Biomarker discovery for disease progression and metastasis in prostate cancer: a multi-omic approach

Sarah Wagner

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of Doctor of Philosophy

February 2019

Copyright statement

Publications

- **Wagner, S.**, Vadakekolathu, J., Tasian, S.K., Altmann, H., Bornhaeuser, M., Pockley, A.G., Ball, G.R., Rutella, S. (2019). A parsimonious 3-gene signature predicts clinical outcomes in an acute myeloid leukemia multi-cohort study (Blood Advances 2019 3:1330-1346).

- Al-Juboori, S. IK., Vadakekolathu, J., Idri, S., **Wagner, S.**, Zafeiris, D., Pearson, J. RD, Almshayakhchi, R., Caraglia, M., Desiderio, V., Miles, A. K., Boocock, D. J., Ball, G. R., and Regad, T. (2019). PYK2 promotes HER2-positive breast cancer invasion (Journal of Experimental & Clinical Cancer Research 2019 38:210).

- Mele, L., la Noce, M., Paino, F., Regad, T., **Wagner, S**., Liccardo, D., Papaccio, G., Lombardi, A., Caraglia, M., Tirino V., Desiderio, V. and Papaccio, F. (2019). Glucose-6-phosphate dehydrogenase blockade potentiates tyrosine kinase inhibitor effect on breast cancer cells through autophagy perturbation. (Journal of Experimental & Clinical Cancer Research 2019 38:160).

- Multhoff, G., Breuninger, S., Stangl, S., Werner, C., Sievert, W., Lobinger, D., Foulds, G., **Wagner, S**., Pickhard, A., Pointek, G., Pockley, A. G. (2018). Membrane Hsp70 - a novel target for the isolation of circulating tumor cells after epithelial-to-mesenchymal transition. (Front Oncol. 2018 Nov 1;8:497).

- Vadakekolathu, J., Rhonda, M., Ross, P. M., **Wagner, S.**, Reeder, S., Demirkan, G., Kuhar, J. R., Ball., G. R., Pockley, A. G., Rutella, S. (2017). Interferon-γ Induces Distinct mRNA and Protein Profiles in Acute and Chronic Myeloid Leukemia. (Blood, 130 (Supp 1), 3945A.).

- **Wagner, S.**, Ball, G., Pockley, G., Miles, A. K (2018). Application of 'omic' technologies in cancer research. (Translational Medicine Reports, 2).

- Zafeiris, D., Vadakekolathu, J., **Wagner, S**., Pockley, A., Ball, G. and Rutella, S. (2017). Discovery and application of immune biomarkers for hematological malignancies. (Expert Rev Mol Diagn. 2017 Nov;17(11):983-1000.).

Acknowledgements

First of all, I would like to thank my supervisory team, Dr David Boocock, Prof Graham Ball and Dr Amanda Miles for their constant support and guidance throughout these wonderful years. A special thanks goes to Amanda, which is not only my supervisor, but also a true friend that helped me through challenging, not only PhD, times. I will never forget this, neither Reggie's little "accident"!

Furthermore, I would like to thank Prof Robert Rees and Prof Graham Pockley for giving me the chance to work in this great research environment and for their constant support and valued advice during my studies. Another big THANKS goes to Stephen Reeder and Anne Schneider. What would the Centre be without you two?! I do not know how often you helped me with your knowledge and experience as well as your well-hidden stocks of reagents and I am very grateful for this.

I would also like to thank Dr Jayakumar Vadakekolathu with his immense knowledge and guidance. I think, I still need to find a technique you are not familiar with ☺. I also want to thank Dr Tarik Regad for sharing with me his valued knowledge on EMT and metastasis and his encouragement to believe in my skills, Dr Clare Coveney for teaching me my first steps into mass spectrometry and for helping me, together with my supervisor Dr David Boocock, in the generation of the MS, and Prof Sergio Rutella for all the experience I gained in the analysis of clinically-derived data. Certainly, all other staff members of the JvG deserve a big THANK YOU for just being great colleagues and friends, especially Dr Gemma Foulds, Dr Stephanie McArdle, Dr Murrium Ahmad and Andrew Marr.

Of course, I don't want to forget all my fellow PhD friends, NTU and visiting students I have had the chance to meet, and to work with during my 4 long years to freedom. A BIG thank goes to the van Geest students Abdullah, Divya, Sarra, Josh, Jenny, Melissa, Rukaia, Luisa, Harish and Marco. You all made my days in the office so much nicer. THANKS. I also want to say MERCI, GRAZIE and THANK YOU to Lucie, Simone and Michael. You are great scientists and friends!

I can feel lucky that I have a great group of friends that I can count on for many years, even if they are far away in good old Germany, so a big DANKE goes to Fabian, Daniel, Michi, Albo, Kevin, Tobi and Caro. I finally made it and hopefully it won't be too long until I will be back on home ground ☺. Special thanks goes to my dear friend Ramona. Even if you are far, far away (in Bavaria). You are always there for me, day and night, and I can't tell you how grateful I am to have a friend like you!

A special DANKE goes to my mum, my Dote, Hanni, and my Opa and Oma. Thank you so much for your constant support since always, during school, university and my PhD studies (I am finally done). VIELEN, VIELEN DANK. I also want to thank the rest of my family, unfortunately it would be too long to list everyone, but you all know I am grateful for your support.

And last, but definitely not least, I want to thank Andrea. My better half. Thank you very much for being there for me during all this time. I know it was not always easy, but I am finally done! GRAZIE DI TUTTO!

Table of Contents

Figures

Tables

Abbreviations

ABC	Avidin-biotin complex
ACN	Acetonitrile
APS	Ammonium persulfate
ATCC	American Tissue Culture Collection
ATP	Adenosine triphosphate
BCa	Breast cancer
bp	base pair
BPH	Benign prostate hyperplasia
BSA	Bovine Serum Albumin
C	Celsius
cDNA	complementary DNA
CI	Confidence interval
CID	Collision induced dissociation
CO^2	Carbon dioxide
cPSA	complexed PSA
CRPC	Castration-resistant prostate cancer
Ct	Cycle time
Da	Dalton
DAB	3,3'-diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DDA	Data dependent acquisition
ddH_2O	double distilled water
DIA	Data-independent acquisition
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DPBS	Dulbecco's phosphate-buffered saline
DPX	Distyrene - plasticiser - xylene
DRE	Digital rectal examination
dT	Deoxythymine
DTT	Dithiothreitol
DU145U	DU145 untreated
DU145T	DU145 treated
E	Epithelial
ECM	Extracellular matrix
EGF	Epidermal growth factor
ELISA	enzyme-linked immunosorbent assay
EMT	Epithelial to mesenchymal transition
ESCC	oesophageal squamous cell carcinoma
EtOH	Ethanol
FA	Formic acid
FC	Fold change
FCS	Foetal calf serum
FDR	False-discovery rate
FFPE	Formalin-Fixed Paraffin-Embedded
Fig	Figure
FITC	Fluorescein isothiocyanate
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
fPSA	free PSA
g	gram
GRCh	Genome Research Consortium human
GS	Gleason Score
h	hours

H_2O_2	Hydrogen peroxide
HCC	hepatocellular carcinoma
HCl	Hydrogen chloride
HGF	hepatocyte growth factor
HPLC	High Performance Liquid Chromatography
HPV	Human papillomavirus
HRP	horseradish peroxidase
HR	Hazard ratio
HRM	Hyper reaction monitoring
ID	Identifier
IDA	Information dependent acquisition
IF	Immunofluorescence
IHC	Immunohistochemistry
kDa	kiloDalton
LC	Liquid chromatography
LNM	lymph node metastasis
M	mesenchymal
m/z	mass to charge ratio
mA	milliampere
MET	Mesenchymal to epithelial transition
MHC	Major histocompatibility complex
min	minutes
ml	millilitre
MMPs	Matrix metalloproteinase
MRI	Magnetic resonance imaging
MRM	multiple reaction monitoring
mRNA	messenger RNA
MS	Mass Spectrometry
n	number
NCBI	National Centre for Biotechnology Information
NCI	National Cancer Institute
ns	not significant
OGP	Octyl β-D-glucopyranoside
OS	Overall survival
OSCC	oral squamous cell carcinoma
p	p-value
P	Proteome
P5B3U	P5B3 untreated
P5B3T	P5B3 treated
PBS	phosphate-buffered saline
PCa	Prostate cancer
PCR	polymerase chain reaction
PSA	Prostate specific antigen
PT	Pathway topology
Q	Quartile
qRT-PCR	quantitative real-time PCR
RFS	Relapse/recurrence-free survival
RIN	RNA Integrity number
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT	Room temperature
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SRM	single reaction monitoring
SWATH	sequential window acquisition of all theoretical fragment ions
T	Transcriptome

Tab	Table
TBS	Tris-buffered saline
TCGA	The Cancer Genome Atlas
TEAB	Triethylamonium bicarbonat
TEMED	Tetramethylethylenediamine
TF	Transcription factor
TGF-β	Transforming growth factor β
TMA	Tissue microarray
TNF	Tumour necrosis factor
TNM	Tumour - Node - Metastasis
TOF	Time of flight
tPSA	total PSA
Tregs	regulatory T cells
Tris	tris(hydroxymethyl)aminomethane
TRUS	Transrectal ultrasound
U	Untreated
UK	United Kingdom
µl	microliter
UTR	Untranslated region
UV	Ultraviolet
V	Volt
WHO	World Health Organisation
x	times

Abstract

Prostate cancer (PCa) is the most common cancer in men and the third most common cause of cancer-related deaths in Europe, which is primarily due to the development of metastasis, which decreases the 5-year survival rate to 30 %. The development of metastasis is the major cause of death in cancer patients and the process highly implicated in the ability of cancer cells to spread is called epithelial to mesenchymal transition (EMT). The aim of this study was to use inducible *in vitro* EMT models for the discovery of novel disease associated biomarkers through the use of multi-omics datasets.

For this, two PCa cell lines were stimulated with transforming growth factor β (TGF-β), resulting in apparent morphological changes indicating a cellular change in the direction of an increased mesenchymal morphology. Induction of EMT was confirmed using quantitative real-time PCR, immunofluorescence staining and western blot analysis. To improve the understanding of underlying changes and for the discovery of novel biomarkers, proteomic and transcriptomic profiles of both models in their induced and non-induced states were generated. Their subsequent integration highlighted 13 potential biomarkers indicative for the process of EMT in PCa and metastasis development. Out of the 13 core markers, four of these were taken forward and further validated using tissue microarrays and the *in silico* analysis of publicly available datasets. The generated results have supported the association of all 4 markers with EMT and disease progression, however two markers were identified to be of particular interest (DPYL3 and SDPR). These two markers have shown significant differences between primary PCa and castration-resistant prostate cancer (CRPC) and Gleason scoring. Furthermore, both of them were shown to be predictive for disease-recurrence. Overall, the generated results have highlighted the successful application of an integrated omics approach for the discovery of novel disease-associated biomarkers for PCa progression.

1. Chapter I – Introduction

1.1 Cancer

1.1.1 Cancer – A brief Overview

Cancer is a general term describing a large, heterogeneous group of diseases in which abnormal cells proliferate without control and develop, in the case of solid tumours, the ability to invade and disseminate to other parts of the body. These changes can be caused through the aberrant regulation of cell growth and resistance to regulatory cell signalling, which was shown to be the main cause of cancer and without intervention, this process can lead to death (Hanahan, Weinberg 2000). Due to its major health impact, a significant amount of research is undertaken in the field of cancer, which has led to an increased understanding of the initiation and development of the disease as well as improving the treatment options available to patients. Unfortunately, cancer is a multifactorial disease and the response to treatment can vary from patient to patient. For this reason, further work is urgently required to characterise cancer mechanisms and to develop tailored, personalised treatment options for each patient (Jackson, Chester 2015).

1.1.2 Cancer: Incidence and mortality

Cancer is a major cause of morbidity and mortality worldwide, and according to recent data, one in two people will develop cancer during their life (Cancer Research UK, 2017a). The most common cancer in the UK (2015) is breast cancer (15 %), followed by prostate (13 %), lung (13 %) and bowel cancer (12 %) (Cancer Research UK, 2018a), which account together for more than half (53 %) of all cancer cases occurring in the UK (Fig.1.1). These statistics highlight the clinical health burden in the population and justify the large amount of research that has been dedicated to its treatment and cure.

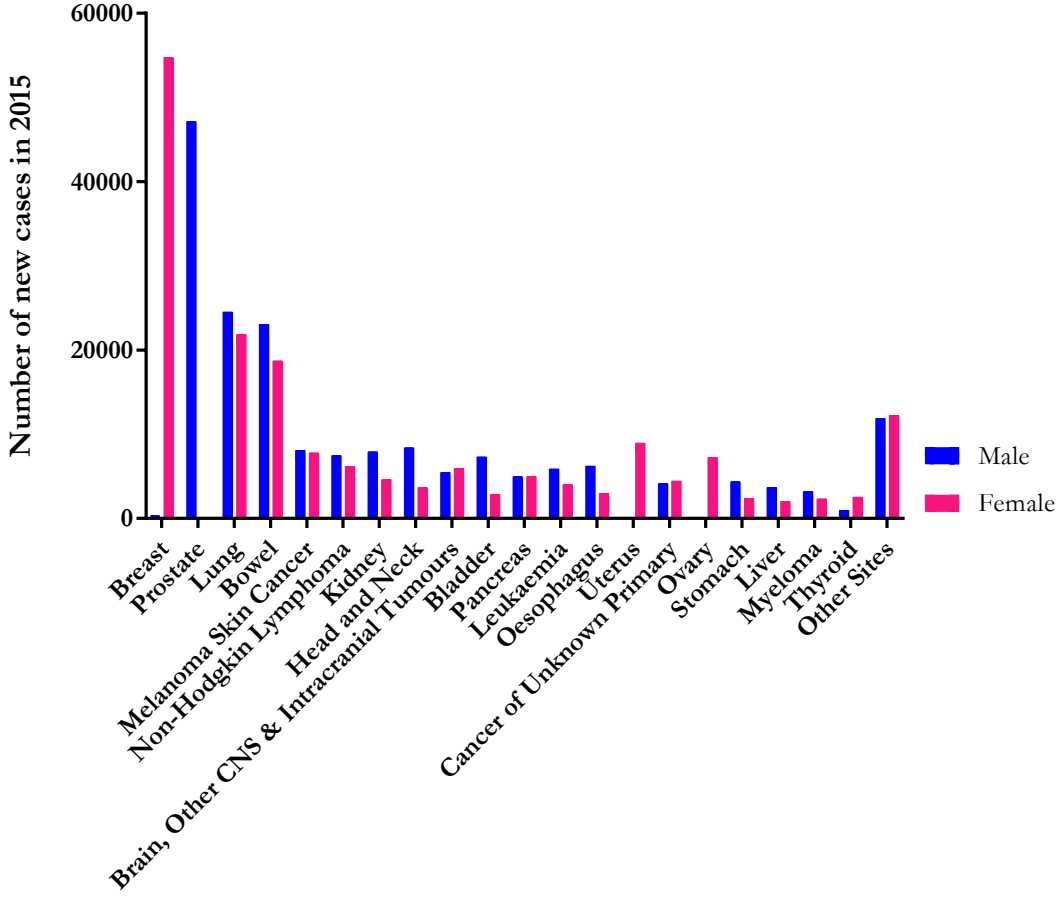


Figure 1.1: Incidence rate of the top 20 cancers in 2015 in the UK (based on a graphic created by Cancer Research UK., Cancer Research UK, 2018a). The blue colour represents the cancer incidence in the male population; the pink colour indicates the cancer incidence in the female population.

Cancer can arise from almost any part and tissue type of the body, and the disease can be classified according to the cell type it resembles or originates. Carcinomas, which have their origin in epithelial cells, are the most common kind of cancers; sarcomas arise from connective tissue such as bone and muscles, whereas lymphomas and leukaemias develop from hematopoietic cells (Cancer Research UK, 2018b). Each cancer type presents different characteristics, such as in their response to treatment or aggressiveness.

Lung cancer is the most common cause of cancer death in the UK (Fig.1.2), and when combining males and females, lung cancer accounts for more than 20 % of all cancer deaths, which is followed by bowel cancer (10 %). Breast and prostate cancer represent the third most common cause of cancer-related deaths, accounting for 7 % of all cases (Cancer Research UK, 2017b). According to the World Health Organisation (WHO), 8.8 million people died from cancer in 2015, which represents nearly one in six of all global deaths (WHO, 2017a). As mentioned before, the mortality rate can vary from cancer type

to cancer type, but in addition to the inherent variability of cancers, two factors influence the survival chances in the majority of cancers. These are the tumour stage at the time of diagnosis and the presence or absence of metastases. Organ confined primary tumours present higher chances of cure, based on the available treatment options and the chances of a complete removal of the tumour, whereas metastatic cancers have a reduced survival rate. This is mainly due to the increased treatment resistance of advanced cancers and the spread of metastatic lesions (Valastyan, Weinberg 2011).

The development of metastases is the main cause of death in cancer patients and it accounts for about 90 % of all cancer-related deaths (Chaffer, Weinberg 2011, Mehlen, Puisieux 2006). The 5-year survival rate of metastasised cancers varies depending on the type of cancer, however the presence of metastasis results overall in a shortened life expectancy with variable survival rates, such as 2.3 % in pancreatic cancer, lung cancer at 4.0 % and breast cancer at 25 % (Heerboth, Housman et al. 2015).

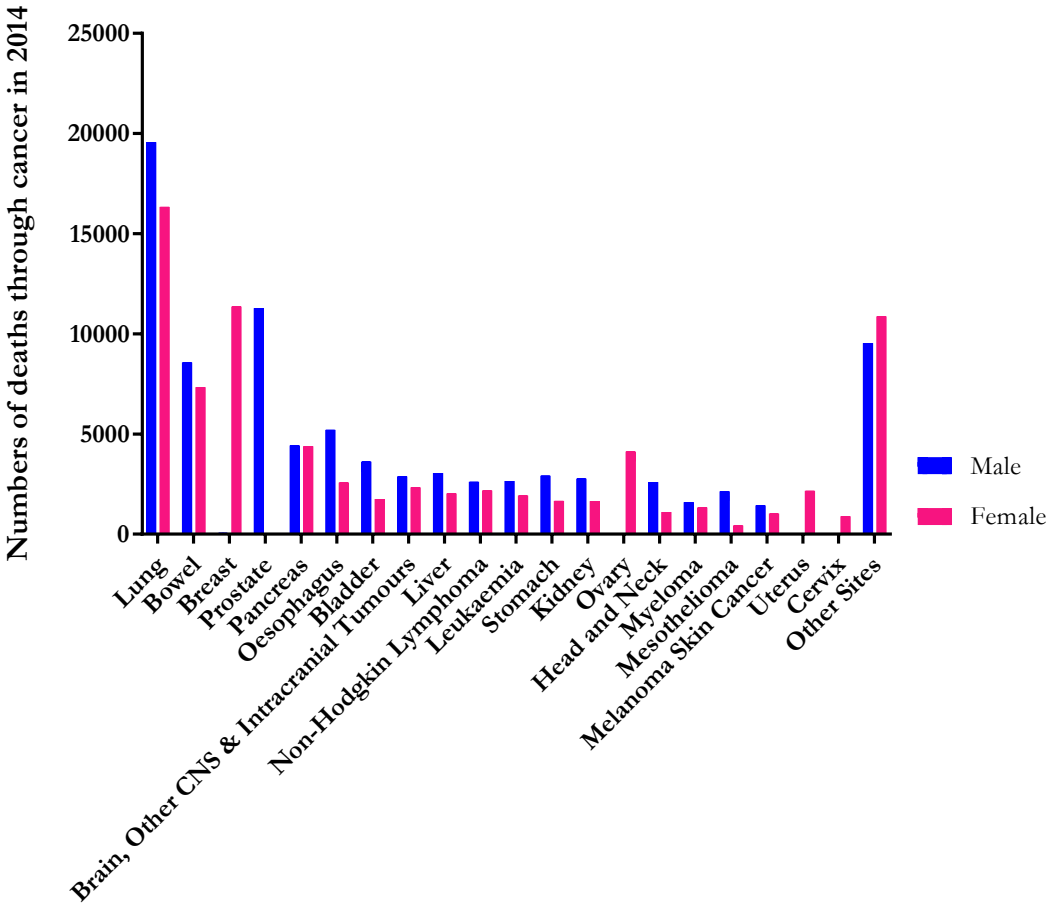


Figure 1.2: Mortality rate for the 20 most common cancers in 2014 in the UK (Based on a graphic created by Cancer Research UK, Cancer Research UK, 2017a). The blue colour represents the cancer mortality in the male population; the pink colour indicates the cancer mortality in the female population.

1.1.3 Hallmarks of Cancer

In 2000, Hanahan and Weinberg have proposed 6 capabilities that a cell has to acquire to become cancerous and to induce tumour growth and development (Fig. 1.3) (Hanahan, Weinberg 2000). During the process of carcinogenesis, healthy cells, which are responsive to regulatory signals from the surrounding microenvironment, change to cells that are non-responsive and are able to grow and invade tissue autonomously, independent from any external signalling (Bertram 2000). This independence is commonly characterised by an increased proliferation rate and a lack of response to apoptotic signals. The previously mentioned hallmarks of cancer include self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of apoptosis, limitless self-renewal capabilities, sustained angiogenesis and the ability to invade tissue and form secondary tumours, so-called metastasis. The acquisition of each of these abilities breaches an anti-cancer defence mechanism present in cells and tissue of the host (Hanahan, Weinberg 2000).

In addition to the original 6 hallmarks, Hanahan and Weinberg proposed in 2011 two further factors, the so-called "emerging hallmarks" in relation to the reprogramming of the energy metabolism in the cell and the ability of cells to evade immune destruction through the host's immune system (Hanahan, Weinberg 2011). Furthermore, Hanahan and Weinberg proposed the enabling characteristics crucial for the acquisition of the proposed hallmarks, comprised of genome instability and mutation, and tumour-promoting inflammation. In summary, this shows that the development of cancer is a highly complex disease induced by multiple steps and changes within the system of healthy cells. In 2017, Fouad & Aanei proposed a more precise definition of the hallmarks of cancer as "*acquired evolutionary-advantageous characteristics that complementarily promote transformation of phenotypically normal cells into malignant ones, and promote progression of malignant cells which sacrificing/exploiting host tissue*" (Fouad, Aanei 2017).

Normally, cells are tightly controlled and their development is based on the interaction with the surrounding cells. External signals determine whether a cell will differentiate, proliferate, migrate or undergo apoptosis. These controls ensure the healthy function of the tissue. Random mutations within genes involved in the control of proliferation and apoptosis influence the cell's fate and the induction of cancer (Giancotti 2014). Faults in the regulatory systems of cells and the generation of cells that are "immune" to external stimuli are commonly caused through genetic aberrations (DNA damage) (Hanahan,

Weinberg 2011, Hanahan, Weinberg 2000). However, variations within the epigenetic landscape, and therefore the reduced or increased expression of genes, can also contribute to the development of cancer (Sharma, Kelly et al. 2010). The DNA damage can be based on single point mutations, in which one or a few base pairs are changed, to large chromosomal aberrations. A crucial aspect of the development of mutations is the further proliferation of the cells carrying genetic mutations.

Figure 1.3: The six hallmarks of cancer proposed by Hanahan & Weinberg (Hanahan, Weinberg 2000), which enable a cell to become cancerous.

This highlights that the ability of cells to acquire these hallmarks cannot be explained through a single cause based on a single genetic abnormality but represents the accumulation of multiple mutations within the genome. Such mutations can in some cases be hereditary, such as is the case for the *BRCA1/2* genes (Pecorino 2012), but most commonly they are induced through DNA damage acquired during the lifetime. Such mutations can be caused by various factors, including chemical, physical and infection-related inducers, which are discussed further in chapter 1.1.4.

1.1.3.1 Self-sufficiency in growth signals

The first hallmark describes the ability of cells to become independent from growth signals of surrounding cells. Healthy cells normally exist in a quiescent state and are strictly regulated through signalling and adhesion molecules expressed by surrounding and adjacent cells. They require activation to change to a proliferative stage, such as growth factors, extracellular matrix (ECM) components and cell-cell adhesion/interaction molecules, which can bind to the transmembrane receptors of the cell (Hanahan, Weinberg 2000). Tumour cells, however, can develop an independence from external growth stimuli and can change into a proliferative stage via three different mechanisms. Firstly, by an increased production of growth signals by the cancer cell itself, which enables the cell to trigger its own proliferation through a positive-feedback loop (Hanahan, Weinberg 2000). Secondly, through an overexpression or structural alteration of growth signalling receptors, which can lead to a hypersensitivity of the cell to external stimuli. Thirdly, cells can have intracellular signalling pathways, normally only active through the binding of growth signals that are continuously activated without the binding of external stimuli (Hanahan, Weinberg 2000, Bertram 2000).

1.1.3.2 Insensitivity to anti-growth signals

As mentioned previously, normal cells are kept in a quiescent state aside from when they are required to proliferate. This means, cells not only need external stimuli to be activated but also signals that keep them in the quiescent state. This state is maintained through the operation of multiple anti-proliferative signals, generated through the activity of tumour suppressor genes. Functionally, tumour suppressor genes can have a suppressing effect on the cell cycle and are able to arrest cells at a certain stage, preventing them from further multiplication. Furthermore, they can promote the apoptosis of cells that have irreparable damage to their DNA. Mutations within tumour suppressor genes result in an alteration or loss of their wild type function, and therefore their ability to suppress and inhibit tumour growth (Muller, Vousden 2013). In cases where this DNA damage is located in both alleles of the tumour suppressor gene, it results in their inactivation and ultimately prevents the cell from being able to regulate the cell cycle and to induce apoptosis.

A well-known tumour suppressor gene is *TP53*, which was shown to be one of the most frequently mutated genes in human cancer (Freed-Pastor, Prives 2012). The gene is commonly activated through cellular stress caused by conditions such as DNA damage

or hypoxic conditions, resulting in the induction of apoptosis or the inhibition of further cell cycle progression of the malfunctioning cell (Freed-Pastor, Prives 2012).

If cells become unresponsive to external signalling, mutated cells can further proliferate, which results in an increasingly unstable condition through the accumulation of additional mutations. Subsequent increases in mutations in tumour suppressor and proto-oncogenes enable cells to eventually develop a malignant cell phenotype (Vogelstein, Kinzler 2004).

1.1.3.3 Evasion and prevention of apoptosis

Healthy tissue growth is maintained through a balance of proliferating and dying cells (Cooper, J. P., Youle 2012). The death of cells presents a natural limitation to the proliferation of genetically damaged cells. Such cell death can be classified into two categories, necrosis and apoptosis (Rock, Kono 2008). Necrosis is the death of a cell through traumatic cell injury, which can be caused through a lack of nutrients or via the direct damage of cellular components.

Apoptosis is the active process used in tissue growth, tissue modelling and controlled cell death induced by cell stress (Pecorino 2012). This process can be triggered through events such as DNA damage, withdrawal of growth cytokines, viral infections and hypoxic conditions, resulting in the activation of intrinsic and extrinsic apoptosis pathways (Bertram 2000). The intrinsic signalling pathway is regulated through the function of so-called sensors and effectors. Sensors screen the extra- and intracellular environment for signs of abnormality, such as DNA damage, and trigger the effectors in case of need. This leads to a downregulation of anti-apoptotic genes and the increased expression of pro-apoptotic genes (Hanahan, Weinberg 2000). Extrinsic apoptosis, mediated through death receptors on the surface of the cell, is commonly (Green, Llambi 2015) induced through cells of the immune system and for the maintenance of homeostasis is induced by external stimuli, resulting in the cell receiving a death signal.

Regardless of the apoptotic pathway triggered, it always results in the release of cytochrome C from the mitochondria. Cytochrome C induces the activity of various caspases, which rapidly degrade cellular organelles and chromatin (Bertram 2000). Most, if not all, tumour cells acquire the ability to prevent this process from taking place by making the cells insensitive to apoptotic stimuli, the upregulation of anti-apoptotic

proteins such as Bcl-2 and the loss of pro-apoptotic proteins such as Bax and Bak (Fouad, Aanei 2017, Hanahan, Weinberg 2000, Green, Llambi 2015).

1.1.3.4 Limitless replicative potential

The proliferation of cells is not only regulated by the interplay of growth and anti-growth stimuli and apoptosis, but also by an intrinsic, cell-autonomous program that limits their ability to multiply to a finite number (Hanahan, Weinberg 2000). Each cell can only divide a certain number of times (approximately 60 to 70) and once this point is reached, the cell stops growing. Cells that have lost their ability to multiply are referred to as senescent cells. This is a natural phenomenon that is correlated to the length of telomeres present at the end of chromosomes. Telomeres have several thousand repeats of short, 6 base pair (bp), sequences of which 50 – 100 bp are lost with each replication and once a critical telomere length is reached, the cells lose their ability to divide further (Fouad, Aanei 2017). This loss occurs based on the inability of DNA polymerases to fully replicate the 3' ends of chromosomes (Hanahan, Weinberg 2000).

For this reason, aside from the independence of growth and anti-growth stimuli, a cancer cell must also be able to overcome the limitation of replication and to become immortal. This is commonly achieved through the maintenance of their telomeres, either through the upregulation of the enzyme telomerase, which can reconstitute the telomere length, or through the activation of a mechanism which prevents the shortening of the telomeres during replication. It has been shown that the upregulation of the enzyme telomerase is present in 85-90 % of human tumours (Fouad, Aanei 2017).

1.1.3.5 Sustained angiogenesis

Angiogenesis is the formation of new blood vessels through branching from pre-existing vessels and is a natural process in healthy tissue formation, which is tightly regulated by a balance of pro- and anti-angiogenic factors. New blood vessels are crucial for the supply of oxygen and nutrients to sustain a healthy cell function. It is also crucial for the survival of the cell and surrounding tissues, which are dependent on their formation and maintenance (Hanahan, Weinberg 2000). In order to expand beyond 1-2 mm in size, a tumour has to ensure a sufficient supply of oxygen and nutrients (Talmadge, Fidler 2010). For this reason, the tumour induces angiogenesis through the secretion of growth factors such as the vascular endothelial growth factor (*VEGF*) and acidic (*FGF1*) and basic

fibroblast growth factor (*FGF2*) and the inhibition of anti-angiogenic factors such as thrombospondin-1 (Hanahan, Weinberg 2000).

1.1.3.6 Emerging hallmark: Reprogramming of energy metabolism

Cancer cells not only have to enable the control of cell proliferation, but they also need to be able to supply the cells with sufficient amounts of energy to fuel the uncontrolled growth and division of cells (Hanahan, Weinberg 2011). Healthy cells normally use aerobic respiration to produce ATP. Here, glucose is broken down through glycolysis into pyruvate in the cytosol, which is then transferred and oxidised in the mitochondria. Under anaerobic conditions, glycolysis is favoured and less pyruvate is transported to the mitochondria meaning that energy is produced through lactic acid fermentation. Cancer cells alter their energy metabolism and seem to favour glycolysis under aerobic conditions. This altered behaviour of cancer cells was discovered by Otto Warburg and is described as the "Warburg effect" or "aerobic glycolysis" (Phan, Yeung et al. 2014). Contradictory to the fact that aerobic energy metabolism is 18-times more efficient compared to aerobic glycolysis, cancer cells are able to sustain and grow through aerobic glycolysis (Hanahan, Weinberg 2011). This is possible through an increased uptake of glucose into the cell, facilitated by the increased expression of glucose transporters on the cell surface, such as GLUT1 (Jones, Thompson 2009). Furthermore, aerobic glycolysis seems to be associated with the upregulation of oncogenes, such as *RAS* and *MYC*, and the suppression of tumour suppressors, such as *TP53* (Hanahan, Weinberg 2011). The use of this approach for energy production enables the use of glycolytic intermediates in various biosynthetic pathways for the generation of nucleosides and amino acids (Hanahan, Weinberg 2011). Moreover, it was shown that this kind of energy metabolism is more frequently found in rapidly dividing cells, such as embryonic tissue.

1.1.3.7 Emerging hallmark: Evasion of immune destruction

The immune system is a host defence mechanism consisting of various biological structures and molecules, which enable the destruction of pathogens and outcomes of infection. The human immune system performs constant surveillance and is responsible for the recognition and elimination of malignant cells. However, some cancerous cells have acquired the ability to avoid the recognition of the immune system thus preventing their eradication (Hanahan, Weinberg 2011).

It has been shown that tumours have a higher infiltration of regulatory T cells (Tregs) (Takeuchi, Nishikawa 2016), which results in immune suppression and a poorer prognosis in many cancers. Furthermore, it is indicated that tumour-derived Tregs have a higher suppressive ability compared to naturally occurring Tregs (Vinay, Ryan et al. 2015). In addition to this, it was also shown that tumour suppressive cytokines are expressed at a higher level in the tumour microenvironment. Such cytokines include TGF-β, Th2 cytokines (IL4/5/6/10/12), chemokines and VEGF (Burkholder, Huang et al. 2014). Tumours are also able to down modulate the machinery of antigen processing, mainly affecting the major histocompatibility complex (MHC) I pathway. This leads to a reduced recognition of the tumour through T lymphocytes and therefore and decreased survival rate (Vinay, Ryan et al. 2015).

1.1.3.8 Tissue invasion and metastasis

The result of successful establishment of a primary tumour mass commonly leads to the invasion of surrounding and adjacent tissue, and the spread of tumour cells to distant sites (Hanahan, Weinberg 2000). This invasion and spread are more commonly referred to as metastasis. Metastases are responsible for 90 % of cancer-related deaths (Sporn 1996). Their process involves multiple steps and is commonly described as the "invasion-metastasis cascade" (Valastyan, Weinberg 2011). It should be noted that the process of metastasis is rarely successful and only 1 in 10 000 cells survives the transport to distant sites (Pecorino 2012). During the process of metastasis, cells detach from the main tumour, enter the blood or lymphatic system (intravasation), circulate through the body, leave the circulatory system (extravasation) and initiate growth at a distant site. For this, the new site needs to offer sufficient nutrients and oxygen. One process highly implicated in the development of metastasis is the epithelial to mesenchymal transition (EMT), which will be discussed further in section 1.3.

1.1.3.9 Enabling characteristics: Genome instability and mutation and tumour-promoting inflammation

In addition to the hallmarks of cancer, which represent steps in the process of cells becoming cancerous, "enabling characteristics" can facilitate the acquisition of hallmarks. One of these characteristics is the development of genomic instability within cells. Through this process, random mutations are acquired, including chromosomal rearrangements. This instability supports the development of mutations that are needed to develop some of the prior mentioned hallmarks of cancer (Hanahan, Weinberg 2011).

Furthermore, within the last few years it has become more apparent that tumours and their microenvironment vary regarding their infiltration with immune cells. An increased immune response within the vicinity of the tumour was initially considered a response of the host to eradicate the tumour; however recent research highlighted the influence and promotion of tumour development through immune infiltration (Hanahan, Weinberg 2011). Such a tumour promoting effect can be caused through the secretion of growth factors by the immune cells and through the release of reactive oxygen species, which can actively function as mutagens for nearby surrounding cells (Hanahan, Weinberg 2011).

1.1.4 Causes of cancer - carcinogenesis

The hallmarks of cancer indicate the steps a tumour is undergoing to successfully invade and grow inside the host. To acquire these abilities, the genetic code of the DNA is altered. This occurs commonly through damage induced to the DNA, whereas single unrepaired mutations will increase the accumulation of further mutations. This enables mutated cells to behave in an autonomous manner and to develop abilities, such as stimuli-independent cell proliferation and tissue invasion, which was described previously.

Over the years, it was shown that cancers caused by hereditary mutations only represent a small proportion of cases. Around 90 to 95 % of all cancers are developed through exogenous factors as part of the environmental exposure, such as smoking, dietary habits and infections, whereas only 5 – 10% of all cancers are induced through hereditary genetic mutations, (Anand, Kunnumakara et al. 2008) such as in the *BRCA1* and *BRCA2* genes (King, Marks et al. 2003). This was further supported by a study, which has highlighted that the chance of being diagnosed with cancer is more strongly influenced by the country you live in rather than by the country you came from (Anand, Kunnumakara et al. 2008), indicating that the cancer risk is highly associated with living standard, life style and dietary behaviour.

One of the most commonly known inducers of cancer are the chemicals within tobacco smoke, that contains at least 40 different carcinogenic compounds, such as formaldehyde or benzene (Hecht 2006). Such carcinogenic compounds can lead to the development of mutations and altered cellular pathways, such as the NF-κB pathway. An active NF-κB pathway leads to the expression of cell proliferative genes and protects the cells from apoptosis (Xia, Shen et al. 2014). Alcohol is another environmental factor that increases

the risk for the development of cancer if consumed regularly. Furthermore, this damage is amplified if alcohol consumption and smoking are performed in parallel. Alcohol not only facilitates the entry of benzopyrene into the oesophagus (Anand, Kunnumakara et al. 2008), but alcohol consumption can also activate the NF-κB pathway (Wang, F., Yang et al. 2015).

Eating habits can also influence the risk and the development of cancer. A heavy consumption of red meat is associated with cancers of the gastrointestinal tract (Bouvard, Loomis et al. 2015). Furthermore, the consumption of charcoaled meat leads to the ingestion of carbon compounds such as pyrolysates and carcinogenic amino acids. Further food additives, such as nitrites and nitrates, also present carcinogens. In close connection to food habits stands the factor of obesity. Obesity is associated with increased risks for the development of various cancers (Calle, Kaaks 2004) such as cancer of the colon and breast. This is caused through the altered regulation and expression of hormones, including insulin and an increased activity of inflammatory pathways, such as JAK/STAT and NF-κB.

Viruses are a main contributor to infection-caused cancers and the most well-known virus is the Human Papilloma Virus (HPV). Various types of this virus were shown to be largely contributing to the development of cervical cancer (Burd 2003). The HPV types 16 and 18 alone cause 70 % of cervical cancers (WHO, 2019b). Based on its large contributions, a vaccine designed against these HPV types has been created to protect women from developing ovarian cancer that occurs due to HPV infection (Shukla, Shirish, Bharti et al. 2009, Wu, Guo et al. 2003).

Further factors that can increase cancer risk and introduced DNA damage are environmental pollutants, such as radiation or air pollutions. There are two main types of radiation that can cause cancer, namely ionising radiation and electromagnetic radiation. Ionising radiation can directly cause damage to the DNA and results in the development of reactive oxygen species (ROS). Such radiation can be the results of atomic fallout, as that caused during the atomic bombing of Japan (Douple, Mabuchi et al. 2011). Ultraviolet (UV) radiation belongs to the group of electromagnetic radiation and is highly associated with the development of skin cancer based on DNA damage and genetic mutations (Narayanan, Saladi et al. 2010). A well-known indoor-air pollutant with strong

implication in the development of lung cancer and mesothelioma is asbestos. Asbestos is a fibrous material, which was commonly used for insulation purposes. Based on its carcinogenic nature, it is now banned from use in many countries. Exposure to asbestos can result in chromosomal aberrations, ani- and polyploidy as well as epigenetic modifications (Pecorino 2012).

1.2 Prostate Cancer

1.2.1 Prostate anatomy

The prostate is a gland of the male reproductive system approximately the size of a large walnut and is surrounded by the prostatic capsule. It is located in the pelvis, surrounds the urethra, sits below the bladder and is anterior to the rectum. Due to its anatomical position, its texture can be examined through rectal examination. The prostate can be divided into three distinct zones called the central, transitional and peripheral zone, which differ both histologically and anatomically (Fig. 1.4) (Aaron, Franco et al. 2016, Lee, Akin-Olugbade et al. 2011). Some publications indicate a fourth zone, which is referred to as the fibromuscular zone (Dunn, Kazer 2011). Each zone differs in their embryological origin and can be distinguished by their histological appearance, biological function and development of pathologic disorders (Lee, Akin-Olugbade et al. 2011). This variant susceptibility to pathological diseases is illustrated by the fact that the majority of PCa (70 to 75 %) develop in the peripheral zone and are defined as adenocarcinomas (Lee, Akin-Olugbade et al. 2011, Kulasingam, V., Diamandis 2008a, Dunn, Kazer 2011). The transitional zone on the other hand is rarely a source of PCa, however it is the exclusive site for the development of benign prostatic hyperplasia (BPH) (Ward, Catto et al. 2001).

Figure 1.4: Schematic representation of prostate anatomy showing three distinct zones within the prostate (Valkenburg, Williams 2011). The prostate is divided into the peripheral, transition and central zone.

1.2.2 Incidence, risk factors and clinical presentation of prostate cancer

Prostate cancer is defined as the development of malignant cell growth in the prostate. It is the most common cancer in men in Europe and the third most common cancer overall. In 2014 around 47 000 men in the UK were diagnosed with PCa (Cancer Research UK, 2017a), whereas the highest incidence rate was shown in men between the ages of 65 and 69 (Fig. 1.5).

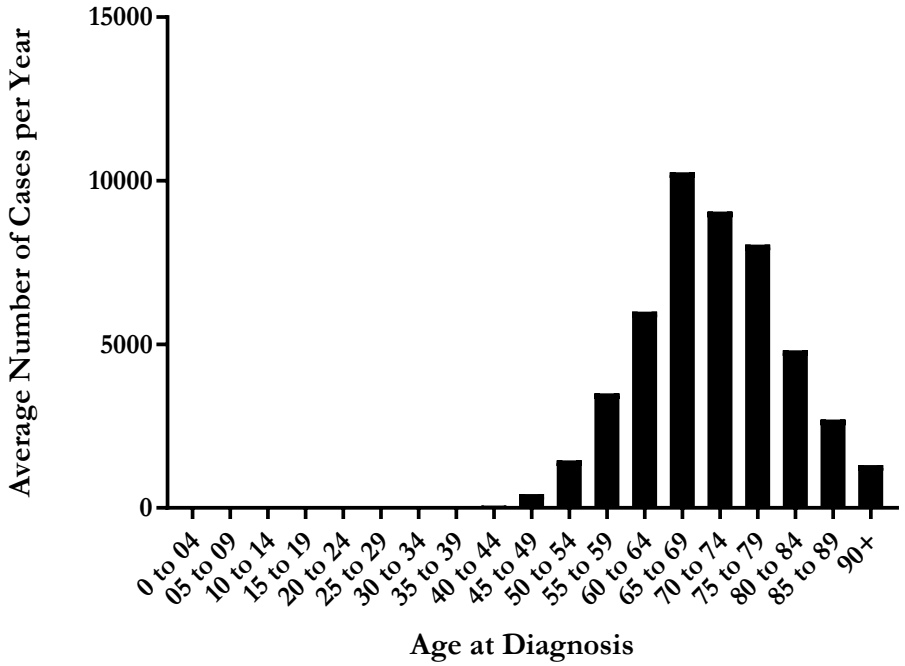


Figure 1.5: Age-specific cases of prostate cancer diagnoses in 2013-2015 in the UK. (Based on a graphic created by Cancer Research UK, Cancer Research UK, 2018c)

In Europe, PCa represents the third most common cause of cancer-related deaths, with a death rate of about 11 300 men in the UK in 2014 (Cancer Research UK, 2017b). The death rates due to prostate cancer highly correlated with age and increase sharply from around age 55 with the highest rate in the 90+ age group (Fig. 1.6) (Cancer Research UK, 2017b). Reduced chances of survival are correlated with the stage of prostate cancer at the time of diagnosis and increased mortality is mainly due to the development of metastasis. About 4 % of PCa patients will develop metastases, which reduces their 5-year survival rate to only 30 % (Thobe, Clark et al. 2011).

Figure 1.6: Age-specific cases of prostate cancer-associated deaths in 2012-2014 in the UK (Based on a graphic created by Cancer Research UK, Cancer Research UK, 2018d).

Recent statistics indicate that 1 in 8 men will develop PCa within their lifetime (Prostate Cancer UK, 2017a), however, not each case will be diagnosed and many will remain without symptoms. Through post-mortem analysis it was shown that many more men die *with* PCa than *from* it (Dunn, Kazer 2011).

There are three major risk factors for the development of PCa; age, ethnic origin and genetic predisposition. Furthermore, having a first-degree relative with PCa increases the risk of developing the disease by 2-fold (Dunn, Kazer 2011). Studies in the US have shown a lifetime risk for the development of PCa is 18.25 % in African-American and 15.25 % in Caucasians. However, the probability of dying through PCa during the lifetime is 2-fold higher in African-American men compared to men from Caucasian descent (Dunn, Kazer 2011).

In PCa, the late onset of symptoms, which are often nonspecific to cancer, increases the difficulty in diagnosis and therefore reduces the chance for the successful prevention of metastasis. Currently employed diagnostic tools are unable to predict whether a tumour will spread or has already spread. For this reason, the identification of biomarkers that

will help to elucidate the aggressiveness of a tumour and its future development will drastically increase the chances of survival for the patient.

1.2.3 Prostate cancer diagnosis and limitations of currently available diagnostic tools

One of the crucial aspects of PCa development is that it remains, in most cases, asymptomatic and often symptoms arise once the tumour is present at an advanced stage or has metastasised (Dunn, Kazer 2011). Despite this, there is currently no routine screening of all men performed. The lack of a robust screening program is based on the limited performance of currently available detection tools, which can lead to false-positive and false-negative results (Slatkoff, Gamboa et al. 2011, Lin, K., Croswell et al. 2011).

For this reason, patients that are initially categorised into a group of increased risk are subjected to further tests. Factors that result in the classification to a high risk group are aspects such as age, ethnicity, family history or obesity. The next step for these patients in commonly the measurement of prostate-specific antigen (PSA) in the serum (Section 1.2.3.1). Elevated levels might indicate the presence of PCa and warrant therefore the further examination using digital rectal examination (DRE) and transrectal ultrasound (TRUS) guided biopsy (Heidenreich, Bastian et al. 2014a) (Section 1.2.3.2).

1.2.3.1 Prostate specific antigen (PSA)

As previously mentioned, the only marker which is currently in routine use in the diagnostic and surveillance of PCa, is called prostate specific antigen (PSA). PSA is encoded by the *KLK3* genes and is also known as kallikrein-3 and gamma-seminoprotein. Other members of the kallikrein family, as well as PSA, are expressed in multiple tissues including many that are steroid hormone regulated (Balk, Ko et al. 2003), such as the prostate. The protein is secreted by the prostate gland and is a major component of seminal fluid. PSA levels are measured in the serum for the diagnosis and surveillance of PCa, however, its utility and contribution to the patient's health and survival are regularly under debate and constant efforts are made for the discovery of better biomarkers (Prensner, Rubin et al. 2012).

As the name already indicates, PSA is not a marker for PCa but is generally associated with the prostate. Increased PSA levels can be associated with PCa but can also be elevated in other prostatic diseases such as benign prostatic hyperplasia (BPH) and

prostatitis (Chiam, Ricciardelli et al. 2014). In addition to this, no defined PSA value is descriptive for a secure positive or negative cancer diagnosis and the measured concentration should be considered a continuous value. Despite this, serum PSA levels of 0 to 4.0 ng/ml are considered to be within the normal range, whereas a concentration above 4.0 ng/ml may warrant follow-up screening, but this would be dependent on further factors such as age, ethnicity and family history (Dunn, Kazer 2011). Normally, DRE and potentially a TRUS biopsy will be performed to secure the diagnosis and to adjust the treatment options available for each patient based on the tumour stage and grade.

A higher PSA concentration suggests an increased likelihood for the presence of PCa, however patients with low PSA levels have also subsequently demonstrated the presence of cancer after follow-up. Due to its variable nature, PSA testing can lead to false-positive diagnoses and over-treatment of healthy patients, as well as overlooked and missed cases of diseased patients. Furthermore, in recent years it was discovered that single nucleotide polymorphisms within *KLK3* influences the serum levels of PSA (Filella, Foj 2016) and it was suggested that a genetic analysis of *KLK3* was performed in addition to routinely performed PSA testing (Filella, Foj 2016).

The increased PSA levels are based on the increased disruption of epithelial cell attachments within the prostate basal membrane and luminal secretions of the tumour cells. Later stage PCa invades stromal layers and the blood circulation through the total loss of glandular organisation (Kulasingam, V., Diamandis 2008a, Romero Otero, Garcia Gomez et al. 2014). This enables cancerous cells to spread throughout the body and can lead to increased chances for the development of metastasis.

PSA exists in various forms throughout its cellular processing from its mRNA to the final protein (Fig. 1.7). Some forms have shown, either alone or in combination, increased specificity and sensitivity for PCa compared to the common measurement of total PSA. ProPSA represents an inactive pro-enzyme which has a 7 amino acid long leader peptide. This leader peptide is cleaved by HK2 or HK4 to produce active PSA. ProPSA is also present in three truncated isoforms, of which [-2]proPSA presents the most stable form. This isoform was shown to be highly associated with PCa compared to BPH and could

be used for early detection as well as the definition of aggressiveness of a tumour (Saini 2016a).

Another use of PSA and its various forms was suggested by using the prostate health index (PHI). This consists of three PSA biomarkers, which are used in the following formula: [-2]proPSA/free PSA)*√PSA (Catalona, W. J., Partin et al. 2011). The PHI is considered to be of use for patients with a PSA range of 4-10 ng/ml and a normal DRE result. It contributes to the reduction of unnecessary biopsies (Saini 2016a).

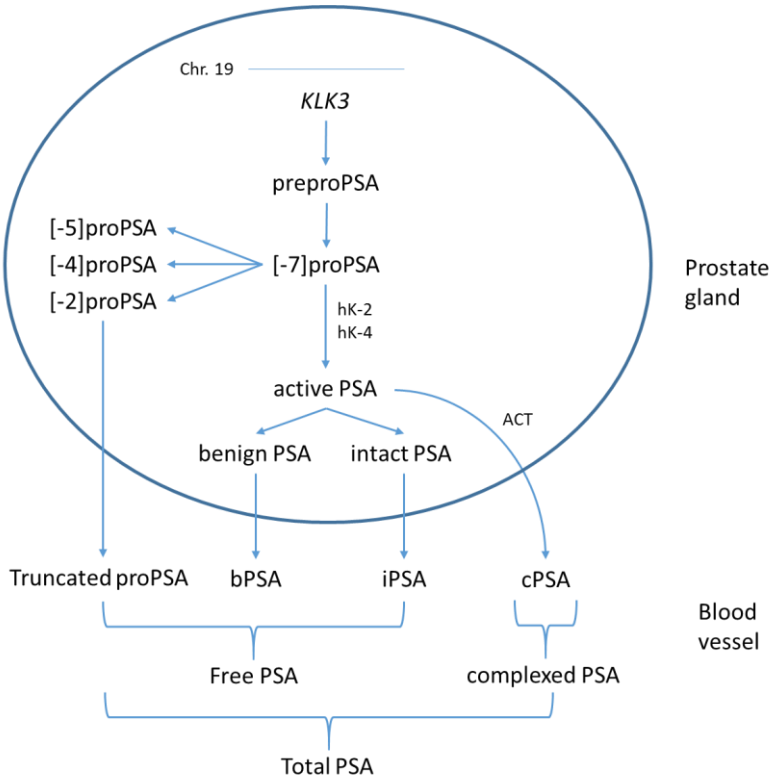


Figure 1.7: Schematic representation of PSA and its derivatives from its mRNA to the final protein, functioning as potential biomarker for PCa detection and monitoring (Saini 2016b). Chr. = Chromosome, *KLK3* = Kallikrein-3, PSA = prostate specific antigen, cPSA = complexed PSA, bPSA = benign PSA, iPSA = intact PSA, ACT = alpha(1)-antichymotrypsin, hk-2 = human kallikrein-2, hk-4 = human kallikrein-4, PSA isoforms = [-5]proPSA, [-4]proPSA, [-2]proPSA. Graph adapted from (Hatakeyama, Yoneyama et al. 2017).

1.2.3.2 Other currently studied potential biomarkers for the detection of PCa

PSA and its various forms are not the only potential biomarkers associated with PCa. Continuous research is focussed on the discovery of novel, more improved biomarkers for the screening and prognosis of PCA.

A well-studied marker for the diagnosis of PCa and differentiation from benign cases is the α-methylacyl coenzyme A racemase (AMACR). This enzyme is highly overexpressed in PCa compared to benign cases. In a study comparing PCa with benign prostate tissue, AMACR was detected in more than 90 % of the PCa cases and in less than 20 % of the benign prostate cases (Jiang, Zhu et al. 2013).

PCA3 is a novel potential biomarker for PCa. It was initially described as DD3 (Bussemakers, van Bokhoven et al. 1999), where it was shown to have an increased expression in prostate cancer tissue compared to non-neoplastic prostate tissue. Furthermore, it could not be quantified in healthy tissue, including the prostate. This marker can be detected in urine, which presents a minimal invasive screening method.

PCA3/DD3 originates from a non-coding region and its function is still mainly unknown (Wang, Y., Liu et al. 2014). However, work on PCA3 in prostate cancer has shown that the inhibition of *PCA3* through transfection leads to a significant decrease in cell viability and growth in LNCaP cells (Ferreira, Palumbo et al. 2012). It also shows a high specificity, ranging between 80 and 90 %, in the diagnosis of PCA, which could potentially lead to a reduction in unnecessary biopsies and over-treatment. (Roobol, Haese et al. 2011). However, dependent on the applied PCA3 score cut-off, lower sensitivities (58 %) and specificities (72 %) were detected (Marks, Fradet et al. 2007). This indicates that the clinical utility of PC3 varies based on the study and the applied cut-off of detection.

Another potential biomarker that is detectable in the urine is the fusion gene TMPRSS2:ERG. It is generally caused by an interstitial deletion at the locus 21q22 and a reciprocal translocation (Gleason 1966, Romero Otero, Garcia Gomez et al. 2014) and is present in about 50 % of PCa cases (Hagglof, Hammarsten et al. 2014). The Transmembrane Protease, Serine 2 (*TMPSS2*) gene transcriptional promotor is strongly regulated through the stimulation with androgens (Lin, B., Ferguson et al. 1999, Roobol, Haese et al. 2011) and the ETS transcription factor (*ERG*), which is an oncogene. Through its function as a transcription factor for genes in the ETS family, it is directly involved in cell proliferation, angiogenesis and differentiation. TMPRSS2:ERG can be detected in the urine after prostate massage and shows a high specificity and sensitivity; however, it shows a low frequency in some populations (Romero Otero, Garcia Gomez et al. 2014).

Reliable screening methods for PCa are urgently needed, but currently available tools lack specificity and sensitivity. A very important aspect in the management of prostate cancer is the ability to distinguish indolent PCa from aggressive disease. Unfortunately, PSA is unable to perform this task. The other presented markers, AMACR, PC3 and TMPRSS2:ERG present potentially useful biomarkers for the detection and disease surveillance of PCa, however, none of these markers have been routinely implemented in clinical settings and further clinical validation is needed.

1.2.3.2 Clinical approaches for the detection and diagnosis of PCa

As mentioned previously, patients that were categorised into a high risk group are subjected to PSA testing. If the test results show abnormal values or if symptoms could be based on the development of PCa, further examinations are performed to confirm the presence or absence of PCa. These further examinations include digital rectal examination (DRE) and biopsy approaches.

1.2.3.2.1 DRE – Digital Rectal Examination

Digital rectal examination (DRE) is a routine test for the screening of PCa, commonly performed after an abnormal PSA test showing elevated blood levels. Here, a rectal examination of the prostate is performed for the feeling of irregularities in size, shape and texture and for the presence of lumps. DRE is only of limited use for small tumours and lacks sensitivity for their detection (Woolf, MD, MPH, Steven H, Rothemich 1999). Furthermore, irregularities, such as an enlarged prostate, are only an indication for the presence of prostate cancer and the results of a DRE should be followed up with further diagnostic testing.

1.2.3.2.2 Biopsy approaches

Biopsies are not routinely applied screening methods, but normally follow elevated PSA concentrations and/or abnormal DRE results. For further clarification as to the potential presence of PCa a biopsy is performed. The guidance of the sample taking can be performed either through ultrasound, such as in the transrectal ultrasound (TRUS) guided imaging technique, and more recently through magnetic resonance imaging (MRI).

TRUS guided biopsies present a sampling error >20 %, which is based on the inability of an ultrasound to image a clear difference between prostate and cancerous tissue (Peltier, Aoun et al. 2015). MRI-guided biopsies enable the identification of lesions within the

prostate, which enables a more targeted sampling. For this reason, MRI-guided biopsies outperform TRUS-guided.

1.2.4 Staging and grading

PCa is the most commonly diagnosed cancer in men and a crucial factor for patient survival is the accurate grading of the tumour. It is vital that this occurs in a uniform manner to ensure an accurate prediction of the tumour behaviour and an optimal selection of treatment (Cheng, Liang, Montironi et al. 2012). Currently, there are two major staging systems applied in clinical practice, the pathological based Gleason scoring and the Tumour-Node-Metastasis (TNM) staging based on clinical parameters. Gleason scoring is based on the histological appearance of tumour material within the prostate, whereas the TNM staging is considered to be clinical staging based on multiple clinically assessed parameters (Dunn, Kazer 2011).

1.3.4.1 Gleason scoring

Gleason scoring was developed by Dr Donald F Gleason (Gleason 1966). A reason for its success and wide application was its successful validation in about 5000 patients. The system is based on the pathological inspection of prostatic tissue sections and the categorisation of carcinoma cells into histological patterns (Humphrey 2004). The system uses five grades, which are used to calculate a score based on the sum of the first and second most prominent patterns. The differentiation of cells decreases from stage one to stage five (Fig. 1.8). Grade one presents well differentiated growth of closely packed, round and uniform acini (Humphrey 2004), whereas grade 5 tissue does not present any glandular differentiation and has lost resemblance to healthy prostate tissue. Gleason scoring is an important tool for the prediction of outcome of PCa patients.

Currently there are some debates about the differences of Gleason 7, which can be generated through 4+3 and 3+4 presented morphologies within the tumour specimen. Studies have shown that disease outcome varies depending on the prominence of stage 4 tissue (Stark, Perner et al. 2009). A study comparing disease outcome in patients with both categories of Gleason 7 has shown a three-fold higher likelihood of lethal PCa in patients with 4+3 compared to 3+4 (Stark, Perner et al. 2009). A further study has also shown a higher risk of cancer related mortality in cases of 4+3 compared to 3+4 (Wright, Salinas et al. 2009).

Another debate is regarding the Gleason score 6 and whether or not it should be categorised as cancer. A change in its category could have serious implications on the patient's treatment and disease outcome. Autopsies on men over 50 have frequently identified the presence of Gleason 6 PCa. On a histological basis, Gleason 6 PCa is fulfilling the histopathological requirements to be defined as PCa; however, some scientists argue that Gleason 6 cancer does not fulfil all 6 hallmarks of cancer (section 1.1.3), and should therefore be treated differently (Carter, Partin et al. 2012, Eggener, Badani et al. 2015). A study on more than 14 000 patients with a Gleason score of six and below, which underwent radical prostatectomy, has shown in only 22 cases an involvement of the lymph nodes (Ross, Kryvenko et al. 2012). This supports the idea that active surveillance offers a favourable option for patients with Gleason 6.

Figure 1.8: Schematic representation of tissue differentiation in PCa across the 5 Gleason scores (PCEC, 2019). Increased Gleason scoring correlated with decreased tissue differentiation.

1.3.4.2 Tumour – Nodes – Metastasis (TNM) Staging

TNM staging (Tab. 1.1) is a pathological staging method for the characterisation of solid tumours and is used to describe the tumour and its current state in more detail. T stands for the primary tumour and its potential invasion into surrounding tissue. N describes lymph node involvement and M stands for metastasis and gives information regarding tumour spread to distant sites (Sobin, Gospodarowicz 2009). The involvement of lymph

nodes is one of the most important prognostic factors and the aggressiveness of PCa is closely linked to the tumour volume (Cheng, Liang et al. 2012).

