIRC3-202* selectively. Surprisingly, this GapmeR did efficiently downregulate *IGFBP5* in melanoma cell lines. Hence, I hypothesize that this experimental inconsistency might indicate dissimilarities in the biological functions of *DIRC3* splice variants across various malignancies. The exact source of this discrepancy remains obscure. Nevertheless, it is noteworthy that *DIRC3* locus contains an antisense gene (*DIRC3-AS1*). The sequence of *DIRC3-AS1* is complementary to the 3' fragment

of *DIRC3-202*, but does not overlap the sequence of another expressed splice variant (*DIRC3-203*). Although expression of *DIRC3-AS1* has not been profiled in TCGA, and it appears to be low in normal thyroid tissue in the GTEx project (median TPM of 0.13), it would be interesting to know whether *DIRC3-AS1* might impact functions of *DIRC3* splice variants.

In conclusion, in spite of the aforementioned experimental dissimilarities, the general findings reported in the melanoma study have reinforced discoveries reported in this thesis: *DIRC3* acts as a tumor suppressor with a cis-regulatory activity towards *IGFBP5*.

5.4. IGF-1 signaling and IGFBP5 in cancer

5.4.1. IGF-1: a multifaceted growth factor

Results of this study indicate that *DIRC3* orchestrates thyroid carcinogenesis at least partially *via* its cis-regulatory effect on *IGFBP5*. IGFBP5 is known to be a prominent actor in the IGF-1 signaling pathway. Indeed, I provide evidence that silencing of *DIRC3* increases the susceptibility of thyroid cancer cells to IGF-1 stimulation.

IGF-1 (also known as somatomedin C) is a peptide structurally related to insulin. IGF-1 is one of most abundant growth factors [95], and it promotes cell proliferation in various tissues in endocrine, paracrine and autocrine fashions [252]. Besides its pro-mitogenic effect, IGF-1 modulates cell motility, differentiation and apoptosis [253]. The endocrine production of IGF-1 is regulated by growth hormone (GH) and it primarily takes place in the liver. The GH/IGF-1 axis promotes somatic growth and development [254]. IGF-1 also contributes to the physiology of thyroid gland. For example, IGF-1 induces *thyroid transcription factor 2* (*TTF-2*) in follicular cells, what in turn promotes production of thyroglobulin [255].

The cellular activity of IGF-1 is mediated by IGF-1R, a cell membrane receptor expressed in almost all cell types. Binding of IGF-1R by the IGF ligands (IGF-1 or IGF-2) leads to the tyrosine phosphorylation in its intracellular domain. The phosphorylated domain recruits several adaptor proteins: SHC adaptor protein 1 (SHC), insulin receptor substrate 1 (IRS1) and insulin receptor substrate 2 (IRS2). These adaptor proteins produce divergent downstream responses. SHC interacts with growth factor receptor-bound-2 (GRB2) and Son-Of-Sevenless (SOS) to activate the MEK pathway, while IRS1 activates the AKT cascade. Consequently, both pathways increase protein synthesis, augment cell proliferation and inhibit apoptosis.

Additionally, IRS2 activates the FAK pathway, involving RAS homolog family member A (RHOA), focal adhesion kinase (FAK) and Rho-associated protein kinase (ROCK), to promote cellular motility *via* alterations in the expression of integrins [252]. The signaling output from IGF-1R is modulated by cross-talk with other RTKs, e.g., epidermal growth factor receptor (EGFR), vascular endothelial growth factor receptor (VEGFR), platelet-derived growth factor receptor (PDGFR), hepatocyte growth factor receptor (MET) and anaplastic lymphoma kinase (ALK). Additionally, TSH may interact with the IGF-1R signaling cascade. For instance, IGF-1 alone did not alter expression of *sodium/iodide symporter* (*NIS*) in thyroid follicular cells, but it was capable to inhibit the stimulatory effect of TSH on the expression of *NIS* [102]. Moreover, the mitogenic effect of TSH on thyroid follicular cells is markedly augmented when TSH is combined with IGF-1 [106]. One of the mechanisms responsible for this synergistic effect is related to forkhead box O (FoxO)-1, a transcription factor regulated by the AKT signaling pathway. Both TSH and IGF-1 promote a rapid exclusion of FoxO-1 from the cell nucleus, consequently reducing expression of certain tumor suppressors under its transcriptomic control (e.g., cyclin-dependent kinase inhibitors) [104].

Given the pro-mitogenic, pro-invasive and anti-apoptotic effect of IGF-1 signaling, it is not surprising that the pathway is heavily involved in carcinogenesis. An interesting example is Laron syndrome, a very rare type of dwarfism linked to mutations in *growth hormone receptor* (*GHR*) and the resultant congenital IGF-1 deficiency. No cases of malignancies were reported in a cohort of few hundred patients affected by the Laron syndrome (up to the age of 85), thus implying that the IGF-1 deficiency confers protection against cancer [254, 256]. Increased expression of IGF-1 and IGF-1R was described in various types of malignancies, e.g., in carcinomas of the lung, breast, prostate, ovary, pancreas, colorectum, and in sarcomas of the bones and soft-tissues. Additionally, elevated expression of IGF-1 and IGF-1R was associated with a higher incidence of distant metastases and inferior prognosis in many cancer types [106].

Expression of IGF-1 and IGF-1R is increased in PTCs and FTCs as compared to the histologically normal thyroid tissue, thyroid adenomas or nodular goiter [20, 103, 108, 257, 258]. Similar observations were made for IGF-2, which was overexpressed by malignant thyrocytes (especially in poorly differentiated tumors) [259]. A recent report indicated that IGF-1R signaling is promoted in some thyroid cancers by the overexpression of *insulin-like growth factor 2 mRNA-binding protein 3* (*IGF2BP3*), a positive posttranscriptional regulator of IGF-2. Elevated expression of *IGF2BP3* was observed in ca. 5% of PTCs, particularly in the tumors without known driver mutations [260]. These PTCs were found to harbor a

chromosomal rearrangement between *thyroid adenoma associated* (*THADA*) in chromosome 2 and *LOC389473* in chromosome 7 (located 12 kb upstream of *IGF2BP3*). It has been proposed that juxtaposition of *IGF2BP3* to *THADA*, a gene actively transcribed in thyroid tissue, could contribute to the overexpression of *IGF2BP3*. This consequently upregulated IGF-2, potentiated IGF-1R signaling, and stimulated cell proliferation, invasiveness and anchorage-independent growth [260]. Another possible mechanism of amplified IGF-1R signaling in thyroid cancers is overexpression of *Linc00210*. *Linc00210* is a lncRNA that serves as a sponge for miR-195-5p, thereby defusing its ability to downregulate *IGF-1R* [261]. Activation of the IGF-1R/AKT pathway by *Linc00210* was shown to increase the proliferation, migration and invasiveness of thyroid cancer cells [261].

Remarkable findings have been recently delivered by the UK Biobank study, which prospectively evaluated associations between incidence of 30 cancer types and the serum concentration of IGF-1. Follow-up of ~500,000 participants with a mean duration of 7.2 years revealed that the serum concentration of IGF-1 was positively associated with the overall cancer risk. Importantly, the relationships were highly diverse for specific cancer types. Out of all analyzed tumor types, the association was strongest for thyroid cancers (HR = 1.22 per 5 nmol/l increment of IGF-1; 95% CI, 1.05–1.42) [107, 109]. Interestingly, this finding is highly compatible with previous reports that described positive associations between the risk of DTC and the individual's tallness or acromegaly (both of which are related to the excessive levels of IGF-1) [20, 23-26, 262-265].

Several *DIRC3* germline variants were found to associate with human height [85, 86]. Given the positive relationship between body height and thyroid cancer risk, and considering the impact of IGF-1 on somatic growth and thyroid oncogenesis, a common molecular mechanism driving both conditions may be assumed. Accordingly, I hypothesize that epidemiological data associating human height with the increased incidence of thyroid cancers can be at least partially explained by shared hereditary variants in *DIRC3*. This link is likely related to the common transcriptomic influence on *IGFBP5*. Consequently, germline variants in *DIRC3* could impact the cellular responses to IGF-1. This hypothesis is partially supported by results presented in our study since the DTC susceptibility variant, rs11693806, impacted expression of *IGFBP5*. Moreover, I showed that downregulation of either *DIRC3* or *IGFBP5* provoked overactivity in the AKT signaling cascade in response to IGF-1 stimulation.

5.4.2. IGFBP5: the modulator of IGF-1/IGF-1R axis

IGFBP5 is the most evolutionarily conserved protein out of six known IGF binding proteins (IGFBPs). IGFBP5 is also the second most abundant member of the family in humans (only after IGFBP3) [253, 266]. IGFBPs are key regulators of the IGF-1/IGF-1R axis and contribute to the fine-tuning of IGF-1 bioavailability. Less than 1% of circulating IGF-1 is free, while the rest is complexed with IGFBPs, or forms ternary complexes consisting of IGF-1, IGFBPs and the acid labile subunit (ALS) glycoprotein. These structures significantly increase the half-life of IGF-1 (10 minutes for monomeric IGF-1, 30 minutes for the binary complexes, and 20 hours for the ternary structures) [95, 253]. Complexing of IGF-1 with IGFBPs increases delivery of IGF-1 to distant tissues. At the same time, however, IGF-1 needs to disembark from the complexes to engage and stimulate IGF-1R.

IGFBP5 has three domains: a conserved N-terminal domain containing the IGF-binding site, an unstructured linker (L-) domain, and a C-terminal domain that includes the thyroglobulin type-I repeat and the nuclear localization sequence (NLS) motifs [266]. N-domain has a high affinity to IGF-1, while the C-domain has much weaker affinity to IGF-1, but potently interacts with the components of ECM (collagens of type 3 and 4, vitronectin, laminin, fibronectin, thrombospondin, osteopontin, hydroxyapatite) [253]. Meanwhile, the central L-domain contains several proteolytic cleavage sites [95, 253]. The functional domains of IGFBP5 have discrete biological roles. For example, the N-terminal fragment exerted anti-proliferative and anti-apoptotic effects, while the C-terminal domain had an anti-migratory role in osteosarcoma cells [267]. Expression of *IGFBP5* is regulated by hormones and growth factors, e.g. prolactin inhibits its production in mammary tissue, while parathyroid hormone induces *IGFBP5* expression in the bones [253]. *IGFBP5* is also influenced by androgens and estrogens in the prostate, ovarian and breast cancers [268-273]. Importantly, expression of *IGFBP5* is stimulated by IGF-1 in many normal and malignant cell types [187-190]. In this instance, the IGF-1/IGF-1R/AKT signaling pathway promotes synthesis of HIF-1, the TF which directly stimulates the hypoxia response element located within the *IGFBP5* promoter [152, 188-190].

In the orthodox paradigm, IGFBP5 counteracts IGF-1 signaling by preventing binding of IGF-1 to IGF-1R. This inhibitory effect is observed on the organismal level, since mice overexpressing *IGFBP5* exhibited a severe reduction in body growth and had an increased perinatal mortality [274]. On the other hand, knockout of *IGFBP5* did not cause substantial

phenotypic alterations in mice, most likely due to the compensatory roles of other IGFBPs [253]. The inhibitory effect of IGFBP5 has been also observed on the cellular level. IGFBP5, either overexpressed endogenously or added exogenously, inhibited the IGF-1-induced cell growth in osteosarcoma cells and abrogated differentiation of osteoblasts [253]. Upregulation of *IGFBP5* also prevented the pro-survival IGF-1 signaling and had a pro-apoptotic influence on neurons, cardiomyocytes and mammary epithelial cells [253, 275, 276].

However, this conventional model is perplexed by data indicating that IGFBP5 may, in fact, promote IGF-1 signaling in specific contexts. For example, IGFBP5 augmented the stimulatory effects of IGF-1 on cell migration and DNA synthesis in vascular smooth muscle cells [187, 277]. Castration led to the upregulation of *IGFBP5* in murine prostate cancer models. This in turn potentiated IGF-1 signaling, elicited androgen-independence and led to cancer progression [278]. Silencing of *IGFBP5* in human colon carcinoma cells reduced mitogenic effect of IGF-1 [279]. Similarly, proliferation of ionocytes in zebrafish could be abrogated by knockout of either *IGF-1R* or *IGFBP5* [280, 281]. Expression of human *IGFBP5* increased ionocyte proliferation in the low calcium environment, while IGFBP5 mutant proteins that were unable to bind IGF-1 did not produce this effect [279]. However, this influence on inonocyte proliferation was context dependent, since IGFBP5 inhibited IGF-1 signaling under the normal calcium conditions [253, 279]. It has been recently shown that this divergent response is likely related to changes in the expression of stanniocalcin 1, level of which is tightly controlled by the calcium level. As discussed above, stanniocalcin 1 is a PAPP-A inhibitor, while PAPP-A functions as a metalloproteinase degrading IGFBP5 [281].

Henceforth, IGFBP5 acts as a molecular switch that either inhibits or promotes IGF-1 signaling in diverse cell and tissue contexts. This diversity can be principally attributed to two mechanisms. Firstly, the aforementioned proteolysis of IGFBP5 causes a large release of stored IGFs and consequently promotes oncogenic effects. Among protease that cleave IGFBP5 are PAPP-A, PAPP-A2, thrombin, elastase, cathepsin G, complement C1s, a disintegrin and a metalloprotease 9 (ADAM 9), ADAM 12S, MMP-1, MMP-2 and prostate serum antigen (PSA) [253, 282]. Notably, it was shown that a protease-resistant form of IGFBP5 was a significantly more potent inhibitor of IGF-1 than the native protein [283]. Accordingly, abundant IGFBP5 may initially inhibit IGF-1 signaling by sequestrating IGFs. However, boosted production of IGFBP5 may promote IGF-1 signaling later on by creating a local reservoir of IGF-1/IGFBP5 complexes. Once the reservoir is degraded, a large amount of IGF-1 is released and robustly stimulates nearby cells. Interestingly, our transcriptome profiling experiment indicated that two

genes implicated into the degradation of IGFBP5 were differentially expressed upon silencing of *DIRC3* or *IGFBP5*: *MMP-1* and *STC1*. This may suggest that these genes participate in feedback mechanisms instigated by alterations in IGF-1 signaling.

Furthermore, the activity of IGFBP5 is modulated by its biding to the components of ECM. Attachment of IGFBP5 to heparin-like glycosaminoglycans protects IGFBP5 from proteolytic degradation [253]. Furthermore, physical interactions of IGFBP5 with certain ECM components (e.g. vitronectin in smooth muscle cells, or hydroxyapatite in the bones) may locally potentiate IGF-1 signaling [253, 284, 285]. On the other hand, interactions of IGFBP5 with fibronectin may inhibit the pathway [286]. The components of ECM may also reciprocally modulate production of IGFBP5, e.g., laminin and type IV collagen reduces, while vitronectin elevates the level of IGFBP5 in the porcine smooth muscle cells [284].

IGFBP5 also appears to function in an IGF-independent way. It was shown that IGFBP5 mutants with a negligible affinity to IGFs abrogated somatic growth in mice [287]. A direct interaction of IGFBP5 with α2β1 integrins was shown to promote attachment of MCF-7 breast cancer cells to vitronectin. This effect was evident for both the native IGFBP5 protein, as well as for the truncated mutant protein devoid of the IGF-biding capability [288]. A non-IGF binding mutant of IGFBP5 promoted migration of vascular smooth muscle cells by interacting with the cell-surface heparan sulfate proteoglycans [277]. Moreover, IGFBP5 is not only secreted extracellularly, but it also localizes to cell nuclei. Its intranuclear transport is driven by the nuclear localization sequence (NLS) motif present in its C-terminal domain. Subsequently, IGFBP5 may interact in the nucleus with certain nuclear receptors: retinoic acid receptor (RAR) and vitamin D receptor (VDR) [95]. While the exact role of these intranuclear interactions remains to be elucidated, it was shown that complexing of IGFBP5 with VDR may reduce responses to 1,25-dihydroxyvitamin D3 in the osteoblast-like cells. This effect was, however, produced only by IGFBP5 that was produced endogenously [253].

In summary, IGFBP5 plays a highly complex role. It primarily acts as a molecular switch regulating the IGF-1 signaling pathway. The ultimate influence of IGFBP5 is dependent on the bioavailability of IGF-1, abundance of IGFBP-degrading proteases and the composition of ECM. Furthermore, IGFBP5 may exert distinct actions when produced endogenously or delivered externally. The protein also has a poorly characterized IGF-independent functions, exerted both extracellularly and in the cell nucleus. Thus, it is hardly surprising that IGFBP5 may produces various, often conflicting, phenotypic effects in different biological models.

5.4.3. *DIRC3* and *IGFBP5* in cancer – a functional dichotomy

Results of this study designate *DIRC3* as a thyroid cancer suppressor with a prominent anti-invasive function. This conclusion is advocated by the pro-invasive phenotype and transcriptomic signature elicited by the suppression of *DIRC3* in thyroid cancer cells. The pro-migratory effect is, however, reversed by overexpression of *IGFBP5*. This indicates that *IGFBP5* governs alterations in the migratory potential produced by *DIRC3*. I also observed that silencing of *DIRC3* protected thyroid cancer cells against the starvation-induced apoptosis, what appears to be consistent with the pro-survival function of IGF-1 signaling [95]. On the other hand, silencing of *DIRC3* limited the MTT reduction rate in cancer cells. This influence appeared to be at least partially independent from *IGFBP5*.

This seemingly paradoxical outcome implies a functional dichotomy of *DIRC3*. Noteworthy, this functional divergence resembles discrepancies in the clinical impact of certain *DIRC3* germline variants. In particular, rs966423[C] constitutes a risk allele for the development of DTCs, however it is the alternative allele, rs966423[T], that associates with augmented patients' mortality [83]. It has been proposed that the former allele may contribute to early steps of tumorigenesis (the cancer initiation), while the latter allele promotes cancer progression and metastases [83]. In light of these findings, it may be hypothesized that *DIRC3* acts as a reversible transcriptomic switch between the pro-proliferative and pro-invasive phenotypes in thyroid cancer cells. This form of functional plasticity is known to be vital in the regulation of tumor progression in various malignancies. It may be also proposed that the dichotomous role of *DIRC3* reflects the functional heterogeneity of *IGFBP5*. Additionally, putative *IGFBP5* - independent mechanisms may further complement the overall biological complexity.

A small study reported that *IGFBP5* was upregulated in PTCs in comparison to unmatched normal thyroid and goiter samples [98]. Our study does not support this observation as neither clinical specimens of DTCs, nor TCGA samples presented significant alterations in expression of *IGFBP5*. Two studies provided indirect evidence on the plausible oncogenic role of *IGFBP5* in PTCs. Liu *et al.* reported that *IGFBP5* was silenced by a tumor-suppressive miR-204-5p, a microRNA found to be frequently downregulated in PTCs. Overexpression of miR-204-5p in PTC cell lines (BCPAP and TPC-1) suppressed their proliferation and induced a cell-cycle arrest. This antioncogenic effect was, however, partially reversed by overexpression of *IGFBP5* [99]. Another study also reported that *IGFBP5* transcripts are targeted by miR-204

[289]. *Urothelial cancer associated 1* (*UCA1*), a lncRNA gene upregulated in PTCs, increased the proliferation, migration and invasiveness of TPC-1 cells. Mechanistically, *UCA1* transcripts sponged miR-204 and consequently upregulated *IGFBP5* [289]. Of note, *UCA1* was a gene that was significantly downregulated in our study in both rs11693806-edited clones of MDA-T32 cell line. This finding could indicate yet another level of interaction between the genotype of *DIRC3* and expression of *IGFBP5*.