Table 1.1: Summary of the TNM classification system adapted from (Sobin, Gospodarowicz 2009)

T – primary tumour

TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1 (T1a, T1b, T1c)	Tumour does not cause any sign of symptoms. Tumour cannot be detected though palpation or digital imaging. However, histopathological analysis can detect the presence of malignant cells.
T2 (T2a, T2b, T2c)	Tumour can be detected with DRE but is still confined within the prostate. Subcategories describe size and penetration of the tumour within the prostate capsule.
T3 (T3a and T3b)	Tumour extends outside the prostate capsule. Subcategories describe the extension into further detail.
T4	Tumour has invaded tissues outside the prostate and seminal vesicles and has spread to nearby organs, such as the bladder or lymph nodes.

N – Regional Lymph Nodes

NX	Regional lymph nodes cannot be assessed.
N0	No metastasis found in regional lymph nodes.
N1	Metastasis found in regional lymph nodes.

M – Distant metastasis

M0	No distant metastasis present
M1 (M1a, M1b, M1c)	Distant metastasis. Subtypes describe the location of the metastases.

Through the combination of the defined TNM stages, Gleason score and PSA levels, the tumour can be attributed to a certain stage and the stages range from I to IV (Cheng, Liang et al. 2012).

1.3.4.3 D'Amico risk classification system

A further risk classification system is the D'Amico scoring. This system is designed to categorise patients into 3 distinct groups of risk for recurrence of PCa after radical prostatectomy through the combination of TNM stage, biopsy Gleason score, and the PSA level prior to surgery (Tab. 1.2) (D'amico, Whittington et al. 1998) (Tab. 1.2).

Table 1.2: D'Amico risk classification system to categorise patients for risk of prostate cancer recurrence after radical prostatectomy . Risk classification is based on Gleason score, TNM stage and PSA levels prior surgery. Adapted from: (D'amico, Whittington et al. 1998)

D'Amico Risk Group	Gleason Score	TNM Stage	Pre-operative serum PSA (ng/ml)
Low Risk	≤6	T1 or T2a	<10
Intermediate Risk	≤6-7	T1 or T2a/b	10-20
High Risk	≤7	T1 or T2a/b/c	>20
	8-10	T1 or T2a/b/c	Any PSA

1.3.5 Treatment options available for PCa

The treatment options vary based on the type of PCa present (Tab. 1.3). Low-risk PCa are patients with clinically confined PCa (T1-T2) and a Gleason score below six, additionally the PSA should be below 10 ng/ml (Heidenreich, Bastian et al. 2014a). These patients are normally subjected to active surveillance, which means that the patients are initially not treated and the development of the disease is checked in regular intervals to ensure the tumour does not progress. In case of disease progression, the treatment is performed with curable intent, through active treatment. Active treatment options include radical prostatectomy, which is the only surgical treatment for localised PCa. Furthermore, localised PCa can be treated using radiation therapy and low-dose-rate brachytherapy.

Table 1.3: Summary of treatment options for PCa according to the stage of PCa. Adapted from (Heidenreich, Bastian et al. 2014a, Heidenreich, Bastian et al. 2014b)

Stage of PCa	**Treatment options**
Localised PCa	Active surveillance
Localised/locally advanced PCa	Radical prostatectomy External Beam Radiation Therapy Permanent seed brachytherapy Cryotherapy High-intensity focused ultrasound
Advanced PCa	Hormone Therapy Chemotherapy

Advanced PCa patients presenting metastatic disease or castrate resistant PCa (CRPC) are commonly subjected to a different treatment regime compared to localised, low risk PCa patients (Heidenreich, Bastian et al. 2014a, Heidenreich, Bastian et al. 2014b). The intention of treatment is commonly focussed on the management of the disease and the improvement of life quality rather than curing the disease. Such treatment includes hormone and chemotherapy. Hormone therapy aims to block the access of PCa cells to dihydrotestosterone. Testosterone is commonly required by PCa cells for their growth and proliferation, however, commonly after a mean time of 2 to 3 years, PCa cells develop an independence from testosterone (Karantanos, Corn et al. 2013). The cancer is then described as CRPC. The treatment choice for these patients is limited and they commonly receive chemotherapy, however a substantial proportion of men do not benefit from this treatment and only small improvements in the overall median survival of CRPC patients can be achieved (Teply, Hauke 2016).

1.3 Metastasis and epithelial to mesenchymal transition (EMT)

The discovery of a molecular pathway involved in the development of metastasis, called epithelial to mesenchymal transition, presented an important step for the understanding of metastasis and for the future discovery of novel prognostic biomarkers (Das, R., Gregory et al. 2014).

1.3.2 Types of epithelial to mesenchymal transition (EMT)

EMT is an evolutionary highly conserved developmental process (Lim, J., Thiery 2012). The first pioneering work on EMT was performed by Elizabeth Hay, who observed the process in chick embryos and described it originally as epithelial-mesenchymal transformation (Greenburg, Hay 1982). Here, polarised epithelial cells, which are attached to a basement membrane and the neighbouring cells, undergo biochemical changes to acquire mesenchymal cell properties. These changes result in an altered gene expression, which leads to increased motility through the degradation of intracellular contacts, increased invasiveness, migratory potential and resistance to apoptotic signals (Kalluri, Weinberg 2009). Morphologically this is visible through the change from a "cobblestone" morphology with highly organised cells to solitary, spindle-shaped cells with a fibroblastic morphology (Thiery, Sleeman 2006). On a molecular level these changes are shown through a reduction of epithelial gene expression and an increase in mesenchymal associated genes. These changes are based on multiple molecular alterations such as the activation of EMT specific transcription factors, an altered expression of cell-surface proteins, additional changes in the expression of cytoskeletal proteins, microRNAs and the production of ECM-degrading proteins. Major genes associated with epithelial cells include E-cadherin (*CDH1*) (Pećina-Šlaus 2003), and genes associated with mesenchymal cells include N-cadherin (*CDH2*) (Zeisberg, Neilson 2009), Vimentin (*VIM*), Fibronectin (*FN1*) (Sudo, Iwaya et al. 2013) and metalloproteases (MMPs) (Lozito, Tuan 2011). The key players involved with the process of EMT are highly conserved and can be found across multiple species (Heerboth, Housman et al. 2015). Furthermore, this process is not unidirectional and it should be highlighted that the cells can reverse the process back into an epithelial morphology. This is called mesenchymal to epithelial transition (MET) (Lim, J., Thiery 2012).

During two consecutive meetings, which took place in 2007 in Poland and in 2008 at Cold Spring Harbor Laboratories, three distinct EMT subtypes were defined (Kalluri, Weinberg 2009). In general, EMT is considered to be a normal and healthy biological process, which plays an important role during embryogenesis (Micalizzi, Farabaugh et al. 2010) and wound healing (Zeisberg, Neilson 2009), but is re-activated during cancer progression (Nieto 2013).

Type I EMT is involved during multiple stages of embryonal development, including the implantation of the embryo and the formation of the placenta. After fertilisation, the embryo undergoes gastrulation, in which the three germ layers are formed. During this process, EMT is occurring when cells migrate into predefined regions of the embryo (Lim, J., Thiery 2012). These migratory cells are then involved in heart morphogenesis, where they undergo multiple cycles of EMT and MET. This process is also initiated during gastrulation to develop multiple compartments of the heart.

Type II EMT is involved in wound healing and organ fibrosis and is mediated and directed by inflammatory cells and fibroblasts (Kalluri, Weinberg 2009). Type II EMT differentiates from Type I EMT through the generation of fibroblasts instead of mesenchymal cells. Here, the process is commonly induced through inflammatory signals or tissue damage and is normally limited with the end occurring at cessation of the healing process (Foroni, Broggini et al. 2012). The third type of EMT is involved in the spread of cancerous cells and the subsequent development of metastasis at distant sites and will be discussed in further detail in the following section (1.3.2.1).

1.3.2.1 Type III EMT: Cancer progression and metastasis – the Invasion-Metastasis Cascade

As mentioned previously the process of EMT is part of a healthy functioning biological system. However, during the development of cancer this process, also described as the "invasion-metastasis cascade", is utilised in the development of metastasis and hence cancer progression (Fig. 1.9) (Kalluri, Weinberg 2009). Here the cells lose their apical-basal cell polarity (Acloque, Adams et al. 2009), invade the surrounding extracellular matrix (ECM) and stromal cell layers, intravasate into blood or lymphatic vessels, through which they are dispersed throughout the body. To sustain this process, the cells must show an increased survival capability, for example through the development of resistance

to anoikis. Anoikis is a form of programmed cell death, which takes place when an anchorage-dependent cell detaches from the associated ECM (Frisch, Screaton 2001). Following resistance to anoikis-induced cell death and circulation throughout the body, the cells extravasate from the circulatory system and into the tissue, where they are able to adapt to the new host microenvironment, colonise and develop secondary tumours (Valastyan, Weinberg 2011). Despite the access of these cells to the lymphatic and blood circulatory systems and their potential to initiate tumour growth in various organs, certain cancer types show preferred sites of secondary tumours, for example PCa metastases are commonly found in the bones, lungs and liver (Bubendorf, L., Schopfer et al. 2000).

Figure 1.9: The invasion-metastasis cascade describing the process of tumour spread from the primary tumour to a distant site (Valastyan, Weinberg 2011)

EMT is a highly conserved process, which presents highly similar molecular processes across all three types of EMT (Foroni, Broggini et al. 2012, Heerboth, Housman et al. 2015). The process can be triggered through intra- and extracellular stimuli, including growth factors such as TGF-β, HGF, EGF and IL-6, and as a response to hypoxia, tumour-stromal cell interactions and chemotherapy (Foroni, Broggini et al. 2012, Kalluri, Weinberg 2009).

The main indicator for the induction of EMT is the loss or reduction of E-cadherin expression, which also represents a hallmark of EMT (Serrano-Gomez, Maziveyi et al. 2016). E-cadherin is a calcium-dependent cell surface protein, which is involved in the cell adhesion of epithelial cells (Liu, F., Gu et al. 2016) and it limits the expression of

mesenchymal genes (Garcia de Herreros 2014). Its reduced expression is caused through the induction of a set of highly conserved transcription factors, so called EMT-TFs, including *SNAI1, SNAI2, TWIST*, *ZEB1* and *ZEB2*. They suppress, upon activation, the expression of epithelial markers and induce genes associated with mesenchymal cells. *SNAI1, ZEB1* and *TWIST* are all individually able to induce EMT in most cell lines. *ZEB1* can directly repress the expression of E-cadherin, whereas *TWIST* is involved downstream in the induction of mesenchymal genes (Lamouille, Xu et al. 2014). In contrast to these, *SNAI1* is involved in both tasks (Garcia de Herreros 2014). Aside from this, the transcription factors show a timely difference in their response. *SNAI1* is the first factor to be induced. In cell culture systems its expression was detectable as early as 30 min after TGF-β treatment, followed by *ZEB1* and other mesenchymal markers after more than 4 h (Garcia de Herreros 2014). The interplay of the EMT-TFs leads to a widespread alteration in gene expression (Tab. 1.4), which is a major field of research. However, the commonly studied genes for the validation of EMT induction are coding for the intermediate filament protein vimentin (*VIM*), the trans-membrane protein N-cadherin *(CDH2)*, the glycoprotein fibronectin (*FN1)* and the cell-adhesion protein E-cadherin (*CDH1).*

Table 1.4: List of genes and proteins associated with the process of EMT, their expression changes and impact on cellular functions. (adapted from (Foroni, Broggini et al. 2012))

Upregulated genes	Downregulated genes	Activation of	Changes in cellular functions
Vimentin (*VIM*)	E-cadherin (*CDH1*)	β-catenin	Increased invasion
N-cadherin (*CDH2*)	Desmoplakin	SMAD2/3	Increased migration
Fibronectin (*FN1*)	Cytokeratin	NF-κβ	Chemotherapy resistance
SNAIL (*SNAI1*)	Occludin	SNAIL	Increased resistance to apoptosis
SLUG (*SNAI2*)	Claudin	SLUG	
TWIST	miRNA200 family	TWIST	
ZEB1/2		ZEB1/2	
Goosecoid (*GSC*)			
FOXC2			
MMP2/3/9			

1.3.3 Transforming growth factor β and the TGF-β superfamily

Transforming growth factor β is a ligand that performs two opposite tasks within cancer development and progression. In early stage tumours, TGF-β functions as a tumour suppressor ligand through the promotion of cell cycle arrest and the induction of apoptosis, however in late stage cancer, TGF-β increases cell motility, invasion and metastasis as well as cell "stemness" (Neuzillet, Tijeras-Raballand et al. 2015). This change is described as the "TGF-β paradox".

The TGF-β superfamily consists of multiple secreted homodimeric signalling proteins (Hinck 2012), including the three isoforms; transforming growth factor β1 (TGFB1), β2 (TGFB2) and β3 (TGFB3). TGF-β1 is the most studied isoform. These three isoforms share about 70 % homology within their sequence (Lebrun 2012), are synthesised in the cell and are secreted in a latent dimeric form into the extracellular matrix. Here, the latent TGF-β is activated through the cleavage by furins and other convertases (Padua, Massagué 2009). The active TGF-β isoform can then induce the TGF-β signalling cascade by binding to surface receptors. The receptors can be grouped in 3 categories, consisting of 7 type I, 5 type II and 1 type III receptor. These receptors are paired in different combinations to form receptor complexes for the members of the TGF-β superfamily (Padua, Massagué 2009). Other members include bone morphogenic proteins (BMPs), which are involved in the embryonic development, and growth and differentiation factors (GDFs) (Hinck 2012).

1.3.4 Transforming growth factor β as an inducer of EMT

As mentioned previously, EMT can be induced through various growth factors. The first described and most commonly used cytokine in the study of EMT is TGF-β (Fig. 1.10) (Serrano, McDonald et al. 2013, Iordanskaia, Nawshad 2011, Raghavan, Smuda et al. 2016).

This image has been removed by the author for copyright reasons

Figure 1.10: TGF-β signalling pathways and resulting genes responses upon activation (Padua, Massagué 2009)

TGF-β present in the microenvironment of a cell can bind to the TGF-β receptor II (TGFBR2), which leads to the formation of a heterotetrameric active receptor complex and results in the unidirectional phosphorylation of a 30-amino-acid regulatory segment called the GS region of the TGF-β receptor I (TGFBR1) (Padua, Massagué 2009). This phosphorylation activates the complex and through this, signalling through SMAD-dependent and SMAD-independent cascades (Pickup, Novitskiy et al. 2013) (Fig.1.11) can occur. TGF-β receptor III (TGFBR3) is not directly involved in the signalling cascade induced through TGF-β, however it functions as a co-receptor by binding TGF-β and presenting it to TGFBR2.

1.3.4.1 Canonical (SMAD-dependent) signalling cascade

One of the induced pathways is the canonical (SMAD-dependent) pathway. The activation of the TGF-β receptor complex leads to the release of the FK506 Binding Protein 1A (FKBP12) (signalling inhibitor) from the active site of TGFβR1. Through the release of FKBP12, the SMAD complex consisting of SMAD2 and SMAD3, also called R-SMAD, can bind to the active site through the support of the SMAD anchoring protein (SARA). This binding leads to the phosphorylation and release of SMAD2/3 from SARA (Attisano, Wrana 2002). The activated SMAD2/3 binds to SMAD4 to form a

heterodimeric complex before translocating into the cell nucleus. Here the complex binds to one of the many DNA binding partners or transcriptional co-activator/repressors, which initiates transcriptional activation or repression of several hundred genes, inducing the previously mentioned EMT-TFs *SNAI*1, *SNAI2* and *ZEB1* (Pickup, Novitskiy et al. 2013, Ikushima, Miyazono 2010).

1.3.4.2 Non-canonical (SMAD-independent) signalling cascade

TGF-β is also known to induce SMAD-independent pathways such as JNK/p38 Pi3K-Akt and Rho-like GTPases (Fig. 1.10) (Ikushima, Miyazono 2010). The JNK/p38 pathway plays an important role in TGF-β induced EMT. Studies have shown that the inhibition of p38 leads to an impairment of changes in the cell shape and the reorganisation of cytoskeletal structures. The binding of TGF-β to TGFBR2 leads to the phosphorylation of PAR6 thereby promoting the degradation of cell junction complexes (Padua, Massagué 2009).

Figure 1.11: SMAD-dependent and SMAD-independent signalling cascade induced through the binding of TGF-β (Adapted from (Pickup, Novitskiy et al. 2013)).

1.3.5 EMT associated biomarkers in cancer

It has been shown that an increased expression of TGF-β1 in PCa cells correlates with disease progression and metastasis (Wikström, Stattin et al. 1998). The investigation of this process and associated pathways might indicate the usefulness of associated genes as markers for metastasis and their clinical utility, however difficulties exists in demonstrating EMT *in vivo*. This is due to two main factors; firstly, it is challenging to distinguish mesenchymal cells of different origin. Currently, it is not possible to define whether a cell is mesenchymal through EMT and exhibits an increased metastatic potential, or whether they are normal stromal cells that exhibit mesenchymal cell properties.

Mesenchymal cells can be identified within the tumour composition; however, it is not possible to define whether these are a natural component of the tumour or cells that underwent the process of EMT (Gao, D., Vahdat et al. 2012). Furthermore, the process of EMT is transient, the cells can change from epithelial to mesenchymal, and then back to an epithelial cell state once it has invaded a secondary site. This leads to the loss of mesenchymal cell properties and the increased expression of epithelial cell markers (Gao, D., Vahdat et al. 2012). A recent study in which 205 tissue specimens of localised PCa were analysed using the common markers for EMT showed no significant association of the expression of any of these markers with clinical outcome. Furthermore, it highlighted the common expression of multiple mesenchymal markers in low-grade tumours (Armstrong, Healy et al. 2016). However, a different study has shown an increase in EMT-TFs correlated with the intensity of tumour progression (Imani, Hosseinifard et al. 2016). In addition, a drastic increase of *ZEB1* was detected in CRPC and was more frequently present in PCa metastasis compared to non-metastatic PCa. It was previously shown that androgen-deprivation could induce EMT in LNCaP cells (cells derived from an androgen-sensitive human prostate adenocarcinoma), which was presented through a reduced adherence as well as increased expression of the mesenchymal markers CDH2 on protein level, as well as *ZEB1* of mRNA level. This process could be reversed through a testosterone rescue, resulting in an increased cell attachment and a reduced expression of CDH2 and *ZEB1* compared to the deprived cell lines (Sun, Y., Wang et al. 2012).

1.4 Biomarker discovery and validation

1.4.1 Biomarkers

A biomarker is a molecule, which can be assessed for the definition of a biological status. This molecule can be descriptive for a healthy or pathogenic process, the response to a treatment or prognostic for the future development of disease (progression) (Ilyin, Belkowski et al. 2004). The National Cancer Institute (NCI) defines a tumour marker as "*A substance found in tissue, blood, or other body fluids that may be a sign of cancer or certain benign (noncancerous) conditions*" (National Cancer Institute, 2018a). Many studies focus on the discovery of novel tumour markers to improve the early detection of cancer, thereby preventing further cancer growth and mortality. Biomarkers can be categorised into six main groups based on their field of application (Tab. 1.5).

Table 1.5: Classes of Biomarkers and their use. Adapted from (Shariat, Scherr et al. 2011) and (Bensalah, Montorsi et al. 2007).

Type of Biomarker	Application of Biomarker
Early Detection	Systematic screening of a population for the identification of subjects with an increased risk for the presence of a disease.
Diagnostic	Used for the assessment and definition of absence or presence of cancer.
Prognostic	Prediction of disease outcome and categorisation of patients into risk groups. Provision on information regarding the clinical development of the disease.
Predictive	Information on the response and effectiveness of a certain treatment.
Therapeutic Target	Identification of patients, which will benefit from a treatment targeting a particular variation of the disease.
Surrogate endpoint	Can be used instead of a clinical endpoint for the measurement of relapse or recurrence and mortality.

A biomarker can be an endogenous marker, which presents an altered expression in malignant tissue, or can be a novel expressed gene induced through the presence of the tumour (Malati 2007, Kulasingam, V., Diamandis 2008a). An ideal biomarker should be detectable through non-invasive sampling, for example in a blood or tissue sample, and show a high specificity and sensitivity for the studied disease. To be successful, the biomarker needs to prove its clinical utility and have a positive impact on patient outcomes (Kulasingam, V., Diamandis 2008a). Overall, it should fulfil the following

criteria "better", "easier" and "cheaper", when compared to currently implemented markers (Bensalah, Montorsi et al. 2007, Kulasingam, V., Diamandis 2008a). Currently proposed biomarkers are commonly lacking the necessary specificity and sensitivity and are therefore rarely validated in clinical studies. Figure 1.12 highlights the available possible sources and available wet-lab methods for the discovery and validation of novel biomarkers (Tab. 1.6).

Table 1.6: Potential biomarker discovery pipeline using multi-omics discovery tools. Adapted from Broad Institute, 2018a

	Phase		
	Discovery	**Verification**	**Validation**
Analysed material	• Cell line model • Mouse model • Clinical specimens	Clinical specimens • Tissue microarrays • Body fluids (e.g. Blood, Urine, Saliva) • Tumour material	Clinical specimens • Body fluids • Tumour material
Analysis method	• Genomic sequencing • RNA sequencing	• Targeted gene sequencing • Quantitative real-time PCR	• Targeted gene sequencing • Quantitative real-time PCR
	• Mass spectrometry analysis	• Immuno-histochemistry • ELISA	• Immuno-histochemistry • ELISA
Number of candidates	>10 000	<100	<10
Number of samples	<100	>100	>1000

1.4.3 Sources for biomarker discovery with a focus on prostate cancer

A crucial aspect for successful biomarker discovery studies are numbers. It is important to have a sufficient sample number, increasing with the complexity and variability of the analysed material (Tab. 1.7). For this reason, a larger sample size is important when using clinical material, compared to more "uniform" cell line material. The choice of sample material depends on the planned experimental approach and the desired outcome of the study. Each material can offer advantages and disadvantages. Cell lines offer a highly controlled and more reproducible environment; additionally, large quantities of material

to analyse and high numbers of replicates are easily generated. On the other hand cell lines represent a highly artificial phenotype and it is very difficult to translate this system to represent what goes on in a biological system.

Clinical material, such as blood (plasma) or tissue specimens, offer a closer relationship to real life than cell lines, however these are difficult to analyse due to their heterogeneous nature and the difficulty in obtaining sufficient replicates to overcome inherent sample complexity and variation. It should be highlighted that the novel biomarker does not necessarily have to be screened in the same clinical material as it was discovered.

Table 1.7: Comparison of the three main sample sources used for the discovery of novel cancer biomarkers (adapted from (Drabovich, Martinez-Morillo et al. 2015))

	Tissue	**Blood**	**Cell lines**
Availability	Limited	Good	Very Good
Non-invasive sampling	No	Yes	Yes
Detection of low abundance molecules	Possible	Difficult	Good
Use in diagnostics	Not suitable for routine screening.	Suitable	Not applicable
Clinical Application	Immunohistochemistry	ELISA	Not applicable

1.4.3.1 Tissue specimens

Tissue specimens can be obtained through surgery or biopsy and can be used for the discovery of tumour markers. Tumour tissue should, in theory, contain a high concentration of tumour specific markers and could therefore represent a valuable source for the discovery of novel biomarkers. However, biopsies, in which the sample material is collected, extract not only cancerous tissue, but also healthy and neoplastic surrounding material, which results in a heterogeneous mixture. Furthermore, only small quantities of sample material can be extracted through the biopsy, which limits the number of analyses and the applicable tools. Laser-capture microdissection can be used to separate cancerous

cells from the surrounding tissue, which enables the focussed profiling of only the cancerous cells. However, this leads to a further reduction of sample material (Kulasingam, V., Diamandis 2008b). Additional problems can arise through the storage of tissue material. Samples are commonly stored as formalin fixed paraffin embedded (FFPE), which can affect the ability to detect some proteins and strongly reduces the quality of RNA and DNA.

Tissue invasion and angiogenesis might allow the tumour to shed molecules into the lymphatic system and the blood stream. Epithelial cancers invade surrounding tissues during their growth. This would enable the use of biomarkers discovered in tissue to be applied through a blood test, especially in the case of tumour markers that are related to disease progression.

1.4.3.2 Blood

Human blood or plasma is the most commonly used sample material in clinical practice. It represents a large, and minimal invasive source for the detection of biomarkers, in which molecules associated with various pathological states can be present. Based on its circulation throughout the body, it has contact to a wide range of organs and tissues and can therefore carry information provided by them. It is estimated that blood contains more than 100 000 different protein variants (Ponomarenko, Poverennaya et al. 2016, Archakov, Lisitsa et al. 2015), however the 20 most abundant proteins account for 99 % of the present protein mass (Anderson, Anderson 2002a). These variations in protein concentrations, which span 10 to 12 orders of magnitude, increase the challenges when analysing blood (Fig. 1.12).

Figure 1.12: The human plasma proteome according to (Anderson, Anderson 2002b)

Current mass spectrometry technology can reach a range of 4-5 orders of magnitude. Potential biomarkers are most likely to be present in the low ng-pg/ml concentration range (Kulasingam, V., Diamandis 2008b). It is therefore likely that highly diluted tumour markers will be missed or defined as background noise. In addition, the detection of circulatory tumour cells can be difficult due to their presence at low concentrations. The presence of identified biomarkers in the serum of patients also depends on the protein stability, its clearance and potential post-translational modifications (Kulasingam, V., Diamandis 2008a). For the routine use of blood for biomarker discovery, improvements in the technology are necessary.

1.4.3.3 Immortalised cell lines

Clinical material presents a high complexity, which can be challenging for the discovery of novel biomarkers. Therefore, a less complex sample material, such as cell lines, offers a good alternative solution for the initial discovery phase.

Immortalised cell lines are taken from a population of cells and are able, through the immortalisation, to be cultured for an increased duration. Cell lines enable the study of cellular processing within the cell and also any interaction or influence on the external system through the study of their secretome (Kulasingam, V., Diamandis 2008b). Despite the reduced cell variation present within a sample, cell lines still present a certain degree

of complexity, however the material is readily available and is easier to process and handle than complex clinical samples. On the other side, this sample source also presents some limitations. Cell lines are highly artificial and lack stimuli through the microenvironment. Many cancer cells lines, including the PCa cell lines DU145 and LNCaP, are generated from metastatic tumour cells, which potentially exhibit an altered cell phenotype compared to the primary tumour. Furthermore, a single cell line represents only one genetic variation of a cancer and is therefore unable to represent the true complexity and variability in a clinical setting (Kulasingam, V., Diamandis 2008b).

1.4.4 Discovery tools (omics)

As described in the previous section, biomarkers can originate from various omic levels, such as genomics, transcriptomics and proteomics. The described complexity of sample material highlighted the need for more thorough analysis; however, some limitations in the analysis are still present, especially in the field of proteomics. To overcome such challenges and to achieve a more complete picture of interactions, not only across one type of omic, but also across biological levels, novel approaches, such as the integration of multi-omics platforms, offer an alternative for the discovery of new biomarkers.

1.4.4.1 Strategies for the discovery of novel cancer biomarkers

1.4.4.1.1 Genomic profiling

The human genome is made of a genetic sequence, which represents the instructions to a functional biological system. The genetic code is comprised of building blocks, the so-called nucleotides, which code for single, coding or non-coding, genes. The completed sequencing of the human genome in 2001 (Venter, Adams et al. 2001) by the Human Genome Project, has reformed the world of science. The human genome enabled a comprehensive search for abnormal sequences, mutations, within the genome, generating a greater understanding of the genetic landscape in diseases, such as cancer. A widely-known example of disease-associated genetic mutations lay within the *BRCA1* (Miki, Swensen et al. 1994) and *BRCA2* (Wooster, Bignell et al. 1995) genes. Mutations within these genes can increase the risk for the development of ovarian and breast cancer. Patients with a known family history of BCa are categorised into high-risk groups and are nowadays regularly screened for potential changes in the sequence of these two genes (Wagner, Ball et al. 2018).

1.4.4.1.2 Transcriptomic profiling

The Transcriptome describes the complete collection of actively transcribed genes within a cell at a set point of time. The transcriptome is comprised of coding and non-coding RNA molecules, of which the first can be translated into proteins. After many years, in which the main attention was focussed on coding genes, in recent years, more and more attention was given to transcripts that do not result in a protein (non-coding RNA) (Khurana, Fu et al. 2016, Shabalina, Spiridonov 2004, Wagner, Ball et al. 2018). These non-coding RNA include microRNAs, and small nucleolar RNAs (Mattick, Makunin 2006). Commonly used techniques for transcriptome profiling of samples include gene expression microarrays and RNA-sequencing.

Microarrays are based on cDNA molecules, spotted on a chip, to which complimentary sequences within a sample of interest can bind (Baldi, Hatfield 2011, Schulze, Downward 2001). Microarrays are commonly used due to their affordability and robustness; however, the approach is limited to *a priori* knowledge of genes. For this reason, RNA-sequencing shows a great advantage over microarrays, based on this independence from sequence knowledge (Wang, Z., Gerstein et al. 2009). Furthermore, RNA-sequencing offers a large range of magnitude in the detection and quantification of RNA molecules. RNA-sequencing platforms can not only analyse coding-RNA, but can also be used to focus on non-coding RNA or a closer look can be taken at active translated genes through the screening of ribosome-bound transcripts (Ingolia, Brar et al. 2012). The understanding of the complete complexity of tumour cells and associated interactions can be achieved through recent advantages that enable the analysis of single cell transcriptomes (Ramskold, Luo et al. 2012). The generated transcriptomic profiles of sample material can provide, to a certain degree, information in the potentially present proteome, however the correlation can be influenced by factors such as half-life time of transcripts and protein, as well as post-translational modifications, which can lead to variations between the transcriptome and proteome (Maier, Guell et al. 2009, Kulasingam, V., Diamandis 2008a).

1.4.4.1.3 Proteomic profiling

The proteome is defined by the entirety of proteins present within a cell at the point of sampling. Currently the human proteome compendium consists of approximately 30 000 proteins, which are represented by around 17 000 genes (Human Proteome Map, 2018). This increased number of proteins compared to the number of genes can be explained through alternative splicing event resulting within one gene that can result in the translation of different protein isoforms (Black 2000). Compared to the generation of transcriptomic profiles, the study of the proteome presents difficulties and limitations resulting in lower numbers of routinely quantified proteins.

A continuous challenge is the complexity and large dynamic range of proteins in lysates, especially material derived from clinical specimens (See section 1.4.3.1). As previously mentioned in section 1.4.4.1.2, also in proteomic studies, it is possible to focus the analysis on a subsection of particular interest, such as proteins associated with distinct compartments (e.g. membrane, cytoplasmic, nuclear). A big improvement in the analysis of proteomes was achieved through the development of data-independent acquisition approach in tandem mass spectrometry (Gillet, Navarro et al. 2012), which enabled the routine quantification of more than 3000 proteins present in one sample. Applied to a high-throughput approach, such numbers of protein quantifications could be achieved within 60 to 120 min per sample. Furthermore, current developments in technology have resulted in a higher mass accuracy, higher detection capability and shorter cycling times (Sciex, 2018a), which further helped to increase the quality of results and the throughput of sample material (Gillet, Navarro et al. 2012, Domon, Aebersold 2006).

1.5 Aims and Objectives of the Study

The underlying questions of this study was whether the use of parallel generated multi-omic profiles of two cell-line derived metastasis models will enable and facilitate the discovery of novel disease-associated biomarkers. In addition to this, the study should also investigate the potentially improved correlation of gene and protein expression data through parallel sample collection and omics profiling and furthermore, if the use of proteomic profiling will contribute to a better understanding of underlying changes. Based on these questions, the study was separated into three separate miles stones represented by the chapter 3, 4 and 5.

Milestone 1 (Chapter 3): Development of two *in vitro* models of EMT and their characterisation using the analysis of EMT markers on a gene and protein expression level. The successful development of both models will present the basis for the generation of matching multi-omic profiles in the following chapter 4.

Milestone 2 (Chapter 4): The previously development models of EMT are used for the generation of matching transcriptomic and proteomic profiles of both cell line models and the validation of their desired phenotype using pathway analyses. Furthermore, the generated profiles will be used for the integration of matching genes and proteins and the analysis of their expression correlation. The successful validation and additional characterisation of underlying changes within the transcriptomic and proteomic profiles of both models will support the further use of these profiles as part of an integrative biomarker discovery approach, which is described in chapter 5.

Milestone 3 (Chapter 5): The omic profiles generated in chapter 4 will we integrated for the identification of a core marker set, followed by the characterisation and validation of a selection of markers in a broader context through the screening of cell lines and clinical specimens. Furthermore, *in silico* analyses are to be used for the identification of an association of clinical parameters with the expression of the selected markers.

2. Chapter II - Materials and Methods

2.1 Materials

2.1.1 Reagents

All reagents were stored according to manufacturer's instruction and used within the defined expiry date.

Cell Culture Media	**Supplier**
Keratinocyte-SFM, with L-glutamine	Gibco Life Technologies
MEM Eagle with Earle's BSS, without L-glutamine	SLS (Lonza)

Cell Culture Media Supplements	**Supplier**
Foetel Calf Serum (FCS)	GE Healthcare Hyclone
L-Glutamine	SLS (Lonza)
Transforming Growth Factor β1 (TGF-β)	Peprotech

Further Cell Culture Reagents	**Supplier**
Dimethyl sulfoxide (DMSO)	Santa Cruz Biotechnology
Dulbecco's phosphate buffered saline (DPBS)	SLS (Lonza)
Trypan Blue solution 0.4 %	Sigma-Aldrich
Trypsin/Versene (T&V)	SLS (Lonza)

Chemical Reagents	**Supplier**
alamarBlue™	Invitrogen
2-mercaptoethanol	Sigma-Aldrich
2-Propanol	Sigma-Aldrich
4x Protogel Resolving Buffer	Geneflow
10x TRIS/Glycine/SDS	Geneflow
10x TRIS/Glycine	Geneflow
Acetonitrile + 0.1 % Formic Acid	Fluka Analytical
Acetonitrile	Fluka Analytical
Ammonium Persulphate (APS)	Geneflow
Bovine Serum Albumin (BSA)	Calbiochem
Chloroform	Sigma-Aldrich
Citric acid	Sigma-Aldrich
Deoxyribonucleotide triphosphate (dNTP)	Promega
Dithiothreitol (DTT)	Melford
Double distilled water (ddH_2O)	Barnstead
DPX mountant for histology	Sigma-Aldrich
Ethanol	Fisher Scientific
Ethanol absolute Electran® molecular biology	VWR Chemicals
Formaldehyde solution (37 %)	Sigma-Aldrich
Haematoxylin	Sigma-Aldrich
Hydrochloric acid (HCl)	Fisher Scientific
Hydrogen peroxidase (H_2O_2)	Sigma-Aldrich

Iodoacetamide	Sigma-Aldrich
Isopropanol	Fisher Chemical
Liquid nitrogen	BOC
Instant Dried Skimmed Milk	Co-operative
Methanol	Fisher Chemical
N-Octyl-Beta-Glycopyranoside (OGP)	Apollo Scientific Limited
Nuclease-free water	Ambion
Oligo(dT)$_{15}$ Primer	Promega
Phosphate buffer saline (PBS) tablets	Oxoid
Presept tablets	Johnson and Johnson
Ponceau S solution	Sigma-Aldrich
Protease Inhibitor	Sigma-Aldrich
Protein Assay Dye Reagent Concentrate	Bio Rad
ProteaseMAX™ Sufactant, Trypsin Enhancer	Promega
Protogel Stacking Buffer	Geneflow
Protogel (30 % Acylamide mix)	Geneflow
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich
Reverse Transcriptase	Promega
RNaseZAP	Ambion
RNasin	Promega
RNA-STAT-60	Amsbio
RT 5x Buffer	Promega
SYBR® Green	BioRad
Teepol	Johnson and Johnson
TEMED	Geneflow
Triethylammonium bicarbonate buffer (TEAB)	Sigma-Aldrich
Trizma (Tris) base	Sigma-Aldrich
Trypsin/Lys-C Mix, Mass Spec Grade	Promega
Tween 20	Sigma-Aldrich
Urea	Melford
Vectashield® Mounting Medium with DAPI	Vector Laboratories
Water with 0.1 % Formic Acid	Fluka Analytical
Xylene	Fisher Scientific

Immunochemical Reagents	**Supplier**
Rabbit anti-human CDH1 (24E10)	Cell Signaling
Rabbit anti-human FN1 (F3648)	Sigma-Aldrich
Rabbit anti-human CDH2 (D4R1H)	Cell Signaling
Rabbit anti-human VIM (D21H3)	Cell Signaling
Rabbit anti-human CRMP4 (DPYL3) (ab101009)	Abcam
Rabbit anti-human FBLI1 (ab154417)	Abcam
Rabbit anti-human P4HA2 (PA5-53530)	ThermoFisher Scientific
Rabbit anti-human SDPR (ab103230)	Abcam
Rabbit anti-human CYCA (ab41684)	Abcam
Goat anti-rabbit IgG Biotin (B8895)	Sigma-Aldrich
Goat anti-rabbit IgG HRP-linked Antibody (7074S)	Cell Signaling
Swine anti-rabbit FITC (F0205)	Dako
Precision Plus WesternC Standards	Bio Rad
Precision Protein StrepTacin-HRP Conjugate	Bio Rad

Kits	Supplier
Avidin/Biotin Blocking Kit	Vector Laboratories
HRM Calibration Kit	Biognosys
RNeasy Mini Kit (250)	Qiagen
R.T.U. Vectastain Universal Elite ABC Kit	Vector Laboratories
Clarity Western ECL Substrate	Bio Rad
CyQUANT® Direct Cell Proliferation Assay Kit	Invitrogen

2.1.2 Consumables and Equipment

All glassware was washed using Teepol, rinsed twice with distilled water before sterilisation using autoclaving.

Laboratory Plastics, Glassware and Sharps	Supplier
Bijou tubes (7 mL)	Starlab
Bioanalyser chips	
Cell culture flasks (T25, T75, T175)	Sarstedt
Cell scraper	
Clear flat bottom 6-well plate, sterile (cell culture)	Sarstedt
Clear flat bottom 96-well plate (protein assay)	Starlab
Cryogenic vials (1.0 mL)	Starlab
Falcon tubes (15 mL, 50 mL)	Sarstedt
Filter tips (10ul, 20ul, 100ul, 200ul, 1000ul)	Starlab
Glass bottles	Duran
Glass coverslips	
Glass slides	
Hypersep™ Spin Tip C_{18}	Thermo Scientific
LC vials & Caps	Chromatography Direct
Micro tubes (0.5 mL, 1.5 mL, 2.0 mL)	Sarstedt
Nitrocellulose membrane	GE Water & Process Techn.
Rotor-Gene Strip Tubes & Caps	Starlab
Pasteur pipettes	Sarstedt
Petri dishes	Sarstedt
Scalpels	SLS
Serological pipettes (5 mL, 10 mL, 25 mL)	Sarstedt
Syringe filter 0.2µm	Sartorius
Syringes (20 mL)	Medicina
Western Blot filter paper	GE Healthcare

Equipment	Supplier
4 °C Fridge	LEC Medical
-20 °C Freezer	LEC Medical
-80 °C Freezer	Panasonic
2100 Bionanalyzer	Agilent Technologies
37 °C/5 % CO_2 Incubator	Scientific Laboratory Supplies
4 °C Centrifuge	Eppendorf
Autoclave	Rodwell
Benchtop vortex mixer	Scientific Industries
Class II Safety Cabinet	Walker

Ekspert™ nanoLC 425	eksigent
Axio Observer.Z1 microscope	ZEISS
Haemocytometer (counting chamber)	Weber
Heating block	Lab-Line
Micropipettes (2 µl, 10 µl, 100 µl, 200 µl, 1000 µl)	Gilson/Starlab
Minispin benchtop centrifuge	Eppendorf
Mixing block	Bioer
Multichannel pipette (300 µl)	Eppendorf
Nanodrop ND-8000 spectrophotometer	Thermo Scientific
Nanopure Diamond water reservoir	Barnstead
Nikon Eclipse Ts100 Light Microscope	Olypmus
PCR workstation cabinet	Grant-Bio
Real-time qPCR thermal cycler	Qiagen
Rocker	Stuart
SCIEX TripleTOF 6600	SCIEX
Syngene G:Box	Syngene
Sonicator	Fisherbrand
Tecan Ultra Microtiter Plate Reader	Tecan
Thermoblock	Biometra
Vacuum concentrator	Eppendorf
Waterbath	Clifton
Weighing Scale	Fisher Scientific

Software	**Company**
AxioVision SE64 Rel.4.8.	ZEISS
BaseSpace (Online)	Illumina
GraphPad Prism v7	GraphPad Software Inc.
Morpheus (Online)	Broadinstitute
NPD.view 2 v2.7.25	Hamamatsu
OneOmics™ (Online)	Illumina/Sciex
Protein Pilot v5.0	Sciex
Peak view v2.1	Sciex
Rotor-GeneQ Series Software v2.3.1	Qiagen
Statistica v13.3	TIBCO
MetaCore v6.37	Clarivate Analytics
Genesys v1.5.4.0	Syngene

2.1.3 Buffers and Gels

2.1.3.1 Cell culture growth media

Growth medium for P5B3	**For 500 ml**
Keratinocyte-SFM, with L-glutamine	487.5 ml
Fetal Calf Serum	12.5 ml (2.5 %)

Growth medium for DU145	**For 500 ml**
MEM Eagle with Earle's BSS	487.0 ml
Fetal Calf Serum	12.5 ml (2.5 %)
L-Glutamine	5 ml (2 mM)

2.1.3.2 Immunofluorescence staining

Blocking buffer	**For 50 ml**
Phosphate Buffer Saline	45 ml
Fetal Calf Serum	5 ml
Tween 20	50 µl

Washing buffer	**For 100 ml**
Phosphate Buffer Saline	100 ml
Tween 20	100 µl

4 % Formaldehyde	**For 40 ml**
Phosphate Buffer Saline	35.7 ml
37 % Formaldehyde	4.3 ml

2.1.3.3 Mass spectrometry analysis

Cell lysis buffer	**For 50 ml**
Urea	28.5 g
Dithiothreitol (DTT)	1 g
N-Octyl-Beta-Glycopyranoside (OGP)	0.5 g
ddH_20	50 ml

Prior use, Protease Inhibitor was added in a dilution of 1:100 to the cell lysis buffer

2.1.3.4 Immunohistochemistry staining

Antigen retrieval buffer	**For 1 L**
Sodium Citrate tribasic dihydrate	2.94 g
ddH_20	1000 ml

pH was adjusted to a pH of 6.0

Blocking buffer for IHC staining using fluorescent 2nd AB	**For 50 ml**
DBPS	50 ml
BSA	5 g
Tween20	50 µl

Phosphate buffer saline	**For 1 L**
Phosphate buffer saline (PBS) tablets	10 x
ddH_20	1000 ml

2.1.3.5 Western blot analysis

4x Laemmli buffer	For 50 ml
1M Tris-HCl pH 6.8	10 ml
Glycerol	20 ml
SDS	4.0 g
β-mercaptoethanol	10 ml
Bromophenolblue	0.1 g
ddH_2O	Up to 50 ml

Running buffer	For 1 L
10x TRIS/Glycine/SDS	100 ml
ddH_20	900 ml

Transfer buffer	For 1 L
10x TRIS/Glycine	100 ml
Methanol	200 ml
ddH_20	700 ml

10x Tris-buffered saline (TBS)	For 1 L
Trizma Base	24.2 g
Sodium chloride	80 g
ddH_20	1000 ml

pH was adjusted to 7.6

Blocking buffer (5 %)	For 50 ml
Instant Dried Skimmed Milk	2.5 g
ddH_20	50 ml

2.1.3.6 Antibodies used throughout this study

2.1.3.6.1 Western blot analysis

Antibody (Host species)	Expected size	Dilution
VIME (Rabbit)	57 kDa	1/1000
CADH1 (Rabbit)	135 kDa	1/1000
CADH2 (Rabbit)	140 kDa	1/1000
FINC (Rabbit)	220 kDa	1/1000
CYPA (Rabbit)	18 kDa	1/1000
Goat anti-rabbit IgG HRP-linked	NA	1/1000
Precision Protein StrepTacin-HRP Conjugate	NA	1/5000

2.1.3.6.2 Antibodies used for immunofluorescence and immunohistochemistry analysis on cells and TMAs

Antibody (Host species)	IF	IHC	IF on TMA
VIME (Rabbit)	1/100	NA	NA
CADH1 (Rabbit)	1/200	NA	NA
FINC (Rabbit)	1/400	NA	NA
CRMP4 (DPYL3) (Rabbit)	NA	1/250	1/250
FBLI1 (Rabbit)	NA	1/100	1/100
P4HA2 (Rabbit)	NA	1/350	1/350
SDPR (Rabbit)	NA	1/50	1/50
CADH1 (Mouse)	NA	NA	1/250
CADH2 (Mouse)	NA	NA	1/500
Goat anti-rabbit IgG Biotin	NA	1/1000	NA
Goat anti-rabbit FITC	1/40	NA	NA
Goat anti-mouse Alexa Fluor 568	NA	NA	1/500
Donkey anti-rabbit Alexa Fluor 488	NA	NA	1/500

2.1.3.7 Quantitative real-time PCR primer used throughout this study

Primer	Gene	Primer 5'-3'	Annealing Temp.	η^*
DPYSL3 F *DPYSL3* R	Dihydropyrimidinase like 3	GGACAACTTCACAGCCATTCCTG GTGCTTGTCACAGCCACGAACT	60°C	95 %
FBLIM1 F *FBLIM1* R	Filamin binding LIM protein 1	CGGCAGAACCTGTTGAGAAAGG ACGTGAAGCACTGGGCATGGTA	60°C	98 %
SDPR F *SDPR* R	Serum deprivation response protein	TCTTCCAGGAGGAAAATGAG CAAATCATCATCTGAGGAGAG	54°C	90 %
P4HA2 F *P4HA2* R	Prolyl 4-hydroxylase subunit alpha 2	CGAATTCTTCACCTCTATTGG GATGTACTCTTTCAGAGACTG	52°C	91 %
VIM F *VIM* R	Vimentin	GAGAACTTTGCCGTTGAAGC GCTTCCTGTAGGTGGCAATC	58°C	N/A
FN1 F *FN1* R	Fibronectin	CAGTGGGAGACCTCGAGAAG TCCCTCGGAACATCAGAAAC	58°C	N/A
CDH1 F *CDH1* R	E-cadherin	TGCCCAGAAAATGAAAAAGG GTGTATGTGGCAATGCGTTC	58°C	N/A
CDH2 F *CDH2* R	N-cadherin	ACAGTGGCCACCTACAAAGG CCGAGATGGGGTTGATAATG	58°C	N/A
TBP F *TBP* R	TATA-box binding protein	TATAATCCCAAGCGGTTTGC GCTGGAAAACCCAACTTCTG	58°C	N/A
ZEB1 F *ZEB1* R	Zinc finger E-box-binding homeobox 1	GGCATACACCTACTCAACTACGG TGGGCGGTGTAGAATCAGAGTC	58°C	N/A
TWIST1 F *TWIST1* R	Twist-related protein 1	GGAGTCCGCAGTCTTACGAG TCTGGAGGACCTGGTAGAGG	58°C	N/A
SNAI1 F *SNAI1* R	Zinc finger protein SNAI1	CCTCCCTGTCAGATGAGGAC CCAGGCTGAGGTATTCCTTG	58°C	N/A
SNAI2 F *SNAI2* R	Zinc finger protein SNAI2	GGGGAGAAGCCTTTTTCTTG TCCTCATGTTTGTGCAGGAG	58°C	N/A

*Primer efficiency

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Routine Cell Culture of prostate cancer cell lines

Two independent prostate cancer cell lines/clones were used during the study (Tab. 2.1). Both cell lines were of androgen-independent prostate cancer, originating from either the primary or the metastatic/secondary tumour site. The growing cells were routinely checked for growth, necessity of media replacement and potential contamination and grown at 37 °C in a humidified atmosphere with 5 % (v/v) CO_2. Cells were further processed at a confluency of 70 – 80 %. For this, the cell medium was removed and the cells washed twice with Dulbecco's phosphate buffered saline (DPBS) for a complete removal of the remaining medium. 5 ml of 0.05 % trypsin mixed with 0.02 % versene was added and incubated at 37 °C and 5 % CO_2 until cells presented a rounded up cell shape, indicating detachment from the surface. Following trypsinisation an equal amount of medium was added to the cells, which were transferred into a collection tube and centrifuged at 240 x g for 5 min. The supernatant was removed and the resulting cell pellet resuspended in 10 ml of fresh medium. Each cell line was used within 10 passages from master stock.

Table 2.1: Table of the two prostate cancer cell lines used during the purpose of study , describing their source, growth conditions and characteristics

Cell line	Description	Source	Growth Condition
P5B3	Single cell clone derived from OPCT-1, a androgen independent primary prostate cancer cell line (T1cN0M0, Gleason 3+3)	ONYVAX	KSFM + 2.5 % (v/v) FCS
DU145	Androgen independent metastatic prostate cancer cell line	American Tissue Culture Collection (ATCC® HTB-81™)	EMEM + 2.5 % (v/v) FCS + 2mM L-Glutamine

The counting of the cells was performed using trypan blue and a haemocytometer. Trypan blue highlights dead cells through blue staining, which were excluded from the cell count. A desired amount of cells was transferred into a fresh cell culture flask. This step was considered as a passage. For the generation of stocks, the cell pellet was resuspended in FCS and DMSO (10:1) (freezing media) and stored until further use at -80 °C. Each vial was gently thawed and immediately transferred into a new tissue culture flask containing fresh media.

2.2.1.2 Treatment with Transforming Growth Factor β – preliminary work (Chapter III)

For the development of inducible EMT models using P5B3, cells were treated with 10 ng/ml TGF-β for 5 days. The cells were initially passaged into a new flask, after 24h the media was exchanged with new supplemented or non-supplemented media. During the stimulation for 5 days, the medium was exchanged every second day with fresh supplemented or non-supplemented medium.

2.2.1.3 Treatment with Transforming Growth Factor β – dataset generation and further work (Chapter III, IV, V)

To keep the response to the TGF-β treatment constant, the cells were seeded at a density of 50 000 cells for P5B3 or 75 000 cells for DU145 into a T175 cell culture flask to prevent the necessity to passage the cells during the treatment. Prior testing on the adequate seeding density identified the required seeding density (data not shown).

The cells were prepared as described above. The treatment length varied from 3 to 10 days, counting from the first addition of supplemented media. P5B3 and DU145 were cells were treated with 10 ng/ml TGF-β. Medium was replaced in regular intervals (Monday, Wednesday, Friday) to maintain a steady level of TGF-β and a healthy nutrient supply.

2.2.1.4 Scratch Assay

Scratch assays are commonly used for the definition of migratory capabilities of cell lines of interest. Here, cells of each model were grown for 10 days at their respective conditions, harvested and counted. 1 x 10^5 cells were transferred into each well of a 24-well plate and incubated for 24 hours. Then media was removed, each well was washed twice with PBS and then fresh FCS-free media was added. Treated cells were further supplemented with

TGF-β. After 24 hours, a scratch was applied to the surface of each well using a sterile 100 µl tip. Media was removed and the wells washed twice with PBS. New FCS-free media (with or without supplementation) was added and the created scratch was imaged and measured on an Axio Observer.Z1 microscope. Prior to the start of the experiment, a line was drawn on the bottom of each well and the images were taken above the line. This enabled accurate, matching measurements over 24h. After 24h, the scratches were imaged and measured again at the identical position. Wound closure was calculated in %. The initial scratch measured at t0 was defined as 100 %. The measurement at 24h defined the % of wound closure.

2.2.2 Molecular biology

2.2.2.1 RNA-extraction

The performed method for the extraction of RNA depended on planned downstream use. RNA was extracted for the targeted analysis of single genes using quantitative real-time PCR and whole transcriptome analysis by RNA-sequencing. Both approaches are described below.

2.2.2.1.1 RNA-extraction for downstream use in quantitative real-time PCR

The cells were grown as previously described (section 2.2.1.2/2.2.1.3). Then they were trypsinised and collected in a collection tube, centrifuged and the resulting cell pellet was, after the complete removal of the remaining medium, immediately frozen in liquid nitrogen. The samples were stored -80 °C until further use.

The extraction was performed using the RNeasy Mini Kit (Qiagen). For the improved quality of extracted RNA and the inactivation of potentially present RNases the RLT buffer was prepared by adding 10 µl of β-mercaptoethanol to 1 ml Buffer RLT.

The generated cell pellet was mixed with 350 to 700 µl of Buffer RLT. In case the result of this step was showing a viscous liquid, the volume of RLT Buffer was increased from 350 to 700 µl. One volume of 70 % EtOH was added to the tube and mixed by pipetting. The sample was transferred onto a provided spin column centrifuged for 15 s at 8000 x g, the flow-through was discarded and 700 µl of Buffer RW1 added. The spin column was centrifuged at 8000 x g for 30 secs and the flow through discarded. 500 µl Buffer RPE was added and the column centrifuged as described before. This step was repeated twice. After the second repeat, the filter column was placed in a new collection tube and the column was centrifuged at full speed. This step secured the complete removal of buffer RPE. For the final elution the filters were placed in 1.5 ml micro tubes, 30 µl of RNase-free water was added onto the surface of the filter, incubated at RT for 10 min and the column was centrifuged for 1 min at 8000 x g. Until further use, the samples were stored at -80 °C.

2.2.2.1.2 RNA-extraction for downstream use in RNA-sequencing

For the analysis of the extracted RNA with RNA sequencing, the extraction of the cells was performed using RNA-STAT-60.

Here, the media was fully removed from the cell culture flask and 1 ml of RNA-STAT-60 was added to a T175 flask. The liquid was spread on the surface by inclining the flasks and after a short incubation, the liquid containing the cells was collected at one corner of the flask using a cell scraper. For homogenising of the samples, the liquid was mixed by pipetting, transferred into a 2 ml micro tube and stored at -80 °C until further use.

Prior to RNA extraction, the samples were fully thawed and equilibrated to room temperature. 200 µl of chloroform was added per 1 ml RNA-STAT-60 used and the samples were mixed by shaking for 60 seconds and rested for 2-3 min at RT. Then, samples were centrifuged at 12 000 x g for 15 min at 4 °C resulting in a red phenol chloroform phase and the clear aqueous phase. The aqueous phase (containing the RNA) was transferred into a new 1.5 ml micro tube and 0.5 ml isopropanol added to each 1 ml of RNA-STAT-60 initially used. The liquid was mixed by pipetting and rested for 8 min at RT. After this, the samples were centrifuged at 12 000 x g for 10 min at 4 °C, resulting in the formation of a white pellet on the bottom of the tube, containing the RNA. The supernatant was removed and the pellet washed with 1 ml of 75 % EtOH by vortexing. Then, the tube was centrifuged at 7500 x g for 5 min at 4 °C, the EtOH was completely removed and the pellet dried under the fume hood. This step was checked regularly to ensure the complete evaporation of the EtOH without a too intensive drying of the samples. As a final step, the samples were resuspended in 100 µl of RNase-free water.

Following the initial RNA-extraction, the samples were further purified using the RNeasy Mini Kit from Qiagen. The RNA was extracted as described above (section 2.2.2.1.1). An additional on-column DNase digestion was performed after the washing step using 700 µl of Buffer RW1. 80 µl of the DNase digestion buffer, prepared according to manufacturer's protocol, was added directly onto the spin column filter and incubated for 15 min at RT. Then the protocol was followed as described above (section 2.2.2.1.1).

2.2.2.2 Quantification of extracted RNA

Independent from the further use of the RNA was each sample initially quantified using the NanoDrop 8000. Samples with a 260/280 value of 1.6 and above were deemed to be sufficient for the downstream use in qRT-PCR. The sample analysis was performed according to manufacturer's protocol.

Samples for use in RNA-sequencing experiments were initially quantified using the NanoDrop 8000 and according to the resulting quantification diluted to a concentration of 500 ng/ml. Following this, the samples were analysed with the Agilent Bioanalyzer using the RNA Agilent Nano Kit with RNA Nano Chips as recommended by manufacturer's protocol (see appendix A2).

2.2.2.3 Reverse transcription

For the use of the extracted RNA in qRT-PCR analysis, it was necessary to convert the RNA into cDNA. This process is called reverse transcription. For this, 1.5 µg of RNA were mixed with 1 µl Oligo dT and adjusted to 10 µl with molecular grade water. The mix was incubated at 70 °C for 5 min and immediately transferred onto ice for a further 5 min. Following this, 5 µl 5x buffer, 1 µl Reverse Transcriptase, 1 µl dNTPs, 0.7 µl RNasin® and 7.3 µl molecular grade water were added to each tube and incubated in a water bath at 40 °C for one hour. For the inactivation of the reaction, the samples were incubated for 5 min at 95 °C and then stored at -20 °C until further use.

2.2.2.4 Quantitative real-time PCR

Quantitative real-time PCR was used for the analysis and measurement of mRNA expression levels of markers of interest. For each reaction, 1 µl of cDNA was mixed with 5.75 µl SYBR® Green, 0.5 µl of the forward and reverse primer at a concentration of 10 µM and 3.75 µl of molecular grade water. The analysis of each sample was performed twice in duplicates on a Rotor-Gene Q real-time PCR cycler manufactured by Qiagen. Samples were repeated according to necessity. The analysis was performed with 40 cycles. The initial denaturation was 5 min at 95°C, followed by 40 cycles. These 40 cycles were based on 10 secs denaturation (95°C), 20 secs annealing (temperature optimised for each gene of interest) and 20 elongation (72°C). After the 40 cycles, a melt curve analysis was performed in every analysis to ensure targeted amplification of the gene of interest and not primer-dimers.

2.2.2.5 Primer efficiency testing

Each primer was tested for its efficiency after purchase and prior to experimental application. For this, a serial dilution of a test sample was generated consisting of a 5-fold dilution each time to create five separate samples/dilutions. The new primer was then analysed as described in section 2.2.2.4, with an estimated annealing temperature. Each measured data point was then assigned to a given concentration resulting in a standard

curve. The x-axis indicates the given concentrations of the used samples compared to the y-axis, which describes the cycling time (Ct) value of each sample (Fig. 2.1). The steepness describes as M is the efficiency of a primer. An m=-3.32 represents 100 % efficiency. If the slope of the standard is too steep, the primer is over-efficient, if the slope is more negative, it gives indications of an inefficient primer. Ideally, the measured efficiency should be 100 % however; a range from 90-110 % was defined as sufficient.

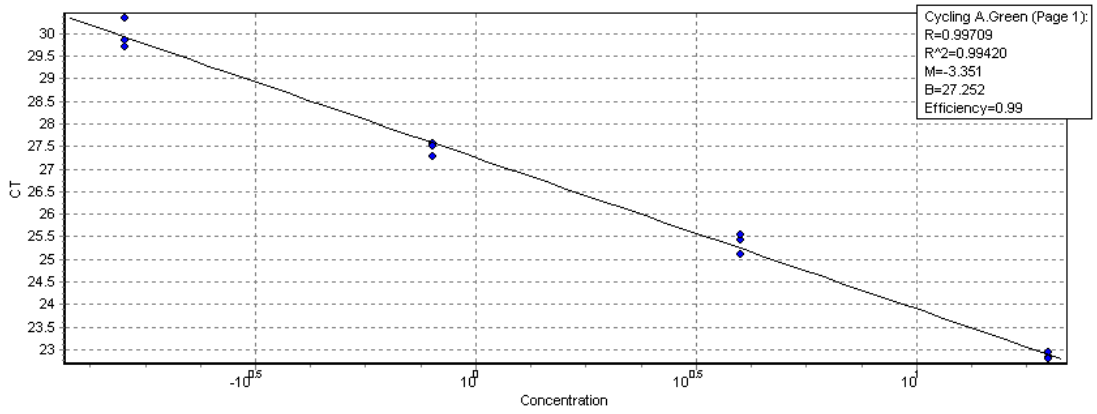


Figure 2.1: Example representation of a generated standard curve for the efficiency testing of a novel primer set.

Table 2.2: Summary of cell lines used throughout the process of this study.

Cell line	Tissue	Type
SKBR3	Breast cancer	Pleural effusion metastasis
MCF7	Breast cancer	Pleural effusion metastasis
MDA231	Breast cancer	Pleural effusion metastasis
MDA453	Breast cancer	Pericardial effusion metastasis
MDA468	Breast cancer	Pleural effusion metastasis
MCF10A	Breast	Fibrocystic disease
OPCT1	Prostate cancer	Primary tumour
P4B6	Prostate cancer	Single cell clone derived from OPCT-1
P4B6B	Prostate cancer	Single cell clone derived from OPCT-1
PC-3	Prostate cancer	Bone metastasis
LNCaP	Prostate cancer	Left supraclavicular lymph node metastasis
SAOS	Osteosarcoma	Primary tumour
P5B3	Prostate cancer	Single cell cline derived from OPCT-1
DU145	Prostate cancer	Central nervous system metastasis

2.2.2.6 RNA-sequencing analysis

The RNA-sequencing analysis was performed by the DeepSeq facility of the University of Nottingham (DeepSeq, 2019). The facility requested 2 µg of RNA with a concentration

of 200 ng/μl and a RNA integrity number (RIN) of 8. Samples were prepared accordingly and submitted to DeepSeq, where they were further processed and analysed.

Solely members of DeepSeq performed the sample processing, however their protocol was described as the following: To ensure the quality and quantity of measured RNA, a repeated measurement of total RNA using the Qubit RNA BR assay kit was performed. This assay is highly selective for RNA and can well tolerate any contaminants within the sample. The fluorescent dye of the assay emits a signal only when it is bound to RNA. Then an additional quality control step using the Agilent 2100 Bioanalyser was performed. Libraries were created using 1 μg of total RNA for each sample and the standard protocol for the IlluminaTruSeq Stranded Total RNA with Ribo-Zero (Human/Mouse/Rat) kit was followed. Library quality control was performed using the 2100 Bioanalyser and High Sensitivity Kit. Libraries were quantified using qRT-PCR, pooled at desired concentrations, denatured and loaded for sequencing according to manufacturer's instructions. Sequencing was performed on Illumina NextSeq500 sequencing platform, and the samples were run over four NextSeq500 High Output v2 150cycle kits to generate 75bp paired-end reads.

2.2.3 Protein biology

2.2.3.1 Protein extraction for use in Western blot and mass spectrometry analysis

For the analysis of the proteomic profile, the cells were grown under the described conditions until 70-80 % confluent and for the defined treatment period. On the day, the remaining media was removed, the cells were washed twice with DPBS and the remaining DPBS was removed as completely as possible. Then 300 - 750 µl of cell lysis buffer was added and the cells were incubated for 2 min. The amount of cell lysis buffer used was dependent on the flask size. A cell scraper was used to detach the cells from the surface and to collect them in one corner of the flask. After homogenisation, the liquid was transferred into a collection tube. The lysed cells were sonicated for 5 min and stored on ice for 5 min; this was performed twice followed by a centrifugation for 10 min at 10 000 x g and at 4 °C. The resulting supernatant was transferred into a new collection tube. Potential remaining cell pellets were stored, together with the supernatant, at -80 °C until further use.

2.2.3.2 Measurement of protein concentration using BioRad DC protein assay

Protein concentration was measured using a protein assay based on the Bradford method and bovine serum albumin (BSA) was used as a standard at the concentrations of 500, 450, 400, 300, 200, 100, 50, 0 µg/ml BSA. The standards were generated by resuspending a BSA stock of 1 mg/ml in undiluted cell lysis buffer, which was then diluted to the described concentration using 1:10 diluted cell lysis buffer.

The utilised cell lysis buffer contained 9.5M urea, for this reason, the samples were diluted at least 1:3. This was necessary to reduce the concentration of urea in the samples to a level that is compatible with the assay. For the dilution, cell lysis buffer diluted 1:10 in H_20 was used. As preparation, the dye reagent was diluted 1:5 with H_20 (working dye). Then 10 µl of each sample was added into one well. The standards were analysed in quadruplicates and the samples in triplicates using a 96-well plate. Then 200 µl of the working dye were added and the absorbance was measured at 570 nm after an incubation of 10 min at RT. The measurement was performed, using a Tecan Ultra Microtiter Plate Reader.

2.2.3.3 SDS-PAGE and Western blot

For the detection of a target protein expression, cell lysates were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. For each sample, 50 µg of protein was used in the experimental procedure.

2.2.3.3.1 SDS-PAGE in reducing conditions

For a successful separation according to molecular weight, the intra- and inter-molecular disulphide bonds within the sample material had to be reduced. This process denaturated the proteins and additionally provides each with a negative charge. This enables the separation of the proteins according to size. For this, 3 volumes of sample were mixed with 1 volume of 4 x Laemmli buffer and incubated at 95°C for 10 min on a heating block. Then the samples were cooled to RT and were ready to use or stored at -80°C until further use.

The identification of single proteins within a sample using Western blot analysis required the initial separation of proteins according to their size. For this SDS polyacrylamide gel electrophoreses (SDS-PAGE) was performed. Here, 1.5 mm thick mini SDS gels with 10-well capacity were prepared (Tab. 2.3). Each gel consisted of two parts, a resolving and a stacking gel. First, the resolving gel was prepared and immediately poured between two glass slides, filling it until 2 cm to the top then directly covered with Isopropanol for a levelled border of the gel.

Table 2.3: Resolving gel preparation for one 1.5 mm gel

Resolving gel (8 ml)	**10 %**
Protogel (30 %)	2.7 ml
Resolving Buffer 4x	2.0 ml
ddH_20	3.3 ml
10 % ammonium persulphate (APS)	80 µl
N,N,N‘,N‘-Tetramethylethylenediamine (TEMED)	8 µl

The gel was left to set for about 15 – 20 min. After complete setting of the gel, the Isopropanol was removed, the space between the glass slides washed twice with H_2O and the remaining H_2O was removed with filter paper. In the meantime, the stacking gel was prepared (Tab. 2.4) and immediately added on top of the resolving gel and after pouring the 10-well combs were inserted. Again, the gel was left to set for about 15-20 min.

Table 2.4: Stacking gel preparation for one 1.5 mm gel

Stacking Gel (3 ml)	**5 %**
Protogel (30 %)	500 µl
Stacking Buffer 4x	800 µl
ddH_20	1.8 ml
10 % ammonium persulphate (APS)	30 µl
N,N,N‘,N‘-Tetramethylethylenediamine (TEMED)	3 µl

The prepared gels were inserted into the appropriate running modules and placed inside the buffer tank. The tank was filled with running buffer. 5 µl of protein ladder was loaded into the first well, followed by the generated sample material. In all experiments, 50 µg of each sample was loaded. Following this, the samples were separated by electrophoresis using a constant voltage of 100 V over 45 min.

2.2.3.4 Western blot

2.2.3.4.1 Protein transfer to nitrocellulose membrane

After the separation by SDS-PAGE, the proteins were transferred from the gel onto a nitrocellulose membrane. For this, the gel was assembled into a so-called “sandwich”. The “sandwich” was composed of the negative electrode, 2 sponges soaked in transfer buffer, a filter paper, the gel, the membrane, another filter paper, 2 sponges and the positive electrode (Fig. 2.2). Prior to use, the membrane was soaked in dH_2O and shortly prior to assembly, transferred into transfer buffer. The filter paper, as well as the sponges, were pre-conditioned in transfer buffer.

Black plate
2 – 3 sponges
Filter paper
Gel
Membrane
Filter paper
2 – 3 sponges
Red plate

Figure 2.2: Schematic representation of "sandwich" assembly for protein transfer onto nitrocellulose membrane

Then, the sandwich was placed into the transfer tank according to the direction of the respective electrode orientation, fully covered with transfer buffer and a constant current of 180 mA was applied for 75 min. The transfer was performed at 4°C. After the transfer, an optional Ponceau Red staining was performed. This staining enabled a quick reversible staining for protein bands and can be performed as a quality control of the transfer. The staining can be easily removed through washes with 1X TSBT.

2.2.3.5.2 Immunoprobing of target proteins on nitrocellulose membrane

After the transfer, the membrane was cut according to the molecular weight of proteins of interests. The selected sections of the membrane were incubated with blocking buffer (1x TBST + 5 % milk) for 1h at RT under constant shaking. This step was performed to prevent non-specific binding of later-used antibodies. After 1 h, the blocking buffer was replaced by fresh buffer containing the desired antibody at the recommended concentration, and the membranes were incubated overnight at 4°C under constant shaking.

On the following day, the primary antibody was removed and the membranes washed three times with 1X TBST for 10 min at RT and under constant shaking. Then, the secondary antibody (host specific) and the conjugate specific for the used molecular weight ladder were diluted in blocking buffer according to the recommended concentration and then added to the membrane, which was then incubated for 2 h at RT under constant shaking. After this, the membrane was washed again three times with 1X TSBT. In the meantime, the Clarity Western ECL Substrate was prepared at a 1:1 dilution. After washing, the membrane was placed onto a dark background, covered with the previously prepared EZ-ECL substrate and the chemiluminescent image acquisition was performed. Exposure times were adapted to target protein quantities, ranging from 1 sec to 5 min.

2.2.4 Mass spectrometry

2.2.4.1 Sample preparation for pilot work (Chapter III)

For the analysis, 25 µg of protein was transferred into single 1.5 ml collection tubes and placed into a vacuum spin concentrator with a temperature of 45 °C until complete evaporation. Each sample was resuspended with 93.5 µl of TEAB (50 mM) and transferred into a new 0.5 ml collection tubes. Here, 1 µl of 0.5 mM DTT was added and the samples incubated for 20 min at 56°C. After this, 2.7 µl of 0.55mM Iodoacetamide was added and incubated at RT for 15 min in the darkness. Finally, 1 µl of 1 % ProteaseMax™ and 2 µl of Trypsin were added and the samples were incubated at 37 °C overnight. After this, the samples were evaporated to dryness (as before) and resuspended in 25 µl of 5 % ACN + 0.1 % FA. A volume of 12 µl of this was transferred into LC vials.

2.2.4.2 Sample preparation for data set generation (Chapter IV + V)

All samples were normalised, by adjusting volume of cell lysis buffer in the sample, to the lowest protein concentration. 50 µg of protein was transferred into single 0.5 ml micro tubes. The cell lysates of the samples were stored in cell lysis buffer, which contained 9.5M Urea. A reduction of the urea concentration to less than 6 M was necessary in order to perform the reduction and alkylation of the proteins, by addition of varying volumes of 50 mM TEAB. The reduction was performed using DTT at a final concentration of 5 mM and an incubation at 37 °C for 30 min. This step was followed by iodoacetamide alkylation at a final concentration of 15 mM and an incubation of 30 min at RT in the dark. Urea concentration was reduced to < 1.2 M to improve Trypsin/Lys-C efficiency. The dilution was performed using 50 mM TEAB. Trypsin/Lys-C was prepared according to manufacturer's instruction and used for the trypsinisation at a ratio of 20:1 protein:trypsin. The samples were then incubated at 37 °C for 16 h on a heated shaking plate. Then Trypsin/Lys-C was added a second time at a ratio of 20:1 (as per manufacturer's optional digestion protocol) to the samples, followed by incubation at 37 °C for 3 h.

2.2.4.3 Sample clean-up using Hypersep™ C18 spin column

The cell lysate samples were desalted and concentrated using a C_{18} spin column. Each column was placed together with the supplied holder into a 2.0 ml micro tube and processed according to manufacturer's protocol following conditioning. The following bullet points describe the procedure. Each time, the described volume was added on top

of the column and then centrifuged at 1073 x g (4000 rpm) in an Eppendorf™ MiniSpin™ for 30 s (time and rpm optimised for the lab, data not shown). The flow through was discarded after each step.

- 3 x 50 µl of 60 % ACN + 0.1 % Formic acid
- 3 x 50 µl of 0.1 % Formic acid
- 50 µl of sample. (Multiple repeats might be necessary, dependent on the amount of sample)
- 3 x 50 µl 0.1 % Formic acid (washing step)
- 3 x 50 µl of 60 % ACN +0.1 % Formic acid (Elution)

Then all samples were placed in a vacuum spin concentrator until the full liquid was evaporated. Finally, the samples were reconstituted in 5 % ACN + 0.1 % Formic acid to a concentration of 2 µg/µl.

2.2.4.4 Mass Spectrometry Analysis

The mass spectrometry analyses were performed by Dr David Boocock and Dr Clare Coveney.

2.2.4.1.1 Pilot work (Chapter III)

All analyses were performed on a SCIEX TripleTOF6600 instrument (Sciex, Warrington, UK). For the creation of a peptide library, 10 µl of each sample was collected in a separate LC vial, 1 µl of HRM peptide mix (selection of non-naturally occurring synthetic peptides in a pooled mix used to calibrate retention time) was added to each LC vial and analysed IDA mode (information or data-dependent acquisition): Top 30; dynamic exclusion 20 s after 2 occurrences, 50 ms accumulation time per target and a cycle time of 1.8 s. The gradient elution was 2-40 % ACN/0.1 % FA over 110 min and 40-80 % ACN/0.1 % FA over 5min. After this, the column was washed at 80 % for 2 min prior to re-equilibration. The total run time was 120 min.

The same samples were then analysed by quantitative SWATH™ (data-independent acquisition developed by SCIEX – Sequential Window Acquisition of all THeoretical ions) mass spectrometry (SCIEX, Warrington, UK) using 40 variable m/z windows, a 40 ms accumulation time and a cycle time of 1.8 secs. The samples were fractionated by online

reversed phase HPLC (YMC 12 nm C18 3µm, 15cm x 300 µm column, 5 µl/min) with a gradient elution of 2-35 % ACN/0.1 % FA over 35 min and 35-80 % ACN/0.1 % FA over 5min. After this, the column washed at 80 % for 5 min prior to re-equilibration. The total run time was 60 min.

2.2.4.1.2 Dataset generation (Chapter IV + V)

The samples for the dataset generation were prepared as described in section 2.2.4.1.1 with the following changes:

For the IDA analysis, the gradient elution was 3-30 % ACN/0.1 % FA over 68 min and 40-80 % ACN/0.1 % FA over 5min. After this, the column was washed at 80 % for 3 min prior to re-equilibration. The total run time was 87 min. For the DIA/SWATH™ analysis, gradient elution was 3-30 % ACN/0.1 % FA over 38 min and 30-40 % ACN/0.1 % FA over 5min. After this, the column was washed at 80 % for 3 min prior to re-equilibration. The total run time was 57 min. Here, the analysis was performed using 100 variable m/z windows, optimised on cell lysate, 25 ms accumulation time and a cycle time of 2.8 secs.

2.2.4.5 Immunohistochemistry of paraffin-embedded tissue sections

Immunohistochemistry was performed for the analysis of the protein expression of selected markers in healthy and diseased tissue sections.
The slides were initially incubated/baked for up to 2 h at 60 °C to ensure and facilitate the complete removal of the paraffin wax and to unmask the antigen epitopes. The slides were then processed according to the following steps.

- Incubation in Xylene (1) for 5 min
- Incubation in Xylene (2) for 5 min
- Incubation in 100 % EtOH (1) for 3 min
- Incubation in 100 % EtOH (2) for 3 min
- Incubation in 70 % EtOH for 3 min

Afterwards, the slides were placed into a bath with running water for 15 min. Citrate buffer was preheated by microwaving it for 10 min, then the slides were slowly dipped into the near-boiling buffer to prevent damage to the slides through the rapid temperature

change. Following this, the slides were microwaved in the citrate buffer for 10 min at maximum intensity. Then the slides were transferred into distilled water, slowly to prevent damage to the slides. The slides were washed in the distilled water for 3 x 5 min, of which after 5 min the water was exchanged. The slides were dried carefully around the tissue sections and placed into a black plastic container. 0.3 % H_2O_2 was added onto of the tissue sections and incubated for 5 min. Following this, the slides were washed for 3 x 5 min in DPBS. In every washing step, the DPBS was exchanged after each 5 min wash cycle.

The slides were dried and 10 % goat serum diluted in DPBS were added onto the tissue sections and incubated for 30 min in the dark. The liquid was tipped off after the incubation and avidin solution was added for 15 min. Following this, the slides were washed for 3 x 5 minutes in DPBS. Then the biotin solution was added to the slides and also incubated for 15 min, then the slides were washed for 3 x 5 minutes in DPBS. In the meantime, the primary antibody was prepared. The concentration optimum of each antibody was defined through dilution series prior the use in relevant tissue sections. The antibody was diluted in 10 % goat serum diluted in DPBS according to the optimised concentration. The primary antibody was added to the sections, after the last wash, then incubated in the dark for 40 min at RT and then transferred to 4 °C overnight.

On the next day, the slides were washed for 3 x 5 min in DPBS, then the secondary antibody was added. The ideal concentration optimised prior final use. The secondary antibody was diluted in a 1.5 % goat serum solution, diluted in DPBS. The secondary antibody was added and the slides incubated for 30 min at RT in the dark. Then the slides were washed for 3 x 5 minutes in DPBS. Following this, ABC buffer was added to the slides and incubated for 30 min at RT in the dark. The slides were washed for 3 x 5 min in DPBS. In the meantime, the DAB reagent was prepared. Here, 2.5 ml dH_20 were mixed with 1 drop of buffer, 2 drops of DAB reagent, and 1 drop of H_2O_2. The DAB reagent was added onto the slides and the slides were observed under the microscope for the development of a staining. The slides were transferred into dH_20 once a sufficient staining was reached. After this, the slides were transferred into running water for 2.5 min and then into dH_20 for 2.5 min. The counterstain and dehydration was performed to the following steps:

- Incubation in Mayer's haematoxylin for 90 secs
- Washing with running water for 1 min
- Incubation in 70 % EtOH for 1 min
- Incubation in 100 % EtOH (2) for 1 min
- Incubation in 100 % EtOH (1) for 2 min
- Incubation in Xylene (2) for 1 min
- Incubation in Xylene (1) for 1 min

Then the slides were left to dry to ensure the complete evaporation of Xylene, before coverslips were fixed in place using DPX.

2.2.4.6 Immunohistochemistry staining of paraffin-embedded prostate cancer tissue sections using fluorescent secondary antibodies

Immunohistochemistry staining on paraffin-embedded tissue microarray slides using fluorescent secondary antibodies was performed in diseased tissue sections.

The slides were initially incubated/baked for up to 2 h at 60 °C. The slides were then processed according to the following steps

.

- Incubation in Xylene (1) for 20 min
- Incubation in Xylene (2) for 20 min
- Incubation in Xylene (3) for 20 min
- Incubation in 100 % EtOH (1) for 3 min
- Incubation in 100 % EtOH (2) for 3 min
- Incubation in 100 % EtOH (3) for 3 min
- Incubation in 70 % EtOH for 3 min

Afterwards, the slides were placed into a bath with running water for 5 min and 3 min in dH_2O. Citrate buffer was preheated by microwaving it for 10 min and the slides were slowly dipped into the near-boiling buffer to prevent damage to the slides through the rapid temperature change. Following this, the slides were microwaved in the citrate buffer for 20 min at maximum intensity. Then the slides were transferred into distilled water, slowly to prevent damage to the slides. The slides were washed in the distilled water for 3 x 2 min, of which after 2 min the water was exchanged. After this, the slides were rinsed with DPBS and placed in a DPBS bath for 3 x 10 min. The slides were dried carefully around the tissue sections and placed into a black plastic container. There, the slides were

blocked against unspecific binding for 1h through the addition of 10 % BSA and 0.1 % Tween20 in DPBS. The liquid was tipped off and the primary antibodies were added, diluted in blocking buffer. The optimal antibody concentration was defined prior use. The samples were incubated in the dark for 60 min at RT and then transferred to 4 °C overnight.

On the next day, the slides were washed for 3 x 10 min in DPBS, then the secondary antibody was added and slides incubated for 1 h at RT in the dark. Then the slides were washed for 3 x 10 minutes in DPBS. The slides were dried, mounting fluid with DAPI was added and covered with a cover slip. The edges were sealed using nail varnish and the slides stored at 4°C until imaging.

2.2.4.7 Immunofluorescence staining of adherent P5B3 and DU145 cells grown on cover slips

Immunofluorescence staining of the cells was performed for the visualisation of the epithelial and mesenchymal marker expression. This was necessary to define the presence/absence of epithelial and mesenchymal cells and for a better characterisation of the cell population. Furthermore, it was also used to confirm the purity of the cell clones according to their phenotypic characteristics and to maintain a high quality standard in the experiments.

The cells were grown for the defined period of time of treatment in flasks and 72 h prior to the staining they were transferred into 24-well plates. The 24-well plate was prior prepared as follows. Cover slips were dipped into 100 % methanol and placed into each well. The cells were added after the complete evaporation of the methanol. This was performed in a hood to ensure sterility was maintained.

After the growth, the media was removed and three wash cycles with DPBS were performed followed by the fixation of the cells with 200 µl of 4 % formaldehyde for 15 min at RT. Next, the formaldehyde was removed and the cells were washed 3x with 250 µl wash solution (100 µl Tween 20 + 100 ml DPBS), which was replaced by 200 µl of blocking solution (10 % ml FCS + 45 ml DPBS + 0.001 % TWEEN 20) and incubated for 60 min at RT. During this time, the primary antibodies were prepared according to manufacturer's recommendation.

After 60 min of incubation, the blocking solution was removed and the primary antibody was added to each well. The plate was then covered in aluminium foil and incubated overnight at 4 °C on a rocker. On the next day, the primary antibodies were removed and each well was washed 3x with 300 µl wash solution for 10 min and placed on a rocker. During this time, the secondary antibody was prepared and after the third washing cycle, 100 µl of it was applied to each well. This step was followed by a 2 h incubation at RT. The plate was covered in aluminium foil and placed on a rocker. Again, the antibody was removed after the 2 h and each well was washed 3x with 300 µl of DPBS for 10 min on a rocker. After the last washing step, the a small drop of mounting fluid with DAPI was placed inside each well and the plate was then stored until imaging at 4 °C.

2.2.5 Experimental layout and generation of sample material from both inducible EMT models

Here, two cell line models of EMT were developed and characterised (Chapter III), and used for the generation of gene and protein expression profiles. To minimise the variation between the sample materials, the cells were grown in parallel and the collection of cell lysates and RNA was performed within 1 hour to minimise protein degradation. Furthermore, each model used media of the same lot number (respectively to cell line) and the TGF-β used was of the same batch. To counteract potential batch effects, samples of both models were generated in two separate treatment rounds, of which half of the sample material of each model was collected in round one, and the other half at round two. Seeded cells remained within the flask during the whole treatment to prevent variation of treatment response based on trypsinisation and potentially induced changes in the expression of the TGF-β receptor. Under normal conditions, cells are seeded at a higher density and are passaged from one flask into a new one in regular intervals, e.g. weekly. However, these passages can influence the cell behaviour. The media changes were performed according to the frequency described in Figure 2.3.

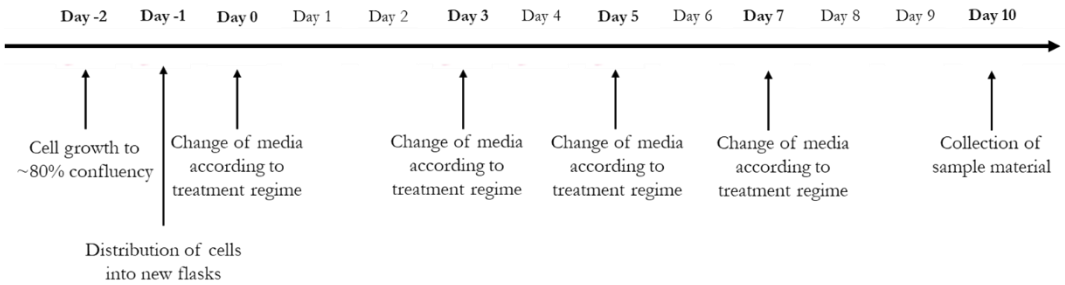


Figure 2.3: Schematic representation of the preparation and treatment regime of both cell lines for the generation of sample material for multi-omics approach and additional experiments conducted for the characterisation of both cell line models.

2.2.6 Processing and filtering of omic data generated through RNA-sequencing and mass spectrometry analyses

2.2.6.1 Processing of RNA-sequencing generated data output

The output generated by DeepSeq were FASTQ-files for each sample was uploaded to Illumina's BaseSpace Sequence platform, where the further processing was performed. The data was aligned using the Homo sapiens (PAR-masked)/hg38 reference genome and the Tuxedo suite with Tophat (Trapnell, Cole, Pachter et al. 2009) and Bowtie 2 (Langmead, Salzberg 2012). In addition, novel transcript assembly was performed. The generated output was presented as FPKM values (Fragments per Kilobase Million) of each gene, which was used for further analysis.

The reference genome was provided by the Genome Research Consortium called "Genome Research Consortium human build 38" (GRCh38). This reference genome, first released in 2013, updated the so far used GRCh19. GRCh19 functioned as a single representation of multiple genomes. GRCh38, however, offers alternate sequences for selected regions, using so-called alternate haplotypes. For this reason, the use of GRChg38 offers a more realistic presentation of the human genome and was therefore selected for the alignment of the generated RNA-sequencing data.

2.2.6.2 Processing of mass spectrometry generated data output

Files generated through IDA analyses of cell lysates (e.g. cytoplasmic, nuclear and membranous fractions of P5B3 and P4B6, as well as a pooled sample of all samples of the dataset generation study) were searched together in Protein Pilot 5.0 with the following parameters; Digestion: Trypsin, Cys Alkylation: Iodoacetamide, ID focus: Biological modifications, Search effort: Thorough ID search effort. The database used was Human Swissprot (Jan 2015). The combined results file (.group) was opened in PeakView 2.1 SWATH microapp (Sciex) and converted to a .txt file. The IDA based library (120 min or 87 min LC run) was aligned to the SWATH data (60 min or 57 min LC run) using the spiked in iRT peptides (HRM kit, Biognosys). The aligned library was extracted from the SWATH data using the OneOmics cloud software suite on the Illumina BaseSpace app platform with the following parameters: 6 peptides per proteins, 6 transitions per peptide, 75 ppm XIC width and a 6 min retention time window. Data was then assembled using the OneOmics Assembler to generate fold change and

confidence data for each protein. Processed data was then downloaded from Illumina BaseSpace as .csv.

2.2.6.3 Filtering of pilot mass spectrometry analysis of P5B3 (Chapter III)

The generated protein expression lists were subjected to a t-test analysis and a p-value of each sample was generated based on a 2-tailed, 2 pairs equal variance analysis. Significant altered proteins <0.05 were selected for further analysis.

2.2.6.4 Initial gene selection and pathway and enrichment analysis (Chapter IV)

The generated p-value of each sample was based on a 2-tailed, 2 pairs equal variance analysis. In addition, the generated p-value was corrected for false discovery using the Bonferroni correction. This correction was used to reduce the likeliness of a type I error (false-positives). In this initial analysis, no filter regarding the fold change was applied.

2.2.6.5 Identification of key marker selection (Chapter V)

The generated p-value of each sample was based on a 2-tailed, 2 pairs equal variance analysis. In addition, the generated p-value was corrected for false discovery using the Bonferroni correction. Furthermore, genes were filtered on a FPKM value of 2 or above in at least one group and an absolute fold change of 2 and above. Protein expression data were filtered for a confidence value of 70 % and above and an absolute fold change of 2 and above (Fig. 2.5).

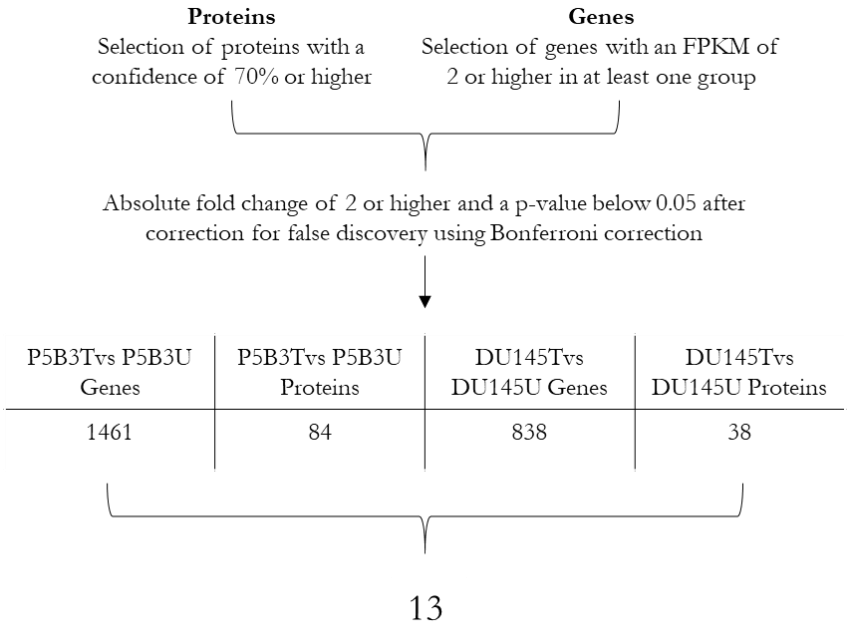


Figure 2.4: Schematic representation of filtering for the identification of 13 core markers

2.2.7 *In silico* analyses of wet-lab derived and publicly available omic datasets

2.2.7.1 Gene Ontology (Chapter III)

The resulting significant proteins (section 3.6.3.1) were subjected to a Gene Ontology analysis. For this, the derived gene list was analysed for the identification of enriched biological processes (http://www.geneontology.org/). Biological processes associated with EMT were selected, affiliated genes selected and presented in a heat map. Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) was used to generate the Venn diagrams of proteins assigned to the different Gene Ontology terms.

2.2.7.2 Heat map Clustering (Chapter III and Chapter IV)

Identified significant markers based on either of the filtering methods described in section 2.2.6.3 and section 2.2.6.4 were applied to MORPHEUS for the generation of a heat map (https://software.broadinstitute.org/morpheus/). Furthermore, the samples of interest were subjected to a clustering approach using Euclidean distance and complete linkage. This clustering approached helped the potential identification of outliers.

2.2.7.3 MetaCore™ analysis (Chapter IV)

A pathway enrichment analysis of significant altered genes and proteins was performed using MetaCore™. The data was uploaded to the cloud-based platform using the gene list and corresponding fold change. A so-called "one click" analysis was performed and enriched pathway maps identified.

2.2.8 Analysis of publically available *in silico* data

Publically available transcriptomic data of cell line and patient-material derived specimens were analysed for further validation. The data was either downloaded from the NCBI Omnibus platform (https://www.ncbi.nlm.nih.gov/geo/), a public accessible repository of generated omic profiles, from CANCERTOOL (Cortazar, A. R., Torrano et al. 2018) (http://web.bioinformatics.cicbiogune.es/CANCERTOOL/), or The Cancer Genome Atlas data portal (https://portal.gdc.cancer.gov/). The origin of the expression data is highlighted in table 2.5. The gene expression of each gene was selected and analysed, either using a t-test, Kaplan-Meier analysis or cox regression analysis. The used method is highlighted in each figure.

Table 2.5: Summary of used public available datasets for the *in silico* validation of novel markers for prostate cancer progression and EMT.

Origin	Sample origin	n[1]	Treatment	Accession
Cell line	A549	3	2 ng/ml TGF-β for 2 weeks	GSE49644[2]
Cell line	HCC287	3	2 ng/ml TGF-β for 2 weeks	
Cell line	NCI-H358	3	2 ng/ml TGF-β for 2 weeks	
Cell line	PANC-1	3	5 ng/ml TGF-β for 5 weeks	GSE23952[2]
Cell line	ARPE-19	3	5 ng/ml TGF-β and 10 ng/ml TNF-α for 60 hours	GSE12548[2]
Patient	benign prostate tissue	28	None	GSE35988[2]
Patient	localised PCa	59	None	
Patient	CPRC/metastasis	35	None	
Patient	Gleason score 6	44	None	TCGA Data Portal[3]
Patient	Gleason score 7	247	None	
Patient	Gleason score 8	64	None	
Patient	Gleason score 9	137	None	
Patient	PCa no recurrence	37	None	Cancertool[4]
Patient	PCa recurrence	42	None	

[1] n = numbers of replicates or patients assigned to the respective group
[2] https://www.ncbi.nlm.nih.gov/geo/
[3] http://web.bioinformatics.cicbiogune.es/CANCERTOOL/
[4] https://portal.gdc.cancer.gov/

2.2.9 Statistical analysis

Error bars describe standard deviation and statistical differences between analysed experimental groups. For this the unpaired t-test was used ($p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = *** and $p \leq 0.0001$ = ****). Figures were generated using GraphPad Prism 7. Univariate cox regression analysis was performed for the identification of predictive capabilities of genes using TIBCO Statistica 13.3. Kaplan-Meier analysis was performed using GraphPad Prism 7. Gene expression values were sorted according to intensity (low to high) and separated into quartiles. Q2 and Q3 were merged based on better separation.

2.2.10 Used online tools and databases

Table 2.6: List of utilised databases and online tools with their use and link

Database/ Online tool	Use	Link
BaseSpace	Processing of RNA-sequencing data	https://basespace.illumina.com/
OneOmics	Processing of Mass spectrometry data	https://sciex.com/applications/life-science-research/oneomics
MetaCore	Pathway enrichment analysis	https://portal.genego.com/
GeneOntology	Identification of enriched biological processes	www.geneontology.org/
NCBI Omnibus	Database for omics-derived datasets	https://www.ncbi.nlm.nih.gov/geo/
TCGA	Database of cancer datasets	https://portal.gdc.cancer.gov/
CANCERTOOL	Online platform for cancer data analysis and source of omics data	http://web.bioinformatics.cicbiogune.es/CANCERTOOL/
MORPHEUS	Generation of heat maps and clustering	https://software.broadinstitute.org/morpheus/

3. Chapter III – Development of two inducible models of epithelial to mesenchymal transition for the study of disease progression in prostate cancer

3.1 Introduction

Prostate cancer (PCa) is the most common cancer in men in Europe and the second most common cause of cancer-related deaths in the United Kingdom (Cancer Research UK, 2017a), primarily due to the development of metastasis. The development of metastasis decreases the 5-year survival rate to only 30 % (Thobe, Clark et al. 2011). At some point during development and growth, cells of the primary tumour gain the ability to spread to distant organs. These characteristics can be acquired through the process of epithelial to mesenchymal transition (Heerboth, Housman et al. 2015). During this, cells change from an epithelial to a mesenchymal-like cell state, invade the blood or lymphatic system, and are distributed throughout the body. At a distant site they undergo mesenchymal to epithelial transition (MET) and initiate the growth of secondary tumours (Kalluri, Weinberg 2009). Epithelial cells are firmly attached cells that are growing in clusters. They are tightly connected by different types of cell junctions, such as tight, gap and adherens junctions, as well as desmosomes. Furthermore, they have an apico-basal polarity. Mesenchymal cells cannot form connective cell layers and they only focally connect to surrounding mesenchymal cells. In culture, mesenchymal cells show a fibroblastic, spindle shape morphology (Thiery, Sleeman 2006), whereas epithelial cells commonly present polygonal shapes, building patches of attached cells.

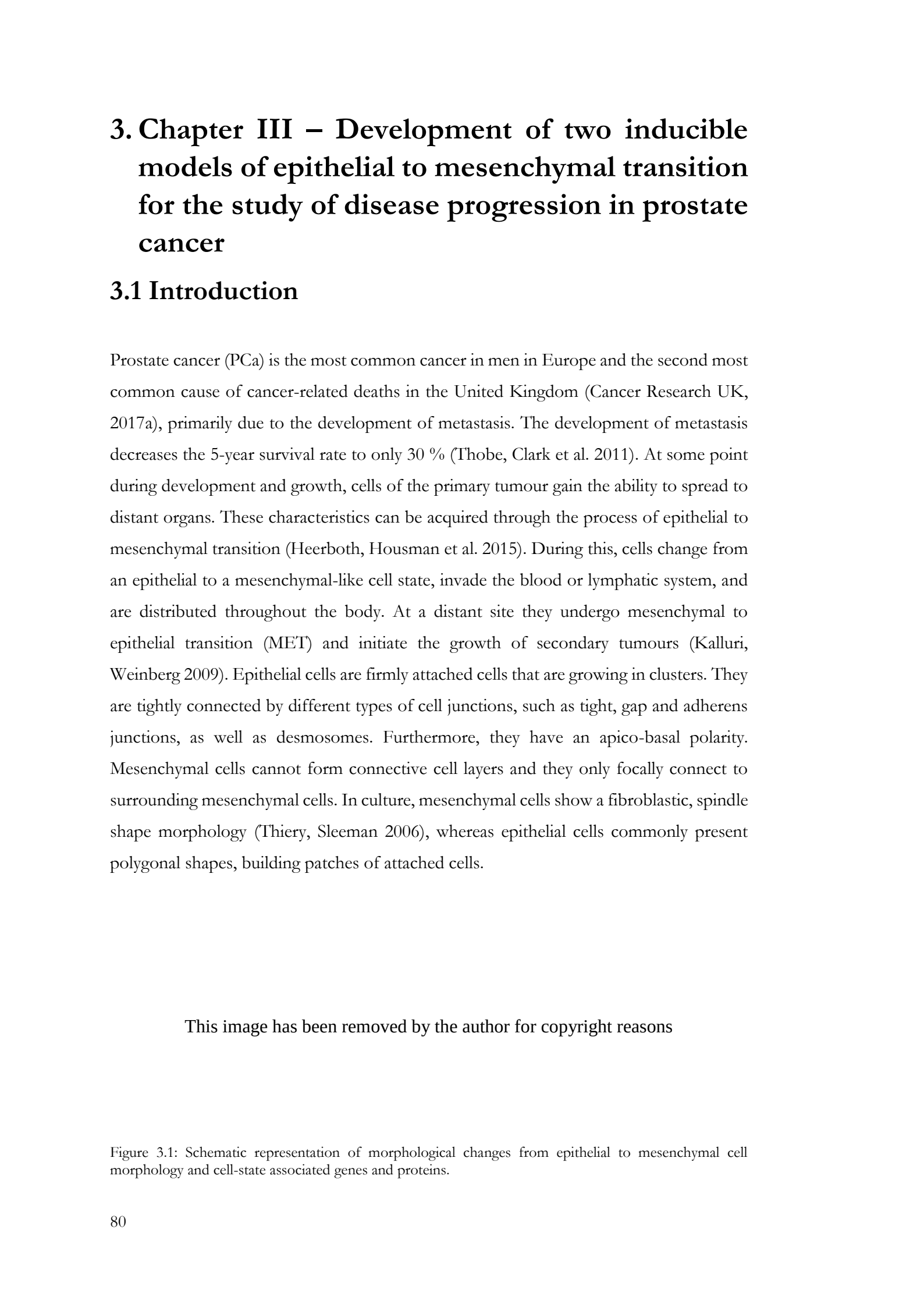


Figure 3.1: Schematic representation of morphological changes from epithelial to mesenchymal cell morphology and cell-state associated genes and proteins.

Certain genes, such as E-cadherin (*CHD1)*, N-cadherin (*CDH2)* and Vimentin (*VIM)* are genes indicative for this process; however, their suitability in clinical use is limited. This is based, on the one hand, on the fact that EMT is a natural process occurring during healthy biological processes, such as wound healing, and on the other hand that the process of EMT can be reversed on cells therefore they do not continuously express mesenchymal (EMT) markers. In addition to this, studies have presented variable results regarding the associated significance of EMT marker expression with survival and disease progression in prostate cancer (Nauseef, Henry 2011). For this reason, the discovery of novel markers indicative for the process of EMT and disease progression is crucial. Such markers could present a potential strategy for routine screening, cancer surveillance and treatment response, as well as potential treatment targets for the suppression and inhibition of cancer spread.

For the discovery and study of novel markers, cell line models present a useful tool for the simulation of EMT *in vitro*. Cell line material can be easily genetically modified and treated with various reagents; furthermore, they present nearly limitless availability. However, many of these cell line models are based on cell lines of metastatic origin, such as MDA-MB-468 (breast cancer) (Bonnomet, Syne et al. 2012), NCI-H358 (lung cancer) (Argast, Krueger et al. 2011) and ARCaP (prostate cancer) (Zhau, Odero-Marah et al. 2008). In this chapter, the development of two inducible cell line models is described, one of which was generated from a primary tumour cell line.

The first model was generated from a single cell clone (P5B3) derived from the parental cell line OPCT-1 (Harner-Foreman, Vadakekolathu et al. 2017). The parental cell line was generated from a primary prostate tumour epithelium and was staged as T1cN0M0 and as Gleason 3+3. Harner-Foreman et al, generated multiple single cell clones from the parental cell line and characterised them with regards to their invasive and metastatic potential and their EMT profile.

P5B3 presents a highly epithelial morphology (Fig. 3.2A+B) and a low EMT profile with a high expression of the epithelial cell marker *CDH1* and less than 0.5 % of *VIM*-positive cells. Furthermore, cells of this single cell clone did not exhibit invasive capabilities and were unable to initiate tumour growth once implanted into mice (Harner-Foreman, Vadakekolathu et al. 2017).

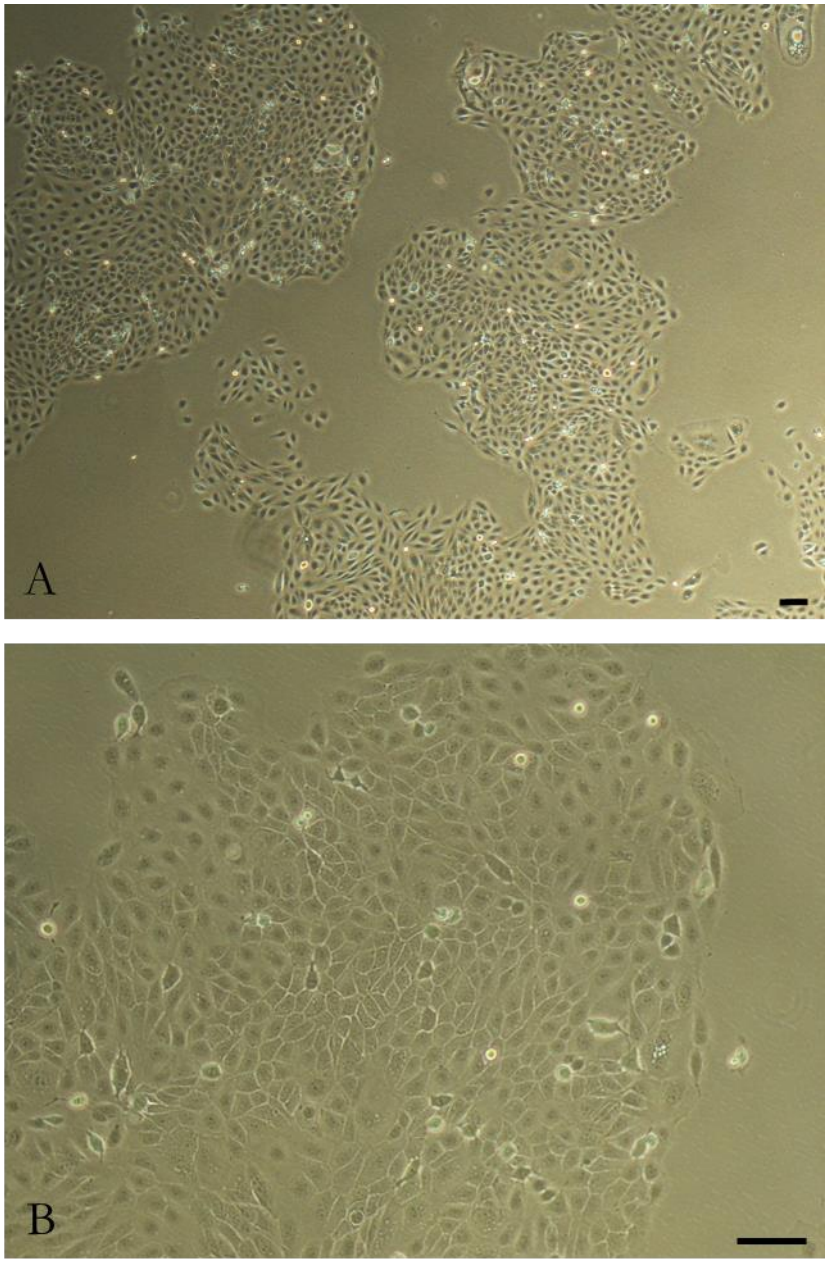


Figure 3.2: P5B3 in its natural state showing a highly epithelial cell morphology with a small proportion of single cells presented at a 4x (A) and 10x (B) magnification The scale bar shows a length of 10 μm

The second model of EMT was generated using the metastatic prostate cancer cell line DU145. DU145 is a commonly-used and well-studied prostate cancer cell line derived from the metastatic tumour site located in the brain (Stone, K. R., Mickey et al. 1978).

The patient from whom DU145 was derived presented a poorly differentiated prostate adenocarcinoma exhibiting metastatic lesions in the central nervous system, liver, lungs and brain (Stone, K. R., Mickey et al. 1978). DU145 is an adherent epithelial cell line (Fig. 3.3A+B), which is hormone-insensitive and hormone-independent (not required for growth) and does not express the prostate specific antigen. Cells injected into nude mice were able to induce tumour growth (ATCC 2018a). The morphology of DU145 can be described as predominantly epithelial. Despite this, it presents a less distinct cobblestone morphology when compared to P5B3.

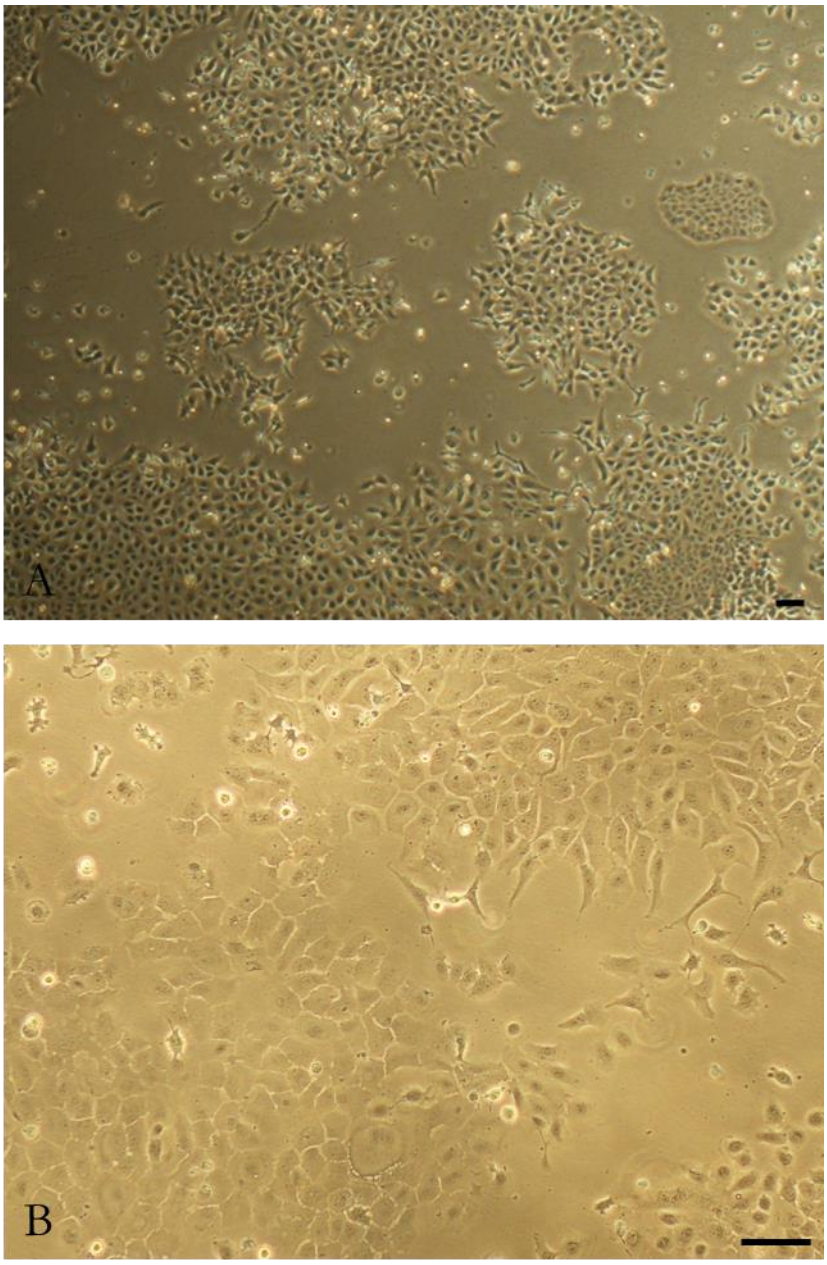


Figure 3.3: DU145 in its natural state showing a highly epithelial cell morphology with a small proportion of single cells presented at a 4x (A) and 10x (B) magnification. The scale bar shows a length of 10 µm.

The aims of this chapter are the development of two inducible *in vitro* models of EMT. The process of the model development will be characterised by multiple checkpoints to be successfully completed prior further proceeding of the experiments.

- Treatment of the cell lines with TGF-β and the observation of morphological changes associated with the induction of a more mesenchymal cell state, highlighted through the development of elongated, potentially solitary, cells.
- Validation of transcriptomic changes associated with a mesenchymal cell state through the use of quantitative real-time PCR analysing most commonly used genes associated with EMT. An increased expression of *CDH2*, *FN1*, *VIM*, *TWIST1*, *SNAI1*, *SNAI2* and *ZEB1* and a reduced expression of *CDH1* will indicate the successful induction of an EMT-like phenotype.
- Validation of proteomic changes to a mesenchymal cell state through the use of Western blot and immunofluorescence analysis for the proteins CADH1, CADH2, FN1 and VIME. Also here, an increased expression of CADH2, VIME and FINC, and a decreased expression of CADH1, supports an induction of an EMT-like phenotype.
- Analysis of potential changes in the invasive behaviour of both cell line models through the performance of scratch assays. The experiment will elucidate potential changes in the cell line behaviour induced through the stimulation with EMT. The induction of EMT can contribute to a more invasive behaviour.

Overall, the successful confirmation of gene and protein expression associated with an induction of EMT will enable the use of both models for the generation of matching transcriptomic and proteomic profiles for the use of data-integration and potential discovery of novel disease-associated biomarkers.

3.2 Results

3.2.1 Development of an inducible model of EMT using a single cell clone derived from a primary prostate cancer cell line using Transforming Growth Factor β

3.2.1.1 Morphological changes induced in P5B3 through the treatment with 10 ng/ml TGF-β for 5 days

Transforming growth factor β, which is a known inducer of EMT *in vitro*, was selected for this study. The untreated cells of P5B3 present a "cobblestone" morphology of epithelial cells tightly attached to each other and the flask surface. Initially, the cells were treated with 10 ng/ml TGF-β for 5 days, which induced morphological changes compared to the untreated control (Fig. 3.4), showing a change from connected island of cells to dispersed elongated cells. The cells developed an elongated cell shape and isolation from surrounding cells. Furthermore, their adherence to the flask surface was reduced. The untreated P5B3 cells did not show any changes in their morphology nor their adherence to the cell culture flask after 5 consecutive days of growth.

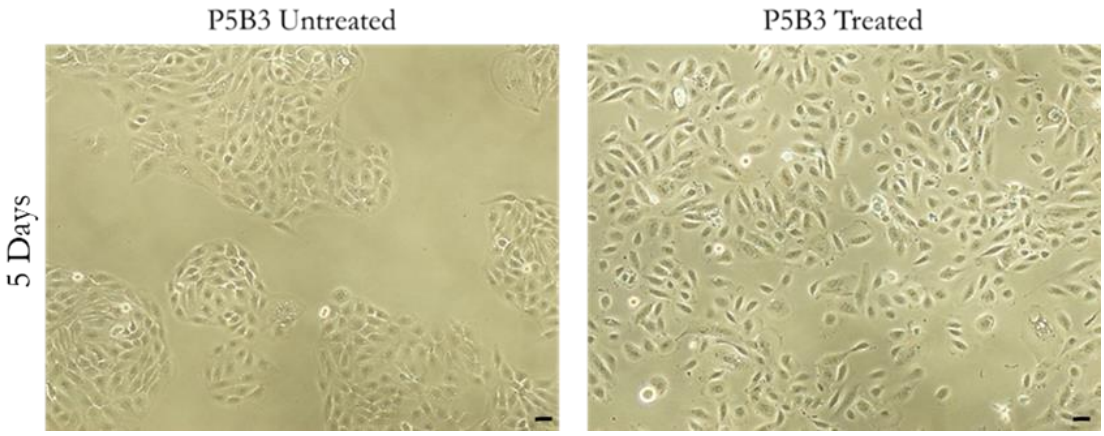


Figure 3.4: Morphological changes of P5B3 after treatment with TGF-β for 5 days with 10 ng/ml TGF-β. The scale bars indicate 10 μm.

3.2.1.2 Gene expression changes induced in P5B3 through the treatment with 10 ng/ml TGF-β for 5 days

After morphological changes were observed through the treatment with TGF-β, the cells were screened for potential changes in the molecular EMT profile. For this, extracted RNA of both conditions was analysed for the following genes: *VIM, CDH1, CDH2, FN1*, and the EMT-Transcription factors (EMT-TFs) *SNAI1*, *SNAI2*, *TWIST1* and *ZEB1*

using quantitative real-time PCR. Figure 3.5 demonstrates the expression changes of these genes at the mRNA level in treated cells compared to the natural P5B3 profile. It could be shown that the treatment induced an expression of the analysed markers associated with a mesenchymal cell state, whereas the epithelial associated gene, *CDH1*, showed a decreased expression. Of all the analysed genes, *VIM* showed the strongest increase with about 1000 times the expression compared to the untreated cells. The other mesenchymal associated genes, *CDH2* and *FN1*, showed the second and third strongest upregulation, respectively. Additionally, the EMT-TFs all showed an increase in their expression, of which *ZEB1* was showing the strongest fold change increase, induced through the treatment. Based on the detected molecular changes indicating morphological changes to an increased mesenchymal phenotype and subsequent induction of EMT, this cell line model was selected for further characterisation and analysis.

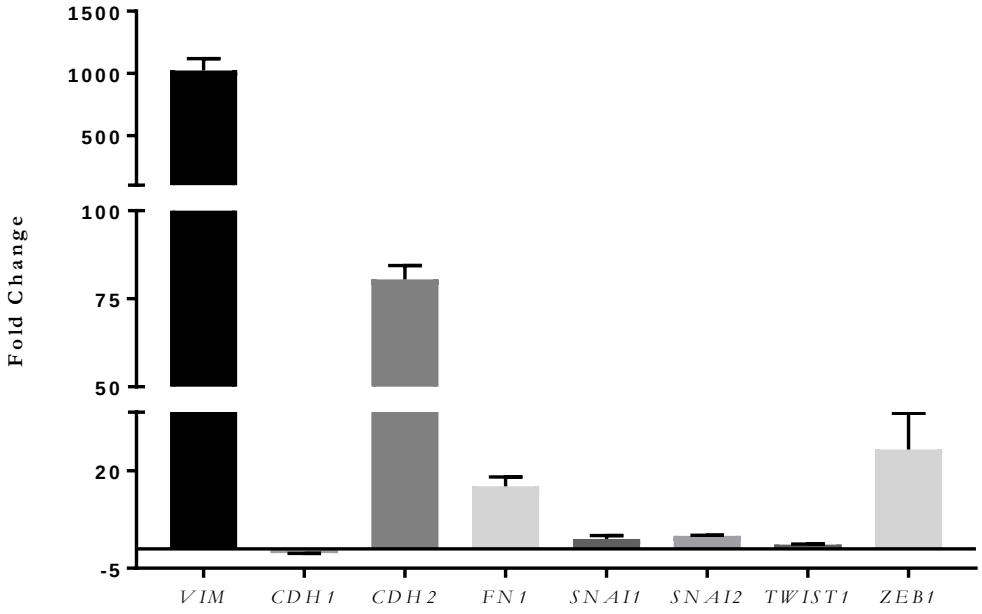


Figure 3.5: Gene expression changes of EMT markers induced in P5B3 upon stimulation with TGF-β. The expression of VIM, CDH1, CDH2, FN1, SNAI1, SNAI2, TWIST1 and ZEB1 was compared between untreated and treated P5B3 cells after incubation with 10 ng/ml TGF-β for 5 days. Results were analysed using the comparative ΔΔCT method (Schmittgen, Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

3.2.1.3 Protein expression changes induced in P5B3 through the treatment with 10 ng/ml TGF-β for 5 days

The molecular changes induced through the treatment of P5B3 indicate the induction of EMT, however, based on the potential variations between gene and protein expression, additional analysis of EMT-associated proteins was performed using immunofluorecense staining. The staining (Fig. 3.6) has shown that P5B3 untreated has a strong expression of CADH1, located at the cell membranes of the cells, whereas no expression of VIME and only low, dispersed FINC expression was detectable at an untreated condition. Upon treatment the expression of CADH1 was strongly reduced and the expression of VIME and FINC strongly increased. The expression of VIME was detected in the cytoplasm, where it comprises, together with the microtubules and microfilaments, the cytoskeleton. Also the expression of FINC was localised in the cytoplasm of the cell. This shows a confirmation of the previously measured molecular changes. The analysis of the EMT-associated proteins confirmed previous findings of the altered gene expression of *CDH1*, *VIM* and *FN1* upon stimulation with TGF-β (Fig. 3.5).

3.2.1.4 Quantitative mass spectrometry analysis of untreated and treated P5B3 cell extracts using 10 ng/ml TGF-β for 5 days

In order to investigate proteomic changes through the stimulation with TGF-β, 25 ug of total protein of each growth condition (n=3) was used and label-free quantitative proteomics was performed on the complete cell lysate. The generated library based on all samples contained 1308 different proteins using a 1 % FDR cut-off. Within this library, only 3 EMT markers, CADH1, VIM and FINC, were identified. The comparison of the protein peak areas of treated and untreated samples have shown significant changes in the expression of VIME and FINC, whereas the decrease in the expression of CADH1 was detected, however this decrease did not present a significant difference (Fig. 3.7).