A wealth of studies has denoted *IGFBP5* to be either oncogene or tumor suppressor. Overexpression of *IGFBP5* was shown to associate with a poor histological differentiation and dismal prognosis in the ovarian, breast, renal and bladder cancers, and in gliomas [95, 290-298]. Expression of *IGFBP5* was higher in invasive breast carcinomas (particularly in the estrogen receptor-positive tumors) when compared to the histologically normal mammary tissue [269, 294, 295]. Significant overexpression of *IGFBP5* was observed in primary breast cancers that metastasized to axillary lymph nodes [292, 299, 300]. Upregulation of *IGFBP5* in breast cancers was also linked to the acquired resistance to PI3K inhibitors [301]. On the other hand, *IGFBP5* inhibited the estradiol-induced growth in MCF-7 cell line [302]. Upregulation of *IGFBP5* inhibited the proliferation, viability and motility in breast cancer cell lines and in murine xenografts [303, 304]. No such influence was observed when the recombinant IGFBP5 protein was applied, hence suggesting that intracrine mechanisms are vital in these activities [303]. It was also shown that upregulation of *IGFBP5* is mandatory to trigger the involution of mammary gland in weaning mammals, or to induce regression of breast cancers treated with Wnt inhibitors [150, 276, 305]. *IGFBP5* was downregulated in the tamoxifen-resistant breast cancer cell lines, cancer xenografts and in the clinical carcinoma specimens [306, 307]. Preoperative treatment with tamoxifen and exemestane (an antiestrogen) upregulated *IGFBP5* in breast cancers, and the magnitude of this increase correlated positively with the reduction in tumor size during anti-estrogen therapy [271, 308]. In fact, knockdown of *IGFBP5* elicited resistance to tamoxifen and downregulated ERα *in vitro* and *in vivo.* The tamoxifen sensitivity was, however, restored with the recombinant IGFBP5 protein [309]. Downregulation of *IGFBP5* was also associated with a shorter survival in breast cancer patients that were treated with tamoxifen [309].

In serous ovarian cancers *IGFBP5* was reported to be either down- or up-regulated, and its overexpression was associated with worse prognosis in patients [291, 298]. Moreover, *IGFBP5* was downregulated in ovarian cancers obtained from patients demonstrating responses or stabilizations during hormonal therapy [272]. Since *IGFBP5* was downregulated by

17β-estradiol in ovarian cancer cells, it has been suggested that the expression of *IGFBP5* may serve as a marker of tumor response during endocrine therapy [272]. Moreover, IGFBP5 protein inhibited the ovarian cancer angiogenesis *in vitro* and *in vivo*. This antiangiogenic influence was related to the inhibition of *VEGF-A* and *MMP-9*, reduced phosphorylation of AKT and ERK, and upregulation of NF-κB in ovarian cancer cells [310, 311].

IGFBP5 was reported to be upregulated in prostate cancers in comparison to benign prostatic hyperplasia [273]. Castration of mice bearing prostate cancers was described to produce conflicting effects: overexpression or downregulation of *IGFBP5* [273, 278]. It has been proposed that upregulation of *IGFBP5* after castration could potentiate the antiapoptotic and mitogenic effects of IGF-1 in prostate cancers. *IGFBP5* also promoted growth and radiosensitivity of prostate cancers, while its silencing decreased the activity of MAPK signaling pathway and induced a cell-cycle arrest [278, 312].

IGFBP5 was upregulated approximately 20-fold in pancreatic cancers in comparison to nonmalignant pancreatic duct specimens. Immunohistochemistry staining of pancreatic cancers revealed a strong focal staining of *IGFBP5* in cancer islets located at the tumor edges [313]. Overexpression of *IGFBP5* in pancreatic cancer cells that were present at the epithelial-stromal boundary was also reported in another study [314]. Similar observations were made for the circulating pancreatic cancer cells. Hence, it has been hypothesized that alterations in *IGFBP5* expression could influence the plasticity of cancer cells – promote a shift from the proliferative to metastatic phenotype [314]. *In vitro* experiments indicated that overexpression of *IGFBP5* downregulated AKT signaling and reduced growth of serum-starved PANC-1 pancreatic cancer cells (*KRAS* mutant, *PTEN* wild-type). In contrast, *IGFBP5* overexpression in serum-starved BxPC-3 cells (*KRAS* wild-type, *PTEN* mutant) promoted cell proliferation by activating AKT and MAPK signaling pathways [315]. Augmented growth of *IGFBP5*-overexpressing BxPC-3 cells was partially linked to the reduced level of p21 [315]. Interestingly, this cyclin-dependent kinase inhibitor was upregulated in our study by silencing of either *DIRC3* or *IGFBP5* in MDA-T32 cells.

Upregulation of *IGFBP5* increased the migratory potential of bladder cancer cell lines [316]. Similarly, IGFBP5 was shown to promote invasiveness of glioblastoma cells, but at the same time *IGFBP5* inhibited their proliferation [317]. Overexpression of *IGFBP5* reduced growth of murine osteosarcoma xenografts, reduced incidence of pulmonary metastases, while silencing of *IGFBP5* promoted osteosarcoma progression [318]. Opposite effects were observed in some *in vitro* studies, in which knockdown of *IGFBP5* in osteosarcoma cell lines

inhibited cell proliferation, migration, invasiveness, anchorage-independent growth, and promoted apoptosis [267, 318, 319]. *IGFBP5* was downregulated in cervical cancers, and in head and neck squamous cell carcinomas (HNSCC) [320-322]. Recombinant IGFBP5 blocked IGF-1 signaling in HNSCC cells, and consequently reduced their proliferation and migration. Similar phenotypes were produced by endogenous overexpression of *IGFBP5*, while knockdown of *IGFBP5* augmented growth of murine HNSCC xenografts [320]. In gastric cancers *IGFBP5* was shown to coordinate up-regulation of p53 and inhibit cell viability and colony formation [323]. Moreover, *IGFBP5* was shown to attenuate proliferation and migration in hepatocellular carcinoma (HCC) due to its inhibitory influence on the IGF-1-AKT signaling pathway [324]. However, another study showed opposite effects, since proliferation of HCC cells was inhibited by the knockdown of *IGFBP5* [325]. In melanoma, overexpression of *IGFBP5* inhibited cell proliferation, anchorage-independent growth, EMT, migration and invasiveness *in vitro*, and reduced tumor growth and metastasis *in vivo* [249]. These effects were attributed to the inhibition of IGF-1R-MEK-ERK signaling pathway and a subsequent downregulation of HIF-1α. Opposite, melanoma-promoting effects were observed once *IGFBP5* was knocked down [326].

In conclusion, roles of *IGFBP5* in malignancies are very diverse. A large share of this heterogeneity can be attributed to the complexity of molecular mechanisms in which IGFBP5 takes part: the IGF-dependent and IGF-independent functions, the modulating effect of extracellular milieu, and the cross-talk with other signaling pathway. Moreover, some studies suggest that functional dichotomy of *IGFBP5* may contribute to the phenotypic plasticity of cancer cells. Accordingly, this concept may also apply to *DIRC3*, the modulator of *IGFBP5* expression. This lncRNA gene may potentially act as a transcriptomic switch: promote tumor growth, but also prevent distant spread of cancer cells. The hypothesis, however, needs to be corroborated in future studies. Furthermore, the prospect of the *IGFBP5*-independent activities should be also explored.

5.5. Mechanistic model of the interplay between *DIRC3* and IGF-1 signaling

Using literature data and the results of this study, I propose a mechanistic model presented in Figure 41. In this paradigm, *DIRC3* produces a cis-regulatory transcript that upregulates *IGFBP5* (the exact molecular mechanism of this influence remains to be elucidated). The resulting upregulation IGFBP5 promotes sequestration of IGF-1 in the extracellular space. This limits accessibility of IGF-1 to IGF-1R. Consequently, the flux from the oncogenic IGF-1R/AKT/mTOR signaling cascade is reduced. Among downstream effectors of mTOR are eukaryotic translation initiation factor 4E binding protein (4E-BP1) and p70 S6 kinase (S6K). Phosphorylation of 4E-BP1 by mTOR complex 1 (mTORC1) decouples 4E-BP1 from eukaryotic translation initiation factor 4E (eIF4E). Hence, the cap-dependent translation of multiple proteins is derepressed. One of the upregulated proteins is HIF-1α. Moreover, the phosphorylation of S6K activates ribosomal protein S6 and promotes protein translation [327]. The upregulated HIF-1 (a heterodimer composed of alpha and beta subunits) binds to DNA sequences termed hypoxia response elements (HREs). HREs are enhancers located within promoters of numerous genes, many of which are critical in oncogenesis. HREs are also present in the promoters of *IGFBP* genes [327]. Hence, HIF-1 directly induces expression of *IGFBP5* [152, 188-190]. Consequently, a negative feedback loop involving IGF-1, IGF-1R/AKT/mTOR cascade, HIF-1 and IGFBP5 is created. Importantly, when expression of *DIRC3* is reduced, *IGFBP5* is also downregulated, and the regulatory loop is disrupted. Consequently, thyroid cancer cells become more sensitive to IGF-1 and experience a stronger activation of the oncogenic AKT signaling.

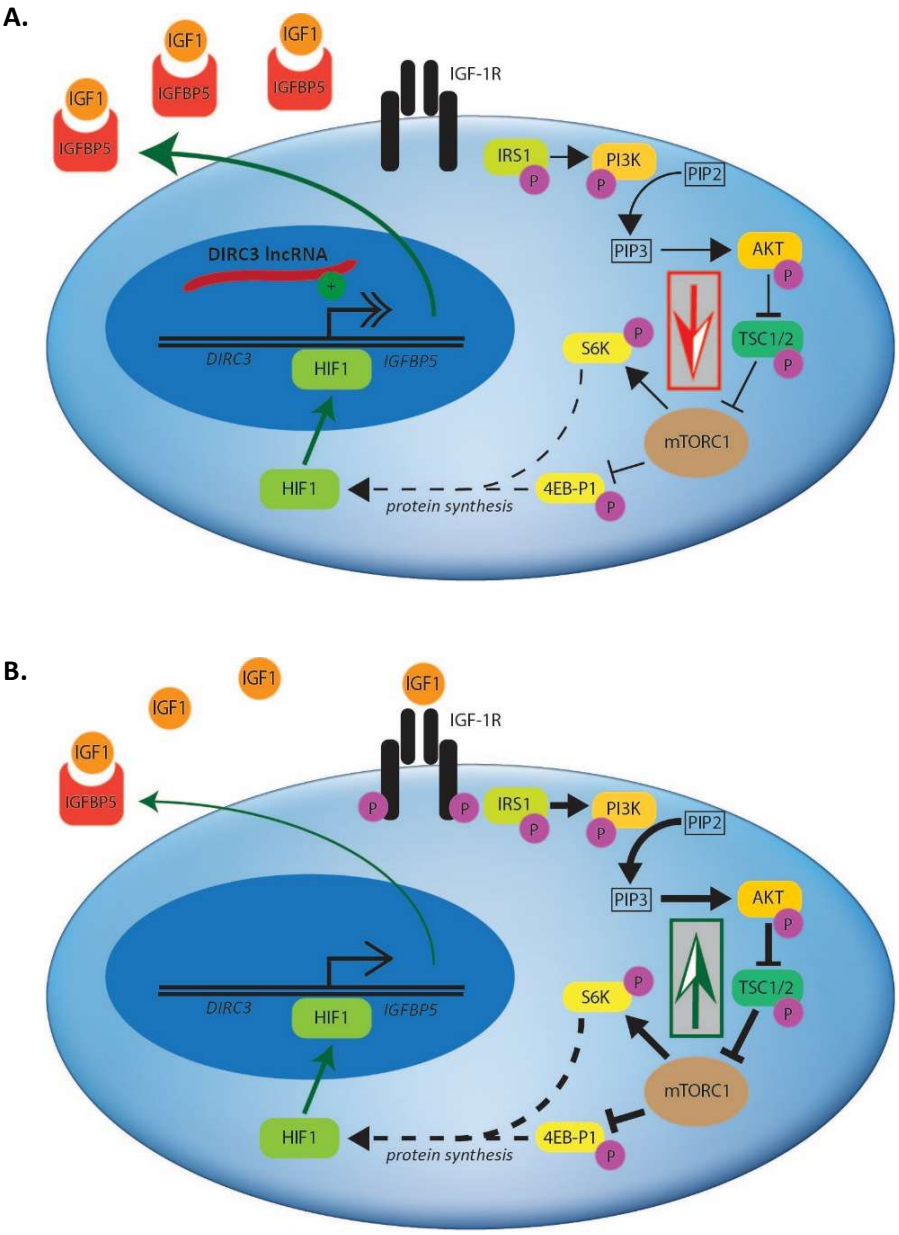


Figure 41. Model of the putative role of *DIRC3* in the IGF-1 signaling pathway in thyroid cancer cells.

The model is based on literature data and the results of this study. **(A).** IGF-1R is stimulated by IGF-1, what promotes a signaling cascade involving insulin receptor substrate 1 (IRS1), PI3K, AKT and mTORC1. Augmented flux in the AKT signaling pathway drives cancer progression. Additionally, mTORC1 upregulates protein synthesis. Among the upregulated proteins are components of hypoxia-inducible factor 1 (HIF-1), a transcription factor promoting expression of many oncogenes. Among transcriptomic targets of HIF-1 is *IGFBP5*, which is consequently upregulated. In thyroid cells that retain an adequate expression of *DIRC3*, this lncRNA further promotes expression of *IGFBP5*. Production and extracellular release of IGFBP5 limits the bioavailability of IGF-1, and thus prevents its excessive stimulatory effect. Accordingly, a negative feedback loop is formed. **(B).** Expression of *IGFBP5* is downregulated in thyroid cancers with low expression of *DIRC3*. Accordingly, the bioavailability of IGF-1 is higher and the oncogenic AKT signaling is augmented. The negative feedback loop is impaired.

Arrow thickness in the schemes indicates the magnitude of signaling flow.

5.6. Study limitations

The strength of this study lies in a wealth of clinical, bioinformatic and experimental data that have equivocally implicated *DIRC3* into thyroid carcinogenesis. Nevertheless, the study has some limitations.

Firstly, the number of DTCs analyzed in the clinical part of this study was rather small, especially in regards to the number of metastatic cancers. Accordingly, the observed clinical associations must be interpreted with a great caution and should be confirmed in large patients' cohorts. Moreover, PTCs constituted ca. 80% of all analyzed cancer cases. While our results were generalized for all DTCs, it also possible that the reported discoveries may hold true only for specific histological cancer types.

Secondly, I observed a possible differential biological role of *DIRC3* splice variants. In particular, silencing of *DIRC3-202* did not change expression of *IGFBP5*, and had no influence on the results of MTT assays. On the other hand, a GapmeR very likely to selectively target *DIRC3-202* was reported to have a phenotypic and transcriptomic influence on melanoma cells [247]. These observations could imply that *DIRC3* splice variants have diverse functions in various tumor types. While I was unable to further address this hypothesis within the scope of this study, future studies may refine this concept.

Thirdly, Western blot experiments have revealed that knockdown of *DIRC3* sensitizes thyroid cancer cells to the stimulatory effect of IGF-1. This outcome strongly mirrored the impact of *IGFBP5* silencing, and was completely reversed by the addition of exogenous recombinant IGFBP5 protein. Accordingly, the IGF-1-sensitizing phenomenon has been attributed to the diminished production of IGFBP5. At this moment, however, I cannot provide direct evidence regarding the level of IGFBP5 protein. Despite my best efforts the enzyme-linked immunosorbent assay (Human IGFBP-5 DuoSet ELISA DY875, R&D Systems) failed to detect IGFBP5 in biological samples (cell lysates and supernatants obtained from various cell lines). I attribute this failure to technical problems with the assay since even high concentrations of the recombinant human IGFBP5 protein produced very faint signals in ELISA. Instead, protein immunoblotting experiments utilizing alternative anti-IGFBP5 antibodies are in progress.

Fourthly, I used Cas9/CRISPR to edit rs11693806, the *DIRC3* germline variant. Phenotypic and transcriptomic alterations generated by genomic editing of rs11693806 were

attributed to the on-target activity of the utilized sgRNA. Successful genomic modification of the variant was documented using the T7EI assay, rhAmp genotyping PCR assay and Sanger sequencing. Moreover, no off-target editing in the protein-coding sequences was detected using WES. However, a full confidence on the editing specificity can be only provided by whole genome sequencing.

Next, CRISPRa produced a robust upregulation of *DIRC3* in thyroid cancer cell lines. Nevertheless, this overexpression had no influence on the level of *IGFBP5* transcripts. This outcome was unexpected given the consistent transcriptomic reciprocity between *DIRC3* and *IGFBP5* observed in the gene silencing and *IGFBP5* overexpression experiments. This perplexing observation has not been elucidated so far. While CRISPRa was expected to recapitulate the endogenous expression of *DIRC3* more faithfully than the plasmid-mediated overexpression (which failed to produce any transcriptomic or functional effects in the melanoma study) [247], it is plausible that CRISPRa was still insufficient to reproduce all aspect of the transcriptomic regulation of *IGFBP5*. More research on this issue should ensue.

Furthermore, I did not plan to provide a definitive mechanism by which *DIRC3* transcripts generate their transcriptomic impact. While our study indicates that *DIRC3* has a cis-regulatory role, more molecular details are anticipated. Functions of lncRNAs are typically related to their interactions with chromatin, proteins and other RNA species. Accordingly, upcoming projects may employ methods such as chromatin immunoprecipitation (ChIP), RNA immunoprecipitation (RIP) or the chromatin conformation profiling techniques. In the light of results of the melanoma and breast cancer studies, it is possible that some functions of *DIRC3* are related to the presence of PREs within its locus. These elements may recruit proteins that are either repressive (as MITF and SOX10 in melanoma) or permissive (as FOXA1 and ERα in breast cancer) in regards to the expression of *IGFBP5*. Since TFs evaluated in the melanoma and breast cancer studies are not functionally relevant in thyroid tissue, and because the defined PREs do not overlap the DTC risk variants, I presume that these particular TFs and PREs are not involved in thyroid oncogenesis. It is, however, plausible that *DIRC3* may control activity of some thyroid-relevant TFs in unrecognized PREs that harbor the DTC risk variants. To elucidate this hypothesis, it is desirable to identify putative biding sites for the thyroid-relevant TFs, assess TF occupancy (e.g., by employing ChIP), and/or evaluate chromatin accessibility with techniques such as DNase-seq or ATAC-seq (Assay for Transposase-Accessible Chromatin sequencing).