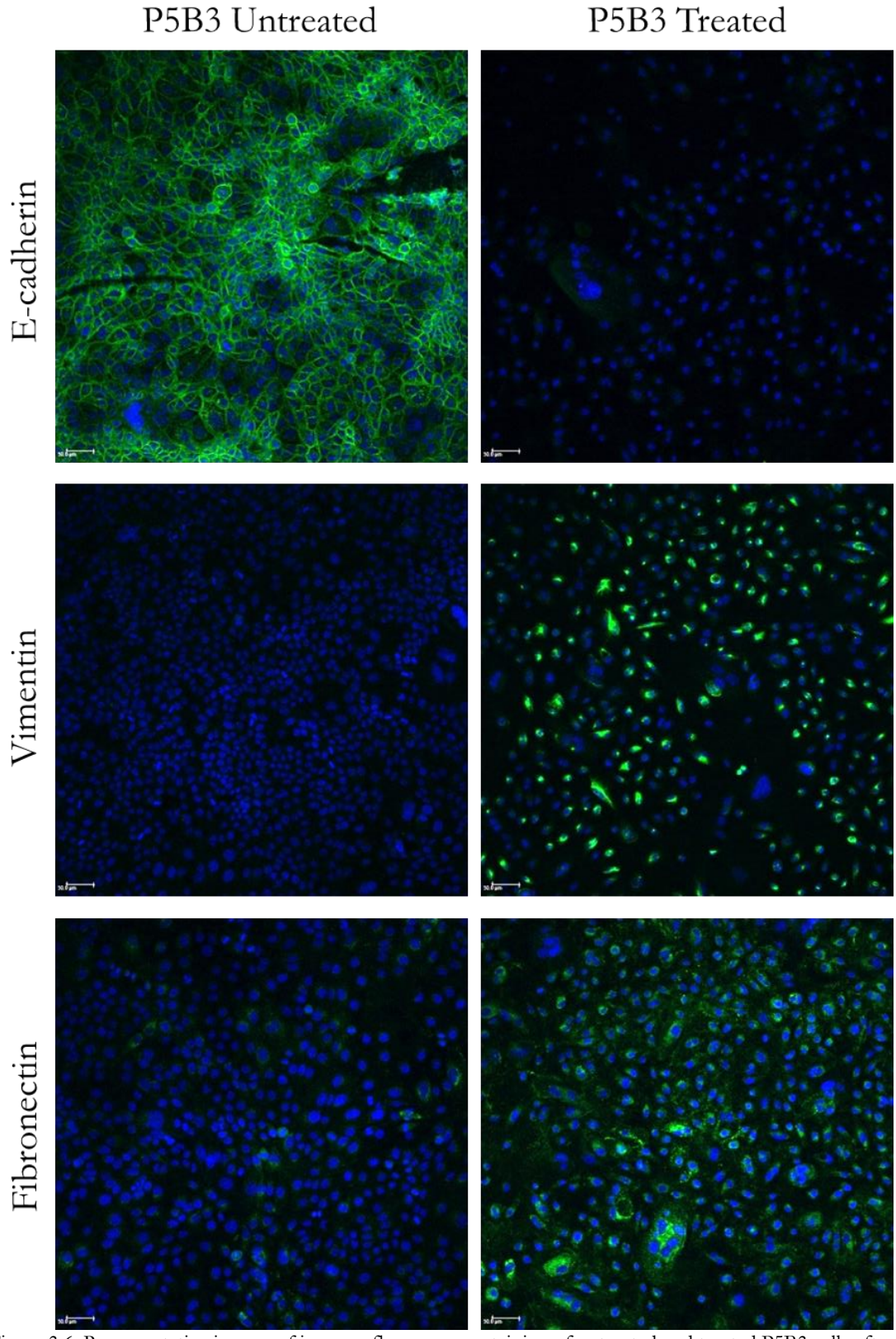


Figure 3.6: Representative images of immunofluorescence staining of untreated and treated P5B3 cells after incubation with 10 ng/ml TGF-β for 5 days. The cells were stained for the mesenchymal marker Fibronectin and Vimentin, as well as the epithelial marker E-cadherin. Staining with DAPI is presented as blue and FITC staining represents staining with the marker of interest. The scale bar shows a length of 50 μm

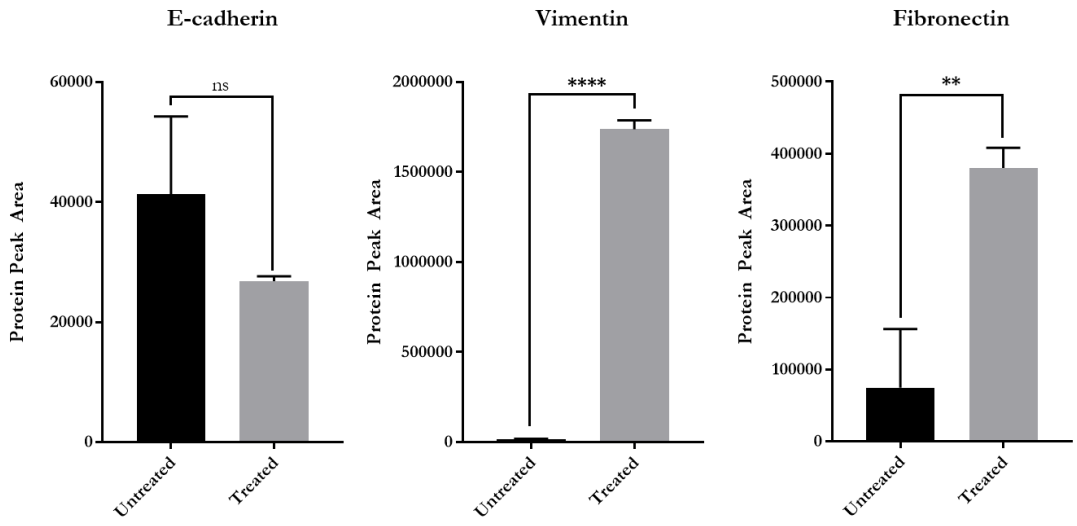


Figure 3.7: Comparison of protein peak areas of E-cadherin (CADH1), Vimentin (VIME) and Fibronectin (FINC) for untreated and treated cells of P5B3 using quantitative mass spectrometry analysis (n=3).

For the further analysis, the list of 1308 proteins was reduced through the application of a significance cut-off of 0.05. 365 proteins showed significant differences between the untreated and treated sample groups. Of these 365 proteins, 195 were additionally showing an absolute change of expression of at least 1.5 fold. These 195 proteins were applied to an enrichment analysis using the enrichment tools supplied by the Gene Ontology Consortium (http://www.geneontology.org/ (Accessed 15.03.18). All together, 71 unique proteins were assigned to Gene Ontology terms widely associated with metastasis (Fig. 3.8A). 33 of these were assigned to "cell adhesion", 24 to "cell migration" and 55 to "tissue development". Figure 3.8B presents the numbers of unique and shared genes of each of the three selected Gene Ontology terms. Furthermore, their expression directionality and their assigned categories are represented in a heat map (Fig. 3.8C).

The terms "cell adhesion", "cell migration" and "tissue development" were selected due to their involvement in the process of EMT. The analysis using Gene Ontology indicated a succesful alteration of epithelial cells into an increased mesenchymal cell state. An example for protein changes in accordance with the induction EMT are the upregulation of migratory proteins, such as ANXA3 (Annexin 3), and ITAV (Integrin Subunit Alpha V) and the reduced expression of cytoskeletal proteins, such as KRT19 (Keratin 19).

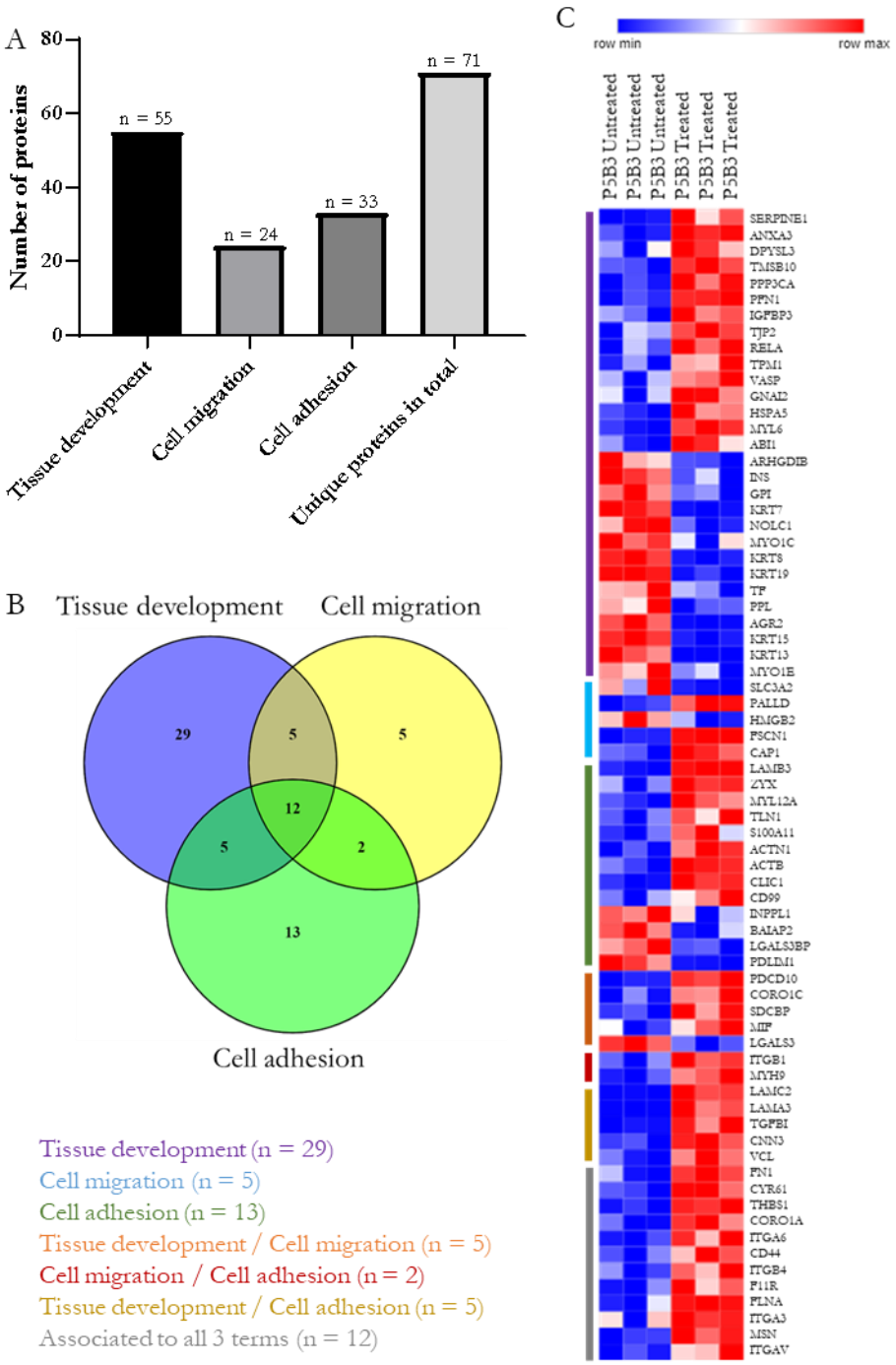


Figure 3.8: Analysis of significant proteins (<0.05) with an absolute fold change of 1.5 and higher using Gene Ontology Consortium (http://www.geneontology.org). 71 proteins were assigned to the biological terms of "cell adhesion", "cell migration" and "tissue development" (A). 12 proteins were detected in all 3 terms (B). The heat map (C) indicates the expression of each protein and the assignment of the proteins to each term. Blue = reduced expression, red = increased expression. The colour coding at the side of the heat map highlights the assigned group (tissue development = purple, cell migration = blue, cell adhesion = green, shared tissue development/cell migration = orange, shared cell migration and cell adhesion = red, shared tissue development/cell adhesion = yellow and detected in all 3 terms = grey).

3.2.1.5 Time-point optimisation of treatment length with TGF-β of P5B3 through the analysis of morphological and gene expression changes

The initial results strongly support the use of this cell line model for the discovery of novel biomarkers associated with the process of disease progression in prostate cancer. Based on this, a time point optimisation experiment was performed in which the length of the treatment was optimised and selected. The treatment length was limited to 10 days, based on the minimal required seeding density of the P5B3 cells for healthy cell growth. For the definition of an optimal time point regarding the successful induction of EMT, morphological observations and molecular changes were analysed using bright field microscopy and qRT-PCR on EMT genes and EMT-TFs.

3.2.1.5.1 Morphological changes in P5B3 over time when treated with 10 ng/ml TGF-β

Cells of P5B3 were treated consecutively for 3, 5, 7 and 10 days with 10 ng/ml TGF-β. During this time, the cells were not passaged and kept in one flask throughout the duration of the experiment. This was done to ensure the uninterrupted treatment with TGF-β. Prior to this, a minimum seeding density was defined as 50 000 cells per T175 flask to ensure the healthy growth of the cells (data not shown).

During the time point experiment, the media was changed every second day in both conditions, untreated and treated. The treated media was supplemented with 10 ng/ml TGF-β in each media exchange. The morphological changes in P5B3 across the time points are shown in Figure 3.9. It can be seen that untreated P5B3 do not alter their morphology throughout the growth on tissue culture plastic for 10 consecutive days. Furthermore, the stimulation of P5B3 with TGF-β led, after 3 days, to morphologically visible changes, which increased throughout the stimulation, showing the clearest difference between treated and untreated cells at day 10 (Fig. 3.9). The treated cells have developed an increased elongated cell shape and have shown a separation from the neighbouring cells, whereas the untreated cells retained the "cobblestone" morphology (Fig. 3.9).

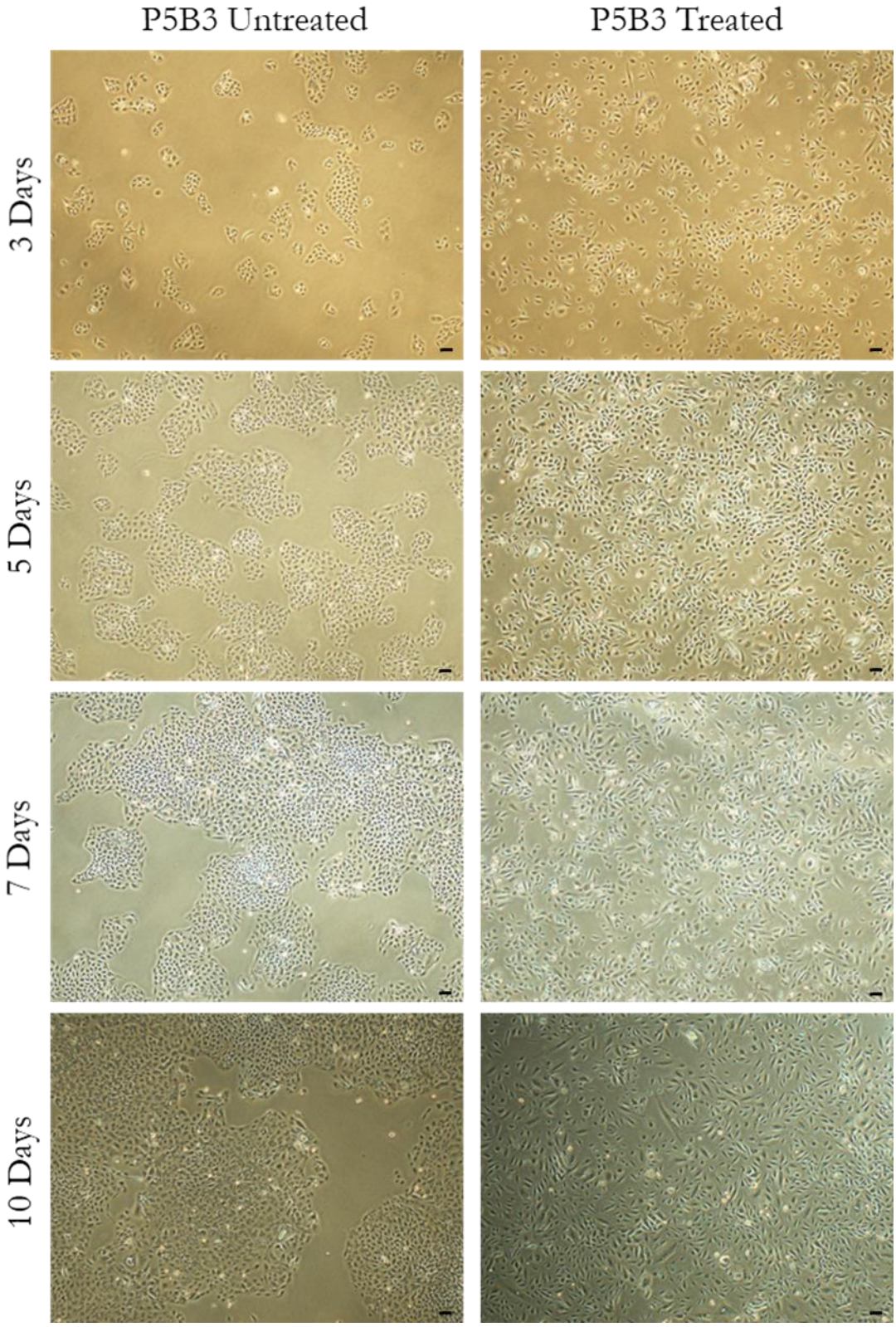


Figure 3.9: Morphological appearance of untreated and treated cells of P5B3 after growth over 10 days. Brightfield images were taken at the timepoints of 3, 5, 7 and 10 days at a 4x magnification. The scale bar indicates 10 μm.

3.2.1.5.2 Gene expression changes in P5B3 over time when treated with 10 ng/ml TGF-β

In addition to the morphological changes observed across the 4 time points, analysis of the gene expression changes of the previously analysed EMT markers; *VIM*, *CDH1*, *CDH2*, *FN1*, *SNAI1*, *SNAI2*, *TWIST1* and *ZEB1*, was performed across the time points of 3, 5, 7 and 10 days (Fig. 3.10). The gene expression changes based on the induced fold change of the days 5, 7 and 10 were compared to the fold change induced after the stimulation for 3 days. Vimentin showed an upregulation after 3 days, however a significant stronger increase could be observed after 5 days of stimulation. The vimentin expression at time point 7 still presents a significant increase compared to the time point of 3 days, however less intense when compared to 5 days (Fig. 3.10A). *CDH1*, the only marker that shows a reduction in its expression is slightly downregulated at the time points 3, 5 and 10, presenting a similar reduction in their expression without any significant differences. The decrease at the time point 7 days presented the strongest and only significant decrease (Fig. 3.10B). *CDH2* (Fig. 3.10C), *FN1* (Fig. 3.10D), *SNAI2* (Fig. 3.10E) and *ZEB1* (Fig. 3.10H) have shown a steady increase in the induced gene expression fold change from day 3 to day 10. *CDH2* and *SNAI2* have presented the strongest fold change increase at day 10, with a more than 150-fold and 6-fold increase in its expression for *CDH2* and *SNAI2*, respectively. The overall analysis highlighted a consistent increase of *CDH2* from time point to time point (Fig. 3.10C). Despite the consistent increase of the *FN1* expression, no significant differences were detected compared to day 3. It highlights a consistent upregulation of *FN1* throughout the length of the stimulation. The expression of the EMT-TF *SNAI1* showed strong variation for the days 5 and 7, and therefore only day 10 presented significant increased expression compared to the first induction at day 3 (Fig. 3.10D). *ZEB1* presented a consistent 20-fold change across the time point 3, 5 and 7 days and sharply rose to a significant fold change of 60 at day 10. The EMT-TF *TWIST1* showed an upregulation of its expression throughout the treatment, with a plateau over 5, 7 and 10 days, however, their overall expression was very low and, for this reason, the fold change analysis showed large variations across repeats and limited significance in their changes was observed (Fig. 3.10G). Significant differences were detected when comparing expressions of 5 and 7 days to 3 days.

Based on the results of this analysis, in which 50 % of the markers showed their strongest upregulation at the time point of 10 days, and the clear morphological changes observed

at this time point, 10 days of treatment were selected for the further experimental procedures.

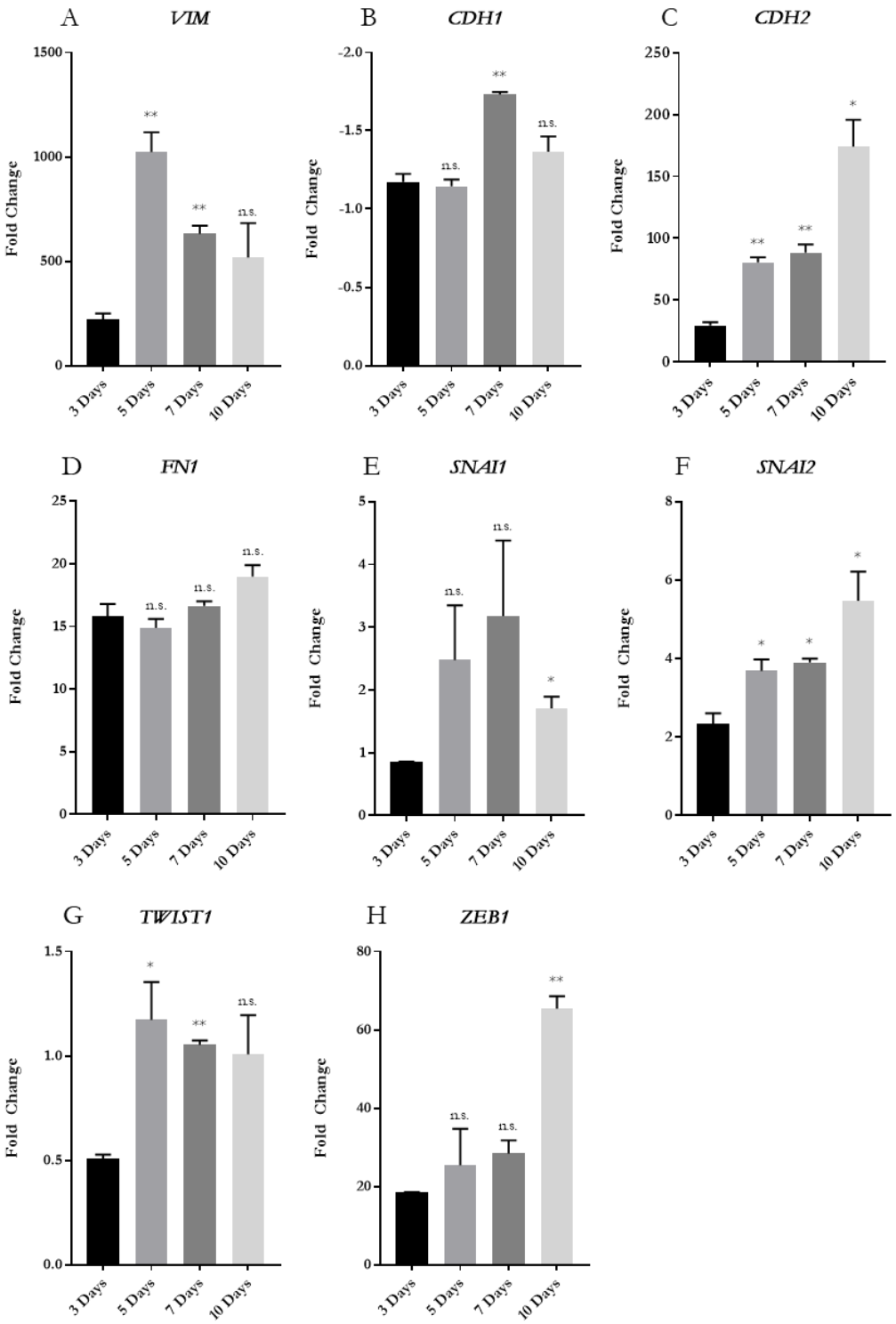


Figure 3.10: Gene expression changes of known EMT markers induced in P5B3 upon stimulation with TGF-β . The gene expression of Vimentin (A), E-cadherin (B), N-cadherin (C), Fibronectin (D), and the EMT-TFs Snail (E), Slug (F), Twist (G) and ZEB1 (H) was measured across four different time points using quantitative real-time PCR and $2^{-\Delta\Delta CT}$ method (Schmittgen, Livak 2008) (n=4). The significance analysis was performed comparing the fold change of days 5, 7 and 10 with the fold change difference of each gene induced after treatment for 3 days. The gene expression was normalised against the TATA-box protein (TBP) gene, which was utilised as reference gene.

3.2.1.6 Protein expression changes in P5B3 after treatment with 10 ng/ml TGF-β for 10 days using Immunofluorescence staining

To further confirm the changes of expression at time point 10 days, immunofluorescence staining of cells was performed. Cells of both treatment conditions were grown on coverslips placed inside 24-well plates and the immunofluorescence staining was performed inside each well. The staining was performed in triplicate across three separate experiments and representative results are shown in Figure 3.11.

Visible changes in their protein expression were detected for all markers based on the comparison of untreated and treated P5B3 cells. Untreated P5B3 cells did not show VIME expression, and the treatment with TGF-β resulted in an increased expression of VIME visible in the cytoplasm. On the other hand, untreated P5B3 cells have shown a strong expression of the epithelial cell marker CADH1 in the membranes of the cells, which was strongly reduced upon treatment. However, a low protein expression remained detectable in the treated cells, indicating a reduction in the protein expression, but not a complete loss. The third studied marker, FINC, could be detected in single, untreated cells, but a strong increase in its expression was observed through the treatment, resulting in its expression in the majority of the cells (Fig. 3.11). In both conditions, the expression of FINC was associated with the cytoplasm.

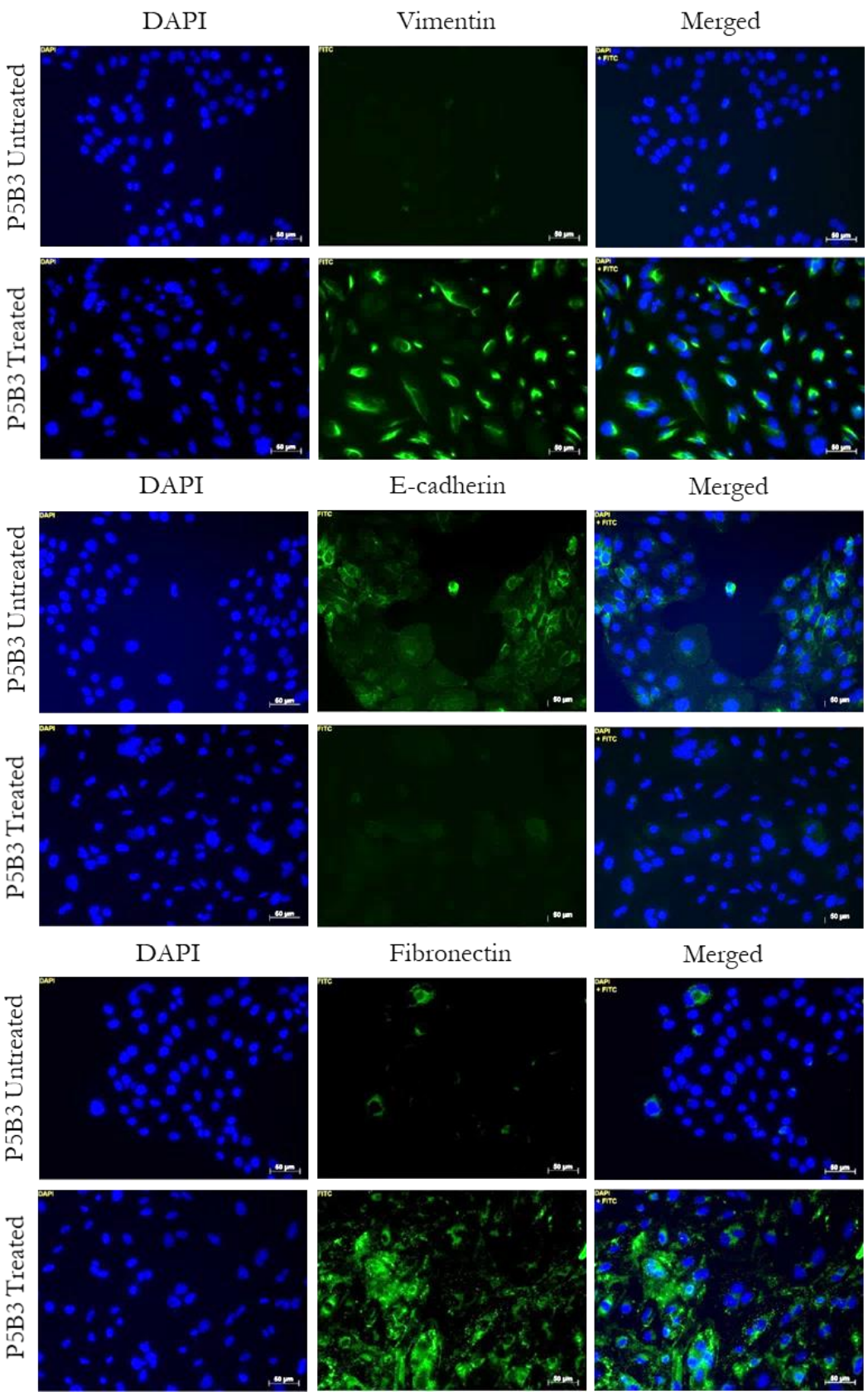


Figure 3.11: Representative images of immunofluorescence staining of untreated and treated P5B3 cells after treatment for 10 days with 10 ng/ml TGF-β showing the EMT marker E-cadherin, vimentin and fibronectin. Staining with DAPI is presented as blue and FITC staining represents staining with marker of interest. The scale bar shows a length of 50μm.

3.2.1.7 Protein expression changes in P5B3 after treatment with 10 ng/ml TGF-β for 10 days using Western blot analysis

As an additional validation of the protein changes induced through the treatment with TGF-β, cell lysates of P5B3 cells in the uninduced and induced cell states were collected and analysed using Western blot analysis (Fig. 3.12). The markers analysed were FINC, VIME, CADH1 and CADH2. Commonly used loading controls are actins or tubulins, which are highly associated with the cytoskeleton. However, during the process of EMT, the cytoskeleton is strongly influenced. To counteract potential bias through this, Cyclophilin A was selected as loading control. In the analysis of the generated sample material, the expression of FINC was consistently upregulated in both biological repeats of the treated compared to the untreated samples. The same was shown for VIME. A reduction in the expression of CADH1 could also be observed, however the intensity of reduction varied between the samples. CADH2 was shown to be upregulated in the treated sample 1 and to a smaller extent also in the treated sample 2, and was not detectable in both untreated samples. Across all samples, the loading control showed a consistent intensity.

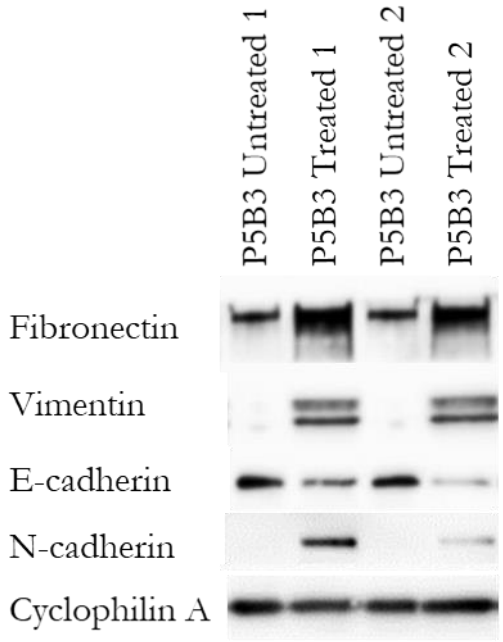


Figure 3.12: Western blot of cell lysates generated from untreated and treated P5B3 cells. Protein analysis of the EMT markers Fibronectin, Vimentin, E-cadherin and N-cadherin. Cyclophilin A was used as loading control. 50 µg of protein was loaded for each sample.

3.2.2 Development of a second inducible model of EMT using a prostate cancer cell line derived from a metastatic site through stimulation with Transforming Growth Factor β

As previoulsy described in section 3.2.1, EMT was sucessfully induced in the single cell clone P5B3. This cell clone was derived from the cell line OPCT-1, which is not as well studied as other prostate cancer cell lines. For this reason, and for the increase in robustness of novel defined biomarkers, it was decided to generate a further cell line model of inducible EMT using the well-studied metastatic prostate cancer cell line DU145.

3.2.2.1 Time point optimisation of treatment length of DU145 through the analysis of morphological and gene expression changes.

3.2.2.1.1 Morphological changes in DU145 over time when treated with 10 ng/ml TGF-β

The treatment of DU145 with TGF-β resulted in a seeminlgy slower reponse over the 10 days treatment compared to the response observed in P5B3, based on the morphological observations over the time course experiment. The first clear changes could be observed after 5 days (Fig. 3.13), whereas P5B3 showed visible changes after 3 days of stimulation. Furthermore, the morphological changes indicated a response of a subpopulation of cells to the stimulation with TGF-β (green cirle. Fig. 3.13), whereas the remaining cells seemed to remain unaffected (red circle, Fig. 3.13)

Also in this cell line, the strongest mesenchymal-like cell morphology could be observed after 10 days, however the morphology varies from P5B3 (Fig. 3.9). DU145 does not develop single cells, but presents grouped islands of elongated cells next to groups of epithelial-like cells. These morphological changes were not observed in the untreated DU145 throughout the growth for 10 consecutive days in the same tissue culture flask.

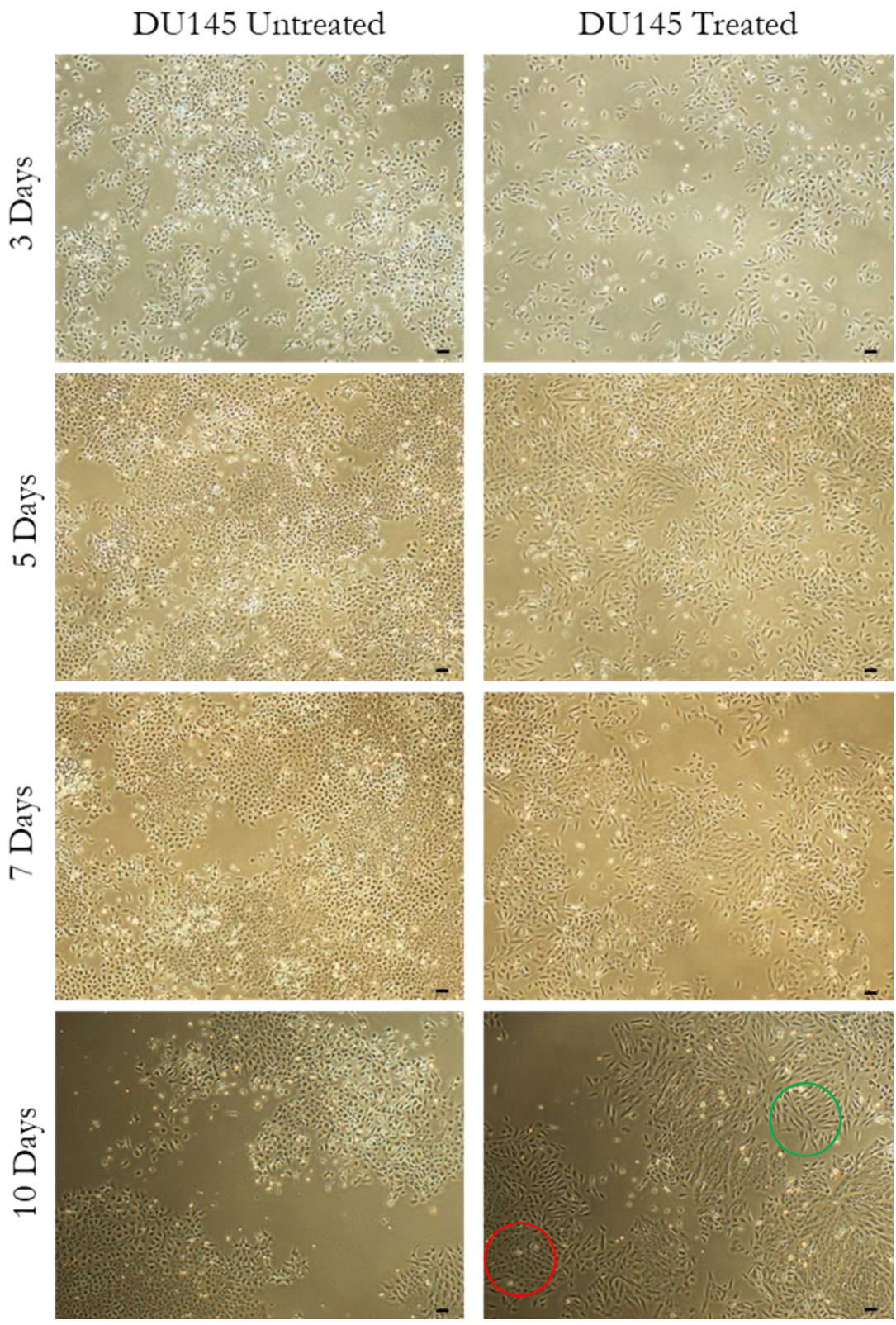


Figure: 3.13: Morphological appearance of untreated and treated cells of DU145 after growth over 10 days. Images were taken at the timepoints of 3, 5, 7 and 10 days at a 4x magnification. The scale bar indicates 10 μm.

3.2.2.1.2 Gene expression changes in DU145 over time when treated with 10 ng/ml TGF-β

The gene expression profiles of EMT markers were studied to further support the observed morphological changes. It has to be noted that no dectection of *CDH2* (Fig. 3.14C) was possible, either in the untreated or treated cell state because the detected Ct values were equal to the performed negative controls or did not cross the defined threshold.

Vimentin showed a 2-fold increase in expression after treatment for 3 days, which further increased across 5 and 7 days. These increases were significantly higher compared to 3 days. After this, the expression of VIM reduced by about 1-fold, presenting a slightly lower fold change compared to 3 days. Despite the small differences between these 2 days, the difference is significant (Fig. 3.14A).

The strongest change of *CDH1* expression was observed after 3 days, with a fold change reduction of about 15-fold. This strength of reduction was highly significantly increased, compared to the redution at the time points of 5, 7 and 10 days, which plateau at a fold change level of about 2 across all 3 time points (Fig. 3.14B). Fibronectin showed no change in expression after treatment for 3 days, however a strong and significant induction after 5, 7 and 10 days, compared to the change after 3 days, was observed (Fig. 3.14D). The EMT-TFs *SNAI1* and *SNAI2* showed a clear increase in their expression, presenting across all time points a fold change increase of about 4, both presenting a peak at day 7 (Fig. 3.14E+F). The TF *TWIST1* however showed a strong increase in expression after 3 days, followed by a visibly reduced upregulation at the time points 5, 7 and 10 days. The increase after 3 days presented variation across the repeats through which no significant differences could be observed compared to the remaining time points (Fig. 3.14G). The TF *ZEB1* showed a consistent upregulation across all 4 time points with the highest, and significant increase compared to day 3 at the time point 10 days (Fig. 3.14H).

Based on the most characteristic EMT morphology after 10 days and the clear detectability of molecular EMT markers at the same time point, 10 days were selected for further studies. The selection of 10 days as a treatment length also enabled an increased comparability to the model of P5B3, because the treatment length of both is identical.

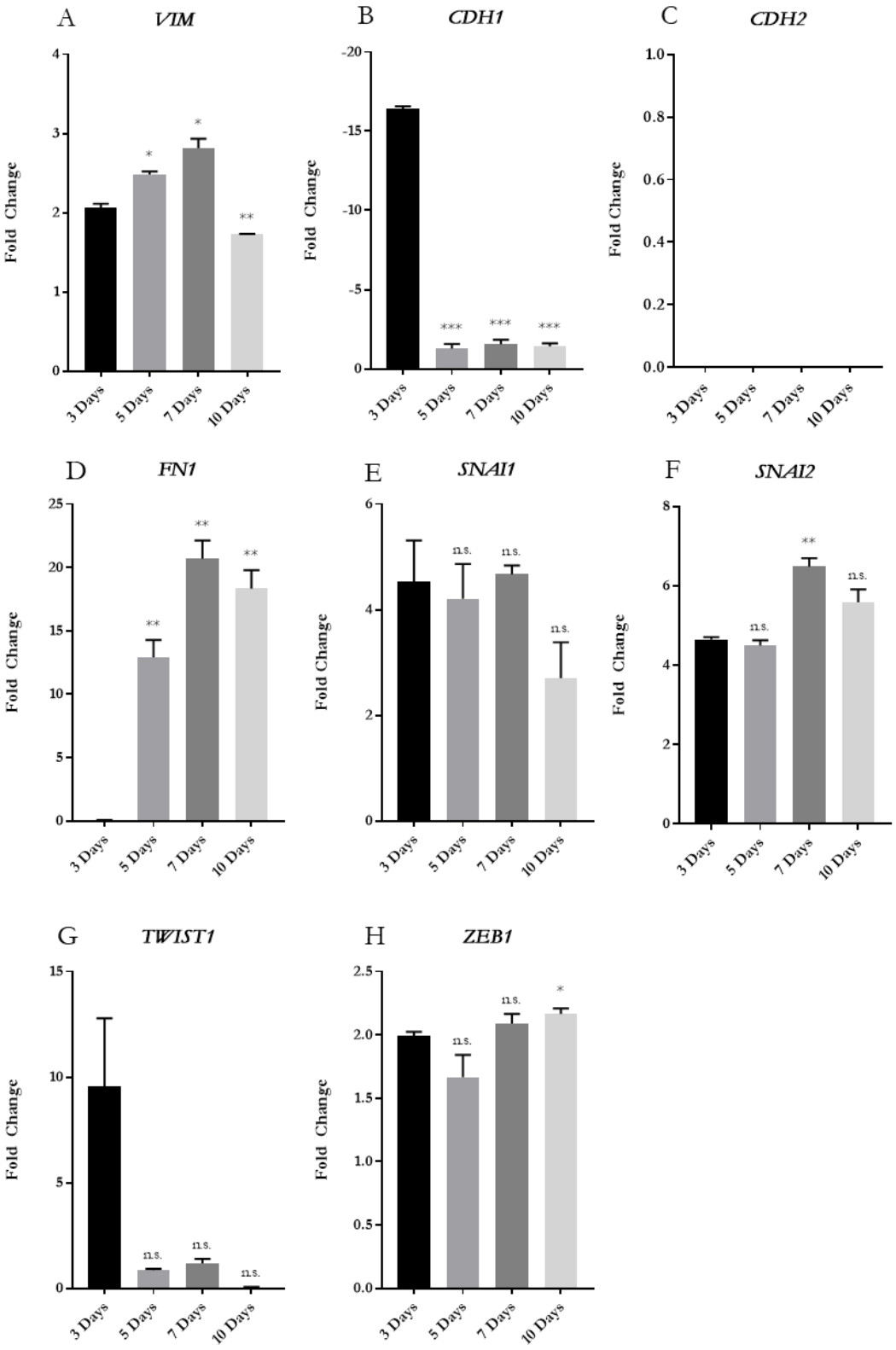


Figure 3.14: Gene expression changes of known EMT markers induced in DU145 upon stimulation with TGF-β . The gene expression of Vimentin (A), E-cadherin (B), N-cadherin (C), Fibronectin (D), and the EMT-TFs Snail (E), Slug (F), Twist (G) and ZEB1 (H) was measured across four different time points using quantitative real-time PCR and $2^{-\Delta\Delta CT}$ method (Schmittgen, Livak 2008) (n=4). The expression of CDH2 (C) could not be detected. The significance analysis was performed comparing the fold change of days 5, 7 and 10 with the fold change difference of each gene induced after treatment for 3 days. The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

3.2.2.2 Protein expression changes in DU145 after treatment with 10 ng/ml TGF-β for 10 days using immunofluorescence staining

To further confirm the changes of expression at time point 10 days, immunofluorescence staining of cells was performed for observation of potential changes in the protein expression of the VIME, CADH1 and FINC (Fig. 3.15). No apparent changes, neither in the protein expression nor the protein localisation could be observed for VIME comparing both conditions. In both conditions, the expression of VIME was localised in the cytoplasm presenting no visual differences in their expression between stimulated and unstimulated cells. The expression of VIME in DU145 was previously shown in independent studies (Bizzarro, Belvedere et al. 2017, Qin, Pan et al. 2014). Untreated DU145 cells showed weak expression of CADH1, which was not detectable after the treatment with TGF-β. The strongest change could be observed in the expression of FINC, which was initially not detected in untreated DU145 cells and strongly increased in its expression after 10 days of treatment.

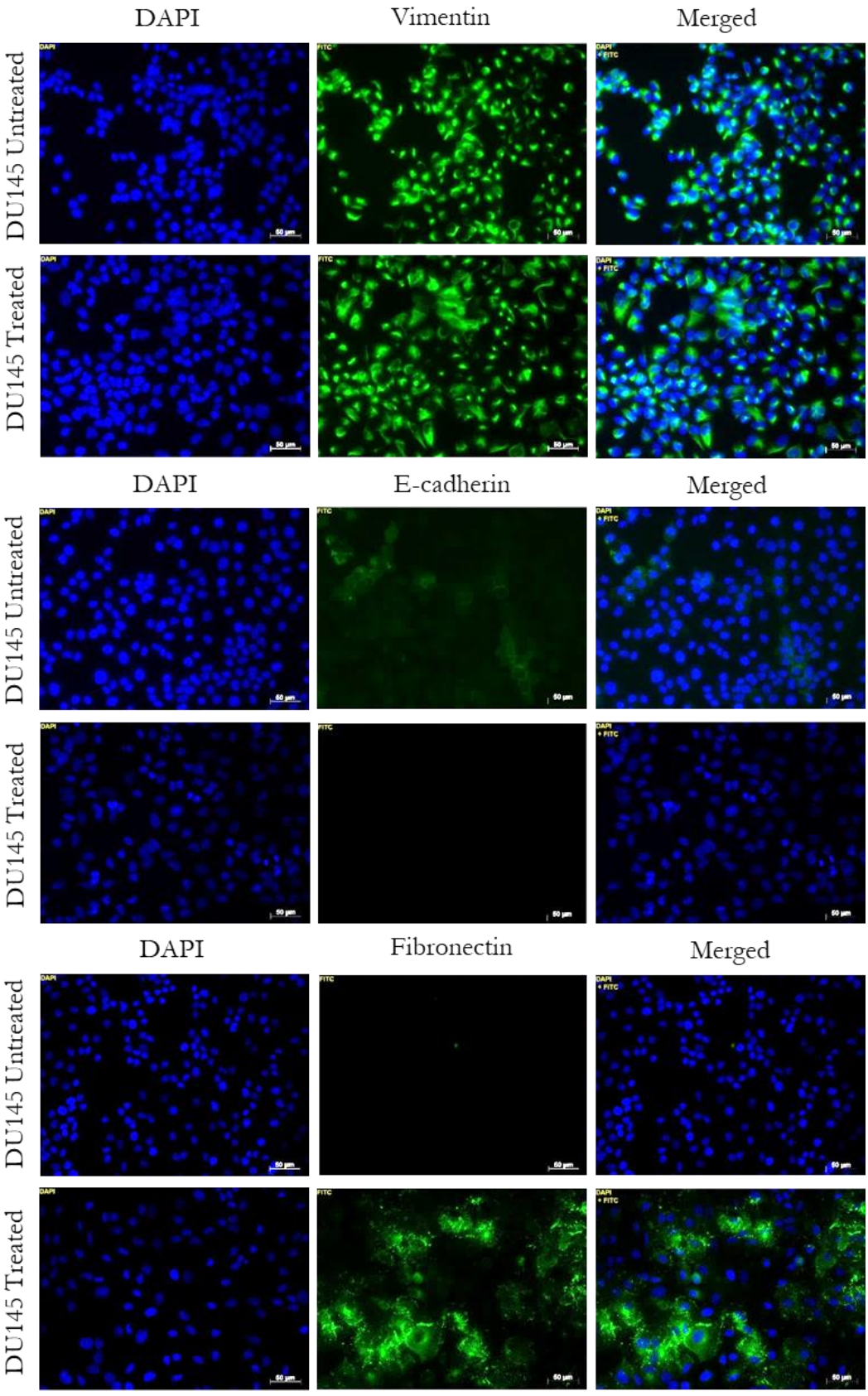


Figure 3.15: Representative images of immunofluorescence staining of untreated and treated DU145 cells after treatment for 10 days with 10 ng/ml TGF-β showing the EMT marker E-cadherin, vimentin and fibronectin. Staining with DAPI is presented as blue and FITC staining represents staining with marker of interest. The scale bar shows a length of 50μm.

3.2.2.3 Protein expression changes in DU145 after treatment with 10 ng/ml TGF-β for 10 days using Western blot analysis

The strong upregulation in the expression of FINC in both treated compared to the untreated samples was also confirmed using Western blot. No clear difference in the expression intensity of VIME was shown. However, the band of VIME was presented as a double band, of which the higher molecular weight band seems to be reduced, whereas the intensity of the lower molecular weight band remains consistent in the treated samples (Fig. 3.16). CADH2 could not be detected through Western blot analysis. A lack of *CDH2* expression was also previously shown in the screening of EMT markers via qRT-PCR (Fig. 3.14).

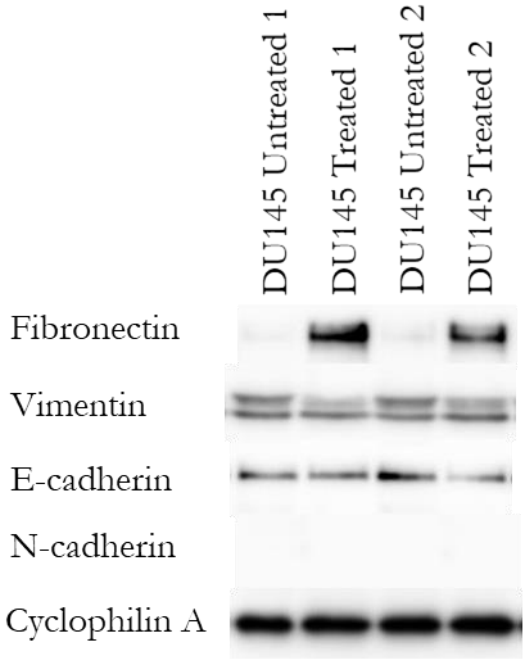


Figure 3.16: Western blot of cell lysates generated from untreated and treated DU145 cells. Protein analysis of the EMT markers fibronectin, vimentin, E-cadherin and N-cadherin. Cyclophilin A was used as loading control. 50 μg of protein was loaded for each sample.

3.2.2.4 Comparison of generated models with the study that defined EMT spectrum

Many studies are focussing on the extreme phenotype of epithelial and mesenchymal cells. However, stimulation and induction of a mesenchymal cell state does not always translate into a complete change to a mesenchymal cell (Lundgren, Nordenskjöld et al. 2009, Hiew, Cheng et al. 2018). Based on these experiences, Huang *et al.* tried to identify molecular characteristics of cells undergoing EMT that will enable the categorisation of cells based on their degree of EMT. For this, they studied the expression profiles of 43 ovarian cancer cell lines and identified four separate phenotypic subgroups: epithelial, intermediate

epithelial (E), intermediate mesenchymal (M) and mesenchymal (Huang, R. Y., Wong et al. 2013), based on their expression of EMT markers, including *VIM*, *CDH1*, *CDH2*, *ZEB1*, *TWIST1* and *SNAI1* (Fig. 3.17). Other genes were studied in addition (e.g. *KRT19*, *ITAG5*, *MMP2* and *ZEB2*) (Huang, R. Y., Wong et al. 2013). These genes were not analysed in the process of this study, instead, other markers, including *FN1* and *SNAI2* were used.

A full epithelial cell subtype was characterised in the study of Huang et al. through high expression of *CDH1* and low expression of all other markers. A full mesenchymal cell subtype, however, presented the lowest expression of *CDH1*, whereas the expression of *VIM* and *TWIST1* was the highest. An intermediate E stage shows a reduction of *CDH1* expression and a slight increase in the expression of *VIM*, *CDH2* and *ZEB1*. The transcription factor *SNAI1* shows a peak in its expression at an intermediate E stage, whereas no differences in the expression of *TWIST1* can be shown between epithelial and intermediate E. An intermediate M phenotype is characterised by an increased expression of *VIM* and *TWIST1*. Furthermore, the highest expression of *CDH2* and *ZEB1* is detected at an intermediate M stage and *CDH1* is more strongly decreased compared to intermediate E.

The comparison of the gene expression changes over time indicates that both models are in the intermediate M stage. This was shown through the highest expression of *ZEB1* (Fig. 3.10H and Fig. 3.14H) at 10 days, whereas a reduction of *ZEB1* is correlated with a full mesenchymal state. Furthermore, for P5B3, *CDH2* also shows the strongest expression across all time points at day 10 (Fig. 3.10C). In addition, a reduction of *CDH2* correlates with full mesenchymal cell state. Unfortunately, *CDH2* was not detectable in DU145 and could therefore not be utilised for a more in-depth characterisation Other markers, such as *TWIST1* (Fig. 3.10G and Fig. 3.14G) and *VIM* (Fig. 3.10A and Fig. 3.14A) do not present conclusive correlations with the published proposal, therefore further factors most likely influenced the expression of these EMT genes.

Figure 3.17: qRT-PCR analysis of selected EMT-associated genes and their expression across 4 distinct subtypes of EMT generated by Huang et al, 2013: epithelial, intermediate epithelial (E), intermediate mesenchymal (M) and mesenchymal. The images were taken from (Huang, R. Y., Wong et al. 2013). The raw data was not available.

3.2.2.5 Analysis of changes in the migratory behaviour of both cell line models upon treatment with 10 ng/ml TGF-β for 10 days using scratch assays

The study of changes in the migratory capabilities of treated cells, which is an important characteristic of cancers developing metastatic potential, was studied through the use of scratch assays. These assays are an easy and commonly used method to analyse the migratory potential of cells. Here, cells were grown in a monolayer to full confluency, and a scratch applied through the layer. The distance between the cellular borders were measured at the same location at time point 0 and 24 h. The percentages of the wound closure of both conditions were calculated and compared.

In P5B3, a complete wound closure could be observed in the induced state of EMT after 24 hours, whereas the wound closure of P5B3 untreated accounted for only about 6 %. These differences present a significant increase in the migratory potential of P5B3 through the treatment with TGF-β (Fig. 3.18).

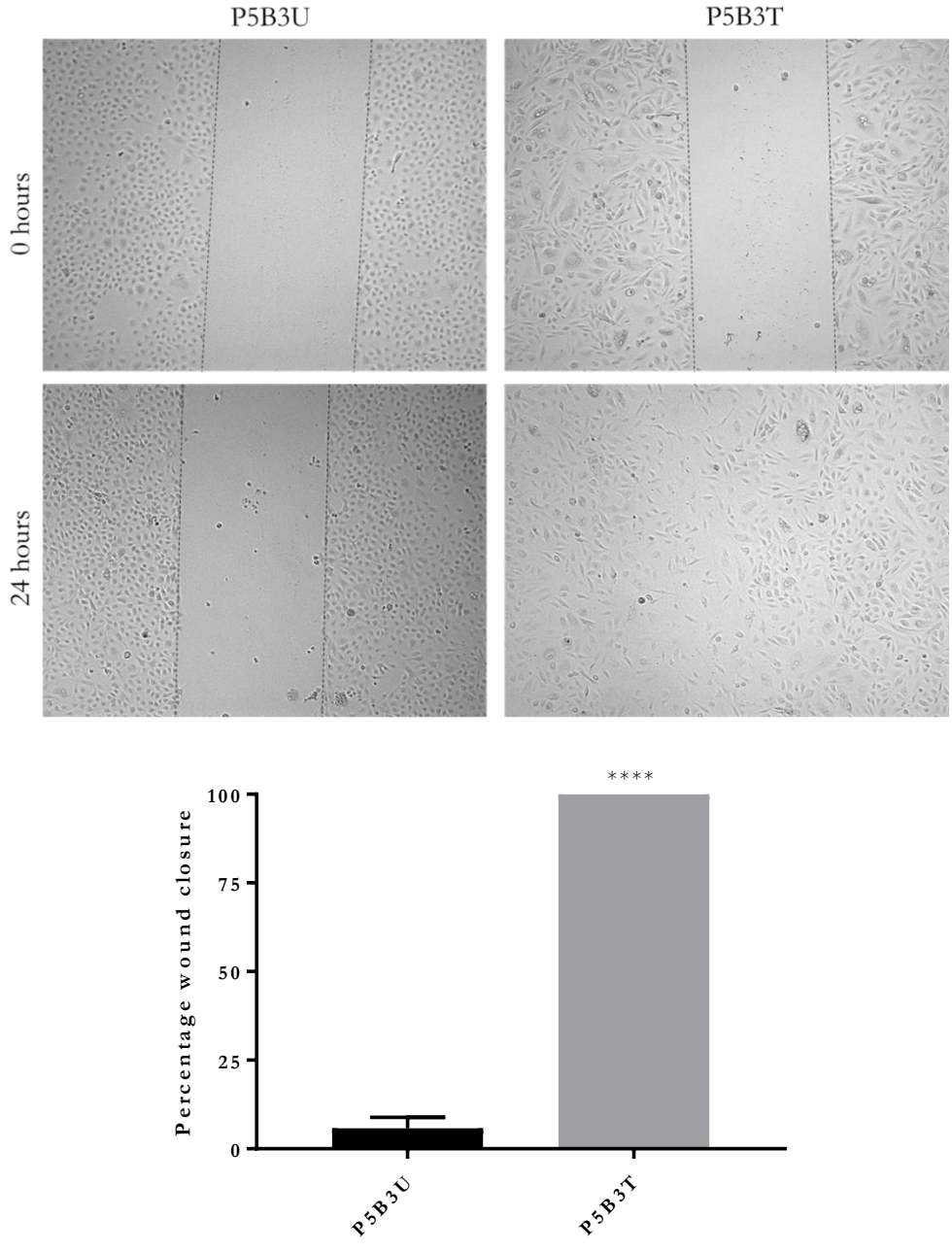


Figure 3.18: Analysis of changes of migratory capabilities in untreated (U) and treated (T) P5B3. Morphological observations and measurements were performed at time point 0 and 24 hours. Measurements were taken at same locations for each well (n=6). A T-test was performed to test for significance on the percentage of wound closure.

The analysis of DU145 untreated and treated showed a wound closure of 28 and 29 %, respectively (Fig. 3.19) over a 24 h period, which did not present a significant difference. Difficulties in the imaging of treated DU145 cells were observed. Through the treatment, the cells changed their morphology resulting in difficulties in imaging using a conventional bright field microscope. In the represented images, black dots are visible within the scratch. These are not cells and are most likely caused by cellular debris or scratches on the tissue culture plastic.

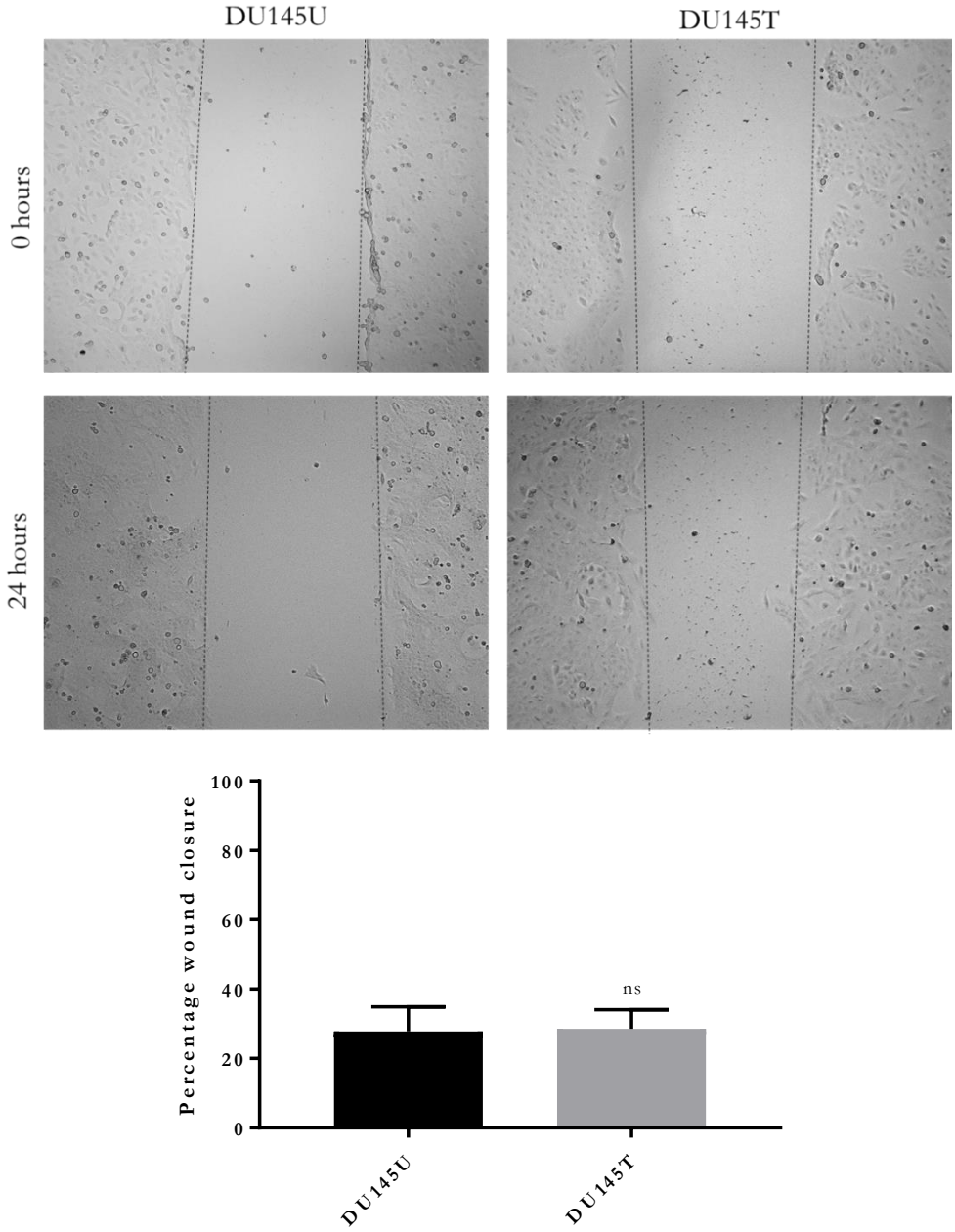


Figure.3.19: Analysis of changes of migratory capabilities in untreated (U) and treated (T) DU145 . Morphological observations and measurements were performed at time point 0 and 24 hours. Measurements were taken at same locations for each well (n=6). A T-test was performed to test for significance on the percentage of wound closure.