Finally, the experimental part of this study employed several validated thyroid cancer cell lines. Nevertheless, future studies should also include *in vivo* experimentations, and ideally shall employ patient-derived xenografts (PDx). These cancer models permit a more accurate recapitulation of the tumor complexity, principally in regards to the interplay between cancer cells and the tumor stroma. This is especially relevant for IGFBP5, since the function of this protein is modulated by ECM. Additionally, experiments in animals could partially reproduce endocrine mechanisms, particularly the effects related to IGF-1. In fact, both mouse and human IGF-1 are structural orthologues and murine IGF-1 is capable to efficiently stimulate human IGF-1R [328, 329].

5.7. Clinical outlook

Results of this study are clinically relevant in multiple ways. Firstly, past studies indicated that germline variants in *DIRC3* impact the risk of developing DTCs. Nevertheless, associations identified for those particular SNVs are modest only (allelic OR ~1.43-1.44) [242, 244]. Accordingly, their clinical usefulness is questionable. Results of this study have linked *DIRC3* and its germline variant (rs11693806) with expression of *IGFBP5*, a gene known to modulate the bioavailability of IGF-1. Recent studies have also indicated that the concentration of circulating IGF-1 has a significant impact on DTC susceptibility (HR = 1.22 per 5 nmol/l increase of IGF-1) [107, 109]. The molecular element connecting these risk factors is *IGFBP5*. Hence, I predict that genotyping of *DIRC3* performed simultaneously with the measurement of circulating IGF-1 could expose a synergistic impact on the DTC risk assessment. I propose that future epidemiological studies should concurrently assess *DIRC3* genotype and the concentration of IGF-1 in the blood to more accurately predict patients' susceptibility to DTC.

Secondly, this study indicates that expression of *DIRC3* in thyroid carcinomas may influence cancer invasiveness. Accordingly, *DIRC3* expression in DTCs may have a prognostic value. To date the only widely recognized molecular prognostic factors in DTCs are somatic mutations in *BRAF* and the *TERT* promoter [30, 33, 330, 331]. A risk-stratification strategy utilizing the measurement of *DIRC3* expression in thyroid cancers may potentially help to further personalize therapy. I foresee that once the reported associations are validated, they may be beneficial to guide clinical decisions: to intensify treatment in the high-risk/*DIRC3-low* cancer patients, or deescalate management of the low-risk/*DIRC3-high* cancers.

Thirdly, I demonstrated that downregulation of *DIRC3* increases the cellular response to IGF-1. Therapeutic inhibition of IGF-1 signaling has been long pursued in clinical oncology, with drugs such as IGF-1R TKIs (AXL1717) and monoclonal antibodies targeting IGF-1R or IGF-1 (figitumumab, teprotumumab, ganitumab, xentuzumab) being tested in a number of preclinical and clinical trials (also recruiting patients diagnosed with advanced DTCs) [106, 252, 332]. While early studies were promising, subsequent phase II and III clinical trials have been largely disappointing. Still, exceptional responders with a long-lasting benefit derived from anti-IGF-1R drugs have been noticed across many trials. It has been proposed that novel predictive factors are necessary to identify patients who could respond to IGF-1R inhibitors [252, 332]. Remarkably, one of such predictors could be the expression of *IGFBP5* in tumors. It has been demonstrated that the level of IGFBP5 corelated inversely with the resistance to IGF-1R-targeting antibodies and TKIs in breast and urinary bladder carcinomas. Similar observations were made elsewhere in a large panel of cancer cell lines (particularly in colon cancer cell lines) [294, 333, 334]. Accordingly, results of our study may provide a rationale for utilizing IGF-1R inhibitors in the RAI-refractory DTCs that downregulate *DIRC3.* In such preclinical or clinical trials, IGF-1R inhibitors might be combined with clinically approved TKIs (such as sorafenib, lenvatinib) to prevent resistance mechanisms mediated by cross-talk with other oncogenic signaling pathways [252].

In summary, results of this study indicate that *DIRC3* has a prominent anti-invasive function in thyroid cancers. The expression of *DIRC3* in DTCs may emerge as a clinically relevant prognostic factor, and may also constitute a predictive marker for novel therapeutic strategies.

CHAPTER VI: CONCLUSIONS

1. *DIRC3* emerges a thyroid cancer suppressor that is frequently downregulated in DTCs. Expression of *DIRC3* may impact the risk of DTC recurrence.

2. *DIRC3* produces nuclear-enriched transcripts that induce expression of *IGFBP5*, a master regulator of the IGF-1 signaling pathway.

3. *DIRC3* silencing increases migration and invasiveness of thyroid cancer cells *in vitro*, but decelerates the increase in number of viable cells in MTT assay.

4. Knock-down of *DIRC3* sensitizes thyroid cancer cells to IGF-1 and promotes AKT signaling pathway.

5. A relatively strong thyroid cancer risk variant, rs11693806, modulates expression of *DIRC3* and *IGFBP5*, and changes the phenotype of thyroid cancer cells.

6. Expression of *DIRC3* may constitute a clinically relevant prognostic factor in DTCs. Thyroid cancers with low expression of *DIRC3* may be more sensitive to IGF-1 stimulation, what consequently indicates a potential new therapeutic avenue.

CHAPTER VII: AUTHOR'S PUBLICATIONS

- Conference abstracts based on results of this PhD thesis:

1. **Wysocki PT**, Kolanowska M, Nowis D. Functional characterization of *DIRC3* long non-coding RNA in differentiated thyroid cancer. *Annals of Oncology*. 2021; Volume 32; S367-S367. European Society for Medical Oncology Congress 2021, Virtual Meeting.

- Author's publications:

1. **Wysocki PT**, Westra WH, Sidransky D, Brait M. Advancing toward a molecular characterization of polymorphous low-grade adenocarcinoma. *Oral Oncol*. 2017 Nov;74:192-193.
2. **Wysocki PT**, Izumchenko E, Meir J, Ha PK, Sidransky D, Brait M. Adenoid cystic carcinoma: emerging role of translocations and gene fusions. *Oncotarget*. 2016 Oct 4;7(40):66239-66254.
3. Ravi R, Noonan KA, Pham V, Bedi R, Zhavoronkov A, Ozerov IV, Makarev E, V Artemov A, **Wysocki PT**, Mehra R, Nimmagadda S, Marchionni L, Sidransky D, Borrello IM, Izumchenko E, Bedi A. Bifunctional immune checkpoint-targeted antibody-ligand traps that simultaneously disable TGFβ enhance the efficacy of cancer immunotherapy. *Nat Commun*. 2018 Feb 21;9(1):741.
4. Ren S, Gaykalova D, Wang J, Guo T, Danilova L, Favorov A, Fertig E, Bishop J, Khan Z, Flam E, **Wysocki PT**, DeJong P, Ando M, Liu C, Sakai A, Fukusumi T, Haft S, Sadat S, Califano JA. Discovery and development of differentially methylated regions in human papillomavirus-related oropharyngeal squamous cell carcinoma. *Int J Cancer*. 2018 Nov 15;143(10):2425-2436.
5. Izumchenko E, Meir J, Bedi A, **Wysocki PT**, Hoque MO, Sidransky D. Patient-derived xenografts as tools in pharmaceutical development. *Clin Pharmacol Ther*. 2016 Jun;99(6):612-21.
6. Brait M, Izumchenko E, Kagohara LT, Long S, **Wysocki PT**, Faherty B, Fertig EJ, Khor TO, Bruckheimer E, Baia G, Ciznadija D, Sloma I, Ben-Zvi I, Paz K, Sidransky D.

Comparative mutational landscape analysis of patient-derived tumour xenografts. *Br J Cancer*. 2017 Feb 14;116(4):515-523.

7. Makarev E, Izumchenko E, Aihara F, **Wysocki PT**, Zhu Q, Buzdin A, Sidransky D, Zhavoronkov A, Atala A. Common pathway signature in lung and liver fibrosis. *Cell Cycle*. 2016 Jul 2;15(13):1667-73.
8. Lee J, **Wysocki PT**, Topaloglu O, Maldonado L, Brait M, Begum S, Moon D, Kim MS, Califano JA, Sidransky D, Hoque MO, Moon C. Epigenetic silencing of S100A2 in bladder and head and neck cancers. *Oncoscience*. 2015 Mar 16;2(4):410-8.
9. Hassan C, **Wysocki PT**, Fuccio L, Seufferlein T, Dinis-Ribeiro M, Brandão C, Regula J, Frazzoni L, Pellise M, Alfieri S, Dekker E, Jover R, Rosati G, Senore C, Spada C, Gralnek I, Dumonceau JM, van Hooft JE, van Cutsem E, Ponchon T. Endoscopic surveillance after surgical or endoscopic resection for colorectal cancer: European Society of Gastrointestinal Endoscopy (ESGE) and European Society of Digestive Oncology (ESDO) Guideline. *Endoscopy*. 2019 Mar;51(3):266-277.
10. **Wysocki PT**. A comprehensive review on posttreatment surveillance in colorectal patients. *Pol Arch Intern Med*. 2021 Mar 30;131(3):276-287.
11. **Wysocki PT**: Nadzór onkologiczny po leczeniu radykalnym raka jelita grubego – przegląd piśmiennictwa i wytycznych praktyki klinicznej. *Med. Prakt*. 2021; 4:37–47
12. Kosakowska EA, Rutkowski A, **Wysocki PT**, Michalski W, Cencelewicz-Lesikow A, Kunkiel M, Jagiello-Gruszfeld A. Evaluation of the efficacy of chemotherapy with capecitabine and oxaliplatin in patients with disseminated colorectal cancer. The impact of primary cancer focus on treatment efficacy. *Oncol Clin Pract*. 2021;17(1):1-5.
13. **Wysocki PT**. Zespół pieczenia jamy ustnej. *Gastroenterologia Kliniczna*. 2017;9(1):7-18.

- Other conference abstracts

1. **Wysocki PT**, Ling S, Shao C, Tan M, Sidransky D, Ha P, Brait M. Highly sensitive detection of *MYB-NFIB* fusion transcripts in adenoid cystic carcinoma. *Cancer Research*. 2016; 76(14). American Association for Cancer Research 106th Annual Meeting 2015. Philadelphia, PA, USA.
2. **Wysocki PT**, Westra WH, Begum S, Sidransky D, Brait M. High throughput detection of *PRKD1* mutations in salivary gland tumors using *in situ* mutation detection in tissue

microarrays. *Cancer Research*. 2016; 75(15). American Association for Cancer Research 107th Annual Meeting 2016. New Orleans, LA, USA.

3. Schubert AD, Izumchenko E, **Wysocki PT**, Sidransky D, Brait M. Improved detection of salivary glands' RNA markers in saliva samples. *Cancer Research*. 2017; 77(13). American Association for Cancer Research 108th Annual Meeting 2017. Washington, DC, USA.
4. Brait M, Izumchenko E, Kagohara L, Long S, **Wysocki P**, Faherty B, Fertig E, Khor T, Bruckheimer E, Baia G, Ciznadija D, Sloma I, Ben-Zvi I, Paz K, Sidransky D. Validation of qBiomarker as an accurate and efficient mutation detection method in a comprehensive analysis of patient-derived tumor xenografts. *Cancer Research*. 2016; 76(14). American Association for Cancer Research 107th Annual Meeting 2016. New Orleans, LA, USA.
5. **Wysocki PT**, Ling S, Tan M, Meir J, Shao C, Sidransky D, Ha PK, Brait M. *MYB* gene family alterations outline salivary gland adenoid cystic carcinoma. AHNS 9th International Conference on Head and Neck Cancer, 2016, Seattle, WA, USA.

CHAPTER VIII: LITERATURE

1. Bray, F., Ferlay, J., Soerjomataram, I.*, et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**, 394-424 (2018).
2. Lim, H., Devesa, S.S., Sosa, J.A.*, et al.* Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *JAMA Otolaryngol Head Neck Surg* **317**, 1338-1348 (2017).
3. Kitahara, C.M. & Sosa, J.A. The changing incidence of thyroid cancer. *Nature Reviews Endocrinology* **12**, 646-653 (2016).
4. Dralle, H., Machens, A., Basa, J.*, et al.* Follicular cell-derived thyroid cancer. *Nature Reviews Disease Primers* **1**, 15077 (2015).
5. Klöppel, G. & Hruban, R. WHO classification of tumours of endocrine organs. (Lyon: International Agency for Research on Cancer2017, 2017).
6. Medas, F., Canu, G.L., Boi, F.*, et al.* Predictive factors of recurrence in patients with differentiated thyroid carcinoma: a retrospective analysis on 579 patients. *Cancers* **11**, 1230 (2019).
7. Ghaznavi, S.A., Ganly, I., Shaha, A.R.*, et al.* Using the American Thyroid Association risk-stratification system to refine and individualize the American Joint Committee on Cancer Eighth Edition disease-specific survival estimates in differentiated thyroid cancer. *Thyroid* **28**, 1293-1300 (2018).
8. Cabanillas, M.E., McFadden, D.G. & Durante, C. Thyroid cancer. *The Lancet* **388**, 2783-2795 (2016).
9. Magreni, A., Bann, D.V., Schubart, J.R.*, et al.* The Effects of Race and Ethnicity on Thyroid Cancer Incidence. *JAMA Otolaryngol Head Neck Surg* **141**, 319-323 (2015).
10. Weeks, K.S., Kahl, A.R., Lynch, C.F.*, et al.* Racial/ethnic differences in thyroid cancer incidence in the United States, 2007-2014. *Cancer* **124**, 1483-1491 (2018).
11. Kwong, N., Medici, M., Angell, T.E.*, et al.* The Influence of Patient Age on Thyroid Nodule Formation, Multinodularity, and Thyroid Cancer Risk. *J Clin Endocrinol Metab* **100**, 4434-4440 (2015).
12. Amundadottir, L.T., Thorvaldsson, S., Gudbjartsson, D.F.*, et al.* Cancer as a Complex Phenotype: Pattern of Cancer Distribution within and beyond the Nuclear Family. *PLoS Medicine* **1**, e65 (2005).

13. Goldgar, D.E., Easton, D.F., Cannon-Albright, L.A.*, et al.* Systematic Population-Based Assessment of Cancer Risk in First-Degree Relatives of Cancer Probands. *JNCI: Journal of the National Cancer Institute* **86**, 1600-1608 (1994).
14. Czene, K., Lichtenstein, P. & Hemminki, K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish family-cancer database. *Int J Cancer* **99**, 260-266 (2002).
15. Oakley, G.M., Curtin, K., Pimentel, R.*, et al.* Establishing a familial basis for papillary thyroid carcinoma using the Utah Population Database. *JAMA Otolaryngol Head Neck Surg* **139**, 1171-1174 (2013).
16. Rashkin, S.R., Graff, R.E., Kachuri, L.*, et al.* Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nature Communications* **11**, 4423 (2020).
17. Fagin, J.A. & Wells, S.A. Biologic and Clinical Perspectives on Thyroid Cancer. *New England Journal of Medicine* **375**, 1054-1067 (2016).
18. Guilmette, J. & Nosé, V. Hereditary and familial thyroid tumours. *Histopathology* **72**, 70-81 (2018).
19. Veiga, L.H., Holmberg, E., Anderson, H.*, et al.* Thyroid Cancer after Childhood Exposure to External Radiation: An Updated Pooled Analysis of 12 Studies. *Radiat Res* **185**, 473-484 (2016).
20. Kitahara, C.M., McCullough, M.L., Franceschi, S.*, et al.* Anthropometric Factors and Thyroid Cancer Risk by Histological Subtype: Pooled Analysis of 22 Prospective Studies. *Thyroid* **26**, 306-318 (2016).
21. Xu, L., Port, M., Landi, S.*, et al.* Obesity and the risk of papillary thyroid cancer: a pooled analysis of three case-control studies. *Thyroid* **24**, 966-974 (2014).
22. Kwon, H., Han, K.-D. & Park, C.-Y. Weight change is significantly associated with risk of thyroid cancer: A nationwide population-based cohort study. *Scientific Reports* **9**, 1546 (2019).
23. Nunney, L. Size matters: height, cell number and a person's risk of cancer. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181743 (2018).
24. Jing, Z., Hou, X., Liu, Y.*, et al.* Association between height and thyroid cancer risk: A meta-analysis of prospective cohort studies. *Int J Cancer* **137**, 1484-1490 (2015).
25. Sado, J., Kitamura, T., Sobue, T.*, et al.* Risk of thyroid cancer in relation to height, weight, and body mass index in Japanese individuals: a population-based cohort study. *Cancer medicine* **7**, 2200-2210 (2018).

26. Kitahara, C.M., Gamborg, M., Berrington de González, A.*, et al.* Childhood Height and Body Mass Index Were Associated with Risk of Adult Thyroid Cancer in a Large Cohort Study. *Cancer Research* **74**, 235-242 (2014).
27. Kakudo, K., Bychkov, A., Bai, Y.*, et al.* The new 4th edition World Health Organization classification for thyroid tumors, Asian perspectives. *Pathology International* **68**, 641-664 (2018).
28. Tallini, G., de Biase, D., Repaci, A.*, et al.* What's new in thyroid tumor classification, the 2017 World Health Organization classification of tumours of endocrine organs. in *Thyroid FNA Cytology* 37-47 (Springer, 2019).
29. Xu, B. & Ghossein, R. Evolution of the histologic classification of thyroid neoplasms and its impact on clinical management. *European journal of surgical oncology* **44**, 338-347 (2018).
30. Xing, M., Alzahrani, A.S., Carson, K.A.*, et al.* Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *JAMA Otolaryngol Head Neck Surg* **309**, 1493-1501 (2013).
31. The Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **159**, 676-690 (2014).
32. Schlumberger, M. & Leboulleux, S. Current practice in patients with differentiated thyroid cancer. *Nature Reviews Endocrinology* (2020).
33. Xing, M., Alzahrani, A.S., Carson, K.A.*, et al.* Association between BRAF V600E mutation and recurrence of papillary thyroid cancer. *J Clin Oncol* **33**, 42-50 (2015).
34. Bai, Y., Kakudo, K. & Jung, C.K. Updates in the Pathologic Classification of Thyroid Neoplasms: A Review of the World Health Organization Classification. *Endocrinology and metabolism (Seoul, Korea)* **35**, 696-715 (2020).
35. Zaballos, M.A. & Santisteban, P. Key signaling pathways in thyroid cancer. *Journal of Endocrinology* **235**, R43 (2017).
36. Goffredo, P., Roman, S.A. & Sosa, J.A. Hurthle cell carcinoma. *Cancer* **119**, 504-511 (2013).
37. Ganly, I., Makarov, V., Deraje, S.*, et al.* Integrated Genomic Analysis of Hürthle Cell Cancer Reveals Oncogenic Drivers, Recurrent Mitochondrial Mutations, and Unique Chromosomal Landscapes. *Cancer Cell* **34**, 256-270.e255 (2018).
38. Filetti, S., Durante, C., Hartl, D.*, et al.* Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **30**, 1856-1883 (2019).