3.3 Discussion

The aim of this chapter was the development and the characterisation of two EMT models based on independent prostate cancer cell lines through the stimulation with transforming growth factor β.

The first part of the work focused on the single, highly epithelial cell clone P5B3, which was derived from the parental cell line OPCT-1. The initial results, based on the stimulation of P5B3 with 10 ng/ml TGF-β, showed morphological changes of the cells, exhibiting elongated cell shape and detachment from surrounding cells (Fig. 3.4). Furthermore, the analysis of gene expression changes confirmed an upregulation of known EMT-induced genes, inducing *VIM*, *FN1*, *CDH2*, *SNAI1*, *SNAI2*, *TWIST1* and *ZEB1* (Fig. 3.5). Furthermore, *CDH1* was downregulated. Based on this cell lysates of both cell line conditions were analysed using quantitative mass spectrometry resulting in a significant enrichment of Gene Ontology terms associated with EMT, including "tissue development", "cell adhesion" and "cell migration" (Fig. 3.8). The process of tissue development is associated with EMT type I and type II. Type I EMT is involved in the formation of new tissue and embryonal structures (Kalluri 2009), whereas in type II EMT the process of wound healing leads to the regeneration of tissues (Stone, R. C., Pastar et al. 2016) and the development of fibrosis (Section 1.3.2). Alterations in cell adhesion and the interaction with the ECM enable the cells to increase their motility and their migratory capabilities (Lamouille, Xu et al. 2014).

Focussing on a selection of significantly altered proteins, including ANXA3, ITAV and KRT19, showed an additional support to the induction of an EMT-phenotype in P5B3 (Fig. 3.8). ANXA3 is a calcium-dependent phospholipid-binding protein, which plays a role in cell differentiation and migration (Du, Liu et al. 2018), furthermore its deregulation was shown in various cancers, such as prostate and hepatocellular carcinoma (Tong, Fung et al. 2015). ITAV functions as the alpha chain V of heterodimeric integrins. Together with CD61 (integrin beta 3) it functions as a receptor for proteins involved in the EMT process, such as fibronectin, vitronectin and metalloproteinases. An increased expression of ITAV was also shown to be involved in disease progression in prostate (Cooper, C. R., Chay et al. 2002) and colorectal cancer (Waisberg, De Souza Viana et al. 2014). Keratin 19 is a member of the keratin family. This protein was downregulated through the treatment with TGF-β. Previous studies have shown that a knockdown of *KRT19* in the

breast cancer cell lines MDA-MB-231 and MCF-7 resulted in an induction of cell profileration, migration and invasion (Saha, Choi et al. 2017) and furthermore, patients with a low expression of KRT19 have shown a significantly worse survival rate compared to patients with a higher KRT19 expression shown in neuroblastoma patients (Nozato, Kaneko et al. 2013). Based on the morphological changes to an increased mesenchymal cell type, as well as molecular and proteomic changes related to the process of EMT, the stimulation of P5B3 with TGF-β was characterised as an induction of an increased mesenchymal phenotype. This hypothesis was further supported by the in-depth analysis of cell lysates. This enabled the characterisation of the P5B3 cell line model in a wider context, disregarding commonly studied EMT markers.

Studies on cell migration and development of metastasis have shown that migration can take place either as single cells or in collective sheets (van Zijl, Krupitza et al. 2011). Furthermore, the degree of EMT seems to affect the cell morphology where it was shown that cells in an intermediate epithelial/mesenchymal cell state can present cell cluster formation (Huang, B., Jolly et al. 2015) with cells showing EMT induced genes or double positive cells presenting epithelial and mesenchymal markers (Fustaino, Presutti et al. 2017). Upon stimulation of both cell lines, morphological changes have presented themselves through the development of single, elongated cells (P5B3) (Fig. 3.9) and grouped elongated cells (DU145) (Fig. 3.13). The morphological differences between both models could highlight potential differences in their stage of EMT, in which P5B3 is further along the EMT pathway compared to DU145. This hypothesis, however, is solely based on morphological observations. Furthermore, it needs to be highlighted that DU145 is a heterogeneous cell line with different cell populations, which can potentially give some indication of the limited morphological response of DU145 to the stimulation with TGF-β. Chunthapong et al. have shown that DU145 cells can be separated into 2 subpopulations, which they named as DU145-E (epithelial) and DU145-F (fibroblastic-like), of which the latter presented a higher invasive phenotype compared to the other (Chunthapong, Seftor et al. 2004). These differences in the phenotype of subpopulations of cells within DU145 might explain and correspond to the differential response of cells within the cell line model of EMT (Fig. 3.13). It might be the case that single, responding cells present a similar EMT-state compared to P5B3, whereas the remaining non-responding cells inhibit strong morphological changes.

The time point optimisation experiment, based on EMT-associated gene expression analysis, has shown strong changes in the majority of the analysed EMT markers and therefore 10 days stimulation was selected for the future work (Fig. 3.10 and Fig. 3.14). The analysis of protein expression in P5B3 untreated and treated for the proteins CADH1, CADH2, VIME and FINC using Western blot (Fig. 3.12) and immunofluorescence analysis (Fig. 3.11) correlated with the measured gene expression. For DU145, the changes in the protein profile with regard to EMT was most clearly visible for FINC, which was strongly expressed in treated cells compared to no visible protein expression based on immunofluorescence staining (Fig. 3.15) and Western blot analysis (Fig. 3.16). The increased expression was also supported by a more than 15-fold upregulation of FN1 after 10 days of stimulation (Fig. 3.14). Vimentin expression was already present in untreated DU145 and no apparent upregulation of this protein could be seen using either method, whereas a 2 to 3-fold upregulation of VIM was observed on a molecular level (Fig. 3.14). Confirming previous qRT-PCR (Fig. 3.14) results, no expression of CADH2 was detectable using Western blot analysis (Fig. 3.16).

Furthermore, the comparison of the gene expression profiles of P5B3 (Fig. 3.10) and DU145 (Fig. 3.14) according to the definitions of EMT subtypes (Fig. 3.17) (Huang, R. Y., Wong et al. 2013) has indicated that both generated models are present at an intermediate mesenchymal cell stage and have not fully converted to a mesenchymal cell stage (Huang, R. Y., Wong et al. 2013). A complete transition to a full mesenchymal phenotype was not achieved, potentially due to the restricted length of treatment or the use of a single cytokine for the induction of EMT. The EMT program can be initiated through various stimuli, including surrounding cells and soluble factors. One of these factors is TGF-β, however this cytokine is only one part of the microenvironment. The combination of multiple cytokines might have supported a full transition to a mesenchymal cell type (Sistigu, Di Modugno et al. 2017).

Overall, the analysis of molecular and protein-based EMT-markers have highlighted a change of expression correlating to a more mesenchymal cell state. Furthermore, morphological changes indicated a response to the stimulation with TGF-β, which presented itself with single, elongated cells of P5B3 (Fig. 3.9) and grouped, elongated cells of DU145 (Fig. 3.13). Based on these analyses, a successful induction of an increased mesenchymal-like phenotype could be confirmed in both cell line models induced

through the stimulation with 10 ng/ml TGF-β. For this reason, both models were selected for the generation of multi-omic profiles based on their molecular and proteomic changes of core-EMT markers.

4. Chapter IV – Generation and characterisation of trancriptomic and proteomic profiles of two inducible models of epithelial to mesenchymal transition

4.1. Introduction

Cancer is a heterogeneous disease and various changes in omic levels, such as the transcriptome and proteome, can lead to changes in pathways that alter downstream processes such as disease development and progression. To date, many one-dimensional studies (Huang, S., Chaudhary et al. 2017), focussing on a single omic level, have been performed characterising cancerous material from cell lines (Bainbridge, Warren et al. 2006, Beck, Schmidt et al. 2011), *in vivo* models (Takaishi, Wang 2007) and real-life tumour tissue made available through clinical studies (Shukla, Sudhanshu, Zhang et al. 2016, Long, Xu et al. 2014). However, despite regular discoveries of novel disease-associated markers and deregulated pathways of significance, novel findings rarely proceed further into clinical trials (Poste 2011). Reasons for this are difficulties in the translation of wet-lab findings into a clinical setting (Drucker, Krapfenbauer 2013) and the fact that many studies are performed studying solely the transcriptome; and such findings do not necessarily translate into changes in the protein expression (Vogel, Marcotte 2012). For this reason, improvements could be made through the use of multi-platform based profiling (Murphy, Murphy et al. 2018, Kulasingam, Vathany, Pavlou et al. 2010). Such an integration of multi-omics data could result in an accelerated discovery of novel biomarkers, potentially presenting more robust and reliable findings that are easier translated into a clinical validation process (Seyhan 2010).

In the case of cancer, the survival-limiting factor is, in the majority of cases, the development of distant metastases (Chaffer, Weinberg 2011). During metastasis changes at various omic levels result in alterations in gene and protein expressions, which enable primary tumours to invade the surrounding tissue, disperse throughout the body and to initiate the growth of secondary tumours at distant sites (Valastyan, Weinberg 2011). Based on the potential alterations of involved features across the different transcriptional and translational steps from a genetic sequence to the proteins, the use of multi-omics profiling presents a promising possibility to derive an increased understanding of changes in the signalling pathway. Such an understanding would potentially not be possible

through the analysis of a single-omic approach (Seyhan 2010). Therefore, the multi-omics study of EMT in prostate cancer could contribute to a better understanding of the underlying changes enabling tumours to invade (Balbin, Prensner et al. 2013).

4.1.1 Summary of commonly used gene expression analysis methods for the generation of transcriptomic profiles

The activation of a gene results in its expression in the form of a so-called messenger RNA (mRNA) and the abundance of mRNAs can give indications in their activity. The comparison of gene expression changes between multiple groups, such as diseased or healthy, can highlight underlying patterns and subtypes relevant for the study of interest, such as cancer. Gene expression profiles can be generated from *in vitro* and *in vivo*, or even patient material, and can be utilised for the discovery and validation of novel markers associated with biological processes or disease states.

Two routinely-used methods for the study of the whole transcriptome are available and these are microarray profiling (Baldi, Hatfield 2011) and RNA-sequencing (Wang, Z., Gerstein et al. 2009). Both methods enable the analysis of coding and non-coding RNA and are routinely and successfully applied in the field of cancer research. An example of the successful application of microarray analysis was shown in the study of Lapointe et al, which profiled 225 prostate tumours for the identification of clinically relevant subtypes of PCa patients (Lapointe, Li et al. 2004). On the other hand, RNA-sequencing analysis was successfully used for the generation of 585 patient-derived gene expression profiles, which resulted in the identification of *PCAT14* as a significant predictor for the development of metastasis, as well as biochemical-progression free survival (Shukla, Sudhanshu, Zhang et al. 2016). Both examples have shown that the study of transcriptomic profiles, independent from the generated platform, can generate meaningful outputs, potentially resulting in future clinically utilised information.

For both methods, RNA is extracted from a specific sample of interest, such as cell line material or tissue sections and cDNA is generated and tagged with either a fluorescence label or a sequencing adaptor. In the case of a microarray analysis, the cDNA material is then hybridised onto an array, which is covered with thousands of pre-defined DNA spots and incubated. During this step, the fluorescent-tagged cDNA can bind to covalent strands of DNA on the chip. After this, non-bound and non-specific bound cDNA molecules are removed during a washing step, and only specific bound cDNA is further

analysed. In the end, the array is scanned and excited with a laser, resulting in the ability to detect fluorescence intensity records for each DNA spot, presenting a single probe ID (Schulze, Downward 2001). *In silico* processing enables the normalisation and quantification of mRNA for each gene of interest. After this, expression intensities can be analysed, for example through the comparison of genes or sample groups.

RNA-sequencing started with the development of a chain-termination sequencing by Dr Frederick Sanger, which is therefore also called Sanger-sequencing (Sanger, Coulson 1975). This method is the gold standard for the sequencing of single genes and is still commonly used for the identification of the genetic sequence of single genes. The second generation of sequencing methods, mainly known as next-generation sequencing (NGS) enables the massive-parallel analysis and quantitation of thousands of genes. In this case, the most commonly used approach is a process called "Sequencing by synthesis" (SBS). Various companies, such as Illumina (Bentley, Balasubramanian et al. 2008) and Applied Biosystems (Voelkerding, Dames et al. 2009) are offering this type of sequencing. In this study, the RNA-sequencing was performed on an Illumina NextSeq500, therefore the sequencing method is described based on the companies' approach.

SBS can be divided into 2 major steps, cluster generation and the actual sequencing. Initially, libraries of cDNA are generated. Adapter regions are added on both sides, then the cDNA is transferred onto a flow cell. This flow cell is a glass slide containing two types of oligos corresponding to one or the other adapter regions previously added to the cDNA. Initial copies of the bound cDNA fragments are generated, and the original template removed. The generated copy is then used to create clusters of identical complementary template molecules based on bridge amplification (Buermans, Den Dunnen 2014). After this, sequencing of the generated strands begins. During the cluster analysis, repeats of both strands are generated and for the first sequencing, one type of molecule is removed and the sequential extension of cDNA copies by fluorescent-tagged nucleotides is performed. The fluorescence tag differs for each single nucleotide, furthermore each nucleotide is attached to a terminator sequence. Every cycle, one nucleotide binds to the cDNA attached to the flow cell, the fluorescence is detected, and the terminator removed. This enables the binding of a new nucleotide to the analysed strand (Buermans, Den Dunnen 2014, Bentley, Balasubramanian et al. 2008). After a predefined number of cycles, the generated strand is removed, and a complete

complementary sequence is generated. This sequence is then used as a template for a second round of sequencing (Buermans, Den Dunnen 2014). The previously described sequencing process is repeated, resulting in so-called paired-end sequencing products. *In silico* processing of the generated reads enables the identification and quantification of RNA molecules in the analysed sample material.

Despite the successful application of both methods, RNA-sequencing offers strong advantages over microarray profiling. These include the unbiased screening of RNA present within the sample, which is limited in microarray analysis by the use of predefined probe sequences (Kukurba, Montgomery 2015). Furthermore, novel transcripts and gene variants at lower abundances can be routinely detected using RNA-sequencing. RNA-sequencing presents a broader dynamic range that can provide a more accurate detection of strong differentially expressed genes (Zhao, S., Fung-Leung et al. 2014, Nookaew, Papini et al. 2012). Microarray analysis shows limitations in the accurate quantification of very low and very highly expressed genes and transcripts (Kukurba, Montgomery 2015). Furthermore, microarray analyses generate gene expression values for multiple probe IDs per gene. These probe IDs cover different sequence segments of each gene and the binding affinity can vary. This commonly results in variations related to their significance and association across genes and can therefore limit the discovery of markers.

4.1.2 Summary of commonly used protein expression analysis methods for the generation of proteomic profiles

Proteomic profiles can be generated using mass spectrometry analysis. The major options available can be categorised into three main approaches. Data-dependent, also called information-dependent acquisition (IDA/DDA) (Fig. 4.1A), targeted proteomics through selected reaction monitoring (SRM), also known as multiple reaction monitoring (MRM) (Fig. 4.1B), and data-independent acquisition (DIA) (Fig. 4.1C) (Sajic, Liu et al. 2015, Hu, Noble et al. 2016, Sidoli, Lin et al. 2015).

All three analysis methods can be performed on tandem mass spectrometers, also known as MS/MS or MS^2. During an MS/MS analysis precursor ions (ions of a defined m/z ratio) are identified in a survey scan (MS1). The ions are then, unfiltered or filtered, selected for further fragmentation (Edmond de Hoffmann, Vincent Stroobant 2007). These fragments are then detected in fragment ion spectra (MS2), matched to a library and the peptides are identified based on their amino acid sequences.

Figure 4.1: Schematic representation of the three major mass spectrometry analysis methods. A = shotgun or data-dependent acquisition, B = selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) and C = data-independent acquisition (DIA), such as SWATH MS (Liu, Yansheng, Huettenhain et al. 2013)

Using data-dependent analysis (DDA), the most abundant ions are selected after the MS1 scan and subjected to further fragmentation and detection in MS2. An advantage of this approach is that it does not require any prior knowledge about the analytes and enables a hypothesis-free analysis (Sidoli, Lin et al. 2015, Aebersold, Mann 2016). Despite this, a DDA approach also presents limitations, mainly based on the sampling of most abundant ions, which can vary in each sample. For this reason, the reproducibility is very limited. Furthermore, the detection of low abundance peptides is difficult, and an accurate quantification of co-eluting peptides is challenging (Sidoli, Lin et al. 2015, Hu, Noble et al. 2016).

The second analysis method is the use of selected reaction monitoring (SRM). Here, a predefined group of previously identified peptides is selected in MS1 and analysed in MS2. This enables a reproducible quantification of targets but requires prior knowledge of the peptides of interest (Hu, Noble et al. 2016). Based on the prior knowledge and its defined selection, an SRM analysis presents a high degree of sensitivity, which enables therefore the detection of low abundance proteins. However, the analysis is restricted to a selection of pre-defined proteins of interest (Aebersold, Mann 2016).

The last analysis method widely used for the analysis of the proteome is called DIA (Sidoli, Lin et al. 2015). Specific DIA acquisition methods are available, such as SWATH-MS (Gillet, Navarro et al. 2012), Shotgun-CID (Purvine, Eppel* et al. 2003) and MS^E (Waters, 2018, (Plumb, Johnson et al. 2006). In this study, the generated protein lysates were analysed using SWATH-MS (Gillet, Navarro et al. 2012). SWATH-MS stands for sequential window acquisition of all theoretical fragment ion mass spectra (Ludwig, Gillet et al. 2018, Gillet, Navarro et al. 2012). Here fragment ion spectra of each precursor ion within a defined m/z window are measured, enabling the generation of multiplexed recordings of all peptides present. The analysis of m/z windows is performed through their cycling across the complete m/z precursor range. In the initially developed DIA approach, the width of the m/z window was defined as an equal width across the complete m/z range, however novel developments enable nowadays the use of variable m/z windows (Zhang, Y., Bilbao et al. 2015, Ludwig, Gillet et al. 2018). These variable m/z are useful for mass regions of higher precursor density or intensity, resulting in increased protein identifications (Zhang, Y., Bilbao et al. 2015). The importance in this approach is the ability to assign the three-dimensional information (retention time, fragment ion m/z and intensity) correctly. This information can be matched to a library, whereas the correct identification and quantification depends on the quality of the previously generated library (Schubert, Gillet et al. 2015). Overall, a DIA approach enables a more in-depth analysis (Borràs, Sabidó 2017) and high-throughput analysis of sample material. Furthermore, an improved quantification of low abundance proteins is possible; however, SRM still presents a better ability for undertaking this task, based on its high sensitivity in the quantitation of targeted proteins and peptides (Hu, Noble et al. 2016).

In the previous chapter, two inducible models of EMT were successfully generated and characterised through the analysis of morphological, gene and protein expression changes. The generated data confirmed the induction of an EMT phenotype, enabling the use of these models for the further scope of the study.

This chapter will describe the generation of matching transcriptomic and proteomic profiles of both cell line models in their "natural" and induced cell state and the use of these profiles for the in-depth characterisation of changes in underlying pathway through the use of Metacore™, a pathway analysis tool. Each cell line and omic profile will be

analysed separately and in combination with their proteomic counterpart. Based on this, the chapter is separated into multiple parts.

- Initially, the generated omics profiles will be used to validate the successful EMT induction through the repeated analysis of the well-studied EMT markers (CDH1, CDH2, VIM, FN1, ZEB1, SNAI1, SNAI2, TWIST1). The repeated comparison of gene and protein expression changes in both cell line models and omic levels will ensure the induction of EMT throughout the dataset generation experiment.

- A selection of significant altered markers (genes or proteins) will be identified and applied using Metacore™ pathway analysis. The selection of these markers and the application of these, will enable the validation of EMT induction and potential identification of additionally affected pathways through the stimulation with TGF-β. This step will be used as additional quality control for the induction of EMT. The pathway analysis will highlight potential off target effects on pathways that might alter the desired phenotypic changes.

- To identify the impact of matching sample collection on the correlation of gene and protein expression, matching markers will be selected and a correlation analysis performed. This analysis will help to highlight potential improvements possible through the parallel extraction of RNA and protein.

The successful performance of these steps will enable the use of these profiles for their integration and the identification of a core marker set, which will be performed in chapter 5.

4.2 Results

The hypothesis behind the study was that matching transcriptomic and proteomic profiles from the same cells in the same condition could facilitate the discovery of novel disease-associated biomarkers (Seyhan 2010) and that markers with a concordant expression on a transcriptomic and proteomic profile could indicate a more robust and reliable biomarker, based on a consistent stability enabling long term detectability. Furthermore, the majority of large patient-derived omic profiles, which are publicly available, have been generated through transcriptomic and genomic analyses, and only limited information was provided on the proteome of these samples. An example for this is "The Cancer Genome Atlas" (Tomczak, Czerwinska et al. 2015), which generated multi-omics profiles of more than 30 cancer types. These profiles cover coding and non-coding transcriptomics, as well as single nucleotide variants and copy number variations. Based on this, the inclusion of quantitative proteomic profiling could increase the implications of detected markers and their potential utility as therapeutic targets, especially since the majority of approved therapeutic drugs target cellular proteins (Landry, Gies 2008).

4.2.1 RNA-sequencing analysis of RNA extracted from treated and untreated P5B3 and DU145 cells

4.2.1.1 Preparation of sample material, analysis and data output

RNA was extracted from four different experimental cell groups; P5B3 untreated, P5B3 treated, DU145 untreated and DU145 treated. A reduced number of biological replicates for both treatment conditions of DU145 (n=9) were used due to space limitations on the analysis platform. As part of the RNA extraction, a DNase treatment was performed for each sample to ensure the complete removal of genomic DNA from the RNA samples. This step was necessary due to the nature of the sequencing approach. The sequencing is performed on cDNA generated from the isolated RNA. Presence of genomic DNA could affect the quality of the generated data and bias the results, since it is not possible to differentiate between reads generated from cDNA and reads generated from genomic DNA.

To confirm the successful removal of genomic DNA, 9 out of 38 samples were randomly selected and used as template for a quantitative real-time PCR analysis. Previously generated cDNA of the same cell line models was used as positive control. The presence of genomic DNA in the analysed samples would result in the detection of the control gene. An absence of genomic DNA is shown through the absence of an amplification product in the randomly selected samples. All tested samples were showing no measurable CT value in the RNA-sequencing samples and therefore a negative result for the presence of genomic DNA (Table. 4.1).

Table 4.1: Representative analysis of 9 randomly selected samples of both cell line models for the testing of the presence of genomic DNA using quantitative real-time PCR (n=2). PCR primers for the reference gene TBP were used for the analysis.

Sample	Cycle Time
P5B3U T8	Not detected
P5B3U T9	Not detected
P5B3U T12	Not detected
DU145U T1	Not detected
DU145T T1	Not detected
DU145T T4	Not detected
P5B3T T13	Not detected
P5B3U T14	Not detected
P5B3T T17	Not detected
Positive control	28.36
Positive control	28.1

The negative results for the detection of genomic DNA in the sample material allowed further quality control of the samples prior to the RNA-sequencing analysis. RNA for use in sequencing approaches has to be of high quality. For the assessment of the quality, the so-called "RNA Integrity Number" (RIN) can be defined. The RIN output is a value between 1 and 10, of which 10 indicates the best quality, representing RNA in the least degraded form (Kukurba, Montgomery 2015). For this study, the cut-off was defined as a RIN of 8 or higher and a concentration of 200 ng/µl per sample, as this was requested by the DeepSeq facility, which further processed the extracted RNA and generated the transcriptomic profile of each supplied sample. Each sample was analysed using the Agilent RNA 6000 Nano Kit with RNA Nano Chips (See Appendix). In Table 4.2 the generated RNA concentrations and RIN values are shown for each analysed sample. All analysed samples have shown a RIN of 10 and a concentration above 200 ng/µl and therefore passed the quality criteria for downstream analysis using RNA-sequencing.

Table 4.2: List of generated samples of both cell line models and treatment conditions and their corresponding RNA concentration (ng/µl) and RNA Integrity Number (RIN), which were downstream subjected to RNA-sequencing analysis.

Sample	ng/µl	RIN		Sample	ng/µl	RIN
P5B3U T7	513.15	10		DU145U T2	411.57	10
P5B3U T8	401.10	10		DU145U T3	402.65	10
P5B3U T9	401.72	10		DU145U T5	573.51	10
P5B3U T10	416.28	10		DU145U T6	517.58	10
P5B3U T11	443.69	10		DU145U T13	419.72	10
P5B3U T12	349.46	10		DU145U T15	424.10	10
P5B3U T13	498.83	10		DU145U T16	353.45	10
P5B3U T16	448.16	10		DU145U T17	414.06	10
P5B3U T17	683.22	10		DU145U T18	335.80	10
P5B3U T18	479.62	10		DU145T T1	317.13	10
P5B3T T7	449.29	10		DU145T T2	445.77	10
P5B3T T8	525.28	10		DU145T T3	393.17	10
P5B3T T9	375.12	10		DU145T T4	427.92	10
P5B3T T10	492.79	10		DU145T T6	479.32	10
P5B3T T11	435.13	10		DU145T T13	585.57	10
P5B3T T12	385.11	10		DU145T T15	359.81	10
P5B3T T13	416.96	10		DU145T T17	395.60	10
P5B3T T16	504.81	10		DU145U T2	411.57	10
P5B3T T17	439.02	10				
P5B3T T18	439.10	10				

The RNA-sequencing analysis and data generation was performed by the DeepSeq facility located at the University of Nottingham, UK (DeepSeq, 2019). The delivered results of their analysis were FASTQ files of each sample and both read directions. FASTQ files are a file format that enables the storage of sequence data in a text format (Cock, Fields et al. 2009). The files were subjected to *in silico* processing using the BaseSpace Sequence Hub of Illumina (BaseSpace, 2019) with the Tuxedo suite (Trapnell, Cole, Roberts et al. 2012). Here, the reads generated in this RNA-sequencing experiment were associated to one of three different sequence types within the genome; so-called exonic, intronic and intergenic regions (Fig. 4.2). The exonic region is comprised of the exons and the untranslated regions (UTR). Untranslated regions can be separated into 5'UTR and 3'UTR, which are located upstream and downstream of the coding regions, respectively, whereas exons present the sequences that code for genes. The other two sequence types are the intragenic and intergenic regions, which are non-coding regions either located within a gene, between the exons, or between genes, outside the coding regions, respectively.

In the analysed samples, the majority of the reads were assigned to the exonic (Fig. 4.2), followed by intronic regions. The least number of reads were assigned to intergenic areas. In P5B3, the percentages of aligned sequences were identical, whereas in DU145, a small reduction in the exonic and a small increase in the number of reads assigned to the intronic region could be observed upon treatment.

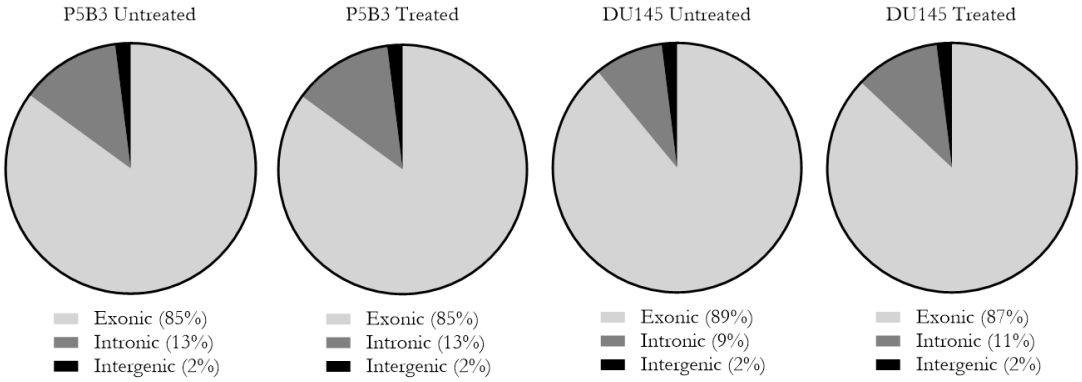


Figure 4.2: Graph indicating the average percent alignment of all reads to exonic, intronic and intergenic regions for the 4 analysed sample sets, namely P5B3 untreated (n=10), P5B3 treated (n=10), DU145 untreated (n=9) and DU145 treated (n=9).

The second output type through the data alignment and processing resulted in a normalised read count per gene and sample. This value is represented by the metric fragments per kilobase of transcript per million (FPKM) mapped reads (Trapnell, C., Williams et al. 2010). This method takes into account the variation of read counts based on the length of a gene. Longer genes will produce a higher number of read counts compared to shorter genes, despite the same expression intensity. For this reason, the count of fragments per gene is divided by its total length. The output value is the previously mentioned FPKM. In total, 26354 genes based on 56891 transcripts were detected within the analysed sample set (Tab. 4.3).

Table 4.3: Summary of detected genes and transcripts within the analysed sample set of untreated (n=10) and treated (n=10) P5B3 and untreated (n=9) and treated (n=9) DU145 cell line samples

	Unique genes	**Transcripts**
RNA-sequencing analysis	26354	56891

4.2.1.2 Validation of EMT gene panel in generated RNA-sequencing profiles

The initial analysis of the generated RNA-sequencing data was focused on the validation of a successful EMT induction. For this, the previously analysed EMT-associated genes (section 3.2.1.2), *VIM*, *CDH1*, *CDH2*, *FN1*, *TWIST1*, *ZEB1*, *SNAI1* and *SNAI2* were selected and their expression compared between the untreated and treated cell line conditions for P5B3 and DU145 (Fig. 4.3). In the sample set of P5B3, 7 out of 8 genes were detected with a significant difference between the untreated and treated cell state, showing an upregulation of *VIM* (Fig. 4.3A), *CDH2* (Fig. 4.3C), *FN1* (Fig. 4.3D), *ZEB1* (Fig. 4.3H), *SNAI1* (Fig. 4.3E) and *SNAI2* (Fig. 4.3F), and a downregulation of *CDH1* (Fig. 4.3B). The expression of *TWIST1* (Fig. 4.3G) has shown no significant difference between untreated and treated cell line samples. A high variability in the expression of this gene was already shown in the initial qRT-PCR analysis, limiting the significance between both cell line conditions.

In DU145, *TWIST1* (Fig. 4.3G) and *CDH2* (Fig. 4.3C) were not detected (nd), however, the remaining 6 markers were significantly deregulated in their expression between untreated and treated conditions. *CDH1* (Fig. 4.3B) was significantly reduced, whereas *VIM* (Fig. 4.3A), *FN1* (Fig. 4.3D), *ZEB1* (Fig. 4.3H), *SNAI1* (Fig. 4.3E) and *SNAI2* (Fig. 4.3F) showed a significant increase. All together in P5B3 and DU145, the expression of significantly deregulated genes was detected according to the expectation of an induced EMT phenotype, meaning that all significant genes, aside from *CDH1,* were upregulated through the stimulation with TGF-β. *CDH1* was downregulated in both cell lines upon treatment.

This analysis confirmed the successful induction of EMT on a transcriptomic level in both models and the desired molecular changes within the samples. This allowed their use in further analyses and biomarker discovery experiments.

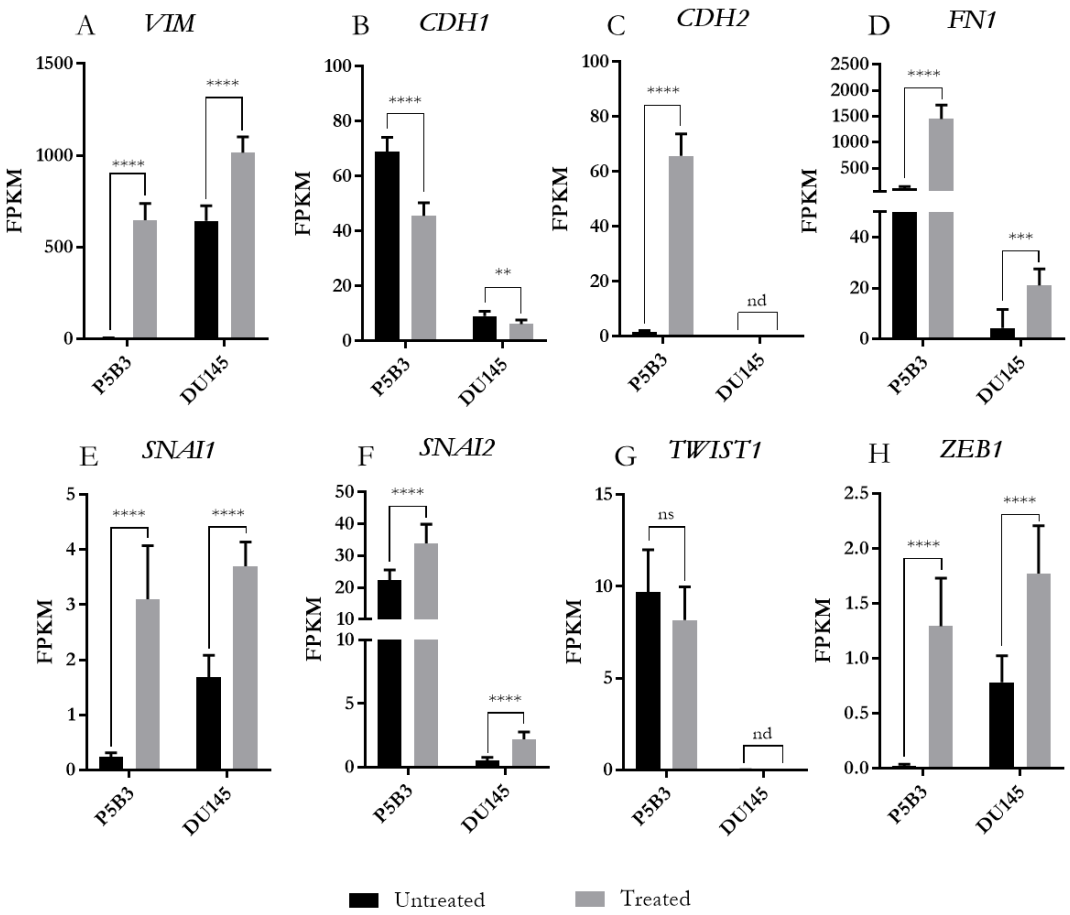


Figure 4.3: Gene expression changes of the EMT markers Vimentin (*VIM*), E-cadherin (*CDH1*), N-cadherin (*CDH2*), Fibronectin (*FN1*), Snail (*SNAI1*), Slug (*SNAI2*), Twist (*TWIST1*) and *ZEB1* across the sample population of untreated and treated P5B3 (n=10 per condition) and DU145 (n=9 per condition) represented in FPKM values.

4.2.1.3 Analysis of RNA-sequencing derived gene expression profiles of both cell line models for the characterisation of underlying pathway changes

For the further downstream analysis, genes that presented a significant difference between the treated and untreated condition after correction for false discovery were selected. The statistical analysis performed is described in Methods (section 2.2.6.4).

4.2.1.3.1 Identification of significant altered genes within the inducible EMT model of P5B3

The analysis of the significantly altered genes detected in P5B3 using the previously described filters resulted in a list of 4575 genes, of which 2787 were up- and 1697 were downregulated (Fig. 4.8A). The 4575 genes were applied to a hierarchical clustering and are presented in a heat map (Fig. 4.4), which has shown a clustering of the samples according to their treatment group, without apparent outliers. This indicated a stable induction state across all samples.

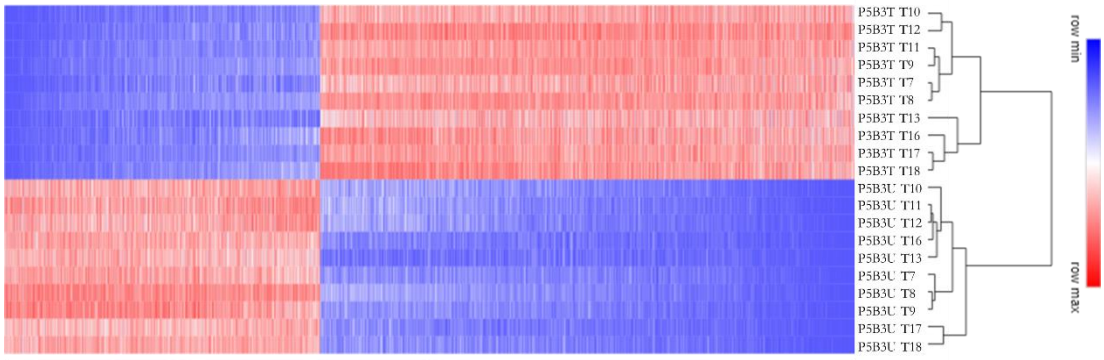


Figure 4.4: Hierarchical clustering of 4575 genes significantly (p-value <0.05) deregulated between untreated and treated P5B3 cells (n=10 per condition) using Euclidean distance and complete linkage.

The analysis of induced changes has highlighted a wide range of expression changes, ranging from fold change increases +1494.53 to fold change decreases of up to -120.44 fold. Within the strongest up- and downregulated genes, markers of EMT and metastasis association were identified. This included upregulated genes such as *CDH11* (+303.40), *VCAN* (+214.21) and *TWIST2* (+178.19) and the downregulated markers *GKN2* (FC = -118.65) and *PSCA* (FC = -103.70).

4.2.1.3.2 Analysis of pathways altered upon stimulation of P5B3 with TGF-β based on significant deregulated genes

For a more detailed analysis of the phenotypic changes induced by the treatment of P5B3, the selected genes and their associated fold changes were applied to the MetaCore™ pathway analysis tool from Clarivate Analytics (https://portal.genego.com/) (Park, A., Lee et al. 2017, Loughran, Leonard et al. 2018). This software enables the association of genes within a given list to defined pathways based on pathway topology. Pathway topology enables the analysis of pathways using not only the detection of markers, but also their expression information, to compute gene level statistics (Khatri, Sirota et al. 2012). The involvement of the genes of a dataset in the described pathways is indicated through a p-value, the corrected p-value and a ratio of detected genes compared to the total number of genes within the pathway. Furthermore, each of the enriched pathways is assigned to a broader category, such as "cell adhesion" or "development".

In the case of the significant altered genes of P5B3, a total of 779 pathways were shown to be significantly enriched, using a cut-off of <0.05 after correction for false discovery (FDR). Within the top 50 most significantly enriched pathways, the majority of pathways were associated with the categories of "Development", followed by "immune response" and "cell adhesion" (Fig. 4.5).

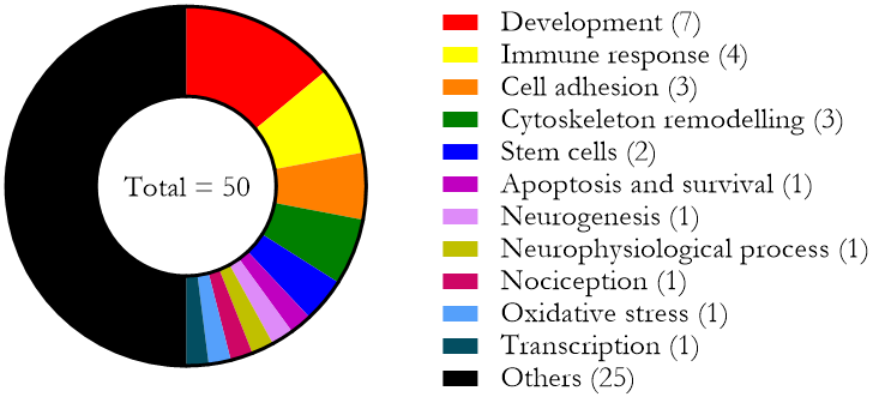


Figure 4.5: Top 50 most significantly enriched pathways based on significant genes in P5B3 grouped by their respective categories (n=4575). List derived from Metacore™ (accessed 02/07/18).

The top 15 most significantly enriched pathways are shown in table 4.4. Of these, 10 pathways are directly associated with TGF-β treatment, the process of EMT or the development of metastasis. The remaining 5 pathways are mainly connected to cytoskeletal rearrangements, which are commonly occurring during the change of

epithelial cells to cells with mesenchymal cell properties (Sun, BO, Fang et al. 2015, Nalluri, O'Connor et al. 2015).

Table 4.4: Top 15 most significant associated pathways of significantly deregulated genes P5B3 sorted by significance after FDR. List derived from Metacore™ (accessed 02/07/18).

Category	**Pathway**	**Total [1]**	**In data[2]**
Development	TGF-β-dependent induction of EMT via RhoA, PI3K and ILK	46	33 (72 %)
Development	Regulation of epithelial-to-mesenchymal transition (EMT)	64	40 (63 %)
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	37 (64 %)
Cell adhesion	ECM remodelling	55	35 (64 %)
Immune response	IL-1 signalling pathway	82	44 (54 %)
Not assigned	ErbB2-induced breast cancer cell invasion	67	38 (57 %)
Not assigned	TGF-β 1-mediated induction of EMT in normal and asthmatic airway epithelium	44	29 (66 %)
Not assigned	TGF-β 1-induced transactivation of membrane receptors signalling in HCC	50	31 (62 %)
Development	TGF-β-dependent induction of EMT via SMADs	35	25 (71 %)
Not assigned	Role of stellate cells in progression of pancreatic cancer	60	34 (57 %)
Not assigned	Stimulation of TGF-β signalling in lung cancer	48	29 (60 %)
Not assigned	Glomerular injury in Lupus Nephritis	92	43 (47 %)
Not assigned	Stellate cells activation and liver fibrosis	70	35 (50 %)
Not assigned	TGF-β-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways	64	33 (52 %)
Not assigned	IGF family, invasion and metastasis in colorectal cancer	33	22 (67 %)

[1]Total: Total number of markers present in the pathway
[2]In data: Number of identified markers of given pathway through the analysis of generated omic profiles

4.2.1.3.3 Identification of significant altered genes within the inducible EMT model of DU145

The dataset of DU145 was applied to the same stringent filters as previously described (Methods). Here, this approach resulted in a list of 2303 significantly altered genes, of which 1324 were up- and 979 were downregulated (Fig. 4.8B). The hierarchical clustering showed a clustering according to treatment group and did not indicate any outliers within the samples set (Fig. 4.6).

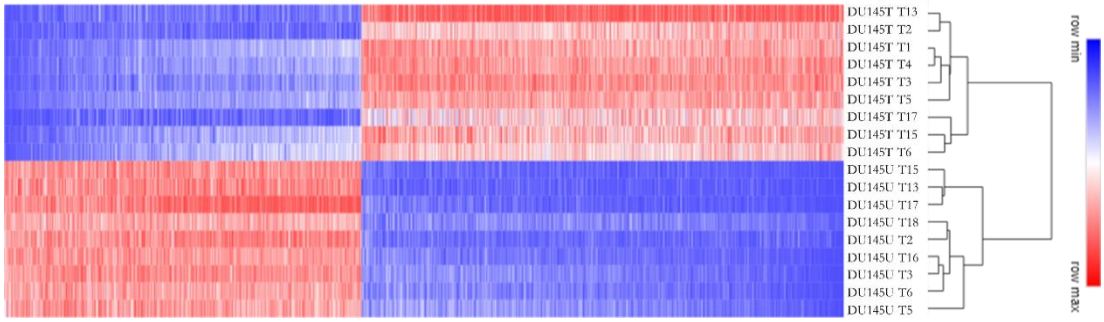


Figure 4.6: Hierarchical clustering of 2303 genes significantly (p-value <0.05) deregulated between untreated and treated DU145 cells (n=10 per condition) using Euclidean distance and complete linkage.

The analysis of induced changes highlighted a wide range of expression changes, ranging from +84.71 to -50.24 fold. Within the strongest up- and downregulated genes, markers of EMT and metastasis association were identified, including *BMP2* (84.71) and *SPOCK1* (70.63) as well as the downregulated markers *KRT32* (-27.18) and *KRT4* (-24.70).

4.2.1.3.4 Analysis of pathways altered upon stimulation of DU145 with TGF-β based on significant deregulated genes

For further characterisation, the significant genes were applied to the MetaCore™ pathway analysis tool. Here, 292 pathways were indicated to be significantly enriched within the supplied gene list. Within the top 50 most significant pathways, the majority were associated with "Cell adhesion", followed by "Development" and "Cytoskeleton remodelling" (Fig. 4.7).

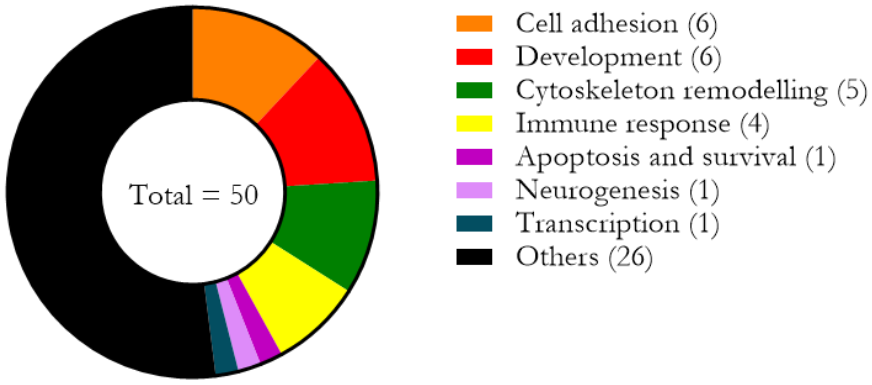


Figure 4.7: Top 50 most significantly enriched pathways based on significant genes in DU145 grouped by their respective categories (n=2303). List derived from Metacore™ (accessed 02/07/18).

.

The top 15 most significant pathways are shown in Table 4.5. A large number of pathways are associated with cytoskeletal changes and interaction of cells with the ECM. However, also pathways involved in the SMAD-dependent and independent signalling activated via the TGF-β receptors were enriched. These results show the successful alteration of the physiological cell state involving cytoskeletal remodelling as well as the induction of EMT.

Table 4.5: Top 15 most significant associated pathways of significantly deregulated genes in DU145 treated compared to DU145 untreated. List derived from Metacore™ (accessed 02/07/18).

Category	**Pathway**	**Total[1]**	**In data[2]**
Cytoskeleton remodeling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	23 (40 %)
Not assigned	TGF-β signalling via SMADs in breast cancer	47	20 (43 %)
Neurogenesis	NGF/ TrkA MAPK-mediated signalling	105	31 (30 %)
Not assigned	B-catenin-dependent transcription regulation in colorectal cancer	36	17 (47 %)
Not assigned	IGF family, invasion and metastasis in colorectal cancer	33	16 (48 %)
Not assigned	TGF-β 1-induced transactivation of membrane receptors signalling in HCC	50	19 (38 %)
Cell adhesion	ECM remodelling	55	20 (36 %)
Not assigned	Insulin-like growth factor family signalling in melanoma	38	16 (42 %)
Cell adhesion	Endothelial cell contacts by non-junctional mechanisms	24	12 (50 %)
Not assigned	Cytoskeleton and adhesion module	64	20 (31 %)
Cytoskeleton remodeling	Integrin outside-in signalling	49	17 (35 %)
Immune response	Function of MEF2 in T lymphocytes	51	17 (33 %)
Not assigned	Causal network (positive)	36	14 (39 %)
Cytoskeleton remodeling	Regulation of actin cytoskeleton nucleation and polymerization by Rho GTPases	46	16 (35 %)
Development	TGF-β-dependent induction of EMT via RhoA, PI3K and ILK	46	16 (35 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Number of identified markers of given pathway through the analysis of generated omic profiles

4.2.1.3.5 Comparison of significant gene expression changes induced in both cell lines models upon stimulation with TGF-β

Both cell line models were treated according to the same treatment regime, including synchronised media changes, TGF-β concentrations and sample collection time. Both cell lines have shown molecular changes associated with epithelial to mesenchymal transition as well as morphological changes associated with a more elongated cell morphology (Chapter III).

To see the similarity of molecular changes within both cell line models, the number of overlapping genes between both cell lines and their expression directionality were investigated. In total 1173 genes were significantly detected in both cell line models, of which 699 genes were upregulated and 365 genes were downregulated in both cell lines (Fig. 4.8C). 109 of the significant genes showed an inverse regulation, which means that an upregulation occurred in one cell line which presented itself as a downregulation in the other cell line, and *vice versa.*

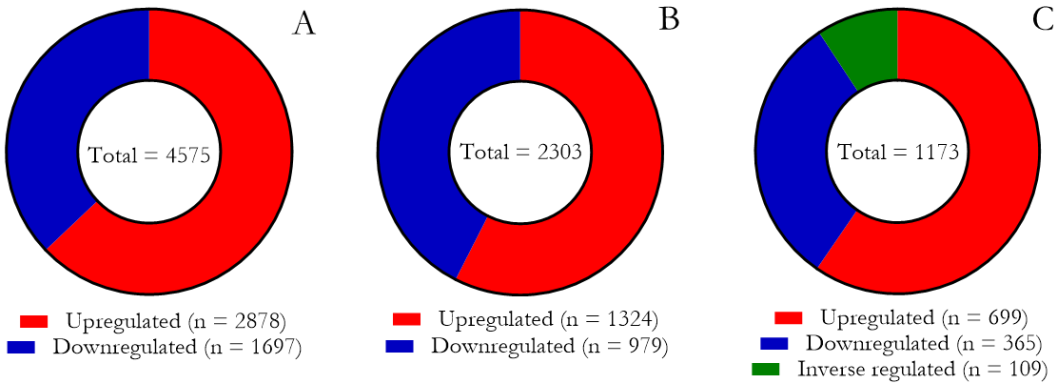


Figure 4.8: Significantly deregulated genes across both cell line models . P5B3 treated with P5B3 untreated (n = 4575) (A) and DU145 treated with DU145 untreated (n=2303) (B) cell lines model (p-value below 0.05 after Bonferroni correction). C represents shared significant genes (n = 1173) between both models. Red indicates an upregulation, blue a downregulation and green an inverse change of expression comparing both models with each other.

Hierarchical clustering was applied for the investigation of a correlation between the cell line models and to infer whether the relationship of the gene expression is stronger between the cell lines or the treatment. The generated heat map (Fig. 4.9) shows a clustering of the treated samples together, with a sub-clustering according to their respective cell line and treatment.

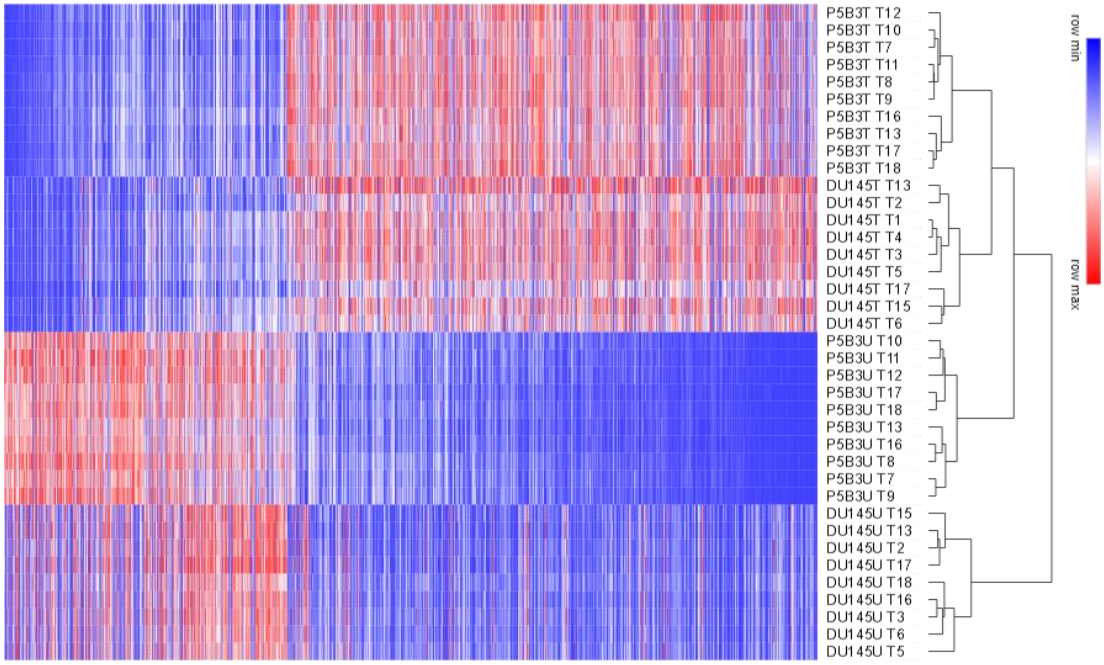


Figure 4.9: Hierarchical clustering of significant genes shared between both cell line models (n = 1173). The clustering was performed using complete linkage and Euclidean distance.

4.2.1.3.6 Identification of shared pathways altered upon stimulation of P5B3 and DU145 with TGF-β based on significant deregulated genes

To analyse further commonalities across the molecular composition of both cell lines, the enriched pathways defined through MetaCore™ were analysed for any overlap. The comparison of the top 15 most enriched pathways [Tab. 4.4 (P5B3) and Tab. 4.5 (DU145)] showed 5 pathways, which were present in both cell line models. Of these 5 pathways, 2 are related to the activation of signalling pathways through TGF-β, furthermore another pathway is involved in the remodelling of the extracellular matrix (ECM) and the IGF family involved in invasion and metastasis in colorectal cancer.

Table 4.6: Shared enriched pathways within the top 15 pathways of both cell lines . Gene numbers per pathway and number of detected genes are shown and the coverage of genes within the pathway through the defined gene list is indicated as a %. List derived from Metacore™ (accessed 02/07/18).

Category	Pathway	Total[1] (Genes)	In data[2] (P5B3)	In data[2] (DU145)
Development	TGF-β-dependent induction of EMT via RhoA, PI3K and ILK	46	33 (72 %)	16 (35 %)
Cell adhesion	ECM remodelling	55	35 (64 %)	20 (36 %)
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	37 (64 %)	23 (40 %)
Not assigned	TGF-β 1-induced transactivation of membrane receptors signalling in HCC	50	31 (62 %)	19 (38 %)
Not assigned	IGF family, invasion and metastasis in colorectal cancer	33	22 (67 %)	16 (48 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Number of identified markers of given pathway through the analysis of generated omic profiles

4.2.2 Mass spectrometry analysis of cell lysates generated from treated and untreated P5B3 and DU145 cells

For the identification of differentially regulated proteins in P5B3 and DU145 through the treatment with TGF-β, a label-free quantitative mass spectrometry analysis was performed on whole cell lysates. Here, the prepared sample material was further processed by Dr David Boocock and Dr Clare Coveney on a Sciex TripleTOF6600 mass spectrometer using two analysis methods enabling the quantitative analysis of the proteome. For this, initially a spectral library, based on data-dependent acquisition, was generated using

pooled sample material of each sample. This library was used downstream for the identification of peaks generated during the data-independent approach (SWATH-MS).

For the library generation, previously generated cytoplasmic and nuclear fractions, extracted according to the Abcam subcellular fractionation protocol (Abcam, 2019) were included for an increased coverage of potentially present peptides. These fractions were previously generated as part of a separate study by Dr Jayakumar Vadakekolathu, follow on work to (Harner-Foreman, Vadakekolathu et al. 2017) (data not shown). The generation of the spectral library resulted in the detection of 2448 proteins, comprising of 27981 peptides. This, however, includes shared peptides, which could be originating from multiple different proteins. These shared peptides, and also proteins whose identification was based solely on shared peptides, were excluded. For this reason, the further analysis was performed with a list of 2197 proteins. Each protein identification was based on up to 6 single unique peptides.

Table 4.7: Summary of detected proteins and peptides within the analysed sample set of both cell line models. Untreated and treated P5B3 (n=10) and untreated (n=9) and treated (n=8) DU145.

Sample	**Proteins (unique peptides)**	**Proteins**	**Peptides**	**Spectra**
Library of pooled samples	2197	2448	27981	63469

4.2.2.1 Validation of EMT protein panel in generated mass spectrometry-derived protein expression profiles

As previously performed in the analysis of the generated RNA-sequencing data, the generated proteomic data was analysed for their EMT profile. For this, the normalised protein peak areas of EMT-associated proteins were selected for both conditions and cell lines and analysed for their differences in intensities. Here, VIME, CADH1 and FINC were detected in both cell line models. Figure 12 shows the expression changes of these proteins in both cell lines. A significant deregulation of VIME, CADH1 and FINC was detected in P5B3, however in DU145 only VIME showed a significant change induced through the treatment (Fig. 4.10).

Overall, the changes indicate an induction of an EMT profile on a proteomic level, however in the case of the cell line model of DU145, these changes were not as clear as in the model of P5B3.

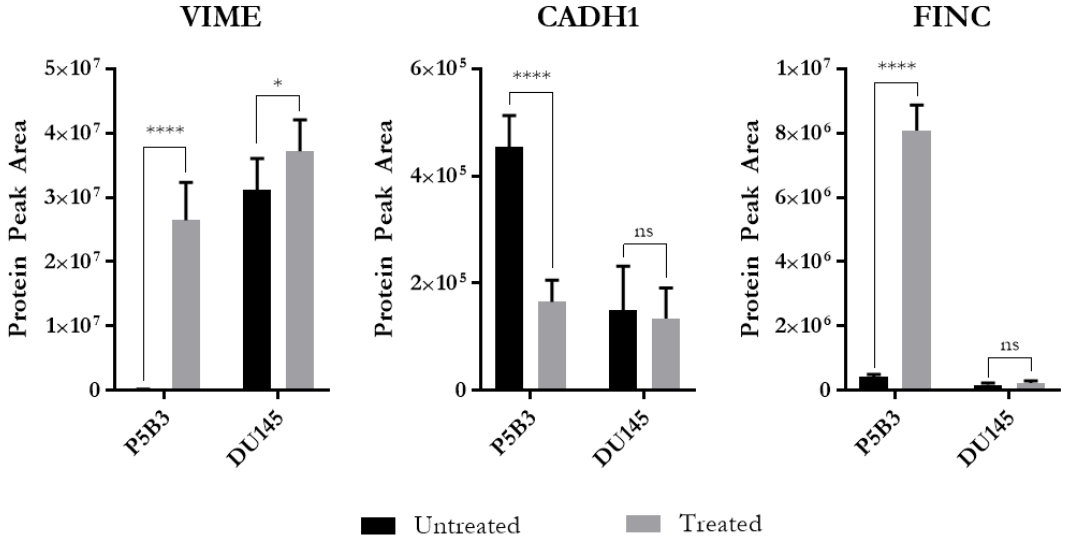


Figure 4.10: Proteomic changes of vimentin, E-cadherin and fibronectin in cell lysates of untreated and treated P5B3 (n=10) and DU145[n= 9 (DU145U), n=8 (DU145T)] analysed with the Sciex TripleTOF 6600 using a data-independent acquisition mode (SWATH MS). The data is presented using the normalised protein peak area of each protein across the sample population.

4.2.2.2 Analysis of mass spectrometry derived protein expression profiles of both cell line models for the characterisation of underlying pathway changes

For the further downstream analysis, proteins that presented a significant difference between the treated and untreated condition after correction for false discovery were selected. The statistical analysis performed is described in Methods (section 2.2.6.4)

4.2.2.2.1 Identification of significant altered proteins within the inducible EMT model of P5B3

The proteins detected in both P5B3 cell states were analysed as described previously. This resulted in the detection of 297 significantly altered proteins, of which 167 were up- and 130 downregulated (Fig. 4.17A). The protein peak areas of these proteins were applied to hierarchical clustering, which confirmed the clustering of the samples according to treatment type (Fig. 4.11) (section 2.2.7.2). Furthermore, it highlighted no distinct outliers within the population. Despite this, the generated heat map highlights variation across the samples of one treatment group.