39. Haugen, B.R., Alexander, E.K., Bible, K.C.*, et al.* 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **26**, 1-133 (2016).
40. Grani, G., Zatelli, M.C., Alfò, M.*, et al.* Real-World Performance of the American Thyroid Association Risk Estimates in Predicting 1-Year Differentiated Thyroid Cancer Outcomes: A Prospective Multicenter Study of 2000 Patients. *Thyroid* **31**, 264-271 (2020).
41. Brose, M.S., Robinson, B., Sherman, S.I.*, et al.* Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncology* **22**, 1126-1138 (2021).
42. Wirth, L.J., Sherman, E., Robinson, B.*, et al.* Efficacy of Selpercatinib in RET-Altered Thyroid Cancers. *New England Journal of Medicine* **383**, 825-835 (2020).
43. Hong, D.S., DuBois, S.G., Kummar, S.*, et al.* Larotrectinib in patients with TRK fusion-positive solid tumours: a pooled analysis of three phase 1/2 clinical trials. *The Lancet Oncology* **21**, 531-540 (2020).
44. Doebele, R.C., Drilon, A., Paz-Ares, L.*, et al.* Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three phase 1 trials. *The Lancet Oncology* **21**, 271-282 (2020).
45. Subbiah, V., Hu, M.I.-N., Gainor, J.F.*, et al.* Clinical activity of the RET inhibitor pralsetinib (BLU-667) in patients with RET fusion+ solid tumors. *Journal of Clinical Oncology* **38**, 109-109 (2020).
46. Brose, M.S., Cabanillas, M.E., Cohen, E.E.W.*, et al.* Vemurafenib in patients with BRAF V600E-positive metastatic or unresectable papillary thyroid cancer refractory to radioactive iodine: a non-randomised, multicentre, open-label, phase 2 trial. *The Lancet Oncology* **17**, 1272-1282 (2016).
47. Falchook, G.S., Millward, M., Hong, D.*, et al.* BRAF inhibitor dabrafenib in patients with metastatic BRAF-mutant thyroid cancer. *Thyroid* **25**, 71-77 (2015).
48. Liu, Z., Hou, P., Ji, M.*, et al.* Highly prevalent genetic alterations in receptor tyrosine kinases and phosphatidylinositol 3-kinase/akt and mitogen-activated protein kinase pathways in anaplastic and follicular thyroid cancers. *J Clin Endocrinol Metab* **93**, 3106-3116 (2008).

49. Hou, P., Liu, D., Shan, Y.*, et al.* Genetic Alterations and Their Relationship in the Phosphatidylinositol 3-Kinase/Akt Pathway in Thyroid Cancer. *Clin Cancer Res* **13**, 1161-1170 (2007).
50. Hou, P., Ji, M. & Xing, M. Association of PTEN gene methylation with genetic alterations in the phosphatidylinositol 3-kinase/AKT signaling pathway in thyroid tumors. *Cancer* **113**, 2440-2447 (2008).
51. Beg, S., Siraj, A.K., Jehan, Z.*, et al.* PTEN loss is associated with follicular variant of Middle Eastern papillary thyroid carcinoma. *Br J Cancer* **112**, 1938-1943 (2015).
52. Ramírez-Moya, J., Wert-Lamas, L. & Santisteban, P. MicroRNA-146b promotes PI3K/AKT pathway hyperactivation and thyroid cancer progression by targeting PTEN. *Oncogene* **37**, 3369-3383 (2018).
53. Guo, F., Hou, X. & Sun, Q. MicroRNA-9-5p functions as a tumor suppressor in papillary thyroid cancer via targeting BRAF. *Oncol Lett* **16**, 6815-6821 (2018).
54. Zhang, J., Du, Y., Zhang, X.*, et al.* Downregulation of BANCR Promotes Aggressiveness in Papillary Thyroid Cancer via the MAPK and PI3K Pathways. *Journal of Cancer* **9**, 1318-1328 (2018).
55. Feng, K., Liu, Y., Xu, L.J.*, et al.* Long noncoding RNA PVT1 enhances the viability and invasion of papillary thyroid carcinoma cells by functioning as ceRNA of microRNA-30a through mediating expression of insulin like growth factor 1 receptor. *Biomed Pharmacother* **104**, 686-698 (2018).
56. Tam, V., Patel, N., Turcotte, M.*, et al.* Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* **20**, 467-484 (2019).
57. Manolio, T.A. Genomewide Association Studies and Assessment of the Risk of Disease. *New England Journal of Medicine* **363**, 166-176 (2010).
58. Saenko, V.A. & Rogounovitch, T.I. Genetic Polymorphism Predisposing to Differentiated Thyroid Cancer: A Review of Major Findings of the Genome-Wide Association Studies. *Endocrinology and metabolism (Seoul, Korea)* **33**, 164-174 (2018).
59. Takahashi, M., Saenko, V.A., Rogounovitch, T.I.*, et al.* The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. *Hum Mol Genet* **19**, 2516-2523 (2010).
60. Gudmundsson, J., Sulem, P., Gudbjartsson, D.F.*, et al.* Discovery of common variants associated with low TSH levels and thyroid cancer risk. *Nature Genetics* **44**, 319-322 (2012).

61. Köhler, A., Chen, B., Gemignani, F.*, et al.* Genome-wide association study on differentiated thyroid cancer. *J Clin Endocrinol Metab* **98**, E1674-1681 (2013).
62. Gudmundsson, J., Thorleifsson, G., Sigurdsson, J.K.*, et al.* A genome-wide association study yields five novel thyroid cancer risk loci. *Nature Communications* **8**, 14517 (2017).
63. Son, H.Y., Hwangbo, Y., Yoo, S.K.*, et al.* Genome-wide association and expression quantitative trait loci studies identify multiple susceptibility loci for thyroid cancer. *Nature Communications* **8**, 15966 (2017).
64. Figlioli, G., Chen, B., Elisei, R.*, et al.* Novel genetic variants in differentiated thyroid cancer and assessment of the cumulative risk. *Scientific Reports* **5**, 8922 (2015).
65. Mancikova, V., Cruz, R., Inglada-Pérez, L.*, et al.* Thyroid cancer GWAS identifies 10q26.12 and 6q14.1 as novel susceptibility loci and reveals genetic heterogeneity among populations. *Int J Cancer* **137**, 1870-1878 (2015).
66. He, H., Li, W., Liyanarachchi, S.*, et al.* Genetic predisposition to papillary thyroid carcinoma: involvement of FOXE1, TSHR, and a novel lincRNA gene, PTCSC2. *J Clin Endocrinol Metab* **100**, E164-172 (2015).
67. Wang, Y., He, H., Li, W.*, et al.* MYH9 binds to lncRNA gene PTCSC2 and regulates FOXE1 in the 9q22 thyroid cancer risk locus. *Proceedings of the National Academy of Sciences* **114**, 474-479 (2017).
68. Jendrzejewski, J., He, H., Radomska, H.S.*, et al.* The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. *Proceedings of the National Academy of Sciences* **109**, 8646-8651 (2012).
69. Jendrzejewski, J., Thomas, A., Liyanarachchi, S.*, et al.* PTCSC3 Is Involved in Papillary Thyroid Carcinoma Development by Modulating S100A4 Gene Expression. *J Clin Endocrinol Metab* **100**, E1370-1377 (2015).
70. Wang, X., Lu, X., Geng, Z.*, et al.* LncRNA PTCSC3/miR-574-5p Governs Cell Proliferation and Migration of Papillary Thyroid Carcinoma via Wnt/β-Catenin Signaling. *Journal of Cellular Biochemistry* **118**, 4745-4752 (2017).
71. Fan, M., Li, X., Jiang, W.*, et al.* A long non-coding RNA, PTCSC3, as a tumor suppressor and a target of miRNAs in thyroid cancer cells. *Exp Ther Med* **5**, 1143-1146 (2013).

72. Jendrzejewski, J., Liyanarachchi, S., Eiterman, A.*, et al.* Fine mapping of 14q13 reveals novel variants associated with different histological subtypes of papillary thyroid carcinoma. *Int J Cancer* **144**, 503-512 (2019).
73. He, H., Li, W., Liyanarachchi, S.*, et al.* The Role of NRG1 in the Predisposition to Papillary Thyroid Carcinoma. *J Clin Endocrinol Metab* **103**, 1369-1379 (2018).
74. Hegde, G.V., de la Cruz, C.C., Chiu, C.*, et al.* Blocking NRG1 and Other Ligand-Mediated Her4 Signaling Enhances the Magnitude and Duration of the Chemotherapeutic Response of Non–Small Cell Lung Cancer. *Science Translational Medicine* **5**, 171ra118-171ra118 (2013).
75. Heining, C., Horak, P., Uhrig, S.*, et al.* NRG1 Fusions in KRAS Wild-Type Pancreatic Cancer. *Cancer Discovery* **8**, 1087-1095 (2018).
76. Zhang, Z., Karthaus, W.R., Lee, Y.S.*, et al.* Tumor Microenvironment-Derived NRG1 Promotes Antiandrogen Resistance in Prostate Cancer. *Cancer Cell* **38**, 279-296.e279 (2020).
77. Sheng, Q., Liu, X., Fleming, E.*, et al.* An Activated ErbB3/NRG1 Autocrine Loop Supports In Vivo Proliferation in Ovarian Cancer Cells. *Cancer Cell* **17**, 298-310 (2010).
78. Bodmer, D., Schepens, M., Eleveld, M.J.*, et al.* Disruption of a novel gene, DIRC3, and expression of DIRC3-HSPBAP1 fusion transcripts in a case of familial renal cell cancer and t(2;3)(q35;q21). *Genes Chromosomes Cancer* **38**, 107-116 (2003).
79. Liyanarachchi, S., Wojcicka, A., Li, W.*, et al.* Cumulative risk impact of five genetic variants associated with papillary thyroid carcinoma. *Thyroid* **23**, 1532-1540 (2013).
80. Wang, Y.L., Feng, S.H., Guo, S.C.*, et al.* Confirmation of papillary thyroid cancer susceptibility loci identified by genome-wide association studies of chromosomes 14q13, 9q22, 2q35 and 8p12 in a Chinese population. *J Med Genet* **50**, 689-695 (2013).
81. Jendrzejewski, J., Liyanarachchi, S., Nagy, R.*, et al.* Papillary Thyroid Carcinoma: Association Between Germline DNA Variant Markers and Clinical Parameters. *Thyroid* **26**, 1276-1284 (2016).
82. Mussazhanova, Z., Rogounovitch, T.I., Saenko, V.A.*, et al.* The Contribution of Genetic Variants to the Risk of Papillary Thyroid Carcinoma in the Kazakh Population: Study of Common Single Nucleotide Polymorphisms and Their Clinicopathological Correlations. *Frontiers in Endocrinology* **11**(2021).

83. Świerniak, M., Wójcicka, A., Czetwertyńska, M.*, et al.* Association between GWAS-Derived rs966423 Genetic Variant and Overall Mortality in Patients with Differentiated Thyroid Cancer. *Clin Cancer Res* **22**, 1111-1119 (2016).
84. Guibon, J., Sugier, P.-E., Kulkarni, O.*, et al.* Fine-mapping of two differentiated thyroid carcinoma susceptibility loci at 2q35 and 8p12 in Europeans, Melanesians and Polynesians. *Oncotarget* **12**(2021).
85. Lango Allen, H., Estrada, K., Lettre, G.*, et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-838 (2010).
86. Wood, A.R., Esko, T., Yang, J.*, et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* **46**, 1173-1186 (2014).
87. Sakaue, S., Kanai, M., Tanigawa, Y.*, et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nature Genetics* **53**, 1415-1424 (2021).
88. Graff, M., Justice, A.E., Young, K.L.*, et al.* Discovery and fine-mapping of height loci via high-density imputation of GWASs in individuals of African ancestry. *The American Journal of Human Genetics* (2021).
89. Dina, C., Bouatia-Naji, N., Tucker, N.*, et al.* Genetic association analyses highlight biological pathways underlying mitral valve prolapse. *Nature Genetics* **47**, 1206-1211 (2015).
90. Michailidou, K., Hall, P., Gonzalez-Neira, A.*, et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nature Genetics* **45**, 353-361, 361e351-352 (2013).
91. Couch, F.J., Kuchenbaecker, K.B., Michailidou, K.*, et al.* Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nature Communications* **7**, 11375 (2016).
92. McVeigh, U., Burke, C., McVeigh, T.*, et al.* Investigating the role of single nucleotide polymorphisms in NRG1 and DIRC3 in predisposition to breast cancer. *British Journal of Surgery*, Vol. 104 21-21 (2017).
93. Fachal, L., Aschard, H., Beesley, J.*, et al.* Fine-mapping of 150 breast cancer risk regions identifies 191 likely target genes. *Nature Genetics* **52**, 56-73 (2020).
94. Dryden, N.H., Broome, L.R., Dudbridge, F.*, et al.* Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. *Genome Res* **24**, 1854-1868 (2014).
95. Baxter, R.C. IGF binding proteins in cancer: mechanistic and clinical insights. *Nature Reviews Cancer* **14**, 329-341 (2014).

96. Wyszynski, A., Hong, C.C., Lam, K.*, et al.* An intergenic risk locus containing an enhancer deletion in 2q35 modulates breast cancer risk by deregulating IGFBP5 expression. *Hum Mol Genet* **25**, 3863-3876 (2016).
97. Ghoussaini, M., Edwards, S.L., Michailidou, K.*, et al.* Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation. *Nature Communications* **4**, 4999 (2014).
98. Stolf, B.S., Carvalho, A.F., Martins, W.K.*, et al.* Differential expression of IGFBP-5 and two human ESTs in thyroid glands with goiter, adenoma and papillary or follicular carcinomas. *Cancer Lett* **191**, 193-202 (2003).
99. Liu, L., Wang, J., Li, X.*, et al.* MiR-204-5p suppresses cell proliferation by inhibiting IGFBP5 in papillary thyroid carcinoma. *Biochem Biophys Res Commun* **457**, 621-626 (2015).
100. Porcu, E., Medici, M., Pistis, G.*, et al.* A meta-analysis of thyroid-related traits reveals novel loci and gender-specific differences in the regulation of thyroid function. *PLoS Genet* **9**, e1003266 (2013).
101. Malinowski, J.R., Denny, J.C., Bielinski, S.J.*, et al.* Genetic variants associated with serum thyroid stimulating hormone (TSH) levels in European Americans and African Americans from the eMERGE Network. *PloS one* **9**, e111301-e111301 (2014).
102. García, B. & Santisteban, P. PI3K is involved in the IGF-I inhibition of TSH-induced sodium/iodide symporter gene expression. *Mol Endocrinol* **16**, 342-352 (2002).
103. Liu, Y.J., Qiang, W., Shi, J.*, et al.* Expression and significance of IGF-1 and IGF-1R in thyroid nodules. *Endocrine* **44**, 158-164 (2013).
104. Zaballos, M.A. & Santisteban, P. FOXO1 controls thyroid cell proliferation in response to TSH and IGF-I and is involved in thyroid tumorigenesis. *Mol Endocrinol* **27**, 50-62 (2013).
105. Schmidt, J.A., Allen, N.E., Almquist, M.*, et al.* Insulin-like growth factor-i and risk of differentiated thyroid carcinoma in the European prospective investigation into cancer and nutrition. *Cancer Epidemiol Biomarkers Prev* **23**, 976-985 (2014).
106. Manzella, L., Massimino, M., Stella, S.*, et al.* Activation of the IGF Axis in Thyroid Cancer: Implications for Tumorigenesis and Treatment. *International Journal of Molecular Sciences* **20**, 3258 (2019).
107. Knuppel, A., Fensom, G.K., Watts, E.L.*, et al.* Circulating Insulin-like Growth Factor-I Concentrations and Risk of 30 Cancers: Prospective Analyses in UK Biobank. *Cancer Research* **80**, 4014-4021 (2020).

108. Lawnicka, H., Motylewska, E., Borkowska, M.*, et al.* Elevated serum concentrations of IGF-1 and IGF-1R in patients with thyroid cancers. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* **164**, 77-83 (2020).
109. Qian, F. & Huo, D. Circulating Insulin-Like Growth Factor-1 and Risk of Total and 19 Site-Specific Cancers: Cohort Study Analyses from the UK Biobank. *Cancer Epidemiol Biomarkers Prev* **29**, 2332-2342 (2020).
110. Ponjavic, J., Ponting, C.P. & Lunter, G. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res* **17**, 556-565 (2007).
111. Quinn, J.J. & Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nature Reviews Genetics* **17**, 47-62 (2016).
112. Statello, L., Guo, C.-J., Chen, L.-L.*, et al.* Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology* **22**, 96-118 (2021).
113. Gupta, R.A., Shah, N., Wang, K.C.*, et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-1076 (2010).
114. Zhou, X., Liu, S., Cai, G.*, et al.* Long Non Coding RNA MALAT1 Promotes Tumor Growth and Metastasis by inducing Epithelial-Mesenchymal Transition in Oral Squamous Cell Carcinoma. *Scientific Reports* **5**, 15972 (2015).
115. Gutschner, T., Hämmerle, M., Eissmann, M.*, et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Research* **73**, 1180-1189 (2013).
116. Wei, J.T., Feng, Z., Partin, A.W.*, et al.* Can urinary PCA3 supplement PSA in the early detection of prostate cancer? *J Clin Oncol* **32**, 4066-4072 (2014).
117. Leucci, E., Vendramin, R., Spinazzi, M.*, et al.* Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* **531**, 518-522 (2016).
118. Iyer, M.K., Niknafs, Y.S., Malik, R.*, et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nature Genetics* **47**, 199-208 (2015).
119. Derrien, T., Johnson, R., Bussotti, G.*, et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775-1789 (2012).
120. Wang, K.C. & Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Molecular Cell* **43**, 904-914 (2011).
121. Engreitz, J.M., Haines, J.E., Perez, E.M.*, et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452-455 (2016).

122. Chen, J., Brunner, A.D., Cogan, J.Z.*, et al.* Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140-1146 (2020).
123. Zhu, H., Lv, Z., An, C.*, et al.* Onco-lncRNA HOTAIR and its functional genetic variants in papillary thyroid carcinoma. *Scientific Reports* **6**, 31969 (2016).
124. Wen, J., Chen, L., Tian, H.*, et al.* Effect of MALAT1 Polymorphisms on Papillary Thyroid Cancer in a Chinese Population. *Journal of Cancer* **10**, 5714-5721 (2019).
125. Schmitt, A.M. & Chang, H.Y. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* **29**, 452-463 (2016).
126. Wang, J., Zhang, X., Chen, W.*, et al.* Regulatory roles of long noncoding RNAs implicated in cancer hallmarks. *Int J Cancer* **146**, 906-916 (2020).
127. Miller, S.A., Dykes, D.D. & Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215 (1988).
128. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).
129. de Kok, J.B., Roelofs, R.W., Giesendorf, B.A.*, et al.* Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Laboratory Investigation* **85**, 154-159 (2005).
130. Piampiani, P., Romei, C., Cosci, B.*, et al.* Identification of housekeeping genes useful for the normalization of rna in studies of gene expression in thyroid carcinomas. in *Endocrine Abstracts*, Vol. 11 (BioScientifica, 2006).
131. Lisowski, P., Pierzchała, M., Gościk, J.*, et al.* Evaluation of reference genes for studies of gene expression in the bovine liver, kidney, pituitary, and thyroid. *Journal of Applied Genetics* **49**, 367-372 (2008).
132. Lonsdale, J., Thomas, J., Salvatore, M.*, et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580-585 (2013).
133. Campbell, P.J., Getz, G., Korbel, J.O.*, et al.* Pan-cancer analysis of whole genomes. *Nature* **578**, 82-93 (2020).
134. Cerami, E., Gao, J., Dogrusoz, U.*, et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* **2**, 401-404 (2012).
135. Kent, W.J., Sugnet, C.W., Furey, T.S.*, et al.* The human genome browser at UCSC. *Genome Res* **12**, 996-1006 (2002).

136. Howe, K.L., Achuthan, P., Allen, J.*, et al.* Ensembl 2021. *Nucleic Acids Research* **49**, D884-D891 (2020).
137. Henderson, Y.C., Ahn, S.H., Ryu, J.*, et al.* Development and characterization of six new human papillary thyroid carcinoma cell lines. *J Clin Endocrinol Metab* **100**, E243-252 (2015).
138. Schweppe, R.E., Klopper, J.P., Korch, C.*, et al.* Deoxyribonucleic acid profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. *J Clin Endocrinol Metab* **93**, 4331-4341 (2008).
139. Zhao, M., Sano, D., Pickering, C.R.*, et al.* Assembly and initial characterization of a panel of 85 genomically validated cell lines from diverse head and neck tumor sites. *Clin Cancer Res* **17**, 7248-7264 (2011).
140. Arroyo-Helguera, O., Anguiano, B., Delgado, G.*, et al.* Uptake and antiproliferative effect of molecular iodine in the MCF-7 breast cancer cell line. *Endocr Relat Cancer* **13**, 1147-1158 (2006).
141. Hall, L.C., Salazar, E.P., Kane, S.R.*, et al.* Effects of thyroid hormones on human breast cancer cell proliferation. *J Steroid Biochem Mol Biol* **109**, 57-66 (2008).
142. Gierlikowski, W., Broniarek, K., Cheda, Ł.*, et al.* MiR-181a-5p Regulates NIS Expression in Papillary Thyroid Carcinoma. *International Journal of Molecular Sciences* **22**, 6067 (2021).
143. Lennox, K.A. & Behlke, M.A. Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Research* **44**, 863-877 (2016).
144. Humphrey, J., Birsa, N., Milioto, C.*, et al.* FUS ALS-causative mutations impair FUS autoregulation and splicing factor networks through intron retention. *Nucleic Acids Research* **48**, 6889-6905 (2020).
145. Mas-Ponte, D., Carlevaro-Fita, J., Palumbo, E.*, et al.* LncATLAS database for subcellular localization of long noncoding RNAs. *RNA (New York, N.Y.)* **23**, 1080-1087 (2017).
146. Cabili, M.N., Dunagin, M.C., McClanahan, P.D.*, et al.* Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol* **16**, 20 (2015).

147. Meng, X., Deng, Y., Lv, Z.*, et al.* LncRNA SNHG5 Promotes Proliferation of Glioma by Regulating miR-205-5p/ZEB2 Axis. *OncoTargets and therapy* **12**, 11487-11496 (2019).
148. Yan, L., Wang, S., Li, Y.*, et al.* SNHG5 promotes proliferation and induces apoptosis in melanoma by sponging miR-155. *RSC Advances* **8**, 6160-6168 (2018).
149. Watts, J.K. & Corey, D.R. Silencing disease genes in the laboratory and the clinic. *The Journal of Pathology* **226**, 365-379 (2012).
150. Marshman, E., Green, K.A., Flint, D.J.*, et al.* Insulin-like growth factor binding protein 5 and apoptosis in mammary epithelial cells. *Journal of Cell Science* **116**, 675-682 (2003).
151. Monget, P., Pisselet, C. & Monniaux, D. Expression of insulin-like growth factor binding protein-5 by ovine granulosa cells is regulated by cell density and programmed cell death in vitro. *Journal of Cellular Physiology* **177**, 13-25 (1998).
152. Ding, M., Bruick, R.K. & Yu, Y. Secreted IGFBP5 mediates mTORC1-dependent feedback inhibition of IGF-1 signalling. *Nature Cell Biology* **18**, 319-327 (2016).
153. Fürstenberger, G. & Senn, H.-J. Insulin-like growth factors and cancer. *The Lancet Oncology* **3**, 298-302 (2002).
154. S, A. FastQC. *Babraham bioinformatics* (2020).
155. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
156. Dobin, A., Davis, C.A., Schlesinger, F.*, et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
157. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
158. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
159. Subhash, S. & Kanduri, C. GeneSCF: a real-time based functional enrichment tool with support for multiple organisms. *BMC Bioinformatics* **17**, 365 (2016).
160. Chavez, A., Scheiman, J., Vora, S.*, et al.* Highly efficient Cas9-mediated transcriptional programming. *Nature Methods* **12**, 326-328 (2015).
161. Bester, A.C., Lee, J.D., Chavez, A.*, et al.* An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance. *Cell* **173**, 649-664.e620 (2018).

162. Konermann, S., Brigham, M.D., Trevino, A.E., *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).
163. Labun, K., Montague, T.G., Krause, M., *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research* **47**, W171-W174 (2019).
164. Doench, J.G., Fusi, N., Sullender, M., *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology* **34**, 184-191 (2016).
165. Gilbert, Luke A., Horlbeck, Max A., Adamson, B., *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647-661 (2014).
166. Forrest, A.R.R., Kawaji, H., Rehli, M., *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462-470 (2014).
167. Radzisheuskaya, A., Shlyueva, D., Müller, I., *et al.* Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. *Nucleic Acids Research* **44**, e141-e141 (2016).
168. O'Brien, A.R., Wilson, L.O.W., Burgio, G., *et al.* Unlocking HDR-mediated nucleotide editing by identifying high-efficiency target sites using machine learning. *Scientific Reports* **9**, 2788 (2019).
169. Liang, X., Potter, J., Kumar, S., *et al.* Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J Biotechnol* **241**, 136-146 (2017).
170. Cho, S.W., Xu, J., Sun, R., *et al.* Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. *Cell* **173**, 1398-1412.e1322 (2018).
171. Aarreberg, L.D., Esser-Nobis, K., Driscoll, C., *et al.* Interleukin-1β Induces mtDNA Release to Activate Innate Immune Signaling via cGAS-STING. *Molecular Cell* **74**, 801-815.e806 (2019).
172. Di Stazio, M., Foschi, N., Athanasakis, E., *et al.* Systematic analysis of factors that improve homologous direct repair (HDR) efficiency in CRISPR/Cas9 technique. *PloS one* **16**, e0247603 (2021).
173. Richardson, C.D., Ray, G.J., DeWitt, M.A., *et al.* Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology* **34**, 339-344 (2016).
174. Liang, X., Potter, J., Kumar, S., *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* **208**, 44-53 (2015).

175. Danecek, P., Bonfield, J.K., Liddle, J.*, et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**(2021).
176. Teer, J.K., Green, E.D., Mullikin, J.C.*, et al.* VarSifter: Visualizing and analyzing exome-scale sequence variation data on a desktop computer. *Bioinformatics* **28**, 599-600 (2011).
177. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W.*, et al.* Integrative genomics viewer. *Nature Biotechnology* **29**, 24-26 (2011).
178. Robinson, J.T., Thorvaldsdóttir, H., Wenger, A.M.*, et al.* Variant Review with the Integrative Genomics Viewer. *Cancer Research* **77**, e31-e34 (2017).
179. Bae, S., Park, J. & Kim, J.-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473-1475 (2014).
180. Concordet, J.-P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research* **46**, W242-W245 (2018).
181. Sun, Q., Hao, Q. & Prasanth, K.V. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. *Trends in genetics : TIG* **34**, 142-157 (2018).
182. Kichaev, G., Bhatia, G., Loh, P.R.*, et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *The American Journal of Human Genetics* **104**, 65-75 (2019).
183. Wojcik, G.L., Graff, M., Nishimura, K.K.*, et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514-518 (2019).
184. Akiyama, M., Ishigaki, K., Sakaue, S.*, et al.* Characterizing rare and low-frequency height-associated variants in the Japanese population. *Nature Communications* **10**, 4393 (2019).
185. N'Diaye, A., Chen, G.K., Palmer, C.D.*, et al.* Identification, replication, and fine-mapping of Loci associated with adult height in individuals of African ancestry. *PLoS Genet* **7**, e1002298 (2011).
186. He, M., Xu, M., Zhang, B.*, et al.* Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. *Hum Mol Genet* **24**, 1791-1800 (2015).
187. Duan, C. & Clemmons, D.R. Differential expression and biological effects of insulin-like growth factor-binding protein-4 and -5 in vascular smooth muscle cells. *J Biol Chem* **273**, 16836-16842 (1998).

188. Fukuda, R., Hirota, K., Fan, F.*, et al.* Insulin-like growth factor 1 induces hypoxia-inducible factor 1-mediated vascular endothelial growth factor expression, which is dependent on MAP kinase and phosphatidylinositol 3-kinase signaling in colon cancer cells. *J Biol Chem* **277**, 38205-38211 (2002).
189. Duan, C., Hawes, S.B., Prevette, T.*, et al.* Insulin-like growth factor-I (IGF-I) regulates IGF-binding protein-5 synthesis through transcriptional activation of the gene in aortic smooth muscle cells. *J Biol Chem* **271**, 4280-4288 (1996).
190. Duan, C., Liimatta, M.B. & Bottum, O.L. Insulin-like growth factor (IGF)-I regulates IGF-binding protein-5 gene expression through the phosphatidylinositol 3-kinase, protein kinase B/Akt, and p70 S6 kinase signaling pathway. *J Biol Chem* **274**, 37147-37153 (1999).
191. Arun, G., Aggarwal, D. & Spector, D.L. MALAT1 Long Non-Coding RNA: Functional Implications. *Non-coding RNA* **6**, 22 (2020).
192. Peng, Y., Fang, X., Yao, H.*, et al.* MiR-146b-5p Regulates the Expression of Long Noncoding RNA MALAT1 and Its Effect on the Invasion and Proliferation of Papillary Thyroid Cancer. *Cancer Biother Radiopharm* **36**, 433-440 (2021).
193. Liu, J., Dong, H., Yang, Y.*, et al.* Upregulation of long noncoding RNA MALAT1 in papillary thyroid cancer and its diagnostic value. *Future Oncol* **14**, 3015-3022 (2018).
194. Zhang, R., Hardin, H., Huang, W.*, et al.* MALAT1 Long Non-coding RNA Expression in Thyroid Tissues: Analysis by In Situ Hybridization and Real-Time PCR. *Endocrine Pathology* **28**, 7-12 (2017).
195. Possieri, C., Locantore, P., Salis, C.*, et al.* Combined molecular and mathematical analysis of long noncoding RNAs expression in fine needle aspiration biopsies as novel tool for early diagnosis of thyroid cancer. *Endocrine* **72**, 711-720 (2021).
196. Covach, A., Patel, S., Hardin, H.*, et al.* Phosphorylated Mechanistic Target of Rapamycin (p-mTOR) and Noncoding RNA Expression in Follicular and Hürthle Cell Thyroid Neoplasm. *Endocrine Pathology* **28**, 207-212 (2017).
197. Huang, J.K., Ma, L., Song, W.H.*, et al.* MALAT1 promotes the proliferation and invasion of thyroid cancer cells via regulating the expression of IQGAP1. *Biomed Pharmacother* **83**, 1-7 (2016).
198. Ye, M., Dong, S., Hou, H.*, et al.* Oncogenic Role of Long Noncoding RNAMALAT1 in Thyroid Cancer Progression through Regulation of the miR-204/IGF2BP2/m6A-MYC Signaling. *Molecular Therapy - Nucleic Acids* **23**, 1-12 (2021).

199. Huang, J.K., Ma, L., Song, W.H.*, et al.* LncRNA-MALAT1 Promotes Angiogenesis of Thyroid Cancer by Modulating Tumor-Associated Macrophage FGF2 Protein Secretion. *J Cell Biochem* **118**, 4821-4830 (2017).
200. Brock, M., Schuoler, C., Leuenberger, C.*, et al.* Analysis of hypoxia-induced noncoding RNAs reveals metastasis-associated lung adenocarcinoma transcript 1 as an important regulator of vascular smooth muscle cell proliferation. *Exp Biol Med (Maywood)* **242**, 487-496 (2017).
201. Shih, C.-H., Chuang, L.-L., Tsai, M.-H.*, et al.* Hypoxia-Induced MALAT1 Promotes the Proliferation and Migration of Breast Cancer Cells by Sponging MiR-3064-5p. *Frontiers in Oncology* **11**(2021).
202. Kölling, M., Genschel, C., Kaucsar, T.*, et al.* Hypoxia-induced long non-coding RNA Malat1 is dispensable for renal ischemia/reperfusion-injury. *Scientific Reports* **8**, 3438 (2018).
203. Sallé-Lefort, S., Miard, S., Nolin, M.A.*, et al.* Hypoxia upregulates Malat1 expression through a CaMKK/AMPK/HIF-1α axis. *Int J Oncol* **49**, 1731-1736 (2016).
204. Laronha, H. & Caldeira, J. Structure and Function of Human Matrix Metalloproteinases. *Cells* **9**, 1076 (2020).
205. Quintero-Fabián, S., Arreola, R., Becerril-Villanueva, E.*, et al.* Role of Matrix Metalloproteinases in Angiogenesis and Cancer. *Frontiers in Oncology* **9**(2019).
206. Patel, A., Straight, A.M., Mann, H.*, et al.* Matrix metalloproteinase (MMP) expression by differentiated thyroid carcinoma of children and adolescents. *J Endocrinol Invest* **25**, 403-408 (2002).
207. Buergy, D., Weber, T., Maurer, G.D.*, et al.* Urokinase receptor, MMP-1 and MMP-9 are markers to differentiate prognosis, adenoma and carcinoma in thyroid malignancies. *Int J Cancer* **125**, 894-901 (2009).
208. Aust, G., Hofmann, A., Laue, S.*, et al.* Human thyroid carcinoma cell lines and normal thyrocytes: expression and regulation of matrix metalloproteinase-1 and tissue matrix metalloproteinase inhibitor-1 messenger-RNA and protein. *Thyroid* **7**, 713-724 (1997).
209. Korem, S., Resnick, M.B. & Kraiem, Z. Similar and divergent patterns in the regulation of matrix metalloproteinase-1 (MMP-1) and tissue inhibitor of MMP-1 gene expression in benign and malignant human thyroid cells. *J Clin Endocrinol Metab* **84**, 3322-3327 (1999).
210. Kameyama, K. Expression of MMP-1 in the capsule of thyroid cancer--relationship with invasiveness. *Pathol Res Pract* **192**, 20-26 (1996).

211. Bumber, B., Marjanovic Kavanagh, M., Jakovcevic, A.*, et al.* Role of matrix metalloproteinases and their inhibitors in the development of cervical metastases in papillary thyroid cancer. *Clin Otolaryngol* **45**, 55-62 (2020).
212. Mizrachi, A., Koren, R., Hadar, T.*, et al.* Expression of MMP-1 in invasive well-differentiated thyroid carcinoma. *Eur Arch Otorhinolaryngol* **268**, 131-135 (2011).
213. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nature Reviews Cancer* **9**, 400-414 (2009).
214. Shamloo, B. & Usluer, S. p21 in Cancer Research. *Cancers* **11**, 1178 (2019).
215. Shi, Y., Zou, M., Farid, N.R.*, et al.* Evidence of gene deletion of p21 (WAF1/CIP1), a cyclin-dependent protein kinase inhibitor, in thyroid carcinomas. *Br J Cancer* **74**, 1336-1341 (1996).
216. Ito, Y., Kobayashi, T., Takeda, T.*, et al.* Expression of p21 (WAF1/CIP1) protein in clinical thyroid tissues. *Br J Cancer* **74**, 1269-1274 (1996).
217. Pickett, C.A., Agoff, S.N., Widman, T.J.*, et al.* Altered expression of cyclins and cell cycle inhibitors in papillary thyroid cancer: prognostic implications. *Thyroid* **15**, 461-473 (2005).
218. Brzeziński, J., Migodziński, A., Toczek, A.*, et al.* Patterns of Cyclin E, Retinoblastoma Protein, and p21 Cip1/WAF1 Immunostaining in the Oncogenesis of Papillary Thyroid Carcinoma. *Clin Cancer Res* **11**, 1037-1043 (2005).
219. Siironen, P., Nordling, S., Louhimo, J.*, et al.* Immunohistochemical expression of Bcl-2, Ki-67, and p21 in patients with papillary thyroid cancer. *Tumour Biol* **26**, 50-56 (2005).
220. Zhao, F., Yang, G., Feng, M.*, et al.* Expression, function and clinical application of stanniocalcin-1 in cancer. *J Cell Mol Med* **24**, 7686-7696 (2020).
221. Argente, J., Chowen, J.A., Pérez-Jurado, L.A.*, et al.* One level up: abnormal proteolytic regulation of IGF activity plays a role in human pathophysiology. *EMBO molecular medicine* **9**, 1338-1345 (2017).
222. Hayase, S., Sasaki, Y., Matsubara, T.*, et al.* Expression of stanniocalcin 1 in thyroid side population cells and thyroid cancer cells. *Thyroid* **25**, 425-436 (2015).
223. Al-Abdallah, A., Jahanbani, I., Mehdawi, H.*, et al.* The stress-activated protein kinase pathway and the expression of stanniocalcin-1 are regulated by miR-146b-5p in papillary thyroid carcinogenesis. *Cancer Biol Ther* **21**, 412-423 (2020).