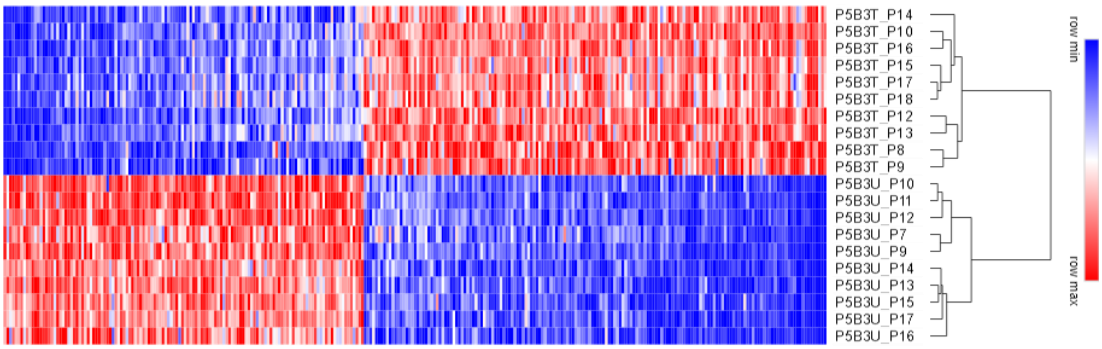


Figure 4.11: Hierarchical clustering of significantly altered proteins in untreated and treated P5B3 (n=10) using complete linkage and Euclidean distance (n=297).

The comparison of protein expression across both conditions showed a wide range of expression changes with fold changes ranging from +184.84 to -39.42. The top 3 most upregulated proteins were VIME (+181.84), FINC (+19.38) and BGH3 (+15.05), whereas the top 3 downregulated proteins were K1C13 (-39.42), LY6D (-6.02) and AGR2 (-5.91).

4.2.2.2.2 Analysis of pathways altered upon stimulation of P5B3 with TGF-β based on significant deregulated proteins

After the application of the significant protein list to the MetaCore™ pathway analysis tool from Clarivate Analytics, 96 pathways were significantly enriched. Within the top 50, "cytoskeletal remodelling" was the most frequent pathway category. The second most common category was "cell adhesion" (Fig. 4.12). This indicates a strong involvement of the protein expression with the morphological changes of the cells and the adjustment of their cellular changes to a more mesenchymal cell state.

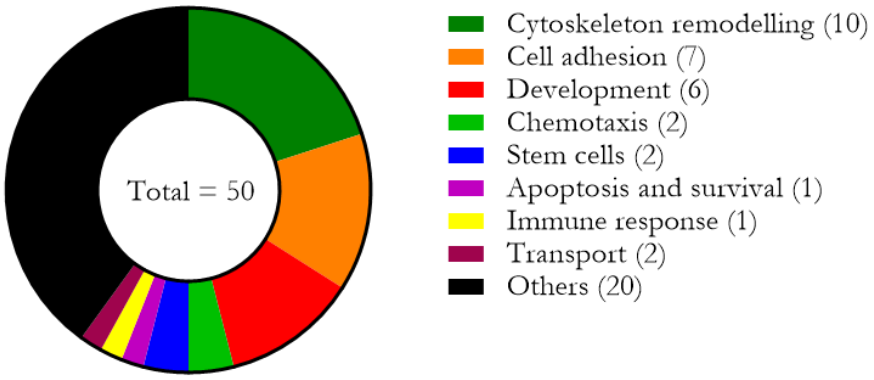


Figure 4.12: Top 50 most significantly enriched pathways based on significant proteins in P5B3 grouped by their respective categories (n=297). List derived from Metacore™ (accessed 02/07/18).

The top 15 most significant pathways enriched through the supplied protein list are shown in Table 4.8. The majority of the defined pathways are associated with "Cytoskeletal remodelling" and "Cellular adhesion", highlighting changes in the cellular morphology through the stimulation with TGF-β. However, also the "EMT induction via the RhoA, PI3K and ILK" pathway is present within the list, showing changes within the system that reach further than morphological alterations, such as reduced adhesion and elongated cell shapes.

Table 4.8: Top 15 most enriched pathways identified through the protein list of P5B3. List derived from Metacore™ (accessed 02/07/18).

Category	**Pathway**	**Total[1]**	**In data[2]**
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	19 (33 %)
Cytoskeleton remodelling	Keratin filaments	36	13 (36 %)
Not assigned	Inhibition of re-myelination in multiple sclerosis: regulation of cytoskeleton proteins	44	12 (27 %)
Development	Regulation of cytoskeleton proteins in oligodendrocyte differentiation and myelination	58	13 (22 %)
Cell adhesion	Histamine H1 receptor signalling in the interruption of cell barrier integrity	45	11 (24 %)
Development	TGF-β-dependent induction of EMT via RhoA, PI3K and ILK	46	11 (24 %)
Cell adhesion	Integrin-mediated cell adhesion and migration	48	11 (23 %)
Not assigned	LRRK2 in neurons in Parkinson's disease	33	9 (27 %)
Not assigned	Cytoskeleton and adhesion module	64	11 (17 %)
Not assigned	Effect of H. pylori infection on gastric epithelial cells motility	43	9 (21 %)
Apoptosis and survival	NGF/ TrkA PI3K-mediated signalling	77	11 (21 %)
Cytoskeleton remodelling	Neurofilaments	25	7 (28 %)
Transport	The role of AVP in regulation of Aquaporin 2 and renal water reabsorption	50	9 (18 %)
Chemotaxis	Inhibitory action of lipoxins on IL-8- and Leukotriene B4-induced neutrophil migration	53	9 (17 %)
Cell adhesion	Gap junctions	30	7 (23 %)

[1] Total: Total number of markers present in the pathway

[2] In data: Number of identified markers of given pathway through the analysis of generated omics profiles

4.2.2.2.3 Comparison of pathways altered upon stimulation of P5B3 with TGF-β based on significant deregulated genes and proteins

In further analyses, the top 15 enriched pathways of both P5B3 lists, genes and proteins, were compared and this highlighted two shared pathways (Tab. 4.9). One was involved in the "Regulation of actin cytoskeleton organisation by the kinase effectors of Rho GTPases" and the other one in "TGF-beta-dependent induction of EMT via RhoA, PI3K and ILK". The pathway "Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases" will be discussed further in subsection 4.4.6.

Table 4.9: Shared enriched top 15 pathways between gene and protein P5B3. List derived from Metacore™ (accessed 02/07/18).

Category	Pathway	Total[1]	In data[2]	Genes[3]	Proteins[4]
Cytoskeleton remodelling	Regulation of actin cytoskeleton organisation by the kinase effectors of Rho GTPases	58	40 (69 %)	37 (64 %)	19 (33 %)
Development	TGF-β-dependent induction of EMT via RhoA, PI3K and ILK	46	33 (72 %)	33 (72 %)	11 (24 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Total number of identified markers of given pathway through the analysis of generated omic profiles
[3] Genes: Total number of genes identified in given pathway through the analysis of generated omic profiles
[4] Proteins: Total number of proteins identified in given pathway through the analysis of generated omic profiles

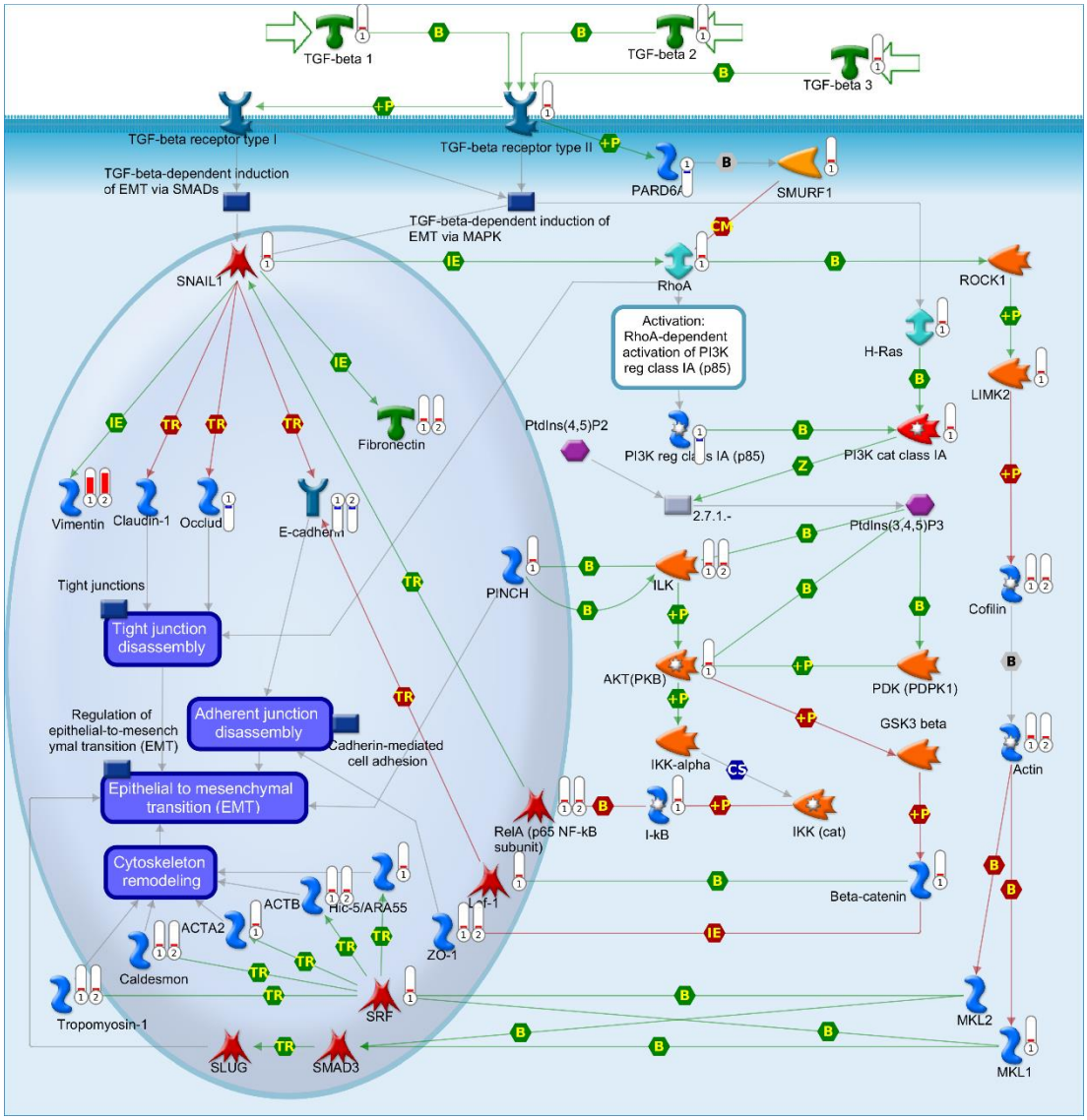


Figure 4.13: Schematic representation of the pathway describing "TGF-beta-dependent induction of EMT via RhoA, PI3K and ILK". Pathway image was generated with MetaCore™ one-click analysis (Metacore™ accessed 02/07/18). Markers covered within either of the lists, genes or proteins, are highlighted with intensity bars representing the induced fold change. Red bars indicate an upregulation and blue bars a downregulation of the gene/protein and the numbers 1 (genes) and 2 (proteins).

Overall, a high number of binding proteins (blue symbols) was observed to be covered by both the genes and proteins identified in the analysis.

4.2.2.2.4 Identification of significant altered proteins within the inducible EMT model of DU145

The proteins detected in both DU145 cell states were analysed as described previously. This resulted in the detection of 187 significantly altered proteins, of which 93 were up- and 94 downregulated (Fig. 4.17B). The protein peak areas of these proteins were applied to hierarchical clustering (complete linkage, Euclidean distance), which confirmed the clustering of the samples according to treatment type (Fig. 4.14), indicating no distinct outliers within the population. Furthermore, the heat map shows a stronger variation between the significant altered proteins across the samples, compared to the previously described transcriptomic changes.

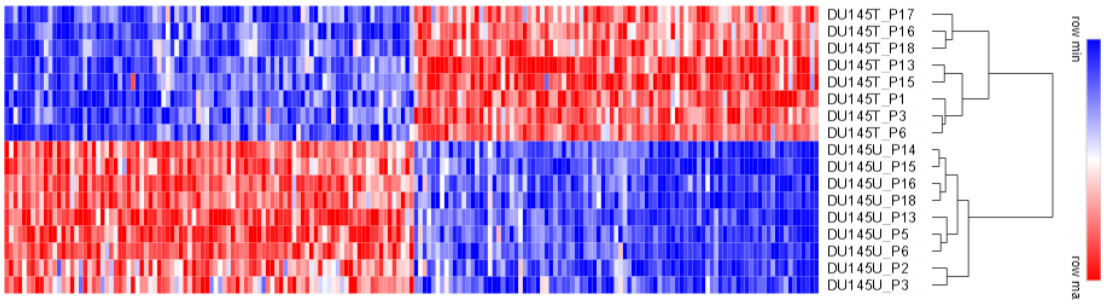


Figure 4.14: Hierarchical clustering of significantly altered proteins in untreated and treated DU145 (n=9 (DU145U), n=8 (DU145T)) using complete linkage and Euclidean distance (n=178).

Also in this analysis, the intensity of protein expression varied greatly between the treated and untreated cells. The two strongest increased proteins were TAGL (8.51) and ITAV whereas one of the most downregulated proteins was TADC2 (-3.66).

4.2.2.2.5 Analysis of pathways altered upon stimulation of DU145 with TGF-β based on significant deregulated proteins

The analysis of significant altered proteins in DU145 with MetaCore™ highlighted 82 significantly enriched pathways. Figure 4.15 shows the associated groups of the top 50 pathways. The top two most frequent categories were associated with "cytoskeleton remodelling" and "cell adhesion", however pathways indicating developmental changes were also enriched.

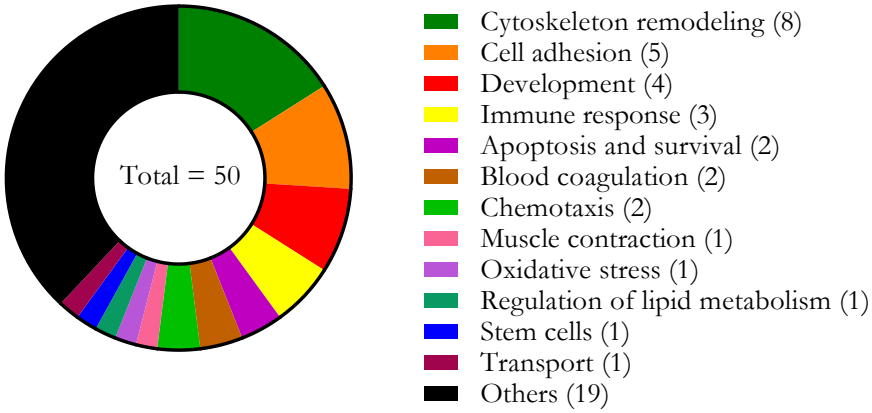


Figure 4.15: Top 50 most significantly enriched pathways based on significant proteins in DU145 grouped by their respective categories (n=187). List derived from Metacore™ (accessed 02/07/18).

Also here the top 15 enriched pathways were identified and are shown in Table 4.10. A strong enrichment of pathways associated with cytoskeleton remodelling and cell adhesion is apparent, which might indicate stronger changes in proteins associated with these processes.

Table 4.10: Top 15 most enriched pathways identified through the protein list of DU145. List derived from Metacore™ (accessed 02/07/18).

Category	**Pathway**	**Total**[1]	**In data**[2]
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	16 (28 %)
Development	Regulation of cytoskeleton proteins in oligodendrocyte differentiation and myelination	58	11 (19 %)
Not assigned	Inhibition of re-myelination in multiple sclerosis: regulation of cytoskeleton proteins	44	10 (23 %)
Cell adhesion	Integrin-mediated cell adhesion and migration	48	10 (21 %)
Cytoskeleton remodelling	Role of PKA in cytoskeleton reorganisation	41	8 (20 %)
Cell adhesion	Histamine H1 receptor signalling in the interruption of cell barrier integrity	45	8 (18 %)
Transport	The role of AVP in regulation of Aquaporin 2 and renal water reabsorption	50	8 (16 %)
Not assigned	Regulation of degradation of deltaF508-CFTR in CF	39	7 (18 %)
Apoptosis and survival	NGF/ TrkA PI3K-mediated signalling	77	9 (12 %)
Cytoskeleton remodelling	Substance P mediated membrane blebbing	16	5 (31 %)
Not assigned	Cytoskeleton and adhesion module	64	8 (13 %)
Not assigned	LRRK2 in neurons in Parkinson's disease	33	6 (18 %)
Chemotaxis	Inhibitory action of lipoxins on IL-8- and Leukotriene B4-induced neutrophil migration	53	7 (13 %)
Not assigned	Possible regulation of HSF-1/ chaperone pathway in Huntington's disease	21	5 (24 %)
Development	MAG-dependent inhibition of neurite outgrowth	37	6 (16 %)

[1] Total: Total number of markers present in the pathway

[2] In data: Number of identified markers of given pathway through the analysis of generated omic profiles

4.2.2.2.6 Comparison of pathways altered upon stimulation of DU145 with TGF-β based on significant deregulated genes and proteins

The top 15 enriched pathways between the significant gene and proteins lists of DU145 were compared and two pathways were shown to be present in both (Tab. 4.11). One is the "regulation of actin cytoskeleton organisation by the kinase effectors of Rho GTPases" and the other one is the "cytoskeleton and adhesion module". The first one will be discussed further in subsection 4.4.6.

Table 4.11: Shared enriched pathways in the top 15 between the gene and protein lists of DU145. List derived from Metacore™ (accessed 02/07/18).

Category	Pathway	Total[1]	In data[2]	Genes[3]	Proteins[4]
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	28 (48 %)	23 (40 %)	16 (28 %)
Not assigned	Cytoskeleton and adhesion module	64	24 (38 %)	20 (31 %)	8 (13 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Total number of identified markers of given pathway through the analysis of generated omic profiles
[3] Genes: Total number of genes identified in given pathway through the analysis of generated omic profiles
[4] Proteins: Total number of proteins identified in given pathway through the analysis of generated omic profiles

The pathway of "cytoskeleton and adhesion module" (Fig. 4.16) identified genes and proteins within the signalling resulting in cytoskeleton remodelling (orange square) and survival. A strong overlay of identified proteins and genes can be observed, highlighting the correlation between the generated transcriptomic and proteomic profiles.

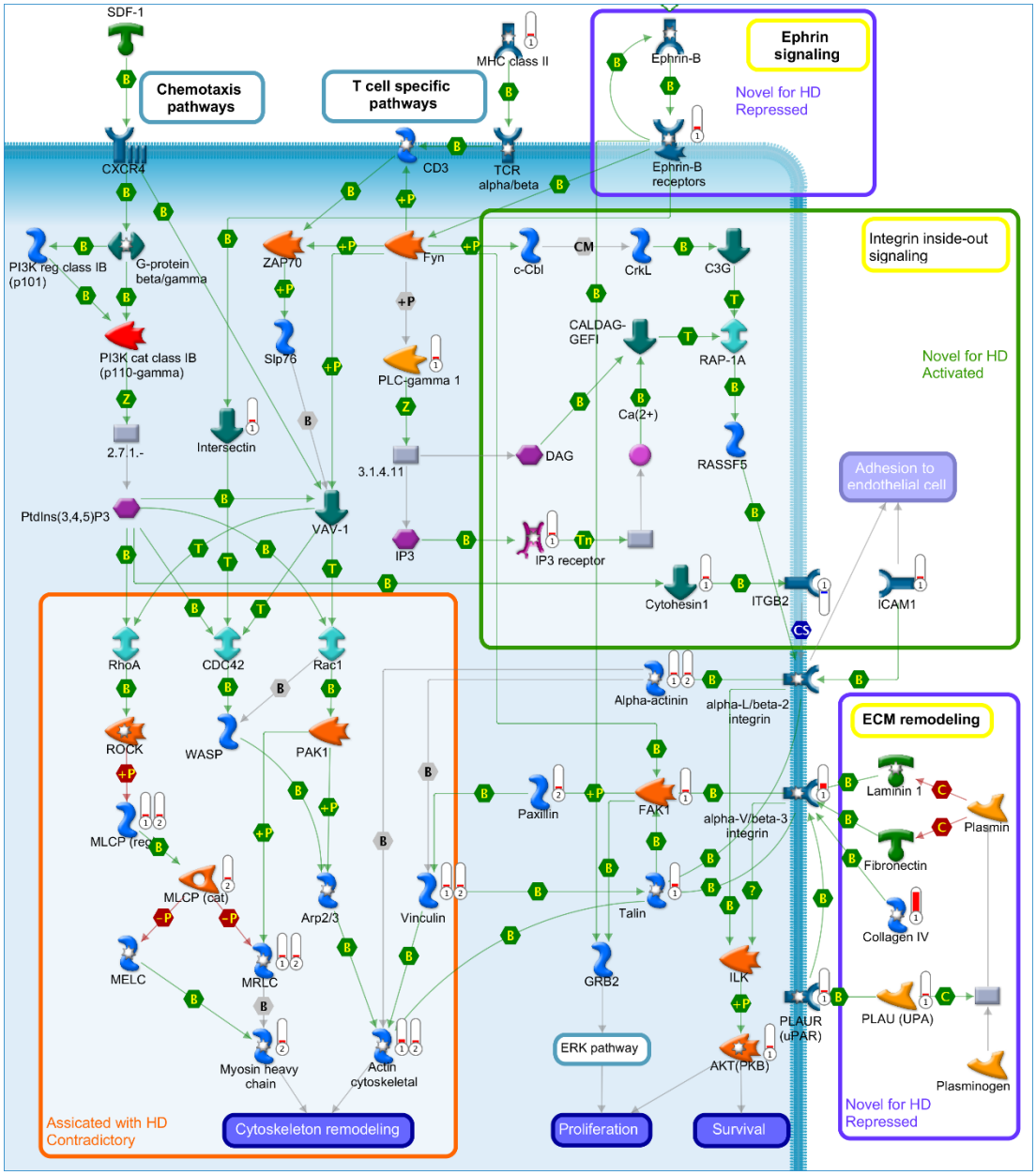


Figure 4.16: Schematic representation of the pathway describing "Cytoskeleton and adhesion module" enriched in both gene and protein lists, of DU145. Pathway image was generated with MetaCore™ one-click analysis (Metacore™ accessed 02/07/18). Markers covered within either of the lists, genes or proteins, are highlighted with intensity bars representing the induced fold change. Red bars indicate an upregulation and blue bars a downregulation of the gene/protein and the numbers 1 (genes) and 2 (proteins). HD=Huntington's disease.

4.2.2.2.7 Comparison of significant protein expression changes induced in both cell line models upon stimulation with TGF-β

To see the similarity of proteomic changes within both cell line models, it was investigated how many significant proteins overlap between both cell lines and whether these overlapping proteins show the same expressional changes through the treatment. In total 89 genes were significantly detected in both cell line models (Fig. 4.17C), of which 55 genes were upregulated and 24 genes were downregulated. 10 of these significant proteins showed an inverse regulation, which means that an upregulation occurred in one cell line, which presented itself in a downregulation in the other cell line, and *vice versa.*

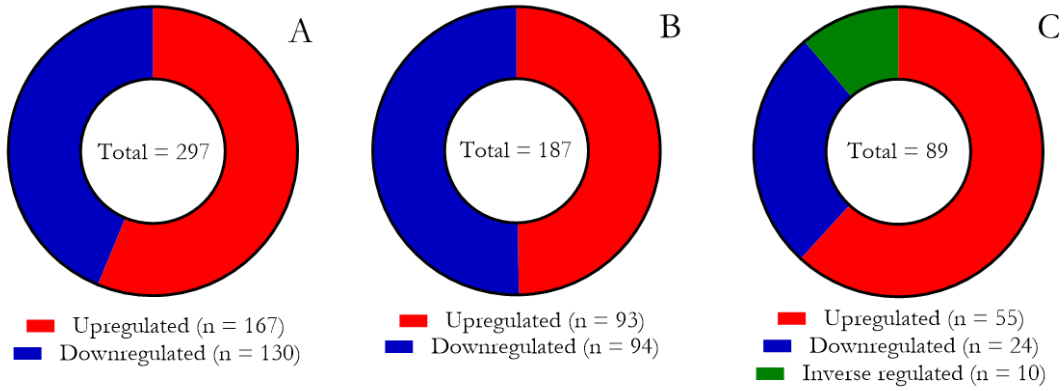


Figure 4.17: Significantly deregulated proteins within both cell lines models. P5B3 (A) and DU145 (B) (p-value below 0.05 after Bonferroni correction). C represents shared significant proteins between both models. Red indicates an upregulation, blue a downregulation and green an inverse regulation between the two models.

For further confirmation of the induction of EMT and the correlation of both cell line models, a hierarchical clustering approach was applied to the significant protein expression data shared between both models. For this, complete linkage and Euclidean distance was used. The generated heat map (Fig. 4.18) shows a clustering of the samples according to the cell type. This means, the samples of P5B3 separate into treated and untreated samples, as well as the samples of DU145, in which treated and untreated samples cluster together. However, within both cell lines, the treatment conditions present further clustering. This might be indicative that the significant protein changes within the cell lines are impacted stronger by the individual response of the cell line to the treatment than the induction of the particular pathways.

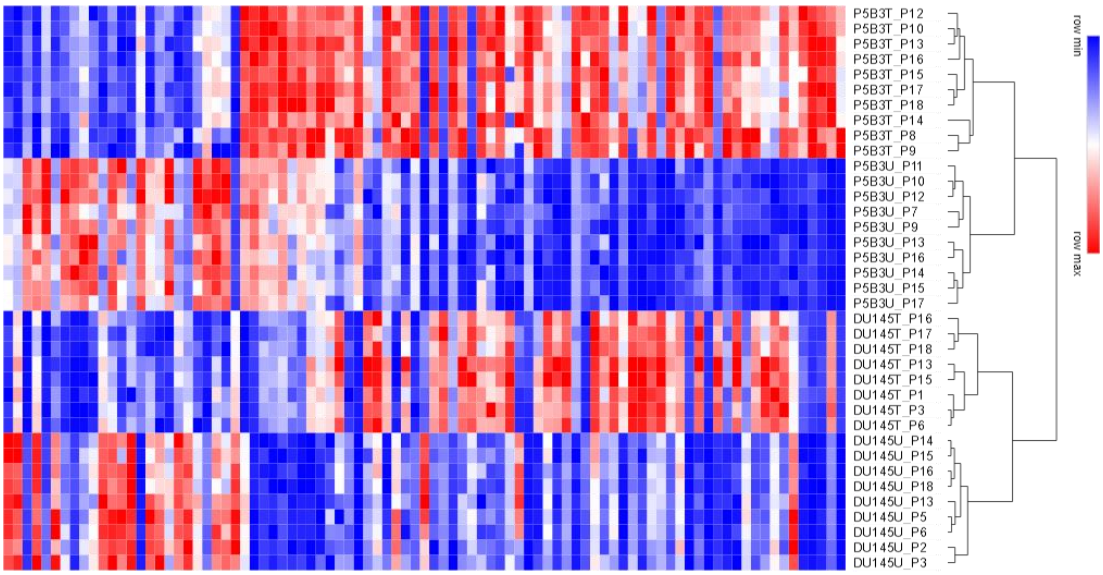


Figure 4.18: Hierarchical clustering of significantly altered proteins shared across both cell line models (n=89) in untreated and treated P5B3T (n=10), P5B3U (n=10) and DU145 (n=9 (DU145U), n=8 (DU145T)) using complete linkage and Euclidean distance.

4.2.2.2.8 Identification of shared pathways altered upon stimulation of P5B3 and DU145 with TGF-β based on significant deregulated proteins

The identification of shared pathways between P5B3 and DU145, identified through the analysis of proteomic profiles, highlighted also here a strong enrichment of pathways associated with the cytoskeleton, and the process of cell adhesion (Tab. 4.12).

Table 4.12: Shared enriched pathways within the top 15 pathways of both cell lines. Protein numbers per pathway and number of detected genes are shown and also indicated as a %. List derived from Metacore™ (accessed 02/07/18).

Category	Pathway	Total[1] (Proteins)	In data[2] (P5B3)	In data[2] (DU145)
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	19 (33 %)	16 (28 %)
Not assigned	Inhibition of re-myelination in multiple sclerosis: regulation of cytoskeleton proteins	44	12 (27 %)	10 (23 %)
Not assigned	LRRK2 in neurons in Parkinson's disease	33	9 (27 %)	6 (18 %)
Cell adhesion	Histamine H1 receptor signalling in the interruption of cell barrier integrity	45	11 (24 %)	8 (18 %)
Cell adhesion	Integrin-mediated cell adhesion and migration	48	11 (23 %)	10 (21 %)
Development	Regulation of cytoskeleton proteins in oligodendrocyte differentiation and myelination	58	13 (22 %)	11 (19 %)
Transport	The role of AVP in regulation of Aquaporin 2 and renal water reabsorption	50	9 (18 %)	8 (16 %)
Not assigned	Cytoskeleton and adhesion module	64	11 (17 %)	8 (13 %)
Chemotaxis	Inhibitory action of lipoxins on IL-8- and Leukotriene B4-induced neutrophil migration	53	9 (17 %)	7 (13 %)
Apoptosis and survival	NGF/ TrkA PI3K-mediated signalling	77	11 (14 %)	9 (12 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Number of identified markers of given pathway through the analysis of generated omic profiles

4.2.2.2.9 Identification of unique shared pathways identified across both omic levels of P5B3 and DU145 following treatment with TGF-β

Overall, one pathway was found to be enriched between both omic levels and cell line models (Tab. 4.13). This pathway is associated with the "Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases". The pathway is shown in Fig. 4.19, whereas the numbers 1 and 2 indicate genes and proteins of DU145, and 3 and 4 indicate genes and proteins of P5B3.

Table 4.13: Single shared pathway between all datasets and cell line models. List derived from Metacore™ (accessed 02/07/18).

Category	Pathway	Total[1]	In data[2] (P5B3)	In data[2] (DU145)
Cytoskeleton remodelling	Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases	58	40 (69 %)	28 (48 %)

[1] Total: Total number of markers present in the pathway
[2] In data: Number of identified markers of given pathway through the analysis of generated omic profiles

Overall it can be seen that a strong concordance of expression across P5B3 and DU145 exists. 10 features within the pathway, namely caldesmon, MRLC, ERM proteins, moesin, vinculin, alpha actin, talin, actin cytoskeletal, F-actin cytoskeletal and filamin A, showed a concordant expression across all 4 omic profiles (P5B3 genes and proteins, DU145 genes and proteins), showing the same expression directionality. However, as illustrated in Fig. 4.19, 3 markers (RAC2, PRK1 and ARPC1) were detected in at least one profile of both cell lines, and all 3 show a reduced expression in DU145, whereas their expression was increased in P5B3.

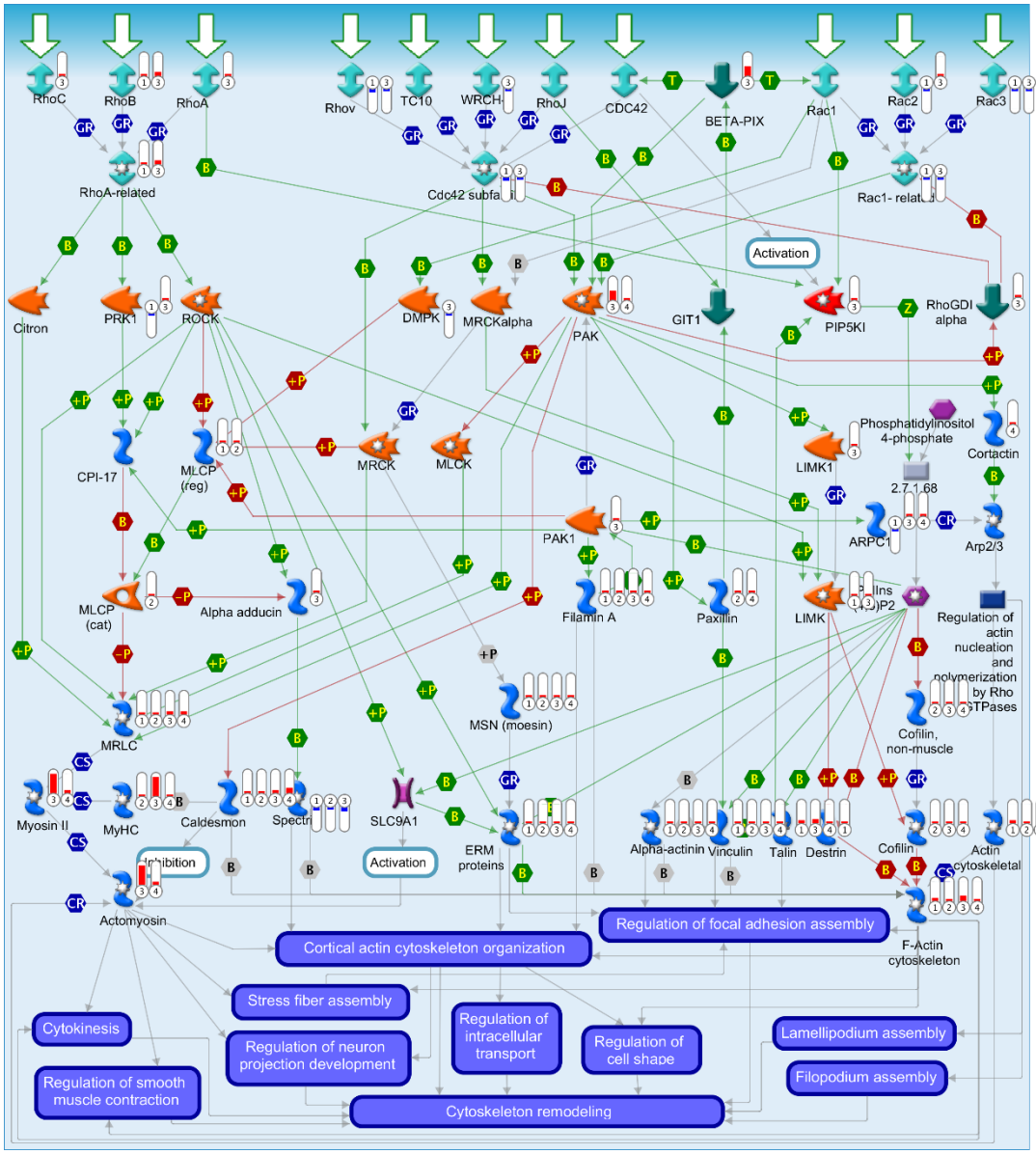


Figure 4.19: Schematic representation of the pathway describing "Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases" enriched in both omic levels of both EMT models. Pathway image was generated with MetaCore™ one-click analysis (Metacore™ accessed 02/07/18). Markers covered within either of the lists, genes or proteins, are highlighted with intensity bars representing the induced fold change. Red bars indicate an upregulation and blue bars a downregulation of the gene/protein. The different omic levels are indicated with numbers 1-4; DU145 genes (1), DU145 proteins (2), P5B3 genes (3), P5B3 proteins (4).

4.2.3 Improved correlation of transcriptomic and proteomic changes through parallel treatment and harvest

A common question is the potential correlation between the gene and resulting protein expression. The R^2 of an analysis can indicate a presence or absence of any correlation between 2 given factors, such as the gene and protein expression. Initial studies comparing the gene and protein expression in yeast have not shown any correlation between both expression levels (Gygi, Rochon et al. 1999), more recent studies in human circulating monocytes and NIH3T3 cells (primary mouse embryonic fibroblasts) have shown significant correlations presenting an R^2 of 0.41 and 0.235 respectively (Schwanhäusser, Busse et al. 2011, Guo, Xiao et al. 2008). Therefore, one major question of the study was whether a matching growth and harvest of proteins and RNA of both models will improve the correlation between these omic levels.

For the analysis of potential correlations between the gene and protein expression levels, only significantly deregulated genes and proteins were selected, which were detected at both omic levels. This was done to reduce the generated noise across the samples. The correlation analysis was performed using Pearson correlation. The applied gene and proteins expression values were min-max normalised across the complete selection of significant genes.

In P5B3, it can be shown that the expression of both untreated (Fig. 4.20A) and treated (Fig. 4.20B) samples show a significant correlation with a p-value below 0.0001 and R^2 above 0.55 for both conditions. A correlation of the fold changes (Fig. 4.20C) shows an even higher correlation of 0.80. However, the outliers VIME and ANXA6 were excluded to calculate the correlation in the fold change without their impact. This resulted in a slight decrease in the correlation to 0.76 (Fig. 4.20D). Both calculations were significant.

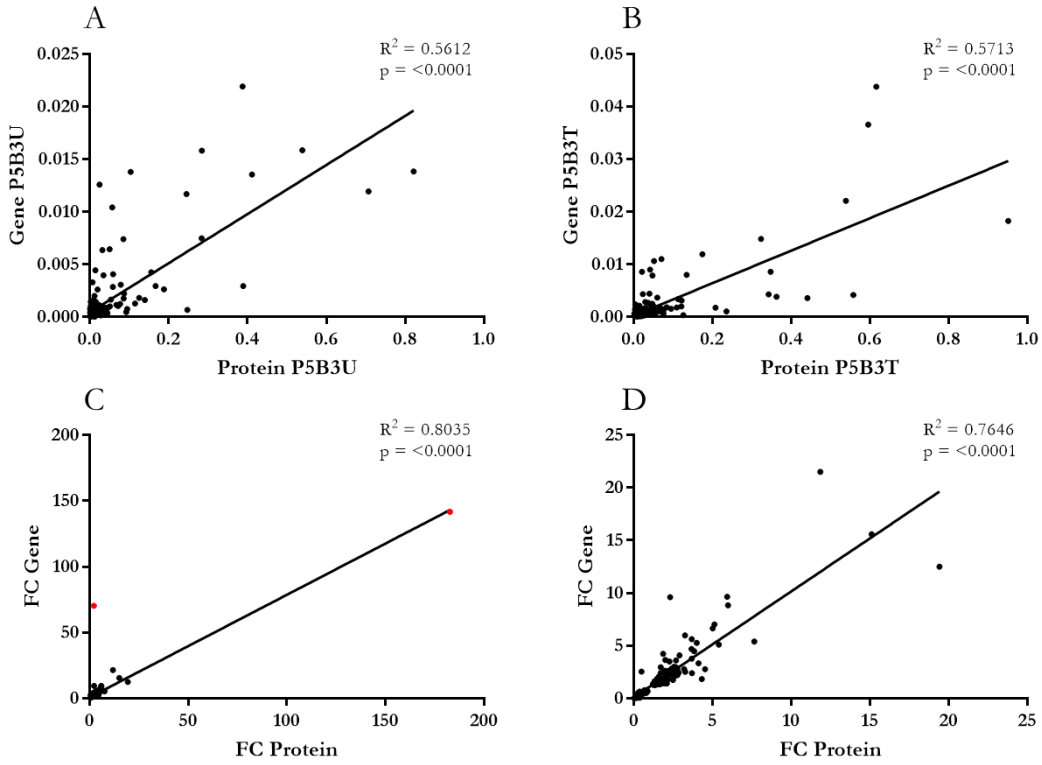


Figure 4.20: Pearson correlation of gene and protein expression between both treatment conditions of P5B3. Untreated P5B3 (n= 10) (A), treated P5B3 (n=10) (B). C shows the correlation of the fold change between gene and protein expression. Two markers were highlighted in C and removed in D. These two markers were excluded due to their outlier nature in order to generate a more realistic correlation analysis.

The same approach was applied to the generated profiles of DU145. Here, limited correlations between the untreated (Fig. 4.21A) and treated (Fig. 4.21B) gene and protein profiles with R^2 of 0.088 and 0.094 respectively were observed. An increase in correlation can be observed through the comparison of fold changes (Fig. 4.21C), which increases the R^2 to 0.80. As previously shown for P5B3, all correlation analyses were statistically significant.

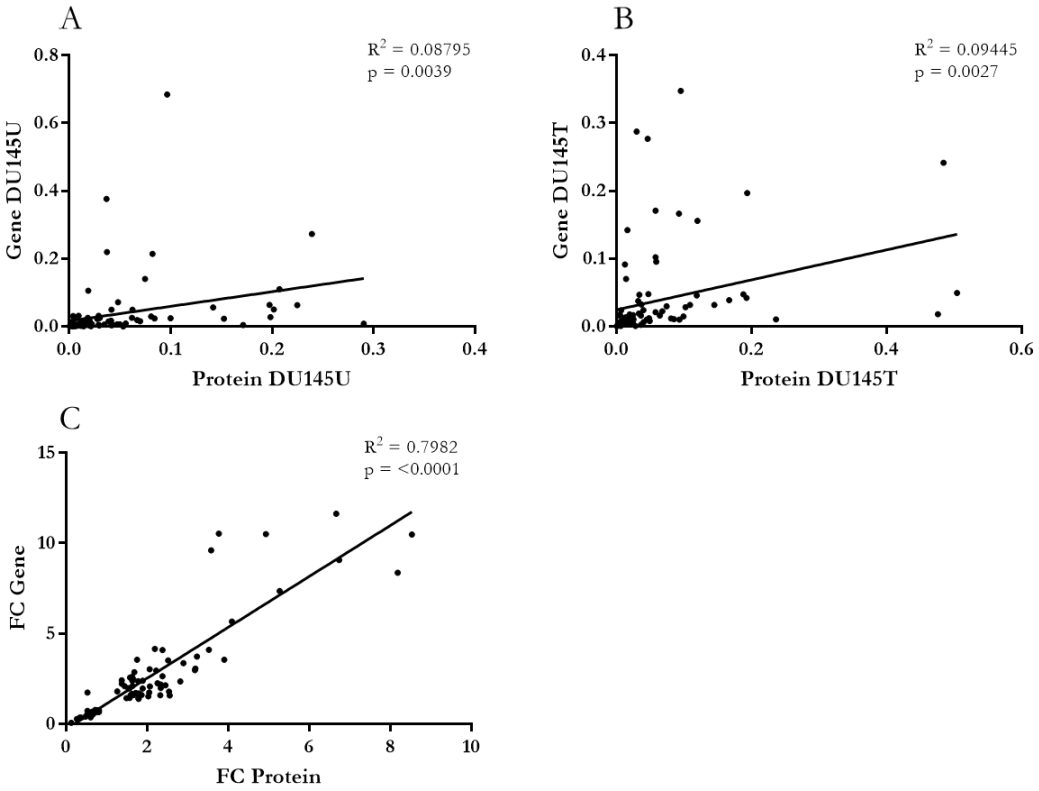


Figure 4.21: Pearson correlation of gene and protein expression between both treatment conditions of DU145. Untreated DU145 (n=9) (A) and treated DU145 (protein: n=8, gene: n=8) (B). C shows the correlation of the fold change between gene and protein expression.

4.4 Discussion

4.4.1 Data quality and considerations through RNA-sequencing and mass spectrometry analysis

The aim of this chapter was the generation and characterisation of matching transcriptomic and proteomic profiles of two models for EMT using the prostate cancer cell lines P5B3 (Harner-Foreman, Vadakekolathu et al. 2017) and DU145.

The analysis of the transcriptome resulted in the detection and quantification of approximately 26 000 genes, of which a larger proportion were significantly altered in P5B3 (n=4575) compared to DU145 (n=2303) (Fig. 4.8A+B). A similar picture was shown at the proteomic level, where approximately 2200 proteins were quantified, and a larger proportion were significantly altered in P5B3 (n=297) compared to DU145 (n=187) (Fig. 4.17A+B). These differences could be based on the response of both cell lines to the stimulation with TGF-β. In P5B3, morphologically all cells responded to the treatment, whereas only a subpopulation of DU145 changed through TGF-β stimulation. Based on the limited response, the intensity of transcriptomic and proteomic changes might have been diluted by the unaltered expression of the non-responsive cells.

Aside from the differences between the two cell lines, a large discrepancy was also shown in the number of detected genes compared to proteins. In an ideal world, matching profiles of translated genes and proteins would present a full coverage of matching gene and protein expression data. This would enable a highly informative characterisation of phenotypic changes and post-translational influences on the steps from gene to protein. However, large differences between the number of detected genes (~26 000) and proteins (~2200) are apparent. Furthermore, with an RNA-sequencing approach, not only coding, but also non-coding genes which don't result in a protein product are detected. Another large contribution to this is based on the limitations of current mass spectrometry methods. Proteins show a large order of magnitude across their expression, and this range can be over more than 10 orders of magnitude, whereas the gene expression ranges are only over 3 to 4 orders of magnitude (Zubarev 2013). Compared to mass spectrometry, RNA-sequencing technology is able to quantify low abundance gene products. The wide spread of protein abundances increases the difficulty in confidently detecting and quantifying all present proteins. However, recent developments in the technology of mass spectrometry analysis based on instrumentation improvements, and advances in

methodologies have enabled nowadays an increased routine quantification of proteins to approximately 4000 to 5000 proteins (Hülsmann, Kravic et al. 2018, He, M., Gou et al. 2018, Bruderer, Bernhardt et al. 2015, Shishkova, Hebert et al. 2016).

Based on the limited number of detected proteins, a correlation analysis was only possible for a small number of significantly detected and matching genes and proteins. Of the identified genes and proteins, the overlap presented a significant correlation in both models and conditions, however in DU145 (Fig. 4.21), the expression showed a large amount of variation and noise across the genes and proteins, for which reason the R^2 remained very low. Therefore, an improved R^2 compared to previously published information (Schwanhäusser, Busse et al. 2011, Guo, Xiao et al. 2008) could only be shown for P5B3 (Fig. 4.20). This increased correlation could be supported by the fact that P5B3 is a highly homogeneous cell line, based on a single cell clone. Therefore, variations due to the heterogeneity of the cell populations are minimal. Based on these findings, the analysis shows that to a certain degree the parallel extraction of genes and proteins can improve the correlation between their expression (Yamasaki, Anderson 2008).

4.4.2 Gene and protein expression changes induced in both EMT models

The analysis of gene expression profiles of both models showed changes to the majority of previously studied EMT genes and EMT-TFs. In P5B3, all genes aside from *TWIST1* (Fig. 4.3G) were significantly deregulated corresponding to an EMT-phenotype (Fig. 4.3). This was also shown in the majority of the genes screened in DU145, however here no gene expression was detected for *CDH2* (Fig. 4.3C) and *TWIST1* (Fig. 4.3G). A lack of detection of *CDH2* was previously shown in the analysis through qRT-PCR and Western blot. These results were confirmed through the RNA-sequencing analysis (Wang, W., Wang et al. 2017, Shankar, Nabi 2015).

As mentioned previously, the use of mass spectrometry resulted in a smaller number of proteins quantified and for this reason, only 3 of the 8 EMT markers were detected at the proteomic level, CADH1, VIME and FINC (Fig. 4.10). Of these, only VIME was significantly upregulated in DU145 upon stimulation, whereas all 3 proteins were significantly changed in P5B3. This analysis confirmed the successful induction of an

EMT-phenotype, based on the previously studied EMT gene profile on a transcriptomic and to a certain degree on a proteomic level.

4.4.2.1 Gene expression changes induced in P5B3

The identification of significantly deregulated genes revealed a large number of highly confident markers (n=4575), which showed a consistent clustering according to their respective group (Fig. 4.4). This clustering, represented in a heat map, also confirmed the absence of outliers, further supporting the consistent stimulation of all 10 replicates (Fig. 4.4). The identification of deregulated genes has highlighted a wide range of expression changes across the significantly altered genes. This included upregulated genes such as *CDH11* (+303.40) and *TWIST2* (+178.19) and downregulated markers such as *GKN2* (FC = -118.65) and *PSCA* (FC = -103.70).

CDH11 (Cadherin-11) belongs to the cadherin superfamily, the same family as the well-known EMT markers *CDH1* and *CDH2*. Studies have shown that it not only belongs to the same family, but also interacts with *CDH2*, a gene upregulated during EMT-induction (Straub, Boda et al. 2003). Its association with poor prognosis in malignancies and rheumatoid arthritis further supports its role in the process of disease progression (Assefnia, Dakshanamurthy et al. 2014). Another EMT-associated marker identified was the transcription factor *TWIST2 (*Twist-related protein 2), which belongs together with *TWIST1* to the Twist superfamily. It promotes EMT and was documented to be involved in breast (Fang, Cai et al. 2011), cervical (Wang, T., Li et al. 2014) and ovarian cancer (Mao, Xu et al. 2013). The downregulated marker *GKN2* (gastrokine 2) is a secretory protein expressed on the gastric surface of mucous cells and has been shown to inhibit the growth and induce apoptosis in gastric cancer cells (Shi, Wang et al. 2014). A further study has also indicated an inhibiting function of *GKN2* on proliferation, migration and invasion of gastric cancer cells (Dai, Zhang et al. 2014). Whereas the other marker *PSCA* (prostate stem cell antigen) was the third most strongly downregulated gene and studies have shown an involvement of this gene with an increase of proliferation and cell cycle progression of PCa cells (Li, E., Liu et al. 2017).

Overall, the identification of these markers within the list of the most deregulated genes in P5B3 has supported their selection for further pathway enrichment analysis, based on the high association with disease progression, poor prognosis and EMT.

4.4.2.2 Gene expression changes induced in DU145

The identification of significantly deregulated genes in DU145 has revealed a large number of highly confident markers (n=2303), which have shown a consistent clustering of the analysed samples according to their respective group (Fig. 4.6). This clustering, represented in a heat map, also confirmed the absence of outliers, further supporting the consistent stimulation of all 9 replicates per treatment group (Fig. 4.6). However, despite their significance, after correction for false discovery, the heat map presented a stronger degree of variability in the expression of these genes across the replicates, compared to P5B3 (Fig. 4.4). Such variation could be caused by the heterogeneity of the cell line and potential variation in the percentage of stimulated and unstimulated cells of each analysed sample. Despite this heterogeneity, the analysis has enabled the identification of deregulated genes with a wide range of expression changes. One of these was the *BMP2* (Bone morphogenic protein 2), which is a secreted ligand of the TGF-β superfamily and is involved in the recruitment and activation of SMAD family members. It was shown to induce EMT in pancreatic cancers (Chen, Liao et al. 2011) and to enhance migration, invasion and metastasis in gastric cancers (Park, Y., Kim et al. 2008). Another gene, which was also the third most upregulated gene, was *SPOCK1* (SPARC/Osteonectin, Cwcv And Kazal Like Domains Proteoglycan 1). It was shown to be induced through TGF-β in breast cancer and correlates with invasion and poor prognosis (Fan, Jeng et al. 2016). Furthermore, it is upregulated in colorectal cancers, promoting the activation of the PI3K/Akt pathway (Zhao, P., Guan et al. 2016). Interestingly, two of the most deregulated markers were both keratins, including *KRT32* and *KRT4*. Despite their limited documentation on EMT association, studies on keratins and disease progression and EMT have shown a strong association. Keratins form major intermediate filaments of epithelial cells, which are downregulated upon EMT induction. Their main function is the promotion of strong adhesion across epithelial cells (Nalluri, O'Connor et al. 2015). One study has suggested that the negative correlation of vimentin and keratin, and their ratio to each other, could be prognostic for postmenopausal breast cancer patients (Thomas, Kirschmann et al. 1999). *KRT4* was shown to be reduced in paired cancerous/non-cancerous tissue of oesophageal squamous cell carcinoma (Uchikado, Inoue et al. 2006), whereas no information on the association of *KRT32* and disease progression was known. However, based on their functionality in epithelial cells, a loss of *KRT32* could be correlated with the induction of EMT and the related loss of epithelial cell characteristics.

Overall, the identification of the above discussed markers, such as *BMP2*, which is highly associated with the TGF-β and EMT induction, supports the selection of EMT-associated genes and is therefore suitable to be used further in pathway enrichment analysis.

4.4.2.3 Topology-based pathway enrichment of significantly altered genes of both cell line models

The significantly altered genes were applied to the Metacore™ pathway analysis tool, which is based on a pathway topology (PT) method. The analysis of significant genes has highlighted the enrichment of pathways involved in "development", "cell adhesion" and "cytoskeletal remodelling" in both, P5B3 (Fig. 4.5) and DU145 (Fig. 4.7). The category of "development" is of interest, since EMT can be categorised into 3 types, of which type I EMT is involved in processes during embryonal development (Kalluri, Weinberg 2009). EMT is a highly conserved process, of which the major genetic components are the same throughout the different types and enrichment of developmental pathways confirms the activation of EMT-associated pathways in the context of cancer. The enrichment of "cytoskeletal remodelling" and "cell adhesion" pathways can be explained by the morphological changes that are induced through the induction of EMT. An epithelial cell with an apico-basal orientation changes into a motile, elongated mesenchymal cell. This process requires major changes in the structure of the cell, resulting in the remodelling of the cytoskeleton. Furthermore, mesenchymal cells present a reduced adhesion to the cell surface, enabling them to present an increased motility, resulting in the alteration of cell adhesion-associated pathways.

The enriched pathways of both cell line models showed the activation of TGF-β-associated pathways, such as the SMAD-dependent and SMAD-independent signalling pathway (Tab. 4.4 and 4.5). Both present routes for the induction of EMT after the binding of TGF-β to the receptor (Derynck, Zhang 2003). In addition, pathways involved with disease progression and the development of metastasis were present as well. Interestingly, in P5B3 a pathway called "Glomerular injury of Lupus nephritis" was shown to be enriched. During the progression of Lupus nephritis, renal tissue is destroyed based on consistent inflammatory processes, resulting in the development of fibrotic tissue. As previously mentioned, EMT is a conserved process which is categorised into three types. This pathway describes EMT type II, which is activated during wound healing and tissue fibrosis (Tennakoon, Izawa et al. 2015, Morishita, Kusano 2011). Overall, P5B3 and DU145 showed an enrichment of common pathways, including the SMAD-independent

signalling pathway (Tab. 4.6). However, the treatment of DU145 visibly presents a stronger effect on morphological changes and cytoskeletal remodelling induced through the process of EMT (Tab. 4.5), whereas P5B3 presents a stronger change based on phenotypic components of EMT (Tab. 4.4). Overall, these results show the successful targeting of the TGF-β pathways, on either a SMAD-dependent or independent route, as well as the alteration of EMT and metastasis-associated pathways.

4.4.3.1 Protein expression changes induced in P5B3

The identification of significantly deregulated proteins in P5B3 has shown a selection of highly confident proteins (n=297), which have shown a consistent clustering according to their respective group (Fig. 4.11). The clustering analysis, presented in a heat map, did not present any obvious outliers (Fig. 4.11). Despite this, compared to the analysis of the altered gene expression, this heat map highlighted a wider range of variations in the expression of the significantly altered proteins across the replicates of each respective group (Fig. 4.11).

Despite the smaller number of significantly identified proteins, the top 3 most induced proteins (VIME, FINC and BGH3) could be confidently associated with the process of EMT. The most upregulated protein was vimentin, presenting a fold change increase of 181.84. This protein, and its corresponding gene, were studied throughout this work as a marker indicative of the induction of EMT. Vimentin is a type III intermediate filament protein, which is a marker of cells with mesenchymal origin and is involved in cell motility (Challa, Stefanovic 2011). The second strongest induced protein is fibronectin with a fold change of 19.38. As with VIME, this protein is also commonly detected and analysed in EMT studies, and also in this study, VIME was used as a marker for the successful induction of EMT upon TGF-β stimulation. FINC is a glycoprotein, which mediates multiple interactions with the extracellular matrix, including cell migration and adhesion (Pankov, Yamada 2002). The third most upregulated marker was BGH3. This protein is also known as the transforming growth factor-beta induced protein ig-h3. It has shown a fold change increase of 15-fold. This protein is commonly induced by TGF-β and is secreted by the ECM (Ween, Oehler and Ricciardelli 2012), furthermore, it is functionally associated with adhesion, migration, proliferation and differentiation (H. J. Kim, et al. 2009). One of the most downregulated proteins was another keratin, keratin 13 (K1C13). Keratins, as previously discussed, are highly associated with epithelial cells and support

their anchorage to neighbouring cells. Through the stimulation, K1C13 was downregulated by nearly 40-fold in P5B3. Previous studies in a human keratinocyte cell line have also shown a downregulation of this protein during EMT induction via TGF-β. Based on this, the study has categorised it as an epithelial marker (Hatta, Miyake et al. 2018). Another marker of interest that presented a downregulation was AGR2 (anterior gradient protein 2). Studies have shown an increased expression of this protein in PCa tissue compared to healthy prostate, however at the same time a loss of this protein was highly associated with disease recurrence of patients with radical prostatectomy (Maresh, Mah et al. 2010).

4.4.3.2 Protein expression changes induced in DU145

The identification of significantly deregulated proteins in DU145 has shown a selection of highly confident proteins (n=187), which have shown a consistent clustering according to their respective group (Fig. 4.14). The clustering analysis, presented in a heat map, did not present any obvious outliers (Fig. 4.14). Despite this, compared to the analysis of the altered gene expression, this heat map highlighted a wider range of variations in the expression of the significantly altered proteins across the replicates of each respective group (Fig. 4.14).

The protein list of DU145 was shown to be the smallest list of significant markers (n=187), however this list contained markers associated with the process of EMT, such as the proteins transgelin, integrin subunit alpha V and Tumor-associated calcium signal transducer 2 precursor, which are discussed below. The strongest induced protein was shown to be transgelin (TAGL). It is an early marker for smooth muscle differentiation and is known to mediate TGF-β induced proliferation (Mitarai, Wada et al. 2017). Furthermore, a correlation of TAGL expression in colorectal cancer cells seems to increase their metastatic potential (Zhou, Fang et al. 2016). This protein was followed by the integrin subunit alpha V (ITAV, also known as CD51) that functions as a receptor for EMT-associated proteins such as *FN1* and thrombospondin 1 (*THBS1*). CD51+ colorectal cancer cells were shown to exhibit traits associated with cancer stem cells, such as enhanced migratory potential, as well as tumour initiation capabilities (Wang, J., Zhang et al. 2017). One of the most downregulated markers was identified as TACD2. TACD2 is a tumour associated calcium signal transducer 2 and studies have shown that the stepwise progression of squamous cell carcinoma is significantly associated with the gradual loss of TAGL expression (Wang, Y., Liu et al. 2014). The presence of known

markers associated with EMT and disease progression has confirmed that the analysis of a small list of significant markers can be attributed successfully to the desired phenotype.

4.4.3.3 Topology based pathway enrichment of significantly altered proteins in both cell line models

The analysis of both lists of significantly altered proteins was applied to the Metacore™ pathway analysis tool. The analysis of significant proteins in P5B3 (n=297) and DU145 (n=187) highlighted a strong enrichment of pathways associated with "cytoskeletal remodelling" and "cell adhesion".

Compared to the analysis of enriched genes, two pathways were shared in P5B3 (Tab. 4.9). One of these was the "TGF-β-dependent induction of EMT via RhoA, PI3K and ILK" (Fig. 4.13), also known as the SMAD-independent signalling. On the supplied graphic representation of this pathway (Fig. 4.13), an overall higher coverage through significant genes compared to significant proteins can be observed. Furthermore, each significant protein is also covered by the corresponding gene, showing identical directionality. In addition to the SMAD-independent pathway, it is also shown that a high coverage of the SMAD-dependent pathway is provided. Both pathways are highly involved in the induction of EMT and present the two major routes for its initiation. The analysis of both genes and proteins has also shown that despite the fact that some markers were solely identified through one omic level, such as SRF (serum response activator), the proteomic analysis, as well as the RNA-sequencing analysis, confidently identified downstream activated markers, such as tropomyosin 1, caldesmon and ACTB (Fig. 4.13). SRF was shown to be upregulated in metastatic gastric cancer cells (X. Zhao, et al. 2014) and in addition, is linked to the development of the mesoderm during embryonal development (Barron, et al. 2005).

The comparison of the enriched pathways of DU145 through gene and protein lists have also shown two shared pathways, of which one was related to the "Cytoskeleton and adhesion module" (Tab. 4.11). The schematic representation of this has shown a frequent identification of proteins involved with the ECM and cytoskeletal remodelling (Fig. 4.16). Furthermore, in the case of DU145, the additional analysis of the proteome has increased the coverage of this pathway by 7 % (Tab. 4.11).
Overall the comparison of the top 15 enriched pathways of both models through their proteomic profiles have shown a high overlap, with the identification of 10 shared

pathways (Tab. 4.12). A reduced level of correlation was identified following the comparison of enriched pathways through the analysis of the gene expression (Tab. 4.6).