224. Dai, D., Wang, Q., Li, X.*, et al.* Klotho inhibits human follicular thyroid cancer cell growth and promotes apoptosis through regulation of the expression of stanniocalcin-1. *Oncol Rep* **35**, 552-558 (2016).
225. Zhang, H., Li, S. & Li, W. Development of a Prognostic Signature Based on Hypoxia-related Genes for Thyroid Cancer. *Research Square* (2021).
226. Xu, F., Xu, H., Li, Z.*, et al.* Glycolysis-Based Genes Are Potential Biomarkers in Thyroid Cancer. *Frontiers in Oncology* **11**(2021).
227. Yeung, H.Y., Lai, K.P., Chan, H.Y.*, et al.* Hypoxia-inducible factor-1-mediated activation of stanniocalcin-1 in human cancer cells. *Endocrinology* **146**, 4951-4960 (2005).
228. Law, A.Y., Ching, L.Y., Lai, K.P.*, et al.* Identification and characterization of the hypoxia-responsive element in human stanniocalcin-1 gene. *Mol Cell Endocrinol* **314**, 118-127 (2010).
229. Ma, X., Gu, L., Li, H.*, et al.* Hypoxia-induced overexpression of stanniocalcin-1 is associated with the metastasis of early stage clear cell renal cell carcinoma. *Journal of Translational Medicine* **13**, 56 (2015).
230. Feldser, D., Agani, F., Iyer, N.V.*, et al.* Reciprocal Positive Regulation of Hypoxia-inducible Factor 1α and Insulin-like Growth Factor 2. *Cancer Research* **59**, 3915-3918 (1999).
231. Xu, X., Lu, Y., Li, Y.*, et al.* Sonic Hedgehog Signaling in Thyroid Cancer. *Frontiers in Endocrinology* **8**, 284-284 (2017).
232. Xu, X., Ding, H., Rao, G.*, et al.* Activation of the Sonic Hedgehog pathway in thyroid neoplasms and its potential role in tumor cell proliferation. *Endocr Relat Cancer* **19**, 167-179 (2012).
233. Agarwal, N.K., Kim, C.-H., Kunkalla, K.*, et al.* Smoothened (SMO) regulates insulin-like growth factor 1 receptor (IGF1R) levels and protein kinase B (AKT) localization and signaling. *Laboratory Investigation* (2021).
234. Oh, S., Shin, S. & Janknecht, R. The small members of the JMJD protein family: Enzymatic jewels or jinxes? *Biochimica et biophysica acta. Reviews on cancer* **1871**, 406-418 (2019).
235. Hübel, C., Gaspar, H.A., Coleman, J.R.I.*, et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am J Med Genet B Neuropsychiatr Genet* **180**, 428-438 (2019).

236. Christakoudi, S., Evangelou, E., Riboli, E.*, et al.* GWAS of allometric body-shape indices in UK Biobank identifies loci suggesting associations with morphogenesis, organogenesis, adrenal cell renewal and cancer. *Scientific Reports* **11**, 10688 (2021).
237. Pulit, S.L., Stoneman, C., Morris, A.P.*, et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* **28**, 166-174 (2019).
238. Zhu, Z., Guo, Y., Shi, H.*, et al.* Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. *J Allergy Clin Immunol* **145**, 537-549 (2020).
239. Milne, R.L., Kuchenbaecker, K.B., Michailidou, K.*, et al.* Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nature Genetics* **49**, 1767-1778 (2017).
240. Michailidou, K., Beesley, J., Lindstrom, S.*, et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nature Genetics* **47**, 373-380 (2015).
241. Baxter, J.S., Johnson, N., Tomczyk, K.*, et al.* Functional annotation of the 2q35 breast cancer risk locus implicates a structural variant in influencing activity of a long-range enhancer element. *The American Journal of Human Genetics* **108**, 1190-1203 (2021).
242. Truong, T., Lesueur, F., Sugier, P.E.*, et al.* Multiethnic genome-wide association study of differentiated thyroid cancer in the EPITHYR consortium. *Int J Cancer* **148**, 2935-2946 (2021).
243. Hińcza, K., Kowalik, A., Pałyga, I.*, et al.* Does the TT Variant of the rs966423 Polymorphism in DIRC3 Affect the Stage and Clinical Course of Papillary Thyroid Cancer? *Cancers* **12**, 423 (2020).
244. Guibon, J., Sugier, P.E., Kulkarni, O.*, et al.* Fine-mapping of two differentiated thyroid carcinoma susceptibility loci at 2q35 and 8p12 in Europeans, Melanesians and Polynesians. *Oncotarget* **12**, 493-506 (2021).
245. Gong, J., Mei, S., Liu, C.*, et al.* PancanQTL: systematic identification of cis-eQTLs and trans-eQTLs in 33 cancer types. *Nucleic Acids Research* **46**, D971-d976 (2018).
246. Kulkarni, O., Sugier, P.E., Guibon, J.*, et al.* Gene network and biological pathways associated with susceptibility to differentiated thyroid carcinoma. *Scientific Reports* **11**, 8932 (2021).
247. Coe, E.A., Tan, J.Y., Shapiro, M.*, et al.* The MITF-SOX10 regulated long non-coding RNA DIRC3 is a melanoma tumour suppressor. *PLoS Genet* **15**, e1008501 (2019).

248. Coe, E.A. Identification and characterisation of MITF-regulated long non-coding RNA candidate regulators of melanoma. *PhD thesis, University of Bath (2020)*.
249. Wang, J., Ding, N., Li, Y.*, et al.* Insulin-like growth factor binding protein 5 (IGFBP5) functions as a tumor suppressor in human melanoma cells. *Oncotarget* **6**, 20636-20649 (2015).
250. Margue, C., Philippidou, D., Reinsbach, S.E.*, et al.* New target genes of MITF-induced microRNA-211 contribute to melanoma cell invasion. *PloS one* **8**, e73473-e73473 (2013).
251. Uhlén, M., Fagerberg, L., Hallström, B.M.*, et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
252. Iams, W.T. & Lovly, C.M. Molecular Pathways: Clinical Applications and Future Direction of Insulin-like Growth Factor-1 Receptor Pathway Blockade. *Clin Cancer Res* **21**, 4270-4277 (2015).
253. Duan, C. & Allard, J.B. Insulin-Like Growth Factor Binding Protein-5 in Physiology and Disease. *Frontiers in Endocrinology* **11**, 100-100 (2020).
254. Werner, H. & Laron, Z. Role of the GH-IGF1 system in progression of cancer. *Mol Cell Endocrinol* **518**, 111003 (2020).
255. Santisteban, P., Acebrón, A., Polycarpou-Schwarz, M.*, et al.* Insulin and insulin-like growth factor I regulate a thyroid-specific nuclear protein that binds to the thyroglobulin promoter. *Mol Endocrinol* **6**, 1310-1317 (1992).
256. Shevah, O. & Laron, Z. Patients with congenital deficiency of IGF-I seem protected from the development of malignancies: a preliminary report. *Growth Horm IGF Res* **17**, 54-57 (2007).
257. Belfiore, A., Pandini, G., Vella, V.*, et al.* Insulin/IGF-I hybrid receptors play a major role in IGF-I signaling in thyroid cancer. *Biochimie* **81**, 403-407 (1999).
258. Karagiannis, A., Kassi, E., Chatzigeorgiou, A.*, et al.* IGF Bioregulation System in Benign and Malignant Thyroid Nodular Disease: A Systematic Review. *In Vivo* **34**, 3069-3091 (2020).
259. Vella, V., Pandini, G., Sciacca, L.*, et al.* A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer. *J Clin Endocrinol Metab* **87**, 245-254 (2002).
260. Panebianco, F., Kelly, L.M., Liu, P.*, et al.* THADA fusion is a mechanism of IGF2BP3 activation and IGF1R signaling in thyroid cancer. *Proceedings of the National Academy of Sciences*, 201614265 (2017).

261. Du, P., Liu, F., Liu, Y.*, et al.* Linc00210 enhances the malignancy of thyroid cancer cells by modulating miR-195-5p/IGF1R/Akt axis. *Journal of Cellular Physiology* **235**, 1001-1012 (2020).
262. Tita, P., Ambrosio, M.R., Scollo, C.*, et al.* High prevalence of differentiated thyroid carcinoma in acromegaly. *Clin Endocrinol (Oxf)* **63**, 161-167 (2005).
263. Dagdelen, S., Cinar, N. & Erbas, T. Increased thyroid cancer risk in acromegaly. *Pituitary* **17**, 299-306 (2014).
264. Wolinski, K., Czarnywojtek, A. & Ruchala, M. Risk of Thyroid Nodular Disease and Thyroid Cancer in Patients with Acromegaly – Meta-Analysis and Systematic Review. *PloS one* **9**, e88787 (2014).
265. Woliński, K., Stangierski, A., Gurgul, E.*, et al.* Thyroid lesions in patients with acromegaly — case-control study and update to the meta-analysis. *Endokrynologia Polska* **68**, 2-6 (2017).
266. Allard, J.B. & Duan, C. IGF-Binding Proteins: Why Do They Exist and Why Are There So Many? *Frontiers in Endocrinology* **9**(2018).
267. Luther, G.A., Lamplot, J., Chen, X.*, et al.* IGFBP5 domains exert distinct inhibitory effects on the tumorigenicity and metastasis of human osteosarcoma. *Cancer Lett* **336**, 222-230 (2013).
268. Krywicki, R.F., Figueroa, J.A., Jackson, J.G.*, et al.* Regulation of insulin-like growth factor binding proteins in ovarian cancer cells by oestrogen. *Eur J Cancer* **29a**, 2015-2019 (1993).
269. Figueroa, J.A., Jackson, J.G., McGuire, W.L.*, et al.* Expression of insulin-like growth factor binding proteins in human breast cancer correlates with estrogen receptor status. *J Cell Biochem* **52**, 196-205 (1993).
270. Naciff, J.M., Khambatta, Z.S., Thomason, R.G.*, et al.* The genomic response of a human uterine endometrial adenocarcinoma cell line to 17alpha-ethynyl estradiol. *Toxicol Sci* **107**, 40-55 (2009).
271. Taylor, K.J., Sims, A.H., Liang, L.*, et al.* Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res* **12**, R39 (2010).
272. Walker, G., MacLeod, K., Williams, A.R.*, et al.* Insulin-like growth factor binding proteins IGFBP3, IGFBP4, and IGFBP5 predict endocrine responsiveness in patients with ovarian cancer. *Clin Cancer Res* **13**, 1438-1444 (2007).

273. Gregory, C.W., Kim, D., Ye, P.*, et al.* Androgen receptor up-regulates insulin-like growth factor binding protein-5 (IGFBP-5) expression in a human prostate cancer xenograft. *Endocrinology* **140**, 2372-2381 (1999).

274. Salih, D.A.M., Tripathi, G., Holding, C.*, et al.* Insulin-like growth factor-binding protein 5 (Igfbp5) compromises survival, growth, muscle development, and fertility in mice. *Proceedings of the National Academy of Sciences* **101**, 4314-4319 (2004).

275. Zhong, J., Deng, J., Ghetti, B.*, et al.* Inhibition of insulin-like growth factor I activity contributes to the premature apoptosis of cerebellar granule neuron in weaver mutant mice: in vitro analysis. *J Neurosci Res* **70**, 36-45 (2002).

276. Ning, Y., Hoang, B., Schuller, A.G.*, et al.* Delayed mammary gland involution in mice with mutation of the insulin-like growth factor binding protein 5 gene. *Endocrinology* **148**, 2138-2147 (2007).

277. Hsieh, T., Gordon, R.E., Clemmons, D.R.*, et al.* Regulation of vascular smooth muscle cell responses to insulin-like growth factor (IGF)-I by local IGF-binding proteins. *J Biol Chem* **278**, 42886-42892 (2003).

278. Miyake, H., Pollak, M. & Gleave, M.E. Castration-induced up-regulation of insulin-like growth factor binding protein-5 potentiates insulin-like growth factor-I activity and accelerates progression to androgen independence in prostate cancer models. *Cancer Research* **60**, 3058-3064 (2000).

279. Liu, C., Xin, Y., Bai, Y.*, et al.* Ca(2+) concentration-dependent premature death of igfbp5a(-/-) fish reveals a critical role of IGF signaling in adaptive epithelial growth. *Science Signaling* **11**(2018).

280. Dai, W., Bai, Y., Hebda, L.*, et al.* Calcium deficiency-induced and TRP channel-regulated IGF1R-PI3K-Akt signaling regulates abnormal epithelial cell proliferation. *Cell Death & Differentiation* **21**, 568-581 (2014).

281. Li, S., Liu, C., Goldstein, A.*, et al.* Calcium state-dependent regulation of epithelial cell quiescence by Stanniocalcin 1a. *Frontiers in Cell and Developmental Biology* **9**, 681 (2021).

282. Maeda, H., Yonou, H., Yano, K.*, et al.* Prostate-specific antigen enhances bioavailability of insulin-like growth factor by degrading insulin-like growth factor binding protein 5. *Biochem Biophys Res Commun* **381**, 311-316 (2009).

283. Imai, Y., Busby, W.H., Jr., Smith, C.E.*, et al.* Protease-resistant form of insulin-like growth factor-binding protein 5 is an inhibitor of insulin-like growth factor-I actions on porcine smooth muscle cells in culture. *J Clin Invest* **100**, 2596-2605 (1997).

284. Zheng, B., Duan, C. & Clemmons, D.R. The effect of extracellular matrix proteins on porcine smooth muscle cell insulin-like growth factor (IGF) binding protein-5 synthesis and responsiveness to IGF-I. *J Biol Chem* **273**, 8994-9000 (1998).

285. Nam, T., Moralez, A. & Clemmons, D. Vitronectin binding to IGF Binding Protein-5 (IGFBP-5) alters IGFBP-5 modulation of IGF-I actions. *Endocrinology* **143**, 30-36 (2002).

286. Xu, Q., Yan, B., Li, S.*, et al.* Fibronectin binds insulin-like growth factor-binding protein 5 and abolishes its ligand-dependent action on cell migration. *J Biol Chem* **279**, 4269-4277 (2004).

287. Tripathi, G., Salih, D.A., Drozd, A.C.*, et al.* IGF-independent effects of insulin-like growth factor binding protein-5 (Igfbp5) in vivo. *Faseb J* **23**, 2616-2626 (2009).

288. Sureshbabu, A., Okajima, H., Yamanaka, D.*, et al.* IGFBP5 induces cell adhesion, increases cell survival and inhibits cell migration in MCF-7 human breast cancer cells. *Journal of Cell Science* **125**, 1693-1705 (2012).

289. Liu, H., Li, R., Guan, L.*, et al.* Knockdown of lncRNA UCA1 inhibits proliferation and invasion of papillary thyroid carcinoma through regulating miR-204/IGFBP5 axis. *OncoTargets and therapy* **11**, 7197-7204 (2018).

290. Wang, H., Wang, H., Zhang, W.*, et al.* Overexpression of IGFBP5, but not IGFBP3, correlates with the histologic grade of human diffuse glioma: a tissue microarray and immunohistochemical study. *Technol Cancer Res Treat* **5**, 195-199 (2006).

291. Wang, H., Rosen, D.G., Wang, H.*, et al.* Insulin-like growth factor-binding protein 2 and 5 are differentially regulated in ovarian cancer of different histologic types. *Modern Pathology* **19**, 1149-1156 (2006).

292. Li, X., Cao, X., Li, X.*, et al.* Expression level of insulin-like growth factor binding protein 5 mRNA is a prognostic factor for breast cancer. *Cancer Sci* **98**, 1592-1596 (2007).

293. Liang, P.I., Wang, Y.H., Wu, T.F.*, et al.* IGFBP-5 overexpression as a poor prognostic factor in patients with urothelial carcinomas of upper urinary tracts and urinary bladder. *J Clin Pathol* **66**, 573-582 (2013).

294. Becker, M.A., Hou, X., Harrington, S.C.*, et al.* IGFBP ratio confers resistance to IGF targeting and correlates with increased invasion and poor outcome in breast tumors. *Clin Cancer Res* **18**, 1808-1817 (2012).

295. Mita, K., Zhang, Z., Ando, Y.*, et al.* Prognostic significance of insulin-like growth factor binding protein (IGFBP)-4 and IGFBP-5 expression in breast cancer. *Jpn J Clin Oncol* **37**, 575-582 (2007).
296. Wang, S., Hong, Q., Geng, X.*, et al.* Insulin-Like Growth Factor Binding Protein 5-A Probable Target of Kidney Renal Papillary Renal Cell Carcinoma. *Biomed Res Int* **2019**, 3210324 (2019).
297. Santosh, V., Arivazhagan, A., Sreekanthreddy, P.*, et al.* Grade-specific expression of insulin-like growth factor-binding proteins-2, -3, and -5 in astrocytomas: IGFBP-3 emerges as a strong predictor of survival in patients with newly diagnosed glioblastoma. *Cancer Epidemiol Biomarkers Prev* **19**, 1399-1408 (2010).
298. Zheng, R., Chen, W., Xia, W.*, et al.* The Prognostic Values of the Insulin-Like Growth Factor Binding Protein Family in Ovarian Cancer. *Biomed Res Int* **2020**, 7658782 (2020).
299. Hao, X., Sun, B., Hu, L.*, et al.* Differential gene and protein expression in primary breast malignancies and their lymph node metastases as revealed by combined cDNA microarray and tissue microarray analysis. *Cancer* **100**, 1110-1122 (2004).
300. Wang, H., Arun, B.K., Wang, H.*, et al.* IGFBP2 and IGFBP5 overexpression correlates with the lymph node metastasis in T1 breast carcinomas. *Breast J* **14**, 261-267 (2008).
301. Wang, W., Lim, K.G., Feng, M.*, et al.* KDM6B Counteracts EZH2-Mediated Suppression of IGFBP5 to Confer Resistance to PI3K/AKT Inhibitor Treatment in Breast Cancer. *Mol Cancer Ther* **17**, 1973-1983 (2018).
302. Hermani, A., Shukla, A., Medunjanin, S.*, et al.* Insulin-like growth factor binding protein-4 and -5 modulate ligand-dependent estrogen receptor-α activation in breast cancer cells in an IGF-independent manner. *Cellular Signalling* **25**, 1395-1402 (2013).
303. Butt, A.J., Dickson, K.A., McDougall, F.*, et al.* Insulin-like growth factor-binding protein-5 inhibits the growth of human breast cancer cells in vitro and in vivo. *J Biol Chem* **278**, 29676-29685 (2003).
304. Akkiprik, M., Hu, L., Sahin, A.*, et al.* The subcellular localization of IGFBP5 affects its cell growth and migration functions in breast cancer. *BMC Cancer* **9**, 103 (2009).
305. Liu, B.Y., Soloviev, I., Huang, X.*, et al.* Mammary tumor regression elicited by Wnt signaling inhibitor requires IGFBP5. *Cancer Research* **72**, 1568-1578 (2012).
306. Hawsawi, Y., Humphries, M.P., Wright, A.*, et al.* Deregulation of IGF-binding proteins -2 and -5 contributes to the development of endocrine resistant breast cancer in vitro. *Oncotarget* **7**, 32129-32143 (2016).

307. Becker, M., Sommer, A., Krätzschmar, J.R., *et al.* Distinct gene expression patterns in a tamoxifen-sensitive human mammary carcinoma xenograft and its tamoxifen-resistant subline MaCa 3366/TAM. *Mol Cancer Ther* **4**, 151-170 (2005).
308. Yamashita, H., Takahashi, S., Ito, Y., *et al.* Predictors of response to exemestane as primary endocrine therapy in estrogen receptor-positive breast cancer. *Cancer Sci* **100**, 2028-2033 (2009).
309. Ahn, B.Y., Elwi, A.N., Lee, B., *et al.* Genetic screen identifies insulin-like growth factor binding protein 5 as a modulator of tamoxifen resistance in breast cancer. *Cancer Research* **70**, 3013-3019 (2010).
310. Hwang, J.R., Cho, Y.J., Lee, Y., *et al.* The C-terminus of IGFBP-5 suppresses tumor growth by inhibiting angiogenesis. *Scientific Reports* **6**, 39334 (2016).
311. Rho, S.B., Dong, S.M., Kang, S., *et al.* Insulin-like growth factor-binding protein-5 (IGFBP-5) acts as a tumor suppressor by inhibiting angiogenesis. *Carcinogenesis* **29**, 2106-2111 (2008).
312. Chen, X., Yu, Q., Pan, H., *et al.* Overexpression of IGFBP5 Enhances Radiosensitivity Through PI3K-AKT Pathway in Prostate Cancer. *Cancer Manag Res* **12**, 5409-5418 (2020).
313. Johnson, S.K., Dennis, R.A., Barone, G.W., *et al.* Differential expression of insulin-like growth factor binding protein-5 in pancreatic adenocarcinomas: identification using DNA microarray. *Mol Carcinog* **45**, 814-827 (2006).
314. Ting, D.T., Wittner, B.S., Ligorio, M., *et al.* Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Reports* **8**, 1905-1918 (2014).
315. Johnson, S.K. & Haun, R.S. Insulin-like growth factor binding protein-5 influences pancreatic cancer cell growth. *World J Gastroenterol* **15**, 3355-3366 (2009).
316. Han, A.L., Veeneman, B.A., El-Sawy, L., *et al.* Fibulin-3 promotes muscle-invasive bladder cancer. *Oncogene* **36**, 5243-5251 (2017).
317. Dong, C., Zhang, J., Fang, S., *et al.* IGFBP5 increases cell invasion and inhibits cell proliferation by EMT and Akt signaling pathway in Glioblastoma multiforme cells. *Cell Div* **15**, 4 (2020).
318. Su, Y., Wagner, E.R., Luo, Q., *et al.* Insulin-like growth factor binding protein 5 suppresses tumor growth and metastasis of human osteosarcoma. *Oncogene* **30**, 3907-3917 (2011).

319. Schneider, M.R., Zhou, R., Hoeflich, A.*, et al.* Insulin-like growth factor-binding protein-5 inhibits growth and induces differentiation of mouse osteosarcoma cells. *Biochem Biophys Res Commun* **288**, 435-442 (2001).
320. Hung, P.S., Kao, S.Y., Shih, Y.H.*, et al.* Insulin-like growth factor binding protein-5 (IGFBP-5) suppresses the tumourigenesis of head and neck squamous cell carcinoma. *The Journal of Pathology* **214**, 368-376 (2008).
321. Hou, X.J., Zhang, Y.Z., Liu, X.*, et al.* Expressions of IGFBP-5, cFLIP in cervical intraepithelial neoplasia, cervical carcinoma and their clinical significances: a molecular pathology. *J Exp Clin Cancer Res* **28**, 70 (2009).
322. Miyatake, T., Ueda, Y., Nakashima, R.*, et al.* Down-regulation of insulin-like growth factor binding protein-5 (IGFBP-5): novel marker for cervical carcinogenesis. *Int J Cancer* **120**, 2068-2077 (2007).
323. Zhang, L., Li, W., Cao, L.*, et al.* PKNOX2 suppresses gastric cancer through the transcriptional activation of IGFBP5 and p53. *Oncogene* **38**, 4590-4604 (2019).
324. Weng, X., Wu, J., Lv, Z.*, et al.* Targeting Mybbp1a suppresses HCC progression via inhibiting IGF1/AKT pathway by CpG islands hypo-methylation dependent promotion of IGFBP5. *EBioMedicine* **44**, 225-236 (2019).
325. Umemura, A., Itoh, Y., Itoh, K.*, et al.* Association of gankyrin protein expression with early clinical stages and insulin-like growth factor-binding protein 5 expression in human hepatocellular carcinoma. *Hepatology* **47**, 493-502 (2008).
326. Maitituoheti, M., Keung, E.Z., Tang, M.*, et al.* Enhancer Reprogramming Confers Dependence on Glycolysis and IGF Signaling in KMT2D Mutant Melanoma. *Cell Reports* **33**, 108293 (2020).
327. Masoud, G.N. & Li, W. HIF-1α pathway: role, regulation and intervention for cancer therapy. *Acta Pharmaceutica Sinica B* **5**, 378-389 (2015).
328. Castro, J., Garcia, R.I., Kwok, S.*, et al.* Functional recovery with recombinant human IGF1 treatment in a mouse model of Rett Syndrome. *Proceedings of the National Academy of Sciences* **111**, 9941-9946 (2014).
329. Zamykal, M., Martens, T., Matschke, J.*, et al.* Inhibition of intracerebral glioblastoma growth by targeting the insulin-like growth factor 1 receptor involves different context-dependent mechanisms. *Neuro-Oncology* **17**, 1076-1085 (2014).
330. Liu, R., Bishop, J., Zhu, G.*, et al.* Mortality Risk Stratification by Combining BRAF V600E and TERT Promoter Mutations in Papillary Thyroid Cancer: Genetic Duet of

BRAF and TERT Promoter Mutations in Thyroid Cancer Mortality. *JAMA Oncology* **3**, 202-208 (2017).

331. Melo, M., da Rocha, A.G., Vinagre, J.*, et al.* TERT Promoter Mutations Are a Major Indicator of Poor Outcome in Differentiated Thyroid Carcinomas. *J Clin Endocrinol Metab* **99**, E754-E765 (2014).
332. Dolgin, E. IGF-1R drugs travel from cancer cradle to Graves'. *Nature Biotechnology* **38**, 385-388 (2020).
333. Pavlicek, A., Lira, M.E., Lee, N.V.*, et al.* Molecular predictors of sensitivity to the insulin-like growth factor 1 receptor inhibitor Figitumumab (CP-751,871). *Mol Cancer Ther* **12**, 2929-2939 (2013).
334. Neuzillet, Y., Chapeaublanc, E., Krucker, C.*, et al.* IGF1R activation and the in vitro antiproliferative efficacy of IGF1R inhibitor are inversely correlated with IGFBP5 expression in bladder cancer. *BMC Cancer* **17**, 636 (2017).
335. Jing, L., Xia, F., Du, X.*, et al.* Identification of key candidate genes and pathways in follicular variant papillary thyroid carcinoma by integrated bioinformatical analysis. *Translational Cancer Research* **9**, 477-490 (2020).
336. Pita, J.M., Figueiredo, I.F., Moura, M.M.*, et al.* Cell cycle deregulation and TP53 and RAS mutations are major events in poorly differentiated and undifferentiated thyroid carcinomas. *J Clin Endocrinol Metab* **99**, E497-507 (2014).
337. Wu, L., Ding, Y., Tong, H.*, et al.* Long noncoding RNA FER1L4 promotes the malignant processes of papillary thyroid cancer by targeting the miR-612/ Cadherin 4 axis. *Cancer Cell Int* **21**, 392 (2021).
338. Hernandez-Prera, J.C., Valderrabano, P., Creed, J.H.*, et al.* Molecular Determinants of Thyroid Nodules with Indeterminate Cytology and RAS Mutations. *Thyroid* **31**, 36-49 (2021).
339. Kebebew, E., Peng, M., Reiff, E.*, et al.* Diagnostic and extent of disease multigene assay for malignant thyroid neoplasms. *Cancer* **106**, 2592-2597 (2006).
340. Kebebew, E., Peng, M., Reiff, E.*, et al.* Diagnostic and prognostic value of angiogenesis-modulating genes in malignant thyroid neoplasms. *Surgery* **138**, 1102-1109; discussion 1109-1110 (2005).
341. Tahara, M., Schlumberger, M., Elisei, R.*, et al.* Exploratory analysis of biomarkers associated with clinical outcomes from the study of lenvatinib in differentiated cancer of the thyroid. *Eur J Cancer* **75**, 213-221 (2017).

342. Trovato, M., Villari, D., Ruggeri, R.M.*, et al.* Expression of CD30 ligand and CD30 receptor in normal thyroid and benign and malignant thyroid nodules. *Thyroid* **11**, 621-628 (2001).
343. Delys, L., Detours, V., Franc, B.*, et al.* Gene expression and the biological phenotype of papillary thyroid carcinomas. *Oncogene* **26**, 7894-7903 (2007).
344. Brennan, K., Holsinger, C., Dosiou, C.*, et al.* Development of prognostic signatures for intermediate-risk papillary thyroid cancer. *BMC Cancer* **16**, 736 (2016).
345. Filho, M.B., Marchi, F., Pinto, C.A.*, et al.* Expression profile of papillary thyroid carcinomas according to cervical lymph node metastasis status. *Annals of Oncology* **27**, vi347 (2016).
346. Lei, S.T., Shen, F., Chen, J.W.*, et al.* MiR-639 promoted cell proliferation and cell cycle in human thyroid cancer by suppressing CDKN1A expression. *Biomed Pharmacother* **84**, 1834-1840 (2016).
347. Wang, S., Chen, Y. & Bai, Y. p21 participates in the regulation of anaplastic thyroid cancer cell proliferation by miR-146b. *Oncol Lett* **12**, 2018-2022 (2016).
348. Yang, H.L., Pan, J.X., Sun, L.*, et al.* p21 Waf-1 (Cip-1) enhances apoptosis induced by manumycin and paclitaxel in anaplastic thyroid cancer cells. *J Clin Endocrinol Metab* **88**, 763-772 (2003).
349. Ito, Y., Kobayashi, T., Takeda, T.*, et al.* Expression of p21 (WAF1/CIP1) protein in clinical thyroid tissues. *Br J Cancer* **74**, 1269-1274 (1996).
350. Brzeziński, J., Migodziński, A., Toczek, A.*, et al.* Patterns of cyclin E, retinoblastoma protein, and p21Cip1/WAF1 immunostaining in the oncogenesis of papillary thyroid carcinoma. *Clin Cancer Res* **11**, 1037-1043 (2005).
351. Zedenius, J., Larsson, C., Wallin, G.*, et al.* Alterations of p53 and expression of WAF1/p21 in human thyroid tumors. *Thyroid* **6**, 1-9 (1996).
352. Shi, Y., Zou, M., Farid, N.R.*, et al.* Evidence of gene deletion of p21 (WAF1/CIP1), a cyclin-dependent protein kinase inhibitor, in thyroid carcinomas. *Br J Cancer* **74**, 1336-1341 (1996).
353. Shen, Y., Dong, S., Liu, J.*, et al.* Identification of Potential Biomarkers for Thyroid Cancer Using Bioinformatics Strategy: A Study Based on GEO Datasets. *Biomed Res Int* **2020**, 9710421-9710421 (2020).
354. Nikolova, D.N., Zembutsu, H., Sechanov, T.*, et al.* Genome-wide gene expression profiles of thyroid carcinoma: Identification of molecular targets for treatment of thyroid carcinoma. *Oncol Rep* **20**, 105-121 (2008).

355. Panagiotou, G., Ghaly, W., Upadhyay, J.*, et al.* Serum Follistatin Is Increased in Thyroid Cancer and Is Associated With Adverse Tumor Characteristics in Humans. *J Clin Endocrinol Metab* **106**, e2137-e2150 (2021).
356. Moretti, S., Nucci, N., Menicali, E.*, et al.* The Aryl Hydrocarbon Receptor Is Expressed in Thyroid Carcinoma and Appears to Mediate Epithelial-Mesenchymal-Transition. *Cancers* **12**, 145 (2020).
357. Fan, R., Dong, L., Li, P.*, et al.* Integrated bioinformatics analysis and screening of hub genes in papillary thyroid carcinoma. *PloS one* **16**, e0251962 (2021).
358. Zhang, S., Wang, Y., Chen, M.*, et al.* CXCL12 methylation-mediated epigenetic regulation of gene expression in papillary thyroid carcinoma. *Scientific Reports* **7**, 44033 (2017).
359. Liu, C., Feng, Z., Chen, T.*, et al.* Downregulation of NEAT1 reverses the radioactive iodine resistance of papillary thyroid carcinoma cell via miR-101-3p/FN1/PI3K-AKT signaling pathway. *Cell Cycle* **18**, 167-203 (2019).
360. Ma, W., Zhao, X., Liang, L.*, et al.* miR-146a and miR-146b promote proliferation, migration and invasion of follicular thyroid carcinoma via inhibition of ST8SIA4. *Oncotarget* **8**, 28028-28041 (2017).
361. Shuwen, H., Xi, Y., Miao, D.*, et al.* Nine Genes Mediate the Therapeutic Effects of Iodine-131 Radiotherapy in Thyroid Carcinoma Patients. *Dis Markers* **2020**, 9369341 (2020).
362. Jeon, M., Kim, W.G., Lim, S.*, et al.* Abstract 2042: Low Slit2 expression is associated with the aggressiveness of papillary thyroid carcinoma. *Cancer Research* **75**, 2042-2042 (2015).
363. Wu, D., Hu, S., Hou, Y.*, et al.* Identification of potential novel biomarkers to differentiate malignant thyroid nodules with cytological indeterminate. *BMC Cancer* **20**, 199 (2020).
364. Puskas, L.G., Juhasz, F., Zarva, A.*, et al.* Gene profiling identifies genes specific for well-differentiated epithelial thyroid tumors. *Cell Mol Biol (Noisy-le-grand)* **51**, 177-186 (2005).
365. Xu, F., Xu, H., Li, Z.*, et al.* Glycolysis-Based Genes Are Potential Biomarkers in Thyroid Cancer. *Front Oncol* **11**, 534838 (2021).
366. Zhang, H., Li, S. & Li, W. Development of a Prognostic Signature Based on Hypoxia-related Genes for Thyroid Cancer. *Research Square* (2021), preprint.

367. Shinohara, S. & Rothstein, J.L. Interleukin 24 is induced by the RET/PTC3 oncoprotein and is an autocrine growth factor for epithelial cells. *Oncogene* **23**, 7571-7579 (2004).
368. Puxeddu, E., Knauf, J.A., Sartor, M.A.*, et al.* RET/PTC-induced gene expression in thyroid PCCL3 cells reveals early activation of genes involved in regulation of the immune response. *Endocr Relat Cancer* **12**, 319-334 (2005).
369. Port, M., Boltze, C., Wang, Y.*, et al.* A radiation-induced gene signature distinguishes post-Chernobyl from sporadic papillary thyroid cancers. *Radiat Res* **168**, 639-649 (2007).
370. Lacroix, L., Lazar, V., Michiels, S.*, et al.* Follicular thyroid tumors with the PAX8-PPARgamma1 rearrangement display characteristic genetic alterations. *The American journal of pathology* **167**, 223-231 (2005).
371. Song, Y.S., Kang, B.-H., Lee, S.*, et al.* Genomic and Transcriptomic Characteristics According to Size of Papillary Thyroid Microcarcinoma. *Cancers* **12**, 1345 (2020).
372. Ruiz-Llorente, S., Carrillo Santa de Pau, E., Sastre-Perona, A.*, et al.* Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. *BMC genomics* **13**, 147-147 (2012).
373. Weinberger, P., Ponny, S.R., Xu, H.*, et al.* Cell Cycle M-Phase Genes Are Highly Upregulated in Anaplastic Thyroid Carcinoma. *Thyroid* **27**, 236-252 (2017).
374. Stokowy, T., Gawel, D. & Wojtas, B. Differences in miRNA and mRNA Profile of Papillary Thyroid Cancer Variants. *International Journal of Endocrinology* **2016**, 1427042-1427042 (2016).
375. Ling, L., Haraguchi, K., Ohta, K.*, et al.* Beta 2-adrenergic receptor mRNA is overexpressed in neoplastic human thyroid tissues. *Endocrinology* **130**, 547-549 (1992).
376. Comiskey, D.F., He, H., Liyanarachchi, S.*, et al.* Characterizing the function of EPB41L4A in the predisposition to papillary thyroid carcinoma. *Scientific Reports* **10**, 19984 (2020).
377. Baris, O., Mirebeau-Prunier, D., Savagner, F.*, et al.* Gene profiling reveals specific oncogenic mechanisms and signaling pathways in oncocytic and papillary thyroid carcinoma. *Oncogene* **24**, 4155-4161 (2005).
378. Hossain, M.A., Asa, T.A., Rahman, M.M.*, et al.* Network-Based Genetic Profiling Reveals Cellular Pathway Differences Between Follicular Thyroid Carcinoma and Follicular Thyroid Adenoma. *International Journal of Environmental Research and Public Health* **17**, 1373 (2020).

379. Phan, A.V. Molecular Signatures of Radiation Response in Human-Derived Anaplastic Thyroid Cancer and Poorly Differentiated Thyroid Cancer. *International Journal of Radiation Oncology, Biology, Physics* **100**, 1381-1382 (2018).
380. Xu, X., Qin, J. & Liu, W. Curcumin inhibits the invasion of thyroid cancer cells via down-regulation of PI3K/Akt signaling pathway. *Gene* **546**, 226-232 (2014).
381. Galdiero, F., Bello, A.M., Spina, A.*, et al.* Identification of BAG3 target proteins in anaplastic thyroid cancer cells by proteomic analysis. *Oncotarget* **9**, 8016-8026 (2018).
382. Ye, M., Dong, S., Hou, H.*, et al.* Oncogenic Role of Long Noncoding RNAMALAT1 in Thyroid Cancer Progression through Regulation of the miR-204/IGF2BP2/m6A-MYC Signaling. *Mol Ther Nucleic Acids* **23**, 1-12 (2021).
383. Zhang, R., Hardin, H., Huang, W.*, et al.* MALAT1 Long Non-coding RNA Expression in Thyroid Tissues: Analysis by In Situ Hybridization and Real-Time PCR. *Endocr Pathol* **28**, 7-12 (2017).
384. Covach, A., Patel, S., Hardin, H.*, et al.* Phosphorylated Mechanistic Target of Rapamycin (p-mTOR) and Noncoding RNA Expression in Follicular and Hürthle Cell Thyroid Neoplasm. *Endocr Pathol* **28**, 207-212 (2017).
385. Ren, Z.F., Du, M.F., Fu, H.*, et al.* MiR-200c promotes proliferation of papillary thyroid cancer cells via Wnt/β-catenin signaling pathway. *Eur Rev Med Pharmacol Sci* **24**, 5512-5518 (2020).
386. Champa, D., Russo, M.A., Liao, X.H.*, et al.* Obatoclax overcomes resistance to cell death in aggressive thyroid carcinomas by countering Bcl2a1 and Mcl1 overexpression. *Endocr Relat Cancer* **21**, 755-767 (2014).
387. You, X., Zhao, Y., Sui, J.*, et al.* Integrated analysis of long noncoding RNA interactions reveals the potential role in progression of human papillary thyroid cancer. *Cancer medicine* **7**, 5394-5410 (2018).
388. You, X., Yang, S., Sui, J.*, et al.* Molecular characterization of papillary thyroid carcinoma: a potential three-lncRNA prognostic signature. *Cancer Manag Res* **10**, 4297-4310 (2018).
389. Chien, M.-N., Yang, P.-S., Lee, J.-J.*, et al.* Recurrence-associated genes in papillary thyroid cancer: An analysis of data from The Cancer Genome Atlas. *Surgery* **161**, 1642-1650 (2017).
390. Russell, J.P., Engiles, J.B. & Rothstein, J.L. Proinflammatory Mediators and Genetic Background in Oncogene Mediated Tumor Progression. *The Journal of Immunology* **172**, 4059-4067 (2004).