4.4.6 Concordantly enriched pathway between P5B3 and DU145 across both omic levels

Interestingly, one pathway was detected across both cell models and both profiles types, despite the differences in the phenotypic response through the stimulation with TGF-β. This pathway was described as "Regulation of actin cytoskeleton organization by the kinase effectors of Rho GTPases". The coverage of P5B3 was higher, compared to DU145 with 69 % versus 48 % respectively (Tab. 4.13), however the analysis showed throughout a larger list of significant markers on both omic analyses in P5B3 compared to DU145, which is most likely reflected here. The schematic representation highlighted a high overlay at the downstream targets of this pathway (Fig. 4.19), with a detection of these targets in at least 3 significant marker lists, such as Cofilin, which was identified in the lists of DU145 proteins, P5B3 genes and proteins.

Rho GTPases are a small family of G proteins with a size ranging from 20 to 40 kDa. Most GTPases are activated through the binding of GTP and inactivated through the binding of GDP (Ridley 2015). The GTPases are involved in the regulation of cell motility cycles and play a role in the changes in the actin cytoskeleton structure (Hanna, El-Sibai 2013). Studies of various cancer models have shown an alteration in the signalling of small Rho GTPases, which present important factors in the initiation and progression of cancer (Ellenbroek, Collard 2007). Furthermore, studies have shown their function in the regulation of the ECM remodelling (Hanna, El-Sibai 2013), and their involvement in the formation of adherence junctions (Jansen, Gosens et al. 2017). It has also been shown that there is crosstalk between Rho GTPases and the TGF-β signalling via several mechanisms using factors such as Rho and Rac1 (Ungefroren, Witte et al. 2018). The consistent alteration of this pathway highlights the underlying changes induced in both cell line models and supports the previous findings describing an EMT-phenotype and their function as models of metastasis.

Overall, the generation of omic profiles have shown that both analysis methods and cell line models enable a characterisation of the desired and induced phenotype, whereas the proteomic analysis has shown an enrichment of cytoskeletal-associated proteins within the list of significant markers. The process of EMT can be characterised by multiple

factors, inducing changes in the gene and protein expression of EMT associated markers (VIME, FINC, CADH1, CADH2 and additional transcription factors). These changes of expression were confirmed in both models. Furthermore, through the induction of EMT, cytoskeletal changes result in the alterations of cytoskeletal associated proteins and adhesion. These changes could also be confirmed through the performed pathway enrichment analyses. Based on this, all analyses have successfully supported and confirmed the induction of EMT and supported the desired phenotype, enabling the use of these datasets for the potential identification of novel disease-associated biomarkers.

In the following chapter, these datasets are further subjected to stringent filtering methods enabling the identification of a core set of EMT markers in both cell line models and omic levels. This set was used for the selection of single markers, which were validated using wet-lab and *in silico* methods.

5. Chapter V – Selection and validation of novel biomarkers of prostate cancer progression and epithelial to mesenchymal transition using integrative data analysis

5.1 Introduction

Over the last 20 years, a single biomarker has been used in the routine clinical testing for prostate cancer of men above 50. This marker, PSA, is secreted by the prostate gland and can be detected through the non-invasive analysis of serum samples (Prensner, Rubin et al. 2012). The introduction of routine analysis of PSA resulted in the increased detection of prostate cancers, including a large proportion of indolent and low stage cancers (Catalona, William J., Smith et al. 1993) and decreased the frequency of high-grade tumours. However, PSA lacks specificity, and it is not possible to define a cut-off PSA-level that enables a secure exclusion of cancer presence (Tanguay, Begin et al. 2002). Various studies have been performed to improve the specificity. For example, Tanguay et al. compared the specificity and sensitivity of total PSA (tPSA), free/total PSA (f/tPSA), and complexed PSA (cPSA) in a cohort of 535 patients, of which nearly 40 % were diagnosed with cancer (Tanguay, Begin et al. 2002). When the regularly used cut-off of tPSA of 4.0 ng/ml was used, a sensitivity and specificity of 87 % and 27 % respectively were measured in the cohort. As a comparison, at a cut-off of 21 %, f/tPSA enabled a maximum sensitivity to specificity combination of 84 % to 50 % respectively. Complexed PSA presented only low specificities with a maximum specificity of 33 % at a sensitivity of 83 % (Tanguay, Begin et al. 2002). Despite improvements in the of use of f/tPSA over the clinically used tPSA, none of the 3 combinations fulfilled the criteria of a suitable new biomarker, which should ideally have a specificity and sensitivity close to 90 %.

Based on the lack of specificity in the tPSA test, an increased number of indolent cases is detected, which means more patients are subjected to "active surveillance", a process that includes routine PSA-level checks every 3 to 6 months and repeated biopsies every 1 to 2 years (Choyke, Loeb 2017). However, active surveillance is still mainly based on regular PSA tests and is often correlated with a strong impact on the mental health of patients (Xu, Neale et al. 2012). Aside from its use for diagnostics after positive DRE-results and the routine screening during active surveillance, PSA is also used as a measurement for

disease recurrence. A biochemical recurrence of increased PSA occurs frequently, however this is often without the actual presence of the disease or any disease-related symptoms (Adhyam, Gupta 2012). Overall, PSA is a routine tool for various prostate cancer related conditions, and is used for detection, prognosis and surveillance, despite showing visible limitations for each of the tasks. Based on its high sensitivity and low specificity, additional markers for follow-up approaches are needed, ideally in the form of biomarker screening that enables the targeted intervention at the required time point (See chapter 1.2.3.1 and 1.2.3.2).

For this reason, large efforts are being made in the discovery of novel disease-associated biomarkers. Over the years, many new biomarkers have been proposed to replace PSA including the α-methylacyl coenzyme A racemase (*AMACR*) (Jiang, Zhu et al. 2013), *PCA3* (Marks, Fradet et al. 2007, Wang, Y., Liu et al. 2014), and the fusion gene TMPRSS2:ERG translocation (Gleason 1966, Romero Otero, Garcia Gomez et al. 2014). Some of these biomarkers and others are commercially available (McGrath, Christidis et al. 2016), however none are applied routinely in a clinical setting, mainly because they do not present a major improvement compared to the established PSA method. This is mainly due to the variation of their specificity and sensitivity based on their cut-off thresholds as well as a limited number of clinical studies validating the suitability of the findings.

There is still an urgent unmet need for the discovery of novel, disease-associated biomarkers of prostate cancer, showing improved specificity and sensitivity compared to current markers. This search is supported by the development and improvement of high-throughput technologies, which has resulted in an exponential increase of new proposed biomarkers of various conditions and disease states; however, despite this, only a small percentage (estimated at 0.1 %) are successfully translated into clinical use (Poste 2011). The limited translation of novel biomarkers can be attributed to multiple factors, including problems in the study design, the utilised platforms for the discovery of proposed markers, and the type of clinical specimens used throughout the study (Goossens, Nakagawa et al. 2015). These factors commonly limit the transferability into a routine clinical setting. All this highlights the fact that despite the increased efforts in the discovery of novel biomarkers, the clinical need for it was not met.

The major clinical concern in prostate cancer is the development of metastasis, which reduces the survival to less than 30 % (Thobe, Clark et al. 2011). For this reason, biomarkers indicative for the development of metastasis or disease progression could improve current active surveillance approaches. Since the life-limiting factor of PCa patients is the development of metastasis, the study of pathways associated with this process could harbour the knowledge necessary to elucidate novel biomarkers.

The aim of this chapter is the increased understanding of the selected biomarkers based on the integration of transcriptomic and proteomic EMT profiles and their further evaluation as potential disease-associated markers in prostate cancer progression. This evaluation will be based on multiple aims, categorised into the further understanding of four markers and their association with EMT and cancer and their evaluation as disease-associated biomarker.

- First, the gene expression of 4 selected markers will be analysed in cell line material. These experiments will be performed to test for correlation of the marker expression with phenotypic characteristics of different cell lines and the potential detectability and applicability of these markers in other cancers (section 5.2.2).

- This will be followed by the screening of healthy tissue RNA. Novel biomarkers always present a new potential drug target, however for this the expression under healthy conditions needs to be identified.

- The initial analyses were focussed on the expression of the respective genes, to further understand their capabilities as biomarkers, a routine method for biomarker screening, immunohistochemistry, will be applied. This method is routinely used as a diagnostic procedure and the successful validation of any of the markers through IHC will support their use as biomarker.

- To overcome the limitations of available models for EMT models, 5 publicly available model data sets were selected and will be analysed gene expression changes induced through the stimulation with EMT-inducing cytokines. The results of these analyses will enable to evaluate and translate the findings of the

EMT-models of this study to a wider context. The use of 2 cell line models could potentially results in the detection of biomarkers with limited use and the successful validation in other cell line models will support the association with the process of EMT.

- Finally, to overcome the limited availability of clinical specimens, publicly available gene expression datasets derived from clinical specimens will be used to further evaluate the capabilities of all four markers with the prediction of clinical conditions, such as Gleason score, disease-recurrence and for the differentiation between localised and advanced prostate cancer.

The experiments performed in this chapter will generate an overview of the characteristics of all four selected markers with different cancerous conditions and their suitability as potential new biomarkers. These results will enable the guidance of future experiments.

5.2 Results

5.2.1 Data integration and selection of a core marker list through the integration of generated omic profiles

The generation of omics profiles commonly results in long lists of potential novel candidates. In this study the analysis of both cell lines models has resulted in the quantification of approximately 26 000 genes and approximately 2000 proteins. To reduce such a number to a potentially more meaningful, and more manageable list, all markers that could potentially be considered were subjected to the following criteria; a p-value below 0.05 after correction for false-discovery using Bonferroni correction and an absolute fold change of 2 and above. Furthermore, the transcriptomic data had to present FPKM values of 2 or more in at least one sample group. This cut-off was selected to ensure the detectability of the marker in routine applications such as quantitative real-time PCR whilst taking into account the variability of human specimens. In addition, the detected proteins had to present a confidence value of at least 70 %. This cut off was selected based on the advice of Dr Stephen Tate, SCIEX Senior Research Scientist and Manager of Software Applications Research, who contributed to the development of the proprietary confidence value (Lambert, Ivosev et al. 2013). Table 5.1 shows the resulting number of significant unique and shared markers based on different comparisons across each model and omic levels. The application of these criteria resulted in the identification of 1461 significant genes and 84 significant proteins for P5B3 and 838 significant genes and 38 significant proteins for DU145 (Tab. 5.1).

Overall, more significantly altered markers were detected in P5B3 compared to DU145 (Tab. 5.1), at both the gene and protein level. Also, when the absolute number of shared markers in both omic analyses was considered, a higher concordant number was detected in P5B3 compared to DU145. Out of the 64 shared significant genes and proteins in P5B3, all shared markers, aside from one (KRT5), presented the same expression directionality. KRT5 demonstrated an upregulation on the gene level and a downregulation on the protein level. In DU145, 29 markers were shared and all of them presented the same directionality (Tab. 5.1).

To answer the question as to whether, after the application of stringent filters, the analysis of proteomic data resulted in the identification of additional markers not identified

through the analysis of transcriptomic data, the lists were compared. In P5B3, the analysis of proteomic data resulted in the discovery of an additional 20 markers, which were uniquely identified to present significant differences at the protein level (Tab. 5.1). In DU145, the protein analysis resulted in 9 additional proteins (Tab. 5.1).

Table 5.1: Identification of significantly differentially regulated markers within all 4 omic datasets and their overlap between cell lines and omic-levels. The two gene expression datasets were subjected to the following criteria: p-value below 0.05 after correction for false-discovery using Bonferroni correction, an absolute fold change of 2 and above and FPKM values of 2 or more in at least one sample group. The protein datasets were filtered based on: a p-value below 0.05 after correction for false-discovery using Bonferroni correction, an absolute fold change of 2 and above and a confidence value of 70 %.

Integration of datasets	**Number of markers**
Significant deregulated genes P5B3	1461
Significant deregulated genes DU145	838
Significant degregulated proteins P5B3	84
Significant degregulated proteins DU145	38
Shared significant deregulated genes and proteins P5B3	64
Shared significant deregulated genes and proteins DU145	29
Shared significant deregulated genes P5B3 and DU145	322
Shared significant deregulated proteins P5B3 and DU145	18
Unique significant deregulated markers (genes and proteins) P5B3	1481
Unique significant deregulated markers (genes and proteins) DU145	847
Shared markers (genes and proteins) both models and omic levels	13

Despite the efforts to reduce the number of genes and proteins, the significant marker selection exceeded the logistics available for routine wet-lab validation approaches. For this reason a core marker set was identified. This core marker set was generated through the integration of markers present on all omics levels and both cell line models. This resulted in a final selection of 13 markers, which are shown in table 5.2, including their respective fold change for each omic level and cell line. The p-value is shown as a representation for all 4 omic level, since all p-values were below 0.0001 (****). Out of the 13 markers, only one marker, SDPR, showed a reduced expression through treatment, whereas the remaining 12 markers presented an increase in their expression. Additionally, the gene and protein expression within and across each model presented the same directionality.

One of the identified markers was BGH3 or *TGFBI* (transforming growth factor β-induced) (Tab. 5.2). This protein is induced by the cytokine TGF-β and its induction can be associated with the successful stimulation of both cell lines with TGF-β. The presence of this marker supports the association of the remaining 12 markers with the process of TGF-β induced EMT (Tab. 5.2). ACTN1 and TUBA4A are directly associated with the cytoskeleton and TPM1, as well as MYL9, and are strongly associated with muscular contractions. Therefore, based on their widespread expression and their associated limitations as potential therapeutic targets, these markers were excluded from further validation. The same decision was made for BGH3, which is also known as *TGFBI*, and the marker TSP1. Both are well known and well-studied markers in cancer and EMT-associated studies, for example in relation to cancer metastasis and renal diseases (Suzuki, Yokobori et al. 2018, Kurpinski, Chu et al. 2009, Brennan, Morine et al. 2012, Hugo 2003, Sweetwyne, Murphy-Ullrich 2012).

Overall it can be seen that the response of both cell lines resulted in the induction of a core set of genes, however the strength of response varied from cell line to cell line. This is clearly visible for example in the change of DPYL3, which shows a very strong upregulation in P5B3 on both the protein and gene levels, whereas SDPR has shown a stronger downregulation in DU145 compared to P5B3 (Tab. 5.2).

Table 5.2: List of 13 markers identified through the integration of both models and all 4 omic profiles. These 13 markers were shared in both cell lines at both the gene and protein levels. FC = fold change. The p-value [p-value (all)] is presented together showing a concordant, highly significant p-value **** across both models and omic profiles.

Protein ID	FC P5B3 Gene	FC P5B3 Protein	FC DU145 Gene	FC DU145 Protein	p-value (all)	Gene/Protein
ACTN1	2.62	2.20	2.08	2.07	****	Actinin alpha 1
DPYL3	21.51	11.79	3.52	2.52	****	Dihydropyrimidinase like 3
FBLI1	2.33	2.24	3.37	2.89	****	Filamin binding LIM protein 1
LMCD1	3.77	3.66	7.36	5.22	****	LIM and cysteine rich domains 1
MYL9	4.48	3.80	2.98	3.17	****	Myosin light chain 9
P4HA2	2.39	2.78	2.36	2.81	****	Prolyl-4-hydroxylase subunit alpha 2
PALLD	5.12	5.38	3.03	2.06	****	Palladin, cytoskeletal associated protein
PDLI7	3.35	4.09	2.15	2.45	****	PDZ and LIM domain 7
SDPR	-5.99	-3.80	-13.61	-7.51	****	Serum deprivation-response protein
BGH3	15.60	15.05	10.50	4.90	****	Transforming growth factor beta induced
TSP1	8.84	5.98	9.61	3.56	****	Thrombospondin 1
TPM1	5.98	3.24	11.63	6.65	****	Tropomyosin 1
TUBA4A	2.81	2.37	4.11	3.52	****	Tubulin alpha 4a

Of the remaining list of 7 markers, the following 4 markers were selected for further verification: DPYL3, FBLI1, SDPR and P4HA2 (Tab. 5.3). The selection of these was based on a literature search and the consolidation of multiple online available resources.

Table 5.3: Final marker selection for further validation presenting the induced fold change for both cell line models and omic level . The p-value [p-value (all)] is presented together showing a concordant, highly significant p-value across both models and omics profiles.

Protein ID	FC P5B3 Gene	FC P5B3 Protein	FC DU145 gene	FC DU145 Protein	p-value (all)	Gene
DPYL3	21.51	11.79	3.52	2.52	****	Dihydropyrimidinase like 3
FBLI1	2.33	2.24	3.37	2.89	****	Filamin binding LIM protein 1
SDPR	-5.99	-3.80	-13.61	-7.51	****	Serum deprivation-response protein
P4HA2	2.39	2.78	2.36	2.81	****	Prolyl-4-hydroxylase subunit alpha 2

5.2.2. Screening of cancerous cell lines for their expression of selected markers

To further analyse the association of these markers with cancer, various cell lines, including breast and prostate cancer, as well as one osteosarcoma cell line were screened. Additionally, a previously developed cell line model of TGF-β treated MCF10A cells was analysed for the involvement of these markers with the TGF-β pathway (data not shown). *DPYSL3* showed overall a low expression across all cell lines compared to the induced state of both models. The expression of *DPYSL3* in PC-3 and SAOS cells was comparable to untreated DU145 cells. Overall, the expression in all BCa cell lines showed levels comparable to the reference gene, whereas the PCa cell line PC-3 showed a similar expression to DU145 untreated and a stronger expression than P4B3 untreated (Fig. 5.1). The expression of *FBLIM1* was shown to be variable across all cell lines, showing the highest relative expression in both cell line models. The expression of *FBLIM1* was induced in MCF10A upon treatment with TGF-β (Fig. 5.1), however the expression in both conditions was comparable to P5B3 untreated. An increased expression was also shown in OPCT-1 and the single cell clones of this cell line: P4B6 and P4B6B (Harner-Foreman, Vadakekolathu et al. 2017) as well as SAOS (Fig. 5.1).

The highest expression of *SDPR* was detected in the breast cancer cell line MDA-MB-231, presenting an increased expression compared to P5B3 and DU145 untreated (Fig. 5.1). A reduction of its expression was shown in the EMT models of P4B3 and DU145, as well as in the stimulated MCF10A cells. All other cell lines presented a very low expression of *SDPR* (Fig. 5.10).

P4HA2 presented a variable expression across the studied cell line samples, showing the highest expression in the TGF-β stimulated MCF10A cells. It was also highly expressed in the single cell clone P4B6B (Harner-Foreman, Vadakekolathu et al. 2017) and the osteosarcoma cell line, SAOS, showing a comparable expression to the stimulated DU145 cells, and an increased expression when compared to the cell line model of P5B3 (Fig. 5.1).

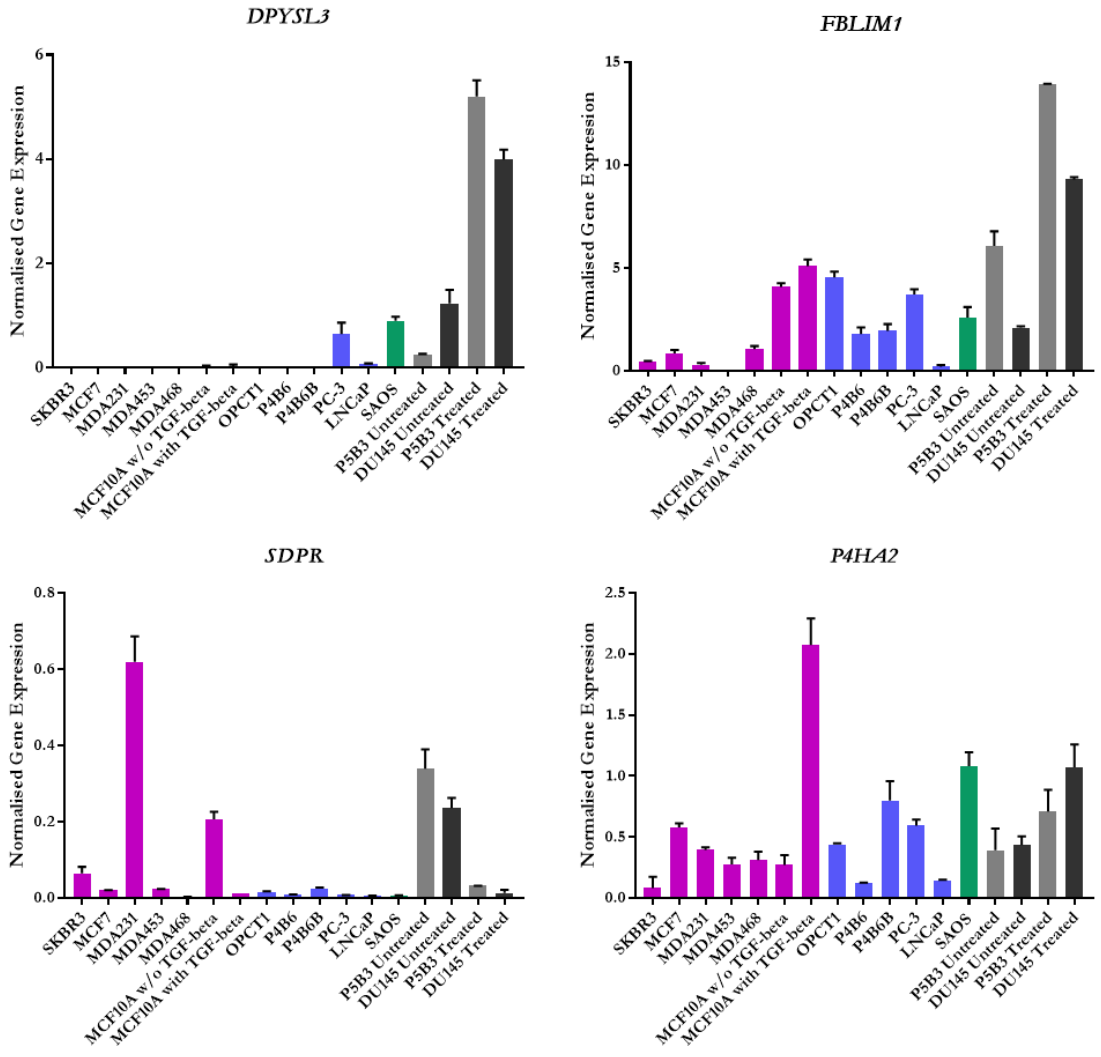


Figure 5.1: qRT-PCR screening of selected markers (DPYSL3, FBLIM1, SDPR and P4HA2) in in-house derived cell line material from various primary and metastatic breast and prostate cancer cell lines as well as one osteosarcoma cell line. Pink represents cell lines associated with breast cancer, blue cell lines associated with prostate cancer and green, the osteosarcoma cell line SAOS. P5B3 is coloured in grey and DU145 in black. Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (TBP) gene, which was utilised as the reference gene. Details on the used cell lines can be found in Table 2.2.

5.2.3 Gene expression analysis of selected markers in healthy tissue RNA

5.2.3.1 Comparison of marker expression in healthy prostate tissue with both cell line models

An ideal biomarker should present an inverse expression in the target tissue compared to the expression in a healthy or non-cancerous state. This means that a marker, whose increased expression is associated with poor survival, should ideally present a low or no expression in healthy, or non-cancerous target tissue. On the other hand, a marker whose loss or reduced expression is associated with negative disease development should ideally present a high expression in healthy tissue. This would enable an easier detection of changes through the development or progression of a disease. For this reason, the gene expression of the 4 selected markers in their treated and untreated condition was compared to their gene expression in healthy prostate tissue (Fig. 5.2).

The expression of *DPYSL3* showed a lower expression level in both untreated cell lines, compared to the healthy tissue control, with a significant difference in P5B3. The expression of *DPYSL3* through the stimulation with TGF-β was significantly increased compared to the healthy tissue in both cell line models (Fig. 5.2A+E).

The expression of *FBLIM1* in the untreated cell lines showed a significant increase in P5B3 and a significant decrease in DU145, whereas the expression was significantly induced in both cell line models upon stimulation (Fig. 5.2B+F).

SDPR showed a consistent expression between the healthy tissue and the unstimulated P5B3 cells (Fig. 5.2C), presenting a significant decrease after the stimulation with TGF-β. The gene expression of *SDPR* in DU145 showed a significantly decreased expression compared to the healthy control (Fig. 5.2G), which was further decreased upon stimulation.

The expression of *P4HA2* showed the lowest expression in the healthy tissue, increasing with the untreated cell lines and showing the highest expression through the stimulation with TGF-β for 10 days in both cell line models (Fig. 5.2D+E). The difference in the expression was significant between the healthy control and P5B3 treated, whereas the

expression in DU145 untreated and treated showed both a significantly increased expression compared to the healthy control (Fig. 5.2E).

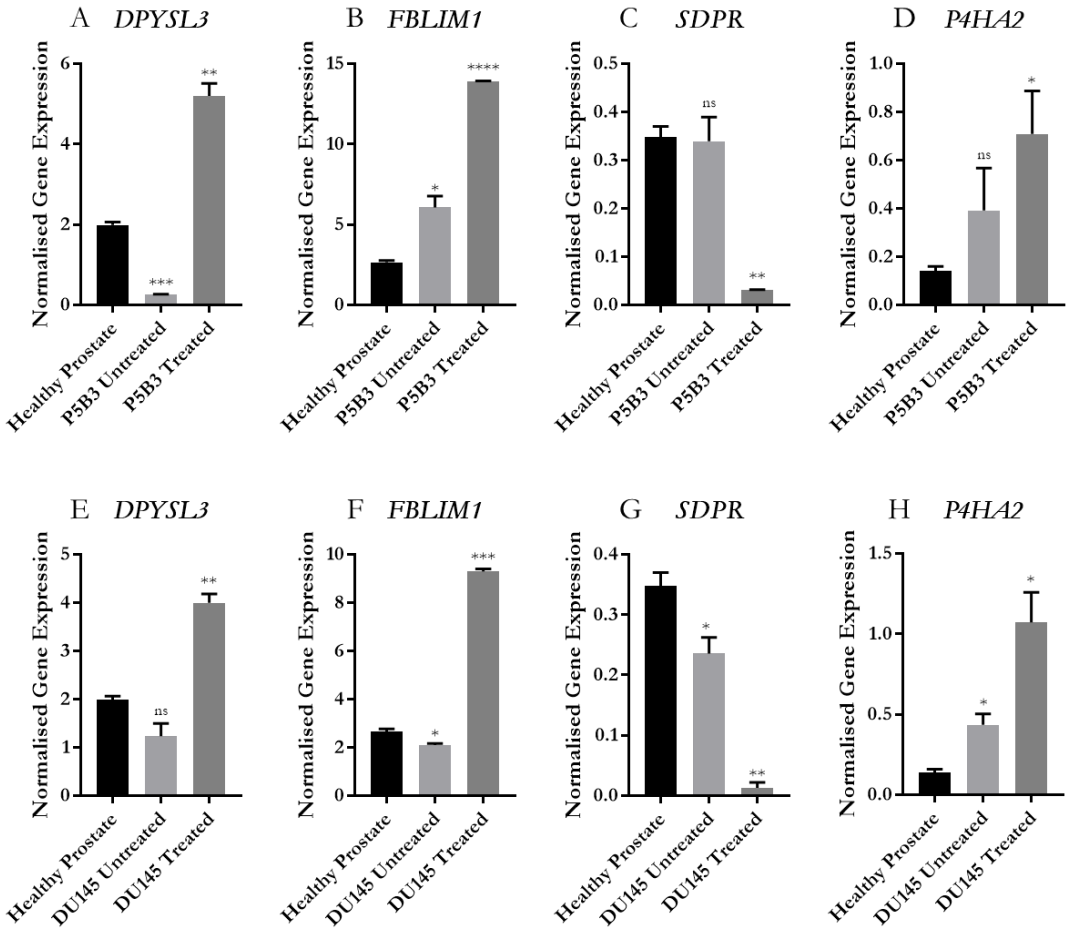


Figure 5.2: Comparison of gene expression of each marker (DPYSL3, FBLIM1, SDPR and P4HA2) in commercially available healthy prostate tissue RNA (Clontech) with the expression in both cell line models in an untreated and treated state. Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

5.2.3.2 Gene expression analysis of all four markers in healthy tissue in comparison to healthy prostate

Some, however not all, novel biomarkers present the option to be utilised as a therapeutic target (Shen 2013). Crucial for this is the information as to whether these markers function as a "messenger" and are only a consequence of underlying changes, or whether they are "driver" markers, such as genes or proteins, that directly influence factors such as tumour growth or disease progression (Shen 2013). For this reason, the measurement of gene and protein expression of novel markers in healthy tissue is crucial to validate the suitability of the studied marker as a therapeutic target.

To evaluate the potential use of the selected markers, RNA extracted from healthy tissue material was screened for their expression and was compared to the expression levels in healthy prostate tissue.

The expression of *DPYSL3* in a healthy tissue panel (Fig. 5.3) showed a significantly higher and overall stronger expression in tissue extracted from the ovary and spinal cord. These were followed by uterus, trachea and retina, however no significant differences compared to the prostate could be detected in the latter two. RNA extracted from breast, colon and skeletal muscle showed a non-significant decreased expression compared to prostate. A significantly lower expression was detected in various tissues (thyroid, spleen, adrenal gland, salivary gland, placenta, thymus, testis), including the essential organs lung, heart, brain, liver and kidney (Fig. 5.3).

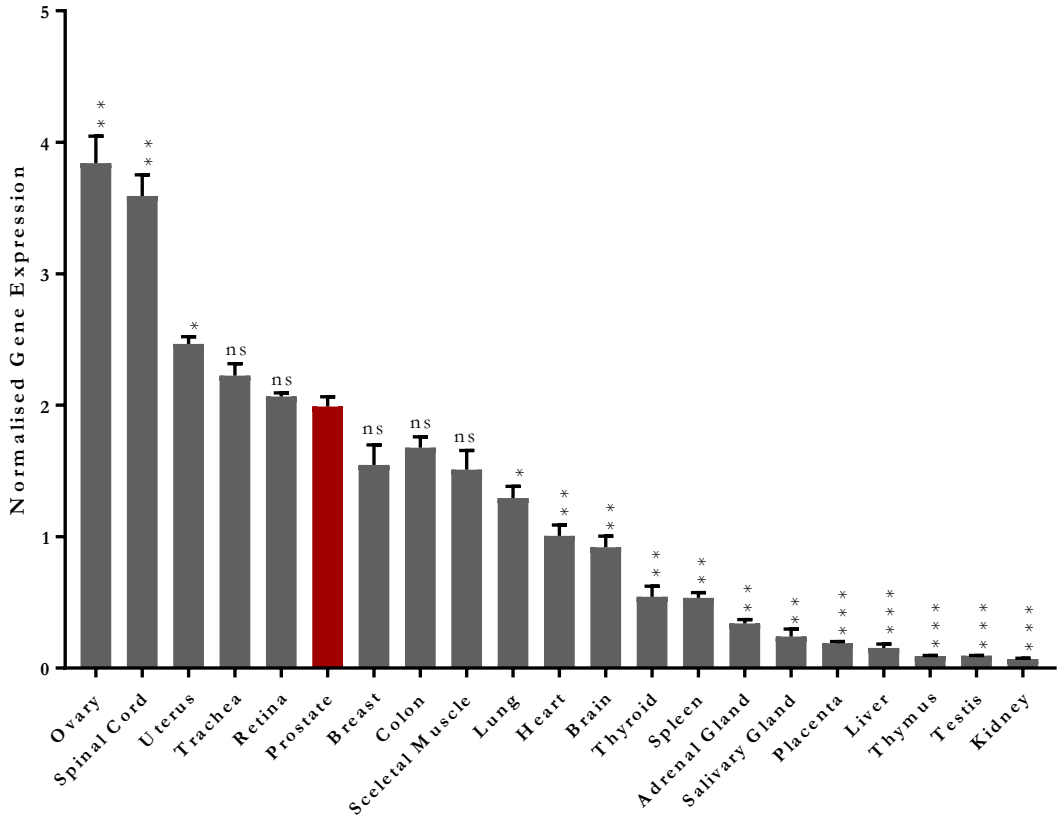


Figure 5.3: Comparison of gene expression of *DPYSL3* in a commercially available healthy tissue RNA panel (Clontech). The expression was compared to the expression in RNA from healthy prostate tissue (Clontech) (red). Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

FBLIM1 showed the highest, significantly increased expression in colon and heart (Fig. 5.4). The expression of *FBLIM1* in healthy prostate tissue was comparable to its expression in uterus, spleen and breast tissue. A significantly lower level of expression was detected in multiple essential organs, including lung, liver, kidney and brain, as well as ovary, placenta, trachea, salivary gland, retina, spinal cord, thymus, skeletal muscle and testis (Fig. 5.4).

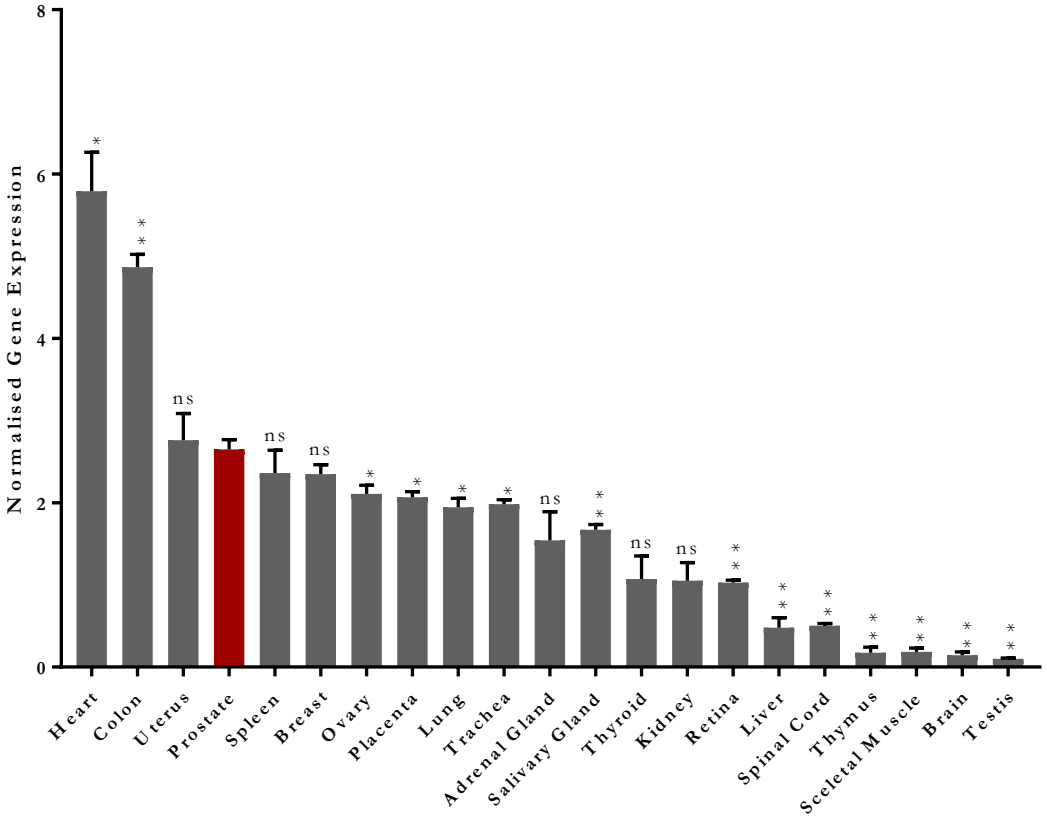


Figure 5.4: Comparison of gene expression of *FBLIM1* in a commercially available healthy tissue RNA panel (Clontech). The expression was compared to the expression in RNA from healthy prostate tissue (Clontech) (red). Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

The serum-deprivation response protein (*SDPR*) was the only marker in this selection that presented a reduced expression upon stimulation with TGF-β. By far the highest, and most significant elevated expression compared to healthy prostate tissue was detected in the spleen, presenting a nearly 8-fold difference. High expression of *SDPR* was also shown in both lung and thyroid, followed by uterus and breast. A comparable expression of *SDPR* in prostate was shown for material extracted from heart, ovary, skeletal muscle, colon, retina and adrenal gland. The lowest expression, showing a significant difference to healthy prostate tissue, was measured in the following organs; spinal cord, placenta, kidney, liver, thymus, trachea, salivary gland, brain and testis (Fig. 5.5).

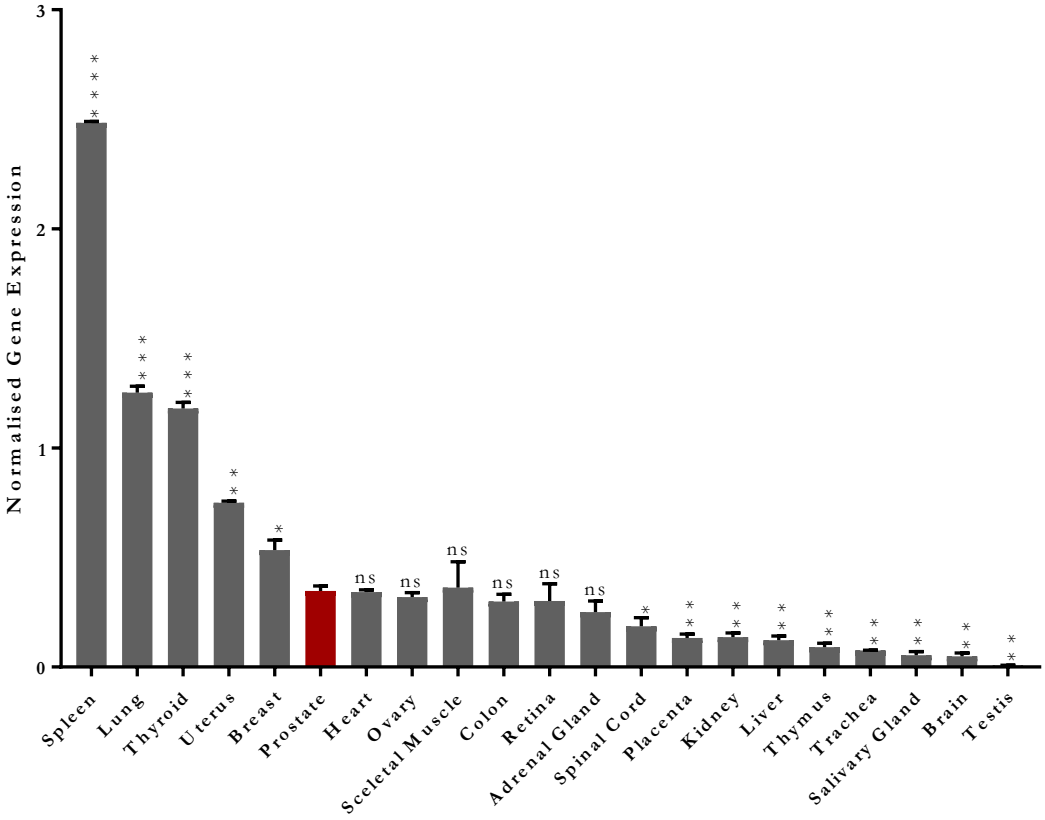


Figure 5.5: Comparison of gene expression of *SDPR* in a commercially available healthy tissue RNA panel (Clontech). The expression was compared to the expression in RNA from healthy prostate tissue (Clontech) (red). Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

P4HA2 showed a strong variation in its expression across the measured sample material, however the overall expression was very low, compared to the used reference gene (*TBP*) (Fig. 5.6). A significantly higher expression of *P4HA2* was detected in material of the salivary gland, followed by heart, kidney, lung, trachea and uterus and a comparable expression to prostate tissue was detected the adrenal gland, thymus, spinal cord and brain. The lowest, and most significantly different expression level, was measured in testis (Fig. 5.6).

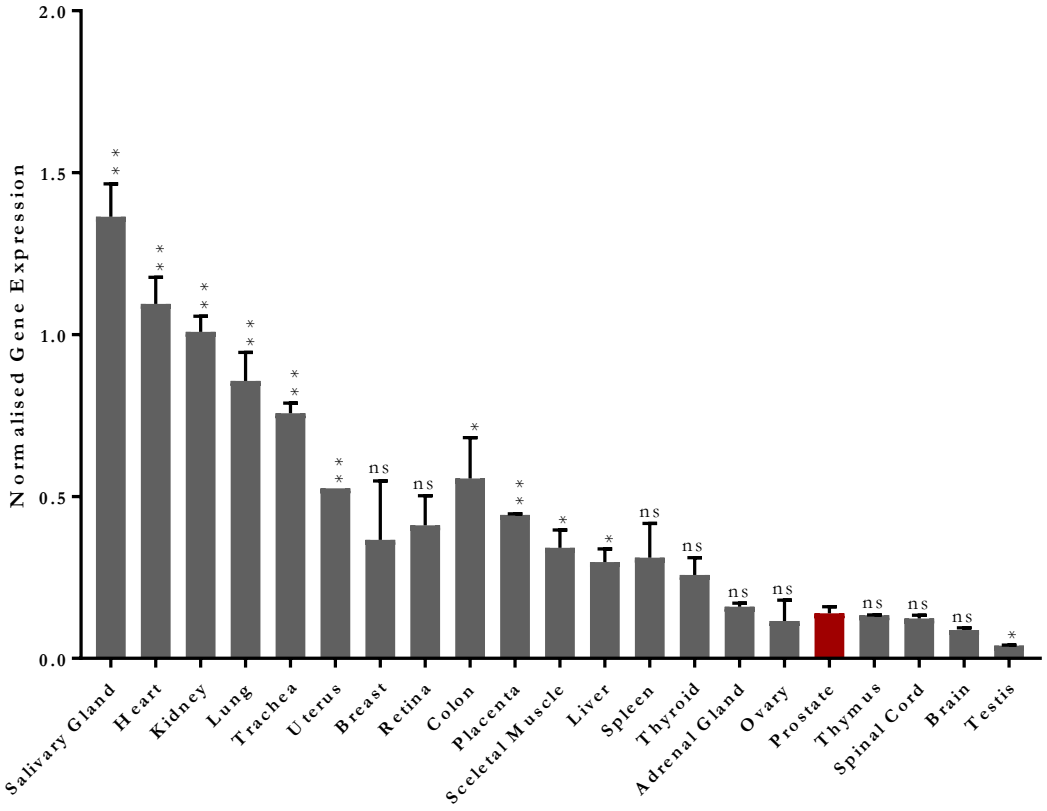


Figure 5.6: Comparison of gene expression of *P4HA2* in a commercially available healthy tissue RNA panel (Clontech). The expression was compared to the expression in RNA from healthy prostate tissue (Clontech) (red). Results were analysed using the comparative $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak 2008) (n=4). The gene expression was normalised against the TATA-box protein (*TBP*) gene, which was utilised as reference gene.

5.2.4 Validation of novel biomarkers using tissue microarray derived from healthy and diseased tissue

The analysis of gene expression in healthy tissue RNA enables an initial overview, however only limited conclusions regarding the protein expression can be made on the basis of this (Vogel, Marcotte 2012). For this reason, the protein expression of all four markers in healthy tissue was analysed using immunohistochemistry staining on commercially available healthy tissues microarrays (US Biomax). Staining intensity was categorised into 4 categories (Fig. 5.7), including 0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining. Localised staining, as for example shown in Figure 5.7 – Staining intensity: 3, resulted in the assignment of the tissue to the higher category and the sample was marked with an * to highlight the focally increased expression.

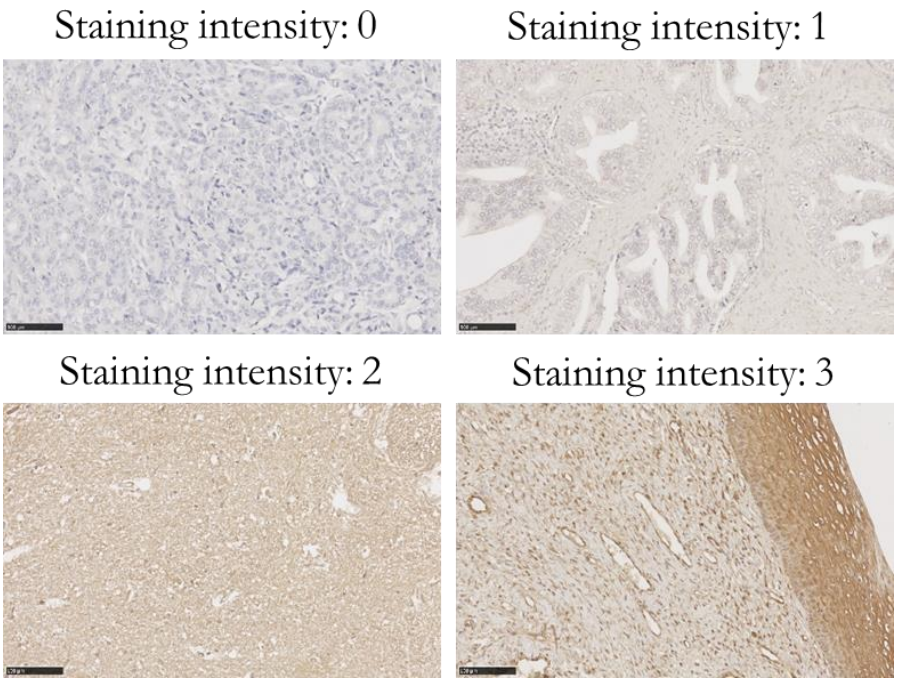


Figure 5.7: Images illustrating scoring method used for healthy and diseased tissue specimens analysed using immunohistochemistry staining . The staining intensity was categorised into 4 intenstities (0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining). Representative images at 20x magnification. Scale bar represents 100 μm.

5.2.4.1 Screening of protein expression in healthy tissue specimens using immunohistochemistry staining on tissue microarrays

The analysis of the selected 4 markers was performed on 3 different TMAs, which was based on their availability. Altogether, all 4 markers could be analysed in the following tissue types: prostate, skin, colon, heart, kidney, liver, lung, brain, pancreas, uterus, ovary and breast. DPYL3 and FBLI1 were furthermore analysed in additional healthy tissue types.

The protein expression of DPYL3 showed a moderate to high expression across all tissues. Locally intensified staining was detected in the samples of skin (Fig. 5.8B), kidney (Fig. 5.8E), breast (Fig. 5.8L), placenta (Fig. 5.9N), stratified muscle (Fig. 5.9O), urethra (Fig. 5.9P), testes (Fig. 5.8Q) and bladder (Fig. 5.9R). This localised expression was mainly found in the glandular structures, such as prostate gland (Fig. 5.8A), acini and ducts in breast tissue (Fig. 5.8L) and uterine glands (Fig. 5.8J). The lowest expression was detected in the lung (Fig. 5.8G), pancreas (Fig. 5.8I), thymus (Fig. 5.9T), spinal cord (Fig. 5.9V) and umbilical cord (Fig. 5.9AA). The localisation of DPYL3 was mainly detected in the cytoplasm with localised presence in the nucleus (Fig. 5.8/5.9).

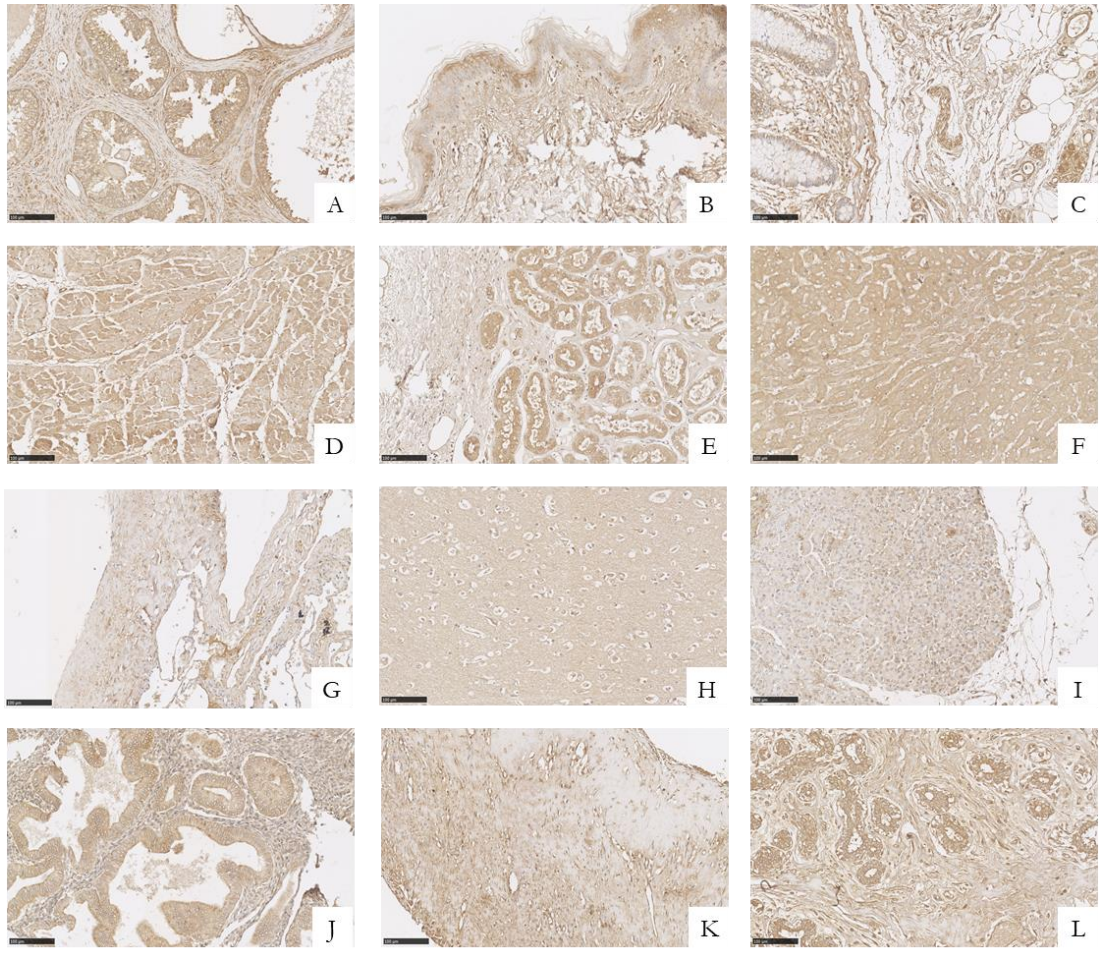


Figure 5.8: DPYL3 protein expression in healthy tissue microarray (US Biomax MNO341). Prostate (A), skin (B), colon (C), heart (D), kidney (E), liver (F), lung (G), brain (H), pancreas (I), uterus (J), ovary (K) and breast (L). Representative images at 20x magnification. Scale bar represents 100 μm.

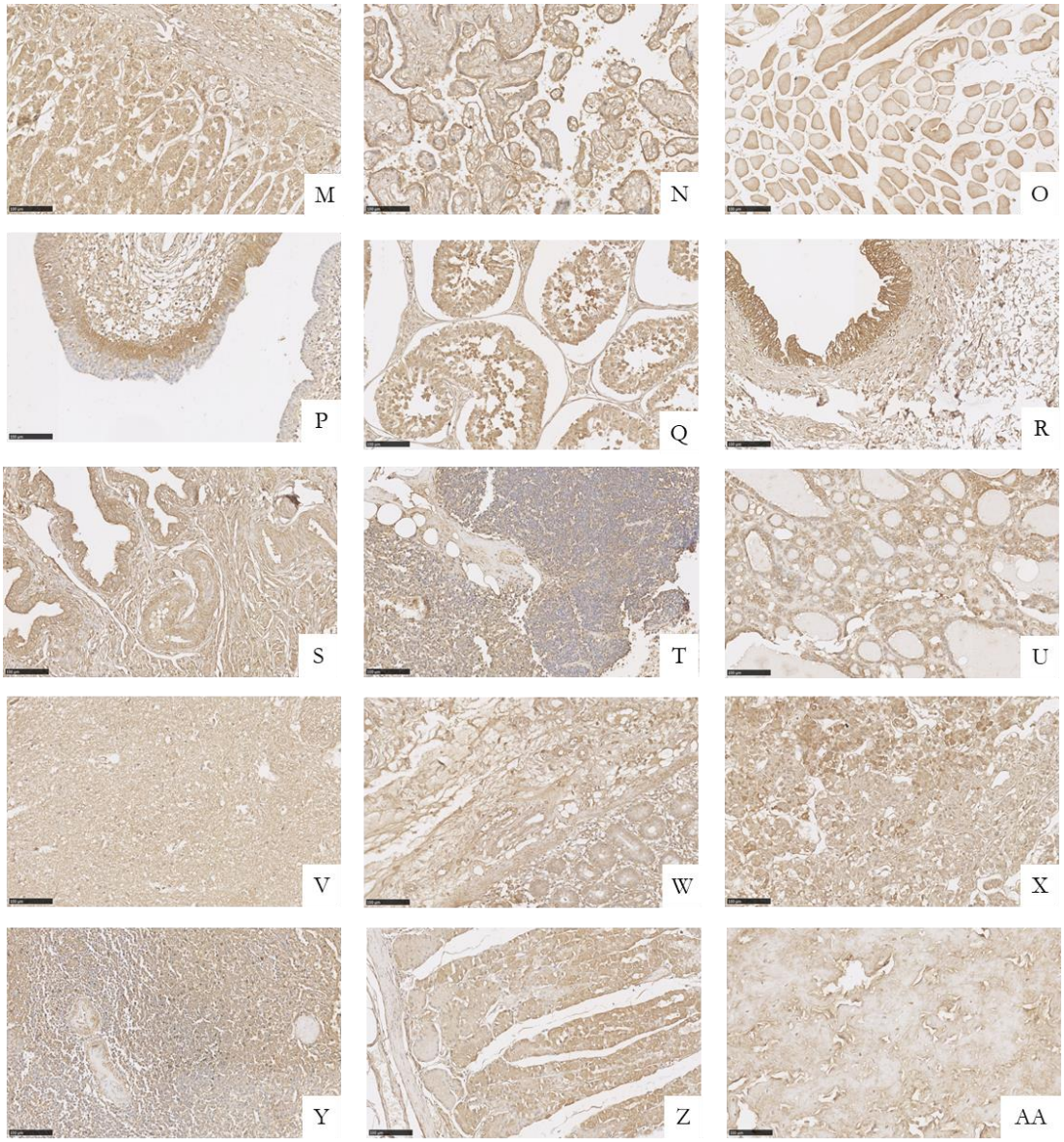


Figure 5.9: DPYL3 protein expression in healthy tissue microarray (US Biomax MNO341). Adrenal gland (M), placenta (N), stratified muscle (O), urethra (P), testes (Q), bladder (R), fallopian tube (S), thymus (T), thyroid (U), spinal cord (V), small intestine (W), pituitary gland (X), spleen (Y), stomach (Z) and umbilical cord (AA). Representative images at 20x magnification. Scale bar represents 100 μm.

The expression of FBLI1 showed an overall low expression in the analysed tissue sections. No staining was detected in pancreatic (Fig. 5.10I) and tonsil (Fig. 5.10O) tissue. Localised staining was shown in the epidermis of the skin (Fig. 5.10B), uterine glands (Fig. 5.10J) and seminiferous tubules located in the testes (Fig. 5.10M). The remaining tissue sections presented a ubiquitous low staining (Fig. 5.10). The expression of FBLI1 was mainly focussed on the cytoplasm of the cell, however nuclear staining was observed in primary spermatocytes located in the seminiferous tubules (Fig. 5.10M).

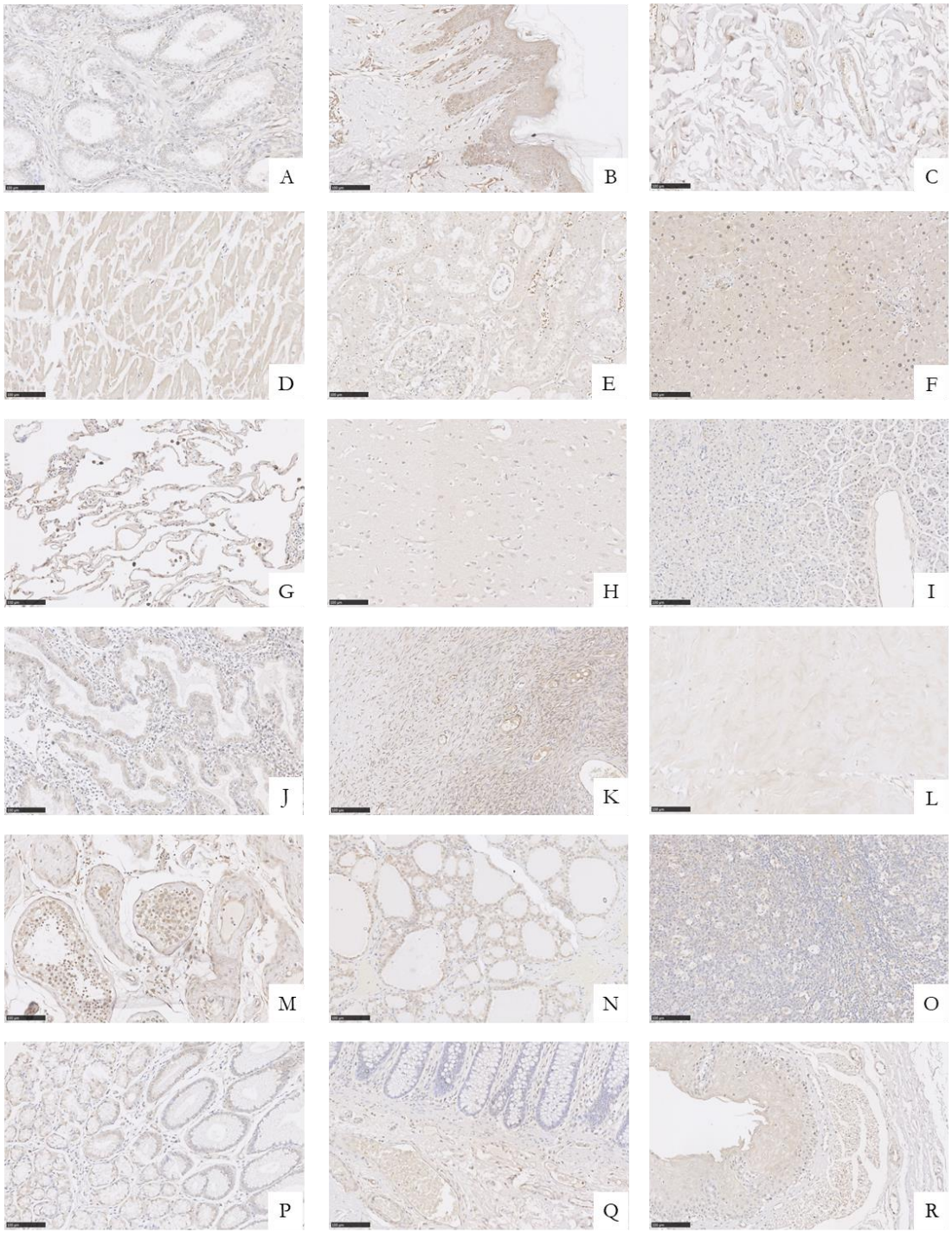


Figure 5.10: FBLI1 protein expression in healthy tissue microarray (US Biomax MNO381). Prostate (A), skin (B), colon (C), heart (D), kidney (E), liver (F), lung (G), brain (H), pancreas (I), uterus (J), ovary (K), breast (L), testes (M), thyroid (N), tonsil (O), stomach (P), small Intestine (Q) and oesophagus (R). Representative images at 20x magnification. Scale bar represents 100 μm.