391. Russell, J.P., Shinohara, S., Melillo, R.M.*, et al.* Tyrosine kinase oncoprotein, RET/PTC3, induces the secretion of myeloid growth and chemotactic factors. *Oncogene* **22**, 4569-4577 (2003).
392. Borrello, M.G., Alberti, L., Fischer, A.*, et al.* Induction of a proinflammatory program in normal human thyrocytes by the RET/PTC1 oncogene. *Proceedings of the National Academy of Sciences* **102**, 14825-14830 (2005).
393. Nakada, T., Sato, H., Inoue, F.*, et al.* The production of colony-stimulating factors by thyroid carcinoma is associated with marked neutrophilia and eosinophilia. *Intern Med* **35**, 815-820 (1996).
394. Vassilatou, E., Fisfis, M., Morphopoulos, G.*, et al.* Papillary thyroid carcinoma producing granulocyte-macrophage colony-stimulating factor is associated with neutrophilia and eosinophilia. *Hormones (Athens)* **5**, 303-309 (2006).
395. Galdiero, M.R., Varricchi, G., Loffredo, S.*, et al.* Potential involvement of neutrophils in human thyroid cancer. *PloS one* **13**, e0199740 (2018).

APPENDIX TABLE.

List of DEGs shared by *DIRC3* and *IGFBP5* in the gene silencing experiment in MDA-T32 cell line. Literature evidence on the gene role in thyroid cancers is described. References to the Table are provided above.

GENE	FULL NAME	EVIDENCE FOR ROLE IN THYROID CANCERS	POSSIBLE ROLE	EXPRESSION UPON *DIRC3* SILENCING
NRXN3	*Neurexin 3-alpha*			↓
COL17A1	*Collagen Type XVII Alpha 1 Chain*	Downregulated in PTCs and anaplastic thyroid cancers as compared to normal thyroid tissue [335, 336].		↑
WNT11	*Wnt Family Member 11*			↓
FER1L4	*Fer-1 Like Family Member 4, Pseudogene (Functional)*	Upregulation of *FER1L4* correlated positively with the lymph node metastasis, extrathyroidal extension and advanced TNM stages in PTCs. Knockdown of *FER1L4* suppressed cell proliferation, migration and invasion in PTCs, whereas ectopic expression of *FER1L4* promoted these processes [337].	Oncogene	↓
ANGPT2	*Angiopoietin-2*	Expression of *ANGPT2* gradually increased from non-neoplastic thyroid tissue to high-risk thyroid cancers. *ANGPT2* expression was useful to distinguish cancers from benign thyroid neoplasms. Expression of *ANGPT2* was predictive of the lenvatinib sensitivity in radioiodine-refractory PTCs [338-341].	Possible oncogene	↑
DBP	*D-box binding PAR bZIP transcription factor*			↓
TNFSF8	*Tumor Necrosis Factor (Ligand) Superfamily, Member 8, CD30L*	The proportion of CD30L+ cells correlated inversely with the PTC aggressiveness [342].		↑
AREG	*Amphiregulin*	Overexpressed in PTCs. Upregulation of *AREG* associated with worse prognosis and increased risk of cervical node metastasis in PTCs [343-345].	Possible oncogene	↑
MGP	*Matrix Gla Protein*			↓
IL1RL1	*Interleukin 1 Receptor Like 1*			↑

ACTL8	*Actin Like 8*			↑
HSD11B1	*Hydroxysteroid 11-Beta Dehydrogenase 1*			↑
CDKN1A	*Cyclin Dependent Kinase Inhibitor 1A, p21*	Inhibits proliferation and arrests the cell cycle. *CDKN1A* deletion is involved in thyroid carcinogenesis. The expression level of *CDKN1A* was useful to distinguish adenomas from PTCs. Progressive loss of p21 expression was observed with advancing clinical stage of DTCs. Expression of p21 was higher in large PTCs and in tumors extending beyond the thyroid capsule. No correlation between the p21 expression, clinical parameters and patients' prognosis could be established in one study [217, 219, 346-352].	Possibly tumor suppressor	↑
BEX1	*Brain Expressed X-Linked 1*	Downregulated in thyroid cancers [353, 354]		↑
FST	*Follistatin*	FST serum level was increased in thyroid cancer patients. Its level correlated positively with the cancer aggressiveness (metastasis, vascular invasion, TNM) [355].		↑
DGKB	*Diacylglycerol Kinase Beta*			↓
CYP1B1	*Cytochrome P450 Family 1 Subfamily B Member 1*	Expression of *CYP1B1* was elevated in PTCs, but was lower in anaplastic cancer as compared to normal thyroid tissue [354, 356, 357].		↓
FBLN7	*Fibulin 7*	Downregulated in PTCs as compared to normal thyroid tissue [357-359].		↓
ST8SIA6	*Alpha-2,8-sialyltransferase*	Downregulated in FTCs. *ST8SIA6* inhibited cell proliferation, migration and invasion in FTCs, while its upregulation decreased the cancer invasiveness [360].	Tumor suppressor	↓
FAM13C	*Family With Sequence Similarity 13 Member C*			↓
LRMDA	*Leucine Rich Melanocyte Differentiation Associated*			↓
ROBO4	*Roundabout guidance receptor 4*	Overexpressed in PTCs. Iodine therapy might exert therapeutic effects by targeting *ROBO4* in thyroid cancer cells [361, 362].		↑

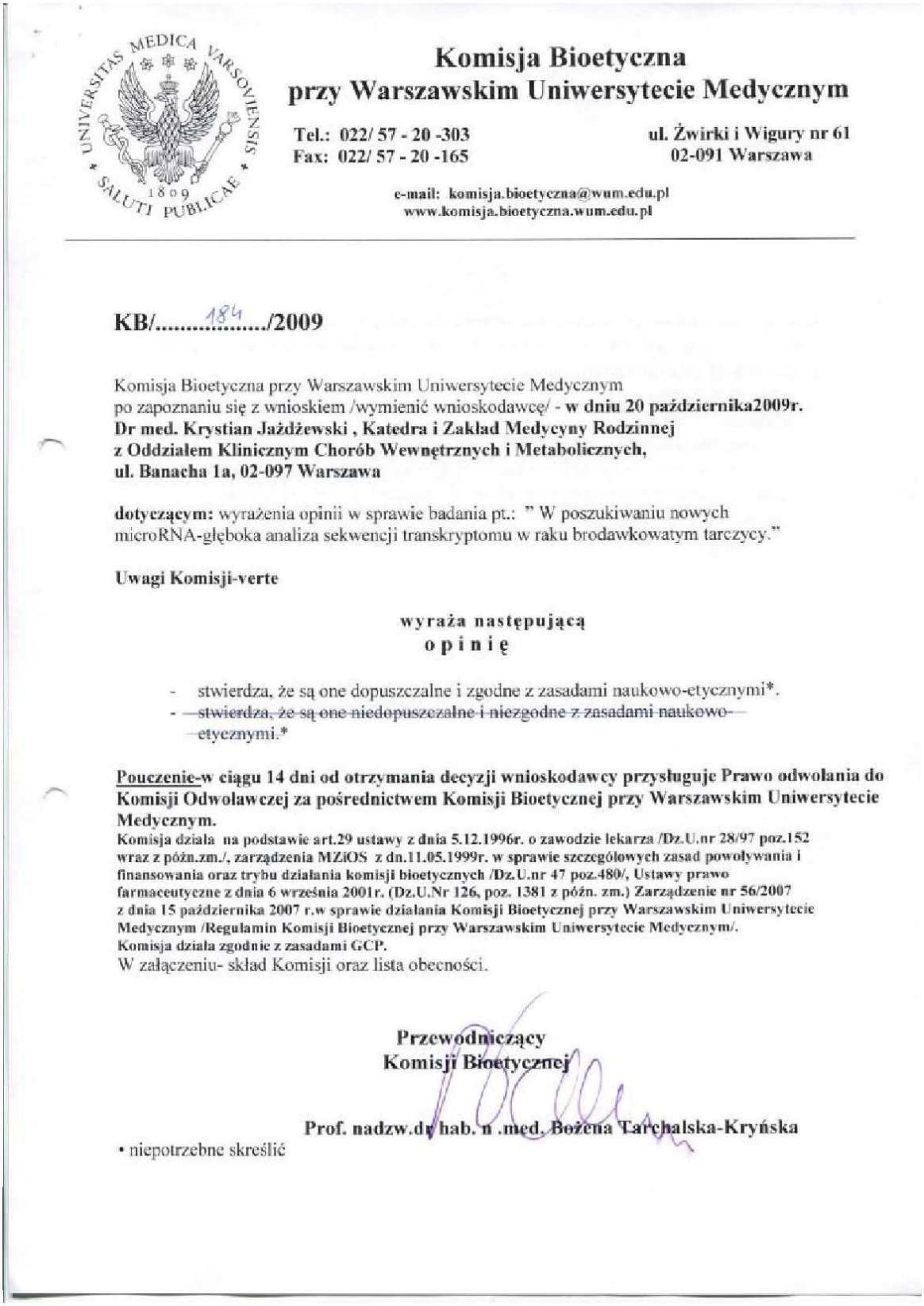
PPP2R2B	*Protein Phosphatase 2 Regulatory Subunit beta*	Downregulated in malignant thyroid nodules [363, 364]		↓
STC1	*Stanniocalcin 1*	Silencing of *STC1* inhibited the thyroid cancer cell proliferation, whereas the recombinant human STC1 enhanced cell proliferation. Altered expression in thyroid tumor cell lines and thyroid tumor tissues. Overexpression was a negative prognostic biomarker in the thyroid cancer patients [222-224, 365, 366].	Oncogene	↑
LY6K	*Lymphocyte Antigen 6 Family Member K*			↑
LRRC38	*Leucine Rich Repeat Containing 38*			↑
IL24	*Interleukin 24*	Involved in the cell cycle regulation and immune suppression. Expression induced by *RET/PTC* oncogene in PTCs. Induction of IL24 by oncogenes may support cell proliferation and tumor growth at the early stages of thyroid carcinogenesis [367, 368].	Possibly oncogene	↑
FAM198B	*Golgi Associated Kinase 1B*			↓
ESM1	*Endothelial Cell Specific Molecule 1*	Expression elevated in the radiation-induced PTCs as compared to sporadic PTCs [369]. Expression of *ESM1* was increased in FTCs harboring *PAX8-PPARγ1* rearrangement. *ESM1* may impact cell adhesion and angiogenesis [370]. *ESM1* was differentially expressed in papillary thyroid microcarcinoma and PTCs [371].		↑
MYO5B	*Myosin VB*	Expression of *MYO5B* was induced by the thyroid transcription factor PAX8. *MYO5B* may regulate the follicular polarity in thyroid cells [372].*MYO5B* was downregulated in anaplastic thyroid cancers [373].		↑
ACTBL2	*Actin Beta Like 2*	Expression of *ACTBL2* was lower in the follicular variant PTCs than in convention PTCs [374].		↓
ADRB2	*Adrenoceptor Beta 2*	Expression of *ADRB2* was higher in DTCs than in normal thyroid tissue [375].		↑
SSC5D	*Scavenger Receptor Cysteine Rich Family Member With 5 Domains*			↓

TNFSF15	*TNF Superfamily Member 15; Vascular Endothelial Cell Growth Inhibitor*	Downregulated by *EPB41L4A* (a validated PTC tumor suppressor). *TNFSF15* was involved in the proliferation and differentiation of thyroid cancer cells [376].		↑
BGN	*Biglycan*	Expression of *BGN* was higher in PTCs than in oncocytic carcinomas [377].		↓
EFHC2	*EF-Hand Domain Containing 2*			↓
OLFML2A	*Olfactomedin Like 2A*	Expression of *OLFML2A* was lower in FTCs than in thyroid adenomas [378].		↓
WDR86	*WD Repeat Domain 86*			↓
SAMD11	*Sterile alpha motif domain-containing protein 11*			↓
INSC	*Inscuteable Spindle Orientation Adaptor Protein*			↑
VSTM1	*V-Set And Transmembrane Domain Containing 1*	*VSTM1* expression impacts the radiation sensitivity of anaplastic and poorly differentiated thyroid cancers [379].		↑
MMP1	*Matrix Metallopeptidase 1*	High *MMP-1* expression was observed in undifferentiated thyroid cancer cell lines. MMP-1 level was significantly higher in follicular thyroid carcinomas than in adenomas. High expression of MMP-1 in DTCs correlated positively with the cancer aggressiveness (laryngotracheal invasion, multifocality, regional and distant metastases). Elevated expression of MMP-1 in DTCs was a risk factor for cervical metastases. No relationship between the MMP-1 expression in DTCs and the cancer invasiveness, metastasis, or recurrence was detected in another study [206-212, 380].	Oncogene	↑
SERPINB2	*Serpin Family B Member 2; Plasminogen Activator Inhibitor 2*	*SERPINB2* inhibited apoptosis in anaplastic thyroid cancer [381].		↑
TRIM15	*Tripartite motif-containing 15*			↓
AL162231.2	-			↓
AL512488.1				↑

AC121764.1	*-*			↓
AC091173.1	*-*			↑
MALAT1	*Metastasis Associated Lung Adenocarcinoma Transcript 1*	*MALAT1* promoted cell proliferation, migration, invasiveness and angiogenesis in thyroid cancers. Knockdown of *MALAT1* reduced cell migration and invasiveness in B-CPAP and K1 thyroid cancer cell lines. *MALAT1* upregulated *IGF2BP2* and enhanced expression of *MYC*, conferring a stimulatory effect on the proliferation, migration and invasiveness of thyroid cancer cells. *MALAT1* was highly expressed in normal thyroid tissues and in thyroid tumors. Expression of *MALAT1* was higher in PTCs and Hürthle cell thyroid neoplasms than in normal thyroid tissue. Its expression correlated positively with the tumor size, lymph node metastases and disease stage in PTCs. *MALAT1* was, however, downregulated in poorly differentiated thyroid cancers and anaplastic thyroid carcinomas [192, 193, 195, 197, 199, 382-384].	Oncogene	↑
TRIL	*TLR4 Interactor With Leucine Rich Repeats*			↓
LINC02407	*-*			↑
LINC00052	*-*			↓
AC107398.3	*-*			↓
AP000894.2	*-*			↓
AC004585.1	*Immune related lncRNA*			↑
DACH1	*Dachshund Family Transcription Factor 1*	Indirect evidence for an anti-proliferative role of *DACH1* in PTCs [385].		↓
FO681492.1	*-*			↓
AC002384.1	*-*			↓
AL662907.1				↓
BCL2A1	*BCL2 Related Protein A*	*BCL2A1* increased the apoptosis resistance in poorly differentiated thyroid cancers [386].		↑
KNOP1P5	*Lysine Rich Nucleolar Protein 1 Pseudogene 5*			↓

STARD4-AS1	*STARD4 Antisense RNA 1*			↓
CYP1B1-AS1	*CYP1B1 antisense RNA 1*	Upregulated in PTCs. Expression of *CYP1B1-AS1* may have a prognostic value [387, 388].		↓
MAF	*MAF BZIP Transcription Factor; V-Maf Avian Musculoaponeurotic Fibrosarcoma Oncogene*			↑
SLC26A9	*Solute Carrier Family 26 Member 9*			↓
MYO7B	*Myosin VIIB*			↑
NPR1	*Natriuretic Peptide Receptor 1*	Low expression of *NPR1* was associated with increased risk of PTC recurrence [389].		↓
CSF2	*Colony Stimulating Factor 2 (GM-CSF)*	Upregulated in PTCs. Expression of *CSF2* was induced by the *RET/PTC3* oncogene [390-395].		↑
MEGF6	*Multiple epidermal growth factor-like domains protein 6*			↓
PART1	*Prostate Androgen-Regulated Transcript 1*			↓
LRRC4C	*Leucine Rich Repeat Containing 4C*	Upregulated in widely-invasive Hürthle cell carcinomas as compared to normal thyroid tissue [37].		↑
WDR31	*WD Repeat Domain 31*			↓
FAM83A	*Family With Sequence Similarity 83 Member A*			↑
N4BP3	*NEDD4 Binding Protein 3*	Downregulated in widely-invasive Hürthle cell carcinomas as compared to normal thyroid tissue [37].		↑
ALPK3	*Alpha Kinase 3*			↓
PTGER2	*Prostaglandin E Receptor 2*			↓
RIMS3	*Regulating Synaptic Membrane Exocytosis 3*			↓
CCL26	*C-C Motif Chemokine Ligand 26*			↑

APPROVAL OF THE BIOETHICS COMMITTEE



Komisja Bioetyczna
przy Warszawskim Uniwersytecie Medycznym

Tel.: 022/ 57 - 20 -303
Fax: 022/ 57 - 20 -165

ul. Żwirki i Wigury nr 61
02-091 Warszawa

e-mail: komisja.bioetyczna@wum.edu.pl
www.komisja.bioetyczna.wum.edu.pl

KB/.....184...../2009

Komisja Bioetyczna przy Warszawskim Uniwersytecie Medycznym
po zapoznaniu się z wnioskiem /wymienić wnioskodawcę/ - **w dniu 20 października2009r.**
Dr med. Krystian Jażdżewski , Katedra i Zakład Medycyny Rodzinnej
z Oddziałem Klinicznym Chorób Wewnętrznych i Metabolicznych,
ul. Banacha 1a, 02-097 Warszawa

dotyczącym: wyrażenia opinii w sprawie badania pt.: " W poszukiwaniu nowych microRNA-głęboka analiza sekwencji transkryptomu w raku brodawkowatym tarczycy."

Uwagi Komisji-verte

wyraża następującą
o p i n i ę

- stwierdza, że są one dopuszczalne i zgodne z zasadami naukowo-etycznymi*.
- ~~stwierdza, że są one niedopuszczalne i niezgodne z zasadami naukowo-etycznymi.*~~

Pouczenie-w ciągu 14 dni od otrzymania decyzji wnioskodawcy przysługuje Prawo odwołania do Komisji Odwoławczej za pośrednictwem Komisji Bioetycznej przy Warszawskim Uniwersytecie Medycznym.

Komisja działa na podstawie art.29 ustawy z dnia 5.12.1996r. o zawodzie lekarza /Dz.U.nr 28/97 poz.152 wraz z późn.zm./, zarządzenia MZiOS z dn.11.05.1999r. w sprawie szczegółowych zasad powoływania i finansowania oraz trybu działania komisji bioetycznych /Dz.U.nr 47 poz.480/, Ustawy prawo farmaceutyczne z dnia 6 września 2001r. (Dz.U.Nr 126, poz. 1381 z późn. zm.) Zarządzenie nr 56/2007 z dnia 15 października 2007 r.w sprawie działania Komisji Bioetycznej przy Warszawskim Uniwersytecie Medycznym /Regulamin Komisji Bioetycznej przy Warszawskim Uniwersytecie Medycznym/.
Komisja działa zgodnie z zasadami GCP.
W załączeniu- skład Komisji oraz lista obecności.

Przewodniczący
Komisji Bioetycznej

Prof. nadzw.dr hab. n .med. Bożena Tarchalska-Kryńska

• niepotrzebne skreślić