The protein expression of SDPR showed a variable expression across the analysed samples, presenting a low to moderate expression in specimens of prostate (Fig. 5.11A), lung (Fig. 5.11B) and breast (Fig. 5.11L), and no staining in ovarian tissue (Fig. 5.11K). Locally increased staining was detected in the epidermis of the skin (Fig. 5.11B) and the endometrium localised in the uterus (Fig. 5.11J). The protein expression in the kidney showed an overall ubiquitous expression with reduced expression in the glomeruli (Fig. 5.11E). A similar presentation was observed in the liver sections (Fig. 5.11F), where a reduced expression is shown in the tissue surrounding the portal tracts. Furthermore, the expression in the lamina propria of the analysed colon section showed a reduced expression compared to the intestinal glands (Fig. 5.11C). The expression of SDPR was mainly found in the cytoplasm of the cells, with limited expression in the nucleus of cells from colon (Fig. 5.11C), skin (Fig. 5.11B) and uterus (Fig. 5.11J).

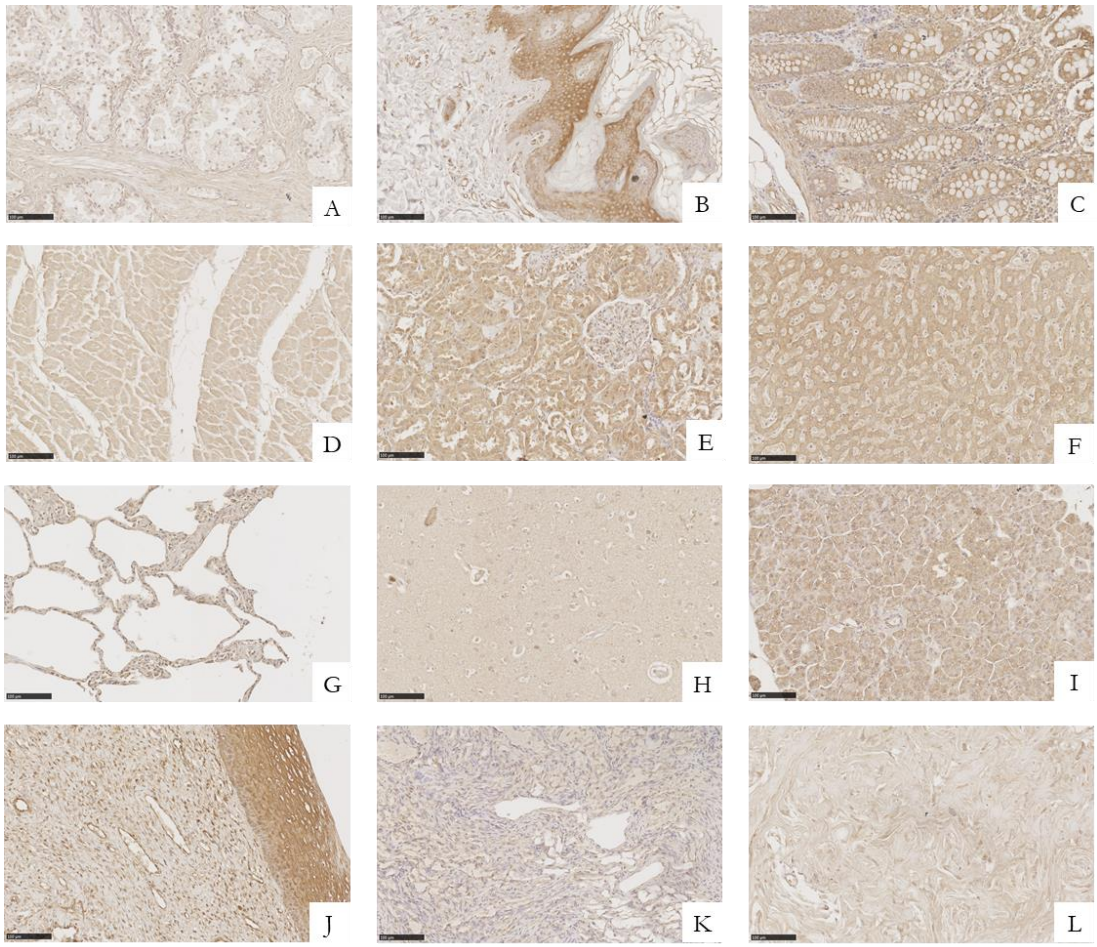


Figure 5.11: SDPR protein expression in healthy tissue microarray (US Biomax BN243c). Prostate (A), skin (B), colon (C), heart (D), kidney (E), liver (F), lung (G), brain (H), pancreas (I), uterus (J), ovary (K) and breast (L). Representative images at 20x magnification. Scale bar represents 100 μm.

In P4HA2, the protein expression was overall ranging from very low to not detectable in all samples (Fig. 5.12). An increased expression was detected in the skin (Fig. 5.12B) presented locally in the epidermal layer. Overall, the expression in kidney (Fig. 5.12E) and liver (Fig. 5.12F) was elevated compared to the other tissue sections, also showing a homogeneously distributed expression. In the section of the kidney, a lower expression in the renal corpuscles could be observed (Fig. 5.12E). In specimens with a visible detection of P4HA2, the protein was localised in the cytoplasm (Fig. 5.12).

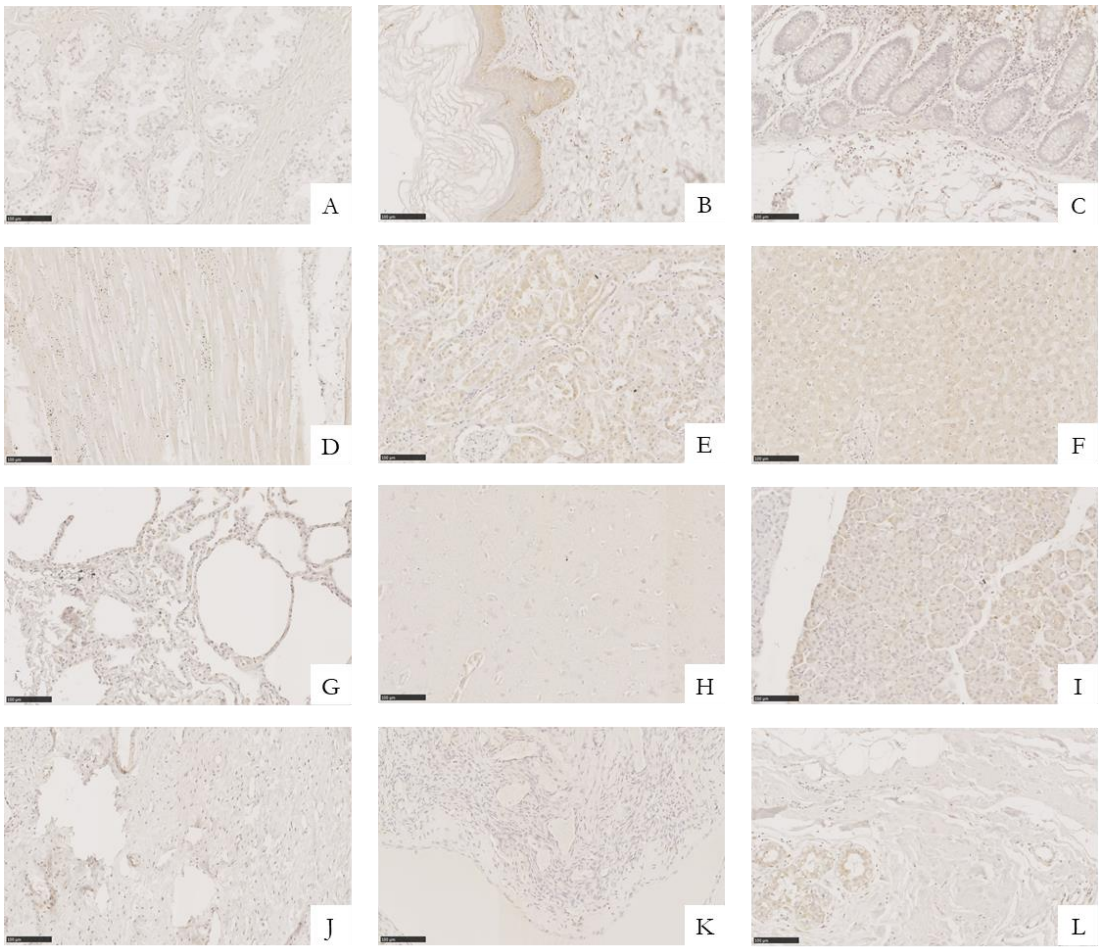


Figure 5.12: P4HA2 protein expression in healthy tissue microarray (US Biomax BN243d). Prostate (A), skin (B), colon (C), heart (D), kidney (E), liver (F), lung (G), brain (H), pancreas (I), uterus (J), ovary (K) and breast (L). Representative images at 20x magnification. Scale bar represents 100 μm.

Table 5.4 presents a summary of the staining intensity detected in the analysed healthy tissue sections. In DPYL3 (Fig. 5.8/5.9), the majority of analysed tissue sections showed a moderate staining intensity, whereas in FBLI1 (Fig. 5.10) the expression was mainly categorised as low staining. The intensity of SDPR presented variability, with tissues mainly assigned to 3 intensity groups (Fig. 5.11). The most frequent staining intensity in P4HA2 (Fig. 5.12) was low staining, followed by no staining and only one tissue section was shown to have a moderate, localised expression (Fig. 5.12B).

Table 5.4: Score summary of immunohistochemistry tissue sections for DPYL3 (Fig. 3.12), FBLI1 (Fig. 3.13), SDPR (Fig. 3.14) and P4HA2 (Fig. 3.15). Table representing staining intensities observed in normal tissue (MNO341 (DPYL3), MNO381 (FBLI1), BN243c (SDPR) and BN243d (P4HA2)). Staining intensities were assigned as previously described (Fig. 5.7) into 4 categories; 0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining used. * indicates localised increased expression. Numbers in cells (column #) represent the total number of individual cores analysed and shaded areas represent the number of cores assigned to this staining intensity.

	DPYL3					FBLI1					SDPR					P4HA2				
* localised	staining intensity					staining intensity					staining intensity					staining intensity				
Tissue type	#	0	1	2	3	#	0	1	2	3	#	0	1	2	3	#	0	1	2	3
Prostate	1			1*		1	1				2					2	2			
Skin	1			1*		1			1*		2				2*	2				2*
Colon	1				1*	1		1			2				2*	2			2	
Heart	1			1		1		1			2			2		2			2	
Kindey	1				1*	1		1			2			2		2			2	
Liver	1				1	1			1		2			2		2			2	
Lung	1		1			1		1			2		2			2			2	
Brain	1			1		1		1			2		2			2			2	
Pancreas	1		1			1	1				2			2		2			2	
Uterus	1			1*		1	1				2				2*	2	2			
Ovary	1		1			1		1			2	2				2	2			
Breast	1			1*		1		1			2		2			2	2			
Adrenal gland	1			1		-					-					-				
Placenta	1			1		-					-					-				
Stratified muscle	1			1		-					-					-				
Urethra	1				1*	-					-					-				
Testes	1				1*	1			1*		-					-				
Bladder	1				1*	-					-					-				
Fallopian tube	1			1		-					-					-				
Thymus	1		1			-					-					-				
Thyroid	1			1		1		1			-					-				
Spinal cord	1			1		-					-					-				
Small intestine	1			1		1		1			-					-				
Pituitary gland	1				1*	-					-					-				
Spleen	1		1			-					-					-				
Stomach	1			1		1		1			-					-				
Umbilical cord	1		1			-					-					-				
Tonsil	-					1	1				-					-				
Oesophagus	-					1		1			-					-				

5.2.4.2 Screening of protein expression in prostate cancer specimens using immunohistochemistry and immunofluorescence on tissue microarrays

The validation of novel biomarkers is commonly performed using tissue microarrays (Hassan, Ferrario et al. 2008) derived from diseased specimens, annotated with clinical parameters, such as Gleason score or tumour stage. Here, a prostate cancer TMA was selected for the screening of all 4 biomarkers in different tumour stages, as well as adjacent healthy prostate tissue (US Biomax PR242b). In this TMA, specimens of 5 patients with Stage II and 5 patients with Stage IV PCa were included in duplicates. Furthermore, there were 4 cores, derived from 2 individuals, of adjacent healthy tissue. In addition to the analysis by IHC, the protein expression was further analysed using immunofluorescence (IF) in the same TMAs. As previously described, the staining intensity through IHC was categorised into 4 categories (Fig. 5.7), including 0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining. Strong localised staining resulted in the assignment of the tissue to the related category, even if the remaining tissue did present a lower staining intensity. An example for this can be seen in Figure 5.7 in the staining intensity 3. Such samples were marked with an * to highlight the focal increased expression, which defined the assigned category.

The analysis of DPYL3 showed a strong expression in healthy prostate tissue, which was shown in the stroma and glands through IHC and mainly in the stroma through IF (Fig. 5.13). Compared to this, adjacent tissue seemed to express a lower intensity of DPYL3 compared to healthy, which was apparent through both analyses. The expression was localised in the cytoplasm of the analysed specimens. Comparing the expression in healthy tissue with stage II PCa, no obvious differences in the expression of DPYL3 could be observed using IHC. However, through IF, a strong expression, mainly localised in the glands, could be observed in 3 out of 5 patients. Stage IV PCa showed none or weak staining through IHC and also the analysis with IF showed a lower expression of DPYL3 in stage IV compared to stage II as well as healthy tissue (Fig. 5.13).

The IHC analysis of FBLI1 in prostate (cancer) sections showed none to faint staining, however below the category of "weak staining". The IF analysis of healthy tissue showed a weak, homogeneous expression of FBLI1 in the stroma (Fig. 5.14). The IF analysis of the tissue sections showed some cores with a strong localised expression of FBLI1, however based on their cell shape and lack of a nucleus, these cells can most likely be assigned to erythrocytes trapped and fixed within the tissue.

The IHC analysis of SDPR in healthy and diseased prostate tissue showed limited differences across the cores (Fig. 5.15). Two patients with stage II PCa showed a moderate staining compared to the low staining detected in the remaining cores. The additional analysis using IF showed a greater range of variability in the expression of SDPR. The strongest expression was observed in the stroma of healthy prostate tissue, with additional staining in the prostate glands, whereas only faint staining was detected in the stroma of adjacent tissue. The expression of SDPR in stage II PCa showed a more homogeneous expression and stronger across the stroma of all 5 patients, compared to stage IV PCa patients. In stage IV PCa the expression overall seems to be reduced and the remaining expression tends to be accumulated in glandular structures and less in the stroma. However, one patient showed a similar stromal expression of SDPR to stage II PCa (Fig. 5.15). The expression of SDPR was localised in the cytoplasm of the cell.

The IHC analysis of P4HA2 in healthy and diseased prostate sections showed no or only faint staining. The faint staining detected was not intense enough to be categorised as "weak staining" (Fig. 5.16). The use of IF staining on the section enabled the detection of P4HA2 expression in these sections. Ubiquitous expression was detected in the stroma of healthy and adjacent healthy tissue, as well as on all 5 stage II tissue sections. The detected expression was comparable to healthy tissue (Fig. 5.16). The expression of P4HA2 in stage IV PCa was not detectable for 2 patients and detectable with a low expression in 2 patients. One patient showed a stronger expression of P4HA2 with focal hotspots of increased expression, which was located in the stroma and the cytoplasm of the cell.

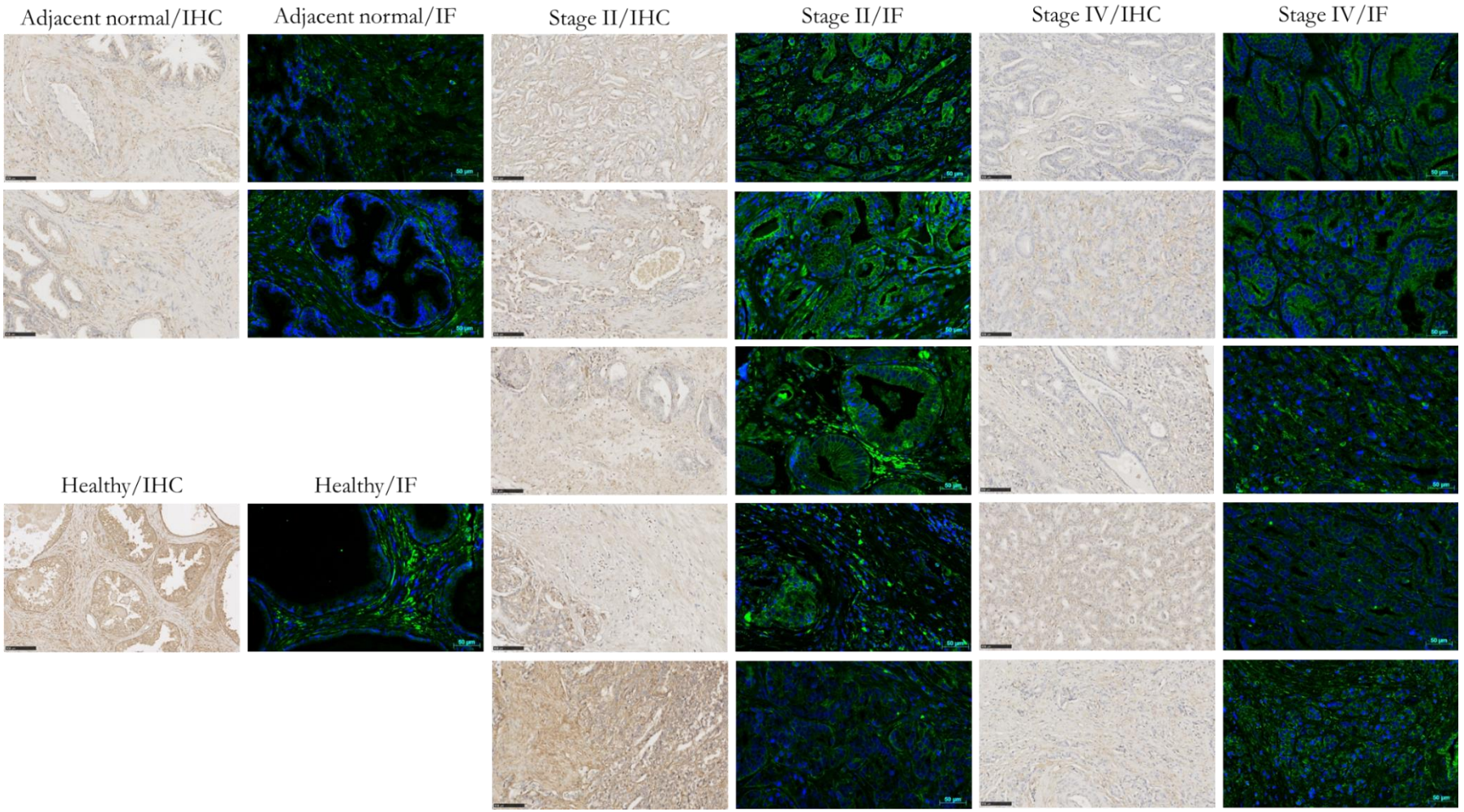


Figure 5.13: DPYL3 expression in healthy prostate, prostate cancer and healthy adjacent prostate tissue samples (PR242B US Biomax). Staining was performed using immunohistochemistry staining (IHC) and immunofluorescence (IF) staining on matching tissue samples. Representative images were taken at 20x magnification. Scale bar represents 100 μm (IHC) and 50 μm (IF). In the IF pictures blue = DAPI staining (cell nucleus) and green = protein of interest, here DPYL3.

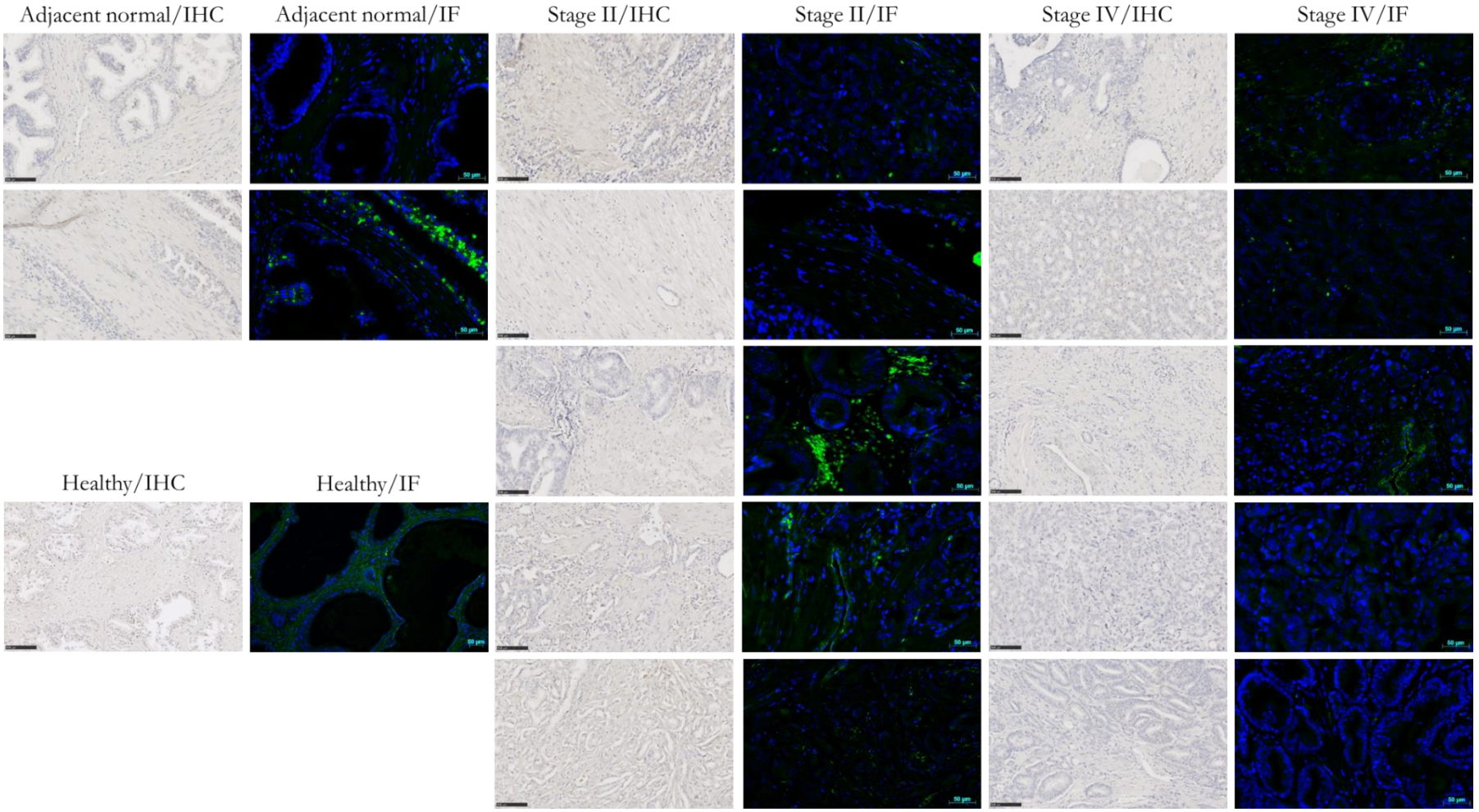


Figure 5.14: FBLI1 expression in healthy prostate, prostate cancer and healthy adjacent prostate tissue samples (PR242B US Biomax). Staining was performed using immunohistochemistry staining (IHC) and immunofluorescence (IF) staining on matching tissue samples. Representative images were taken at 20x magnification. Scale bar represents 100 μm (IHC) and 50 μm (IF). In the IF pictures blue = DAPI staining (cell nucleus) and green = protein of interest, here FBLI1.

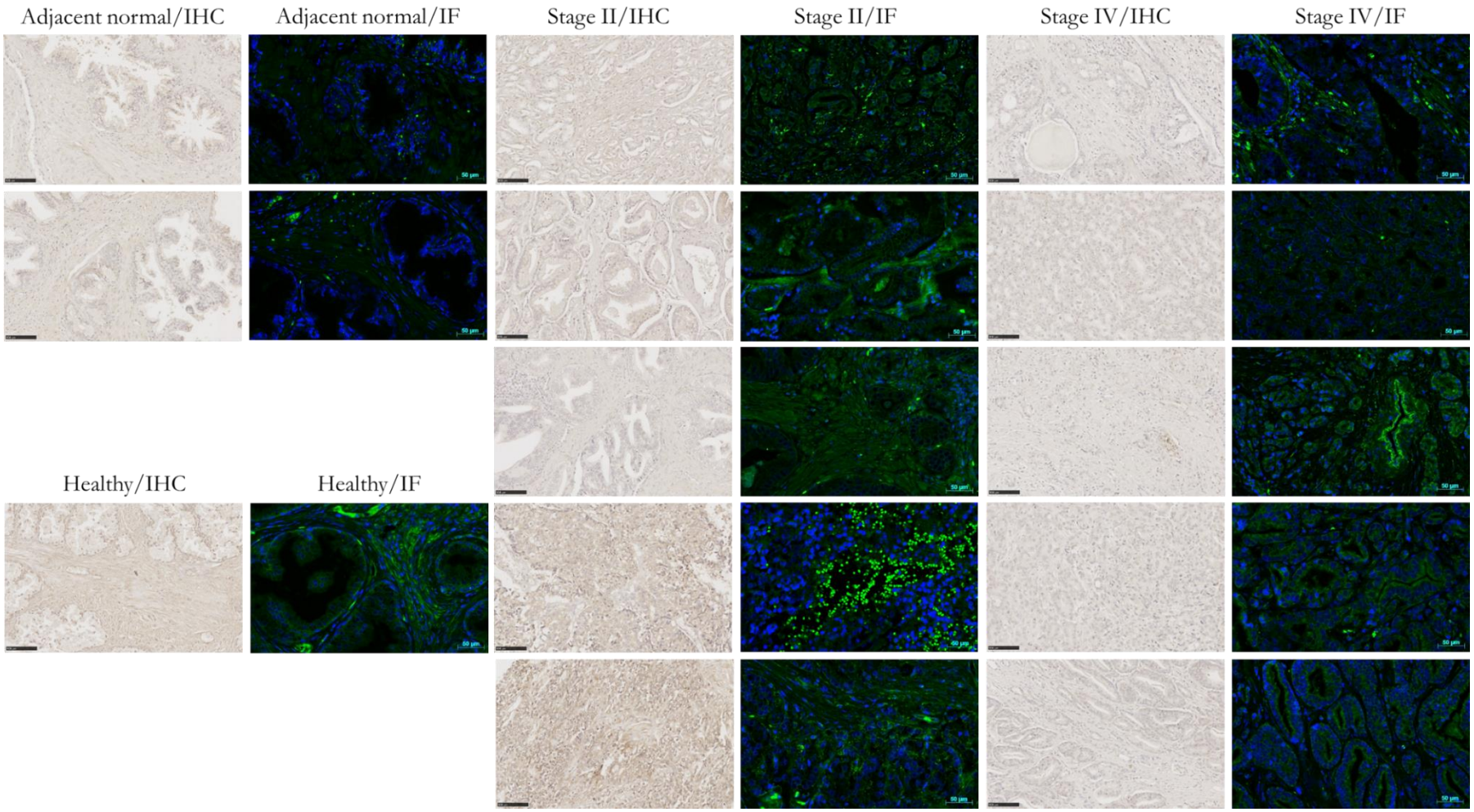


Figure 5.15: SDPR expression in healthy prostate, prostate cancer and healthy adjacent prostate tissue samples (PR242B US Biomax). Staining was performed using immunohistochemistry staining (IHC) and immunofluorescence (IF) staining on matching tissue samples. Representative images were taken at 20x magnification. Scale bar represents 100 μm (IHC) and 50 μm (IF). In the IF pictures blue = DAPI staining (cell nucleus) and green = protein of interest, here SDPR.

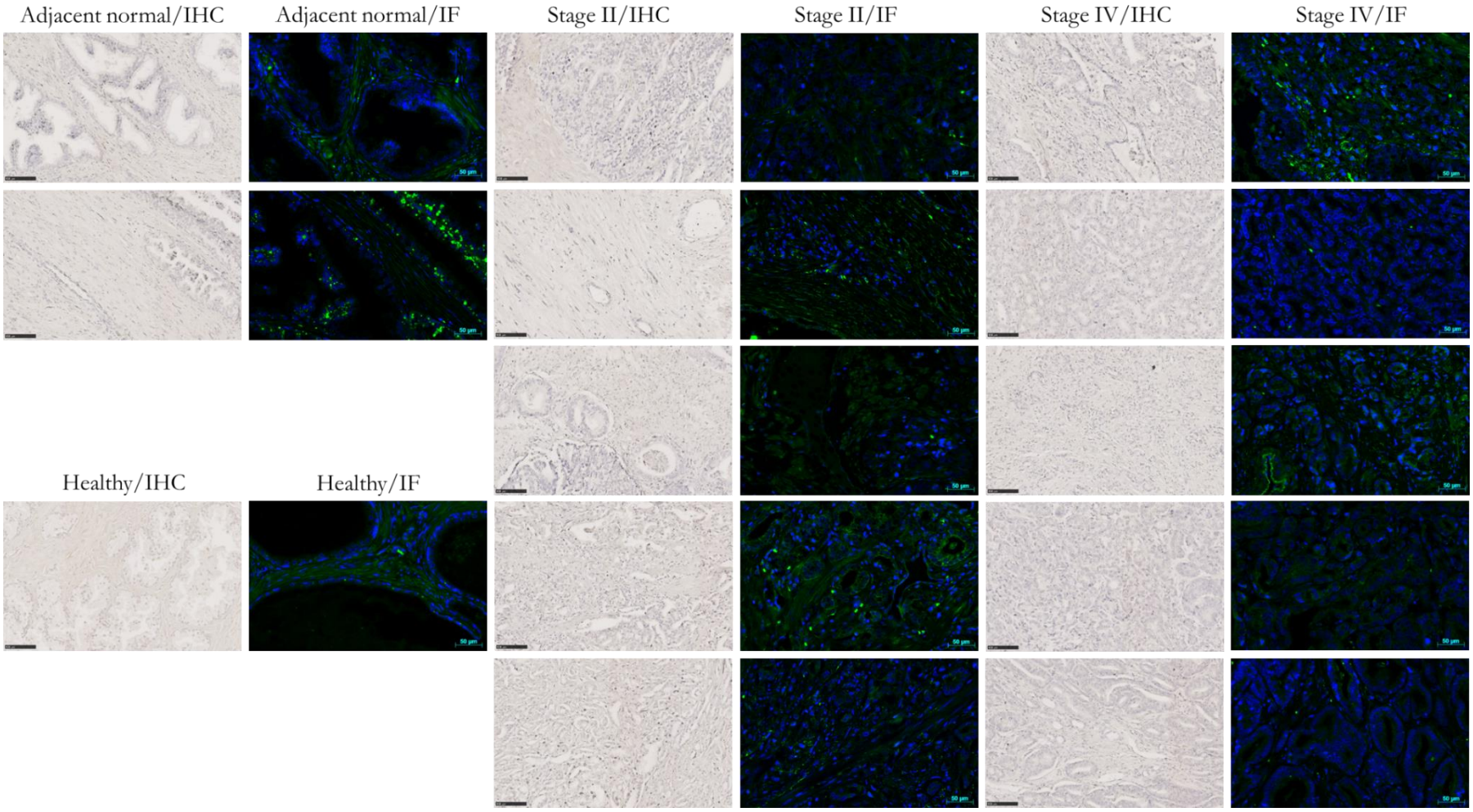


Figure 5.16: P4HA2 expression in healthy prostate, prostate cancer and healthy adjacent prostate tissue samples (PR242B US Biomax). Staining was performed using immunohistochemistry staining (IHC) and immunofluorescence (IF) staining on matching tissue samples. Representative images were taken at 20x magnification. Scale bar represents 100 μm (IHC) and 50 μm (IF). In the IF pictures blue = DAPI staining (cell nucleus) and green = protein of interest, here P4HA2.

Table 5.5: Score summary of immunohistochemistry tissue sections for DPYL3 (Fig.3.12), FBLI1 (Fig. 3.13), SDPR (Fig. 3.14) and P4HA2 (Fig. 3.15). Table representing staining intensities observed in normal prostate tissue taken from the previously analysed healthy tissue TMAs (MNO341 (DPYL3), MNO381 (FBLI1), BN243c (SDPR) and BN243d (P4HA2)) and a TMA comprised of adjacent normal, Stage II and Stage IV prostate cancer tissue (PR242b). Staining intensities were assigned as previously describes (Fig. 5.7) into 4 categories; 0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining used. * indicates localised increased expression. Numbers in cells (column #) represents the total number of individual cores analysed and shaded areas represent number of cores assigned to this staining intensity.

	DPYL3					**FBLI1**					**SDPR**					**P4HA2**				
* localised	staining intensity					staining intensity					staining intensity					staining intensity				
Tissue type	#	0	1	2	3	#	0	1	2	3	#	0	1	2	3	#	0	1	2	3
Healthy tissue	1			1*		1	1				2		2			2	2			
Adjacent normal	4		4			4	4				4		4			4	4			
Stage II	10		8	2		10	10				10		6	4		10	10			
Stage IV	10	6	4			10	4				10		10			10	10			

5.2.6 *In silico* validation of selected markers in publicly available datasets

Publicly available datasets have the potential to be mined for quick and easy biomarker validation; furthermore, they overcome the limitations based on the availability of sample material. *In silico* validation has previously been successfully performed in various cancers, including gastric cancer (Szász, Lánczky et al. 2016) and non-small cell lung cancer (Yu, Xu et al. 2015).

5.2.6.1 *In silico* validation of selected markers in publicly available datasets of TGF-β induced cell lines models

The use of publicly available data sets also enables the study of markers of interest in additional cell line models. For this, 5 independent models of EMT, generated in 3 studies, were selected and the expression of *DPYSL3*, *FBLIM1*, *SDPR* and *P4HA2* was analysed. In the first study, three lung cancer cell lines (A549, HCC827 and NCI-H358) were treated with 2 ng/ml TGF-β for 3 weeks (Sun, Yuting, Daemen et al. 2014). The second study stimulated a pancreatic cancer cell line (PANC-1) with 5 ng/ml TGF-β for 5 days (Maupin, Sinha et al. 2010) and the third study treated a cell line derived from healthy retinal pigmented epithelium (ARPE-19) with 5 ng/ml TGF-β together with 10 ng/ml TNF-α for 60 hours (Takahashi, Nagano et al. 2010). Three of the analysed cell lines were derived from the primary tumour, one from the metastatic site and one was generated from healthy tissue (Tab. 5.6). All 5 cell lines are characterised as epithelial, adherent cells.

Table 5.6: Summary of analysed cell lines for the *in silico* validation of *DPYSL3*, *FBLIM1*, *SDPR* and *P4HA2*.

Cell line	Disease	Type	Morphology
A549[1]	Lung cancer	Primary tumour – lung	Epithelial
HCC827[1]	Lung cancer	Primary tumour – lung	Epithelial
NCI-H358[1]	Lung cancer	Metastasis – alveolus	Epithelial
PANC-1[2]	Pancreatic cancer	Primary tumour – pancreas/duct	Epithelial
ARPE-19[3]	Healthy tissue	Healthy – retina - eye	Epithelial

[1](Sun, Yuting, Daemen et al. 2014)
[2](Maupin, Sinha et al. 2010)
[3](Takahashi, Nagano et al. 2010)

All three lung cancer cell lines showed a significant increase in the expression of *DPYSL3* upon stimulation with TGF-β (Fig. 5.17A-C), whereas the intensity of induction varied. The treatment of PANC-1 showed a slight increase in the expression; however, this increase did not show a significant change (Fig. 5.17D). Compared to this, treatment of ARPE-19 showed both a significant and the most intense increase in the expression of *DPYSL3* when the stimulation of all 5 cell lines was compared (Fig. 5.17E).

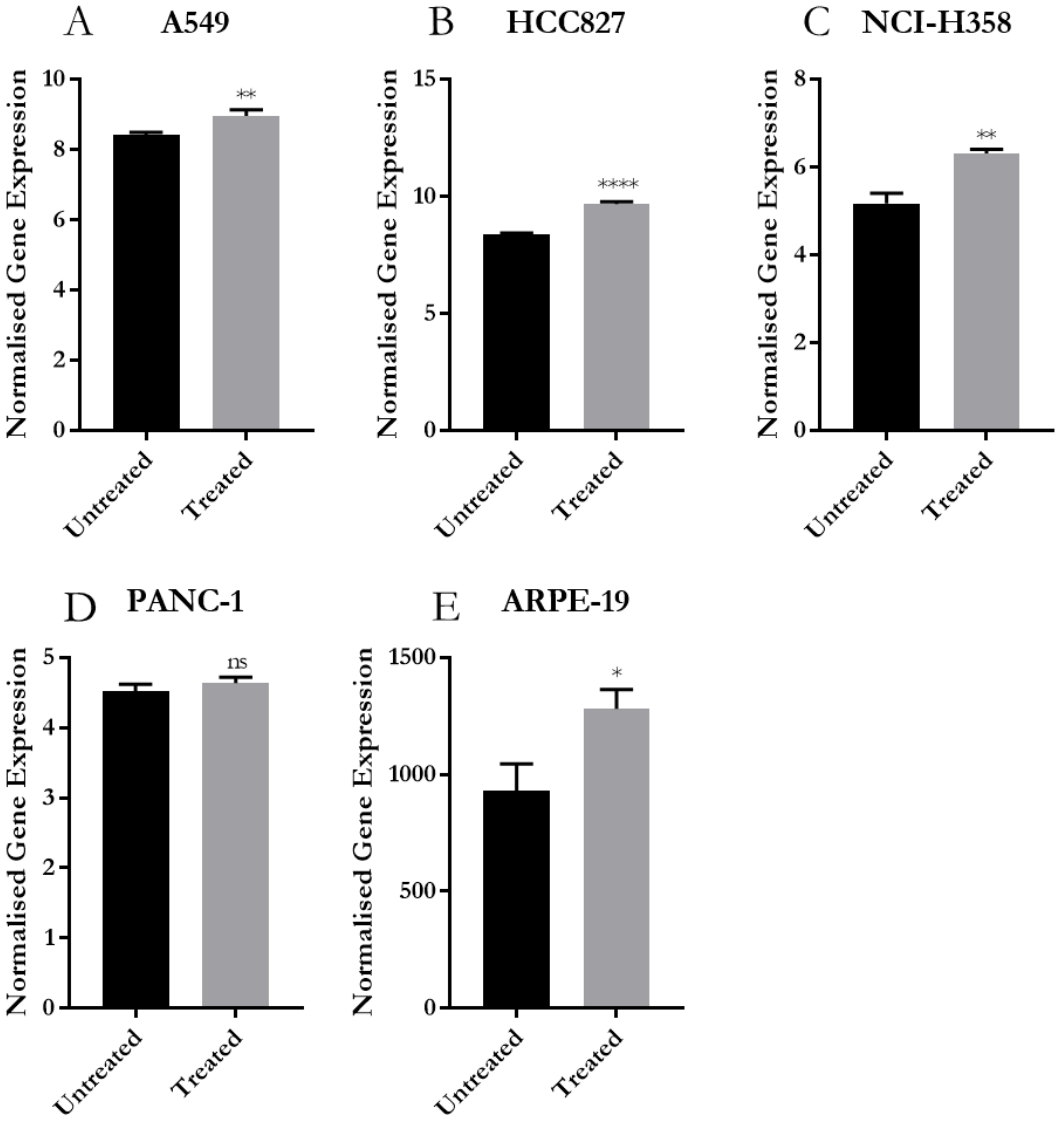


Figure 5.17: *In silico* gene expression analysis for *DPYSL3* generated from cell-derived whole transcriptome analyses of EMT-induced cell lines. A549, HCC827 and NCI-H358 (GSE49644) (Sun, Yuting, Daemen et al. 2014), PANC-1 (GSE23952) (Maupin, Sinha et al. 2010) and APRE-19 (GSE12548) (Takahashi, Nagano et al. 2010). The profiles were generated in triplicates per condition using Affymetrix Human Genome U133 Plus 2.0 Array

The expression of *FBLIM1* showed a significant change in all 3 lung cancer cell lines; however, the expression was upregulated upon stimulation in A549 and NCI-H358 (Fig. 5.18A+C), whereas the expression was decreased in HCC827 (Fig. 5.18B). A faint increase in the expression was detected in PANC-1 upon stimulation, but the difference was not significant (Fig. 5.18D). *FBLIM1* was shown to be reduced in APRE-19, but with a high degree of variation. Based on this, the decreased expression did not present a significant difference (Fig. 518E).

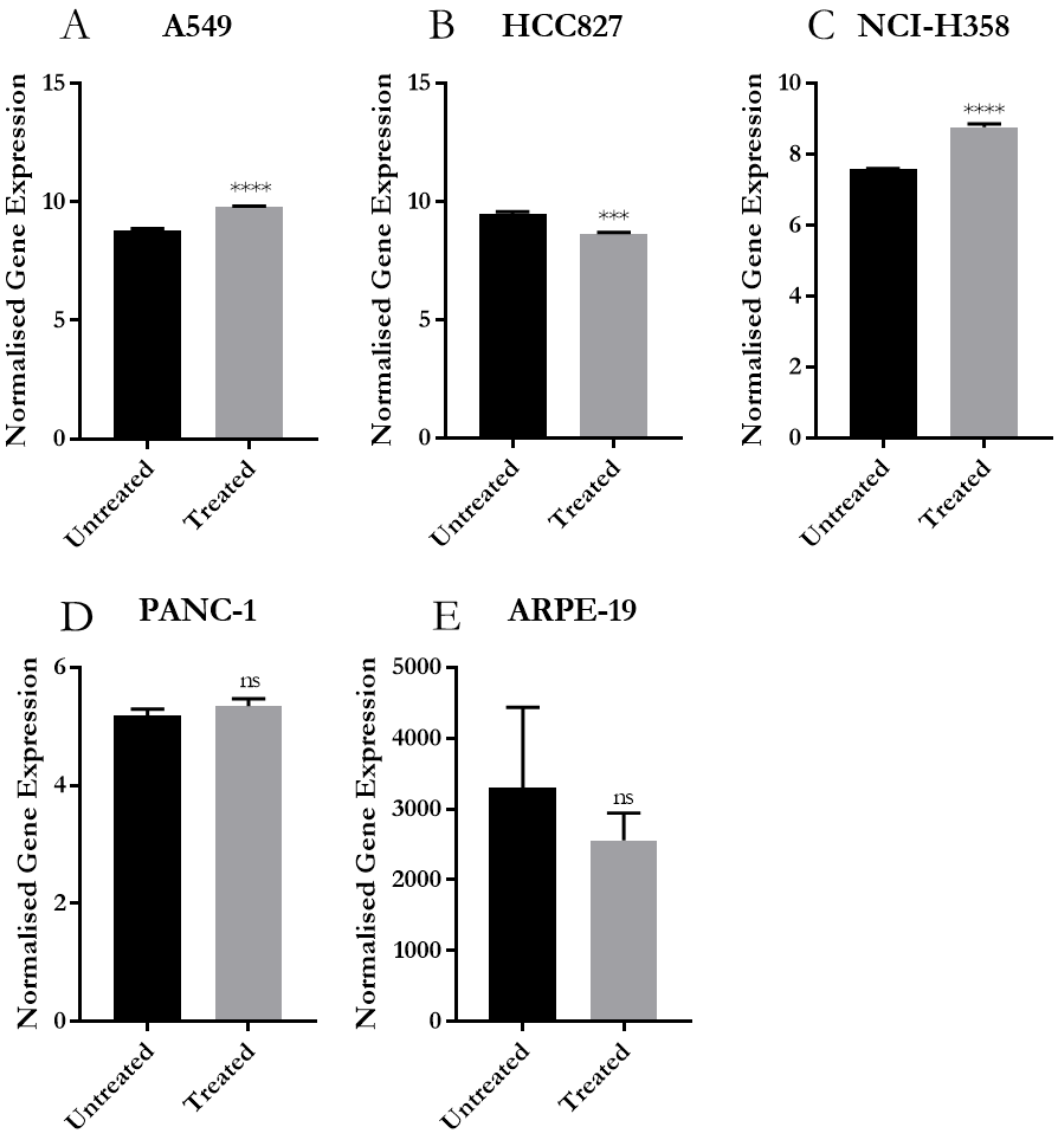


Figure 5.18: *In silico* gene expression analysis for *FBLIM1* generated from cell-derived whole transcriptome analyses of EMT-induced cell lines. A549, HCC827 and NCI-H358 (GSE49644) (Sun, Yuting, Daemen et al. 2014), PANC-1 (GSE23952) (Maupin, Sinha et al. 2010) and APRE-19 (GSE12548) (Takahashi, Nagano et al. 2010). The profiles were generated in triplicates per condition using Affymetrix Human Genome U133 Plus 2.0 Array.

The analysis of *SDPR* showed a consistent, significant decrease in its expression upon stimulation (Fig. 5.19A-E) in all 5 cell lines. The strongest reduction was observed in ARPE-19 with approximately 9-fold decrease, followed by A549 and HCC827, with a fold change reduction ranging from -3 to -2, respectively (Fig. 5.19A+B). The least intense decrease was observed in the pancreatic cell line PANC-1 (Fig. 5.19D).

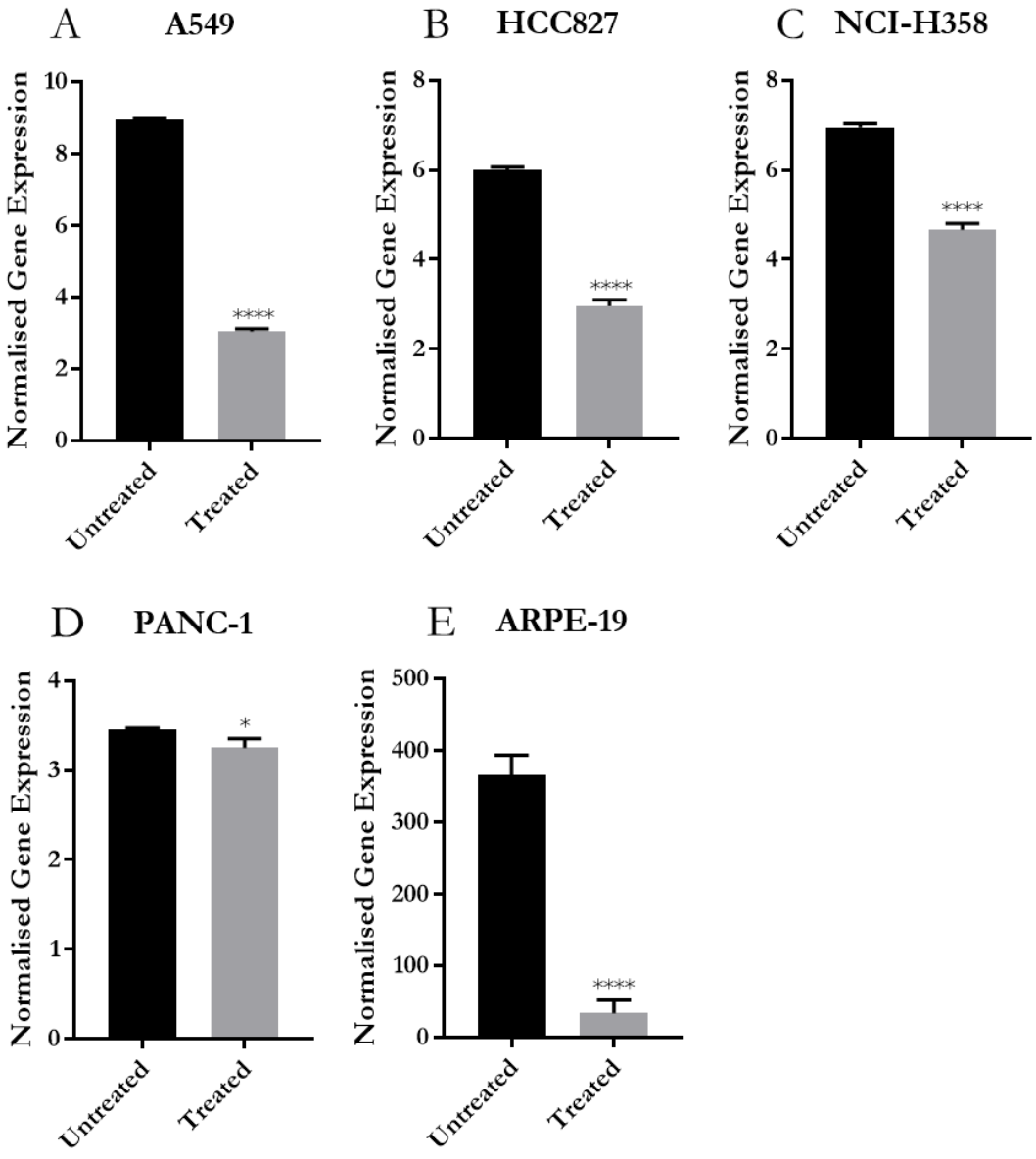


Figure 5.19: *In silico* gene expression analysis for *SDPR* generated from cell-derived whole transcriptome analyses of EMT-induced cell lines. A549, HCC827 and NCI-H358 (GSE49644) (Sun, Yuting, Daemen et al. 2014), PANC-1 (GSE23952) (Maupin, Sinha et al. 2010) and APRE-19 (GSE12548) (Takahashi, Nagano et al. 2010). The profiles were generated in triplicates per condition using Affymetrix Human Genome U133 Plus 2.0 Array.

In all 5 analysed cell line models, a significant alteration in the expression of *P4HA2* was shown. A significant increase was detected in all cancerous cell lines, namely A549 (Fig. 5.20A), HCC827 (Fig. 5.20B), NCI-H358 (Fig. 5.20C) and PANC-1 (Fig. 5.20D), whereas a significant reduction was observed in the healthy tissue derived cell line ARPE-19 (Fig. 5.20E).

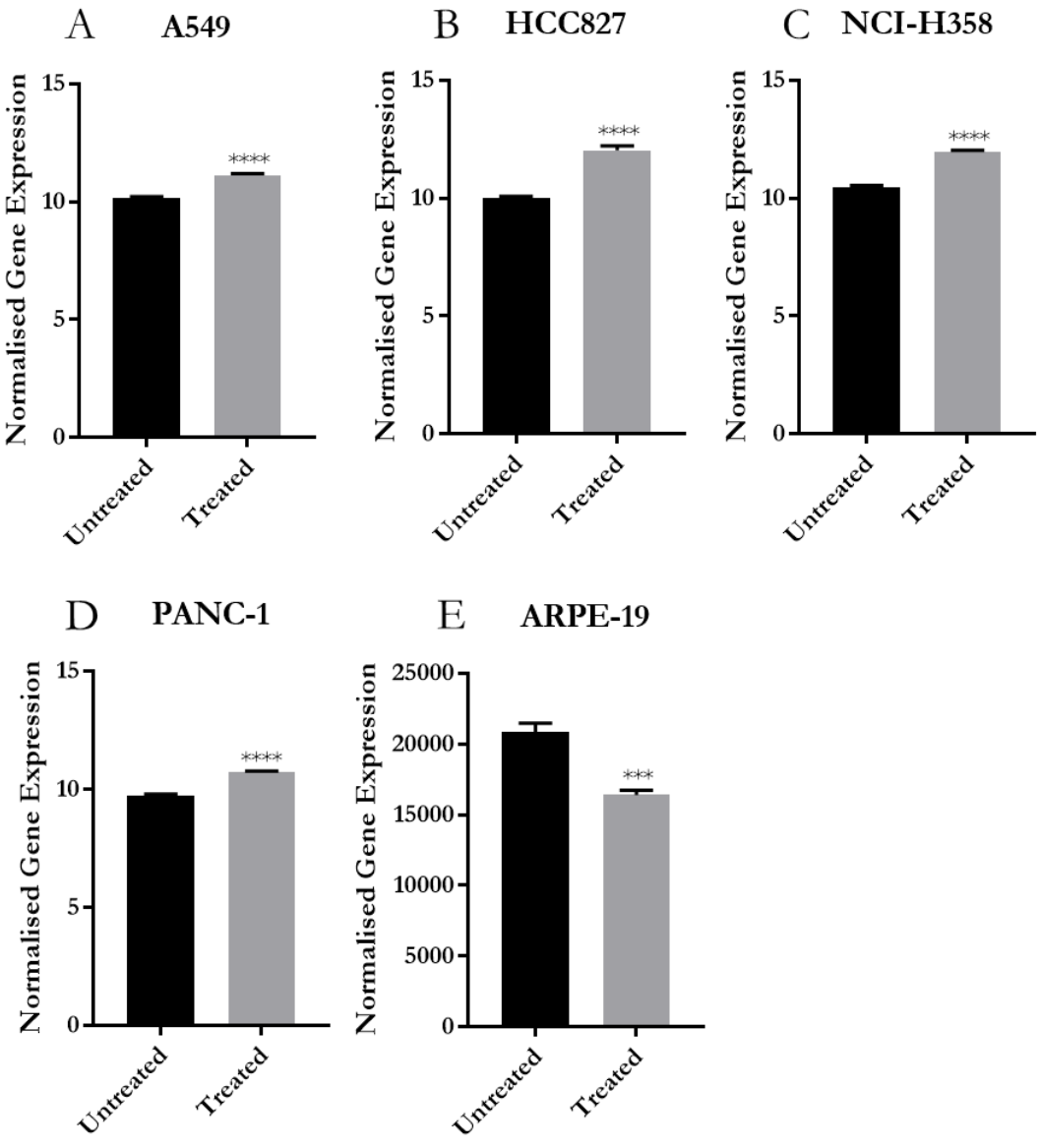


Figure 5.20: *In silico* gene expression analysis for P4HA2 generated from cell-derived whole transcriptome analyses of EMT-induced cell lines. A549, HCC827 and NCI-H358 (GSE49644) (Sun, Yuting, Daemen et al. 2014), PANC-1 (GSE23952) (Maupin, Sinha et al. 2010) and APRE-19 (GSE12548) (Takahashi, Nagano et al. 2010). The profiles were generated in triplicates per condition using Affymetrix Human Genome U133 Plus 2.0 Array.

5.2.6.2 *In silico* validation of selected markers in publicly available datasets of patient derived transcriptomic profiles

The analysis of tissue microarrays enables the study of protein expression in patient material and can give indications on the expression intensity and protein localisation in healthy and diseased tissue specimens. However, the available cores on a single slide present only a snapshot of a few cases and tissue areas, and potential clinical associations of novel markers can be missed. To overcome the limitations of tissue microarrays, an *in silico* analysis of patient-derived gene expression profiles was performed in addition to the previously performed wet lab validation (Chapter 5.2.4.2).

5.2.6.2.1 Comparison of benign tissue with primary PCa and CRPC

The utilised datasets were generated previously in a study on the lethal landscape of castration-resistant prostate cancer (Grasso, Wu et al. 2012). The sample material was categorised into normal (benign prostate tissue) (n = 28), localised PCa (n = 59) and castration-resistant prostate cancer/metastasis (n = 35) (Grasso, Wu et al. 2012).

In the analysed data, the expression of *DPYSL3* decreased significantly with disease progression, showing the highest expression in non-cancerous prostate tissue, a significant lower expression in primary PCa and a further, significant decrease in CRPC (Fig. 5.21A).

When investigating *FBLIM1*, a significantly lower expression was observed in primary PCa compared to healthy tissue and CRPC. The expression of *FBLIM1* in healthy and CRPC specimens presented a comparable intensity and did not show any significant difference (Fig. 5.21B).

A similar expression pattern to *DPYSL3* was also observed in *SDPR*, (Fig. 5.21C), which showed a significant reduction from healthy tissue, via primary PCa and CRPC. The reduction of its expression was shown to be significantly different across all 3 conditions (healthy, localised PCa and CRPC) (Fig. 5.21C).

The comparison of *P4HA2* expression in normal tissue with primary PCa and CPRC tissue showed significant differences. Initially, the expression of P4HA2 was slightly decreased in primary PCa compared to healthy, whereas the expression in CRPC was

increased by about 1.5 and 2-fold when compared to healthy tissue and primary PCa, respectively (Fig. 5.21D).

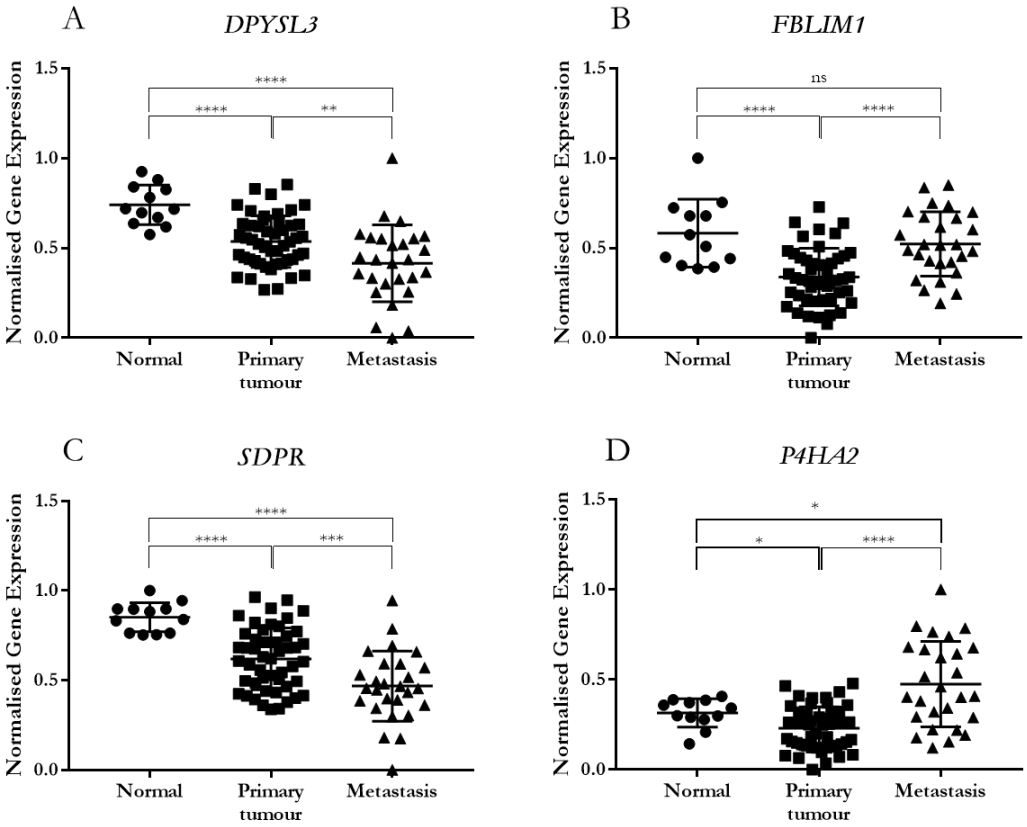


Figure 5.21: Gene expression of *DPYSL3* (A), *FBLIM1* (B), *SDPR* (C) and *P4HA2* (D) in normal, primary tumour and CRPC tissue generated from patient-derived whole transcriptome analyses (Grasso, Wu et al. 2012). The data is publicly available under the following accession number: GSE35988. The profiles were generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F. Gene expression was normalised using min-max normalisation. The sample material was categorised into normal (benign prostate tissue) (n = 28), localised PCa (n = 59) and castration-resistant prostate cancer/metastasis (n = 35).

5.2.6.2.2 Comparison of the gene expression of all 4 markers across different Gleason scores

A second dataset was generated as part of "The Cancer Genome Atlas" TCGA- project. In this project, large sample numbers of various cancers were selected and analysed on multiple omic levels, including the genome and transcriptome. Here, the gene expression profiles of the 4 markers of interest were selected and their expression compared across four different Gleason scores (Abeshouse, Ahn et al. 2015), GS6 (n = 44), GS7 (n = 247), GS8 (n = 64) and GS9 (n = 137).

A significant difference could be observed in *DPYSL3* expression across GS7, GS8 and GS9 compared to GS6, whereas the expression of *DPYSL3* was reduced with the increase in Gleason score (Fig. 5.22A). The same pattern was also observed for *SDPR* (decrease with increased Gleason Grade), in which the differences across the Gleason scores were also shown to be significant (Fig. 5.22C). *FBLIM1* and *P4HA2* showed only limited association with defined Gleason scoring, presenting for all or the majority of the comparisons no significant differences (Fig. 5.22B+D).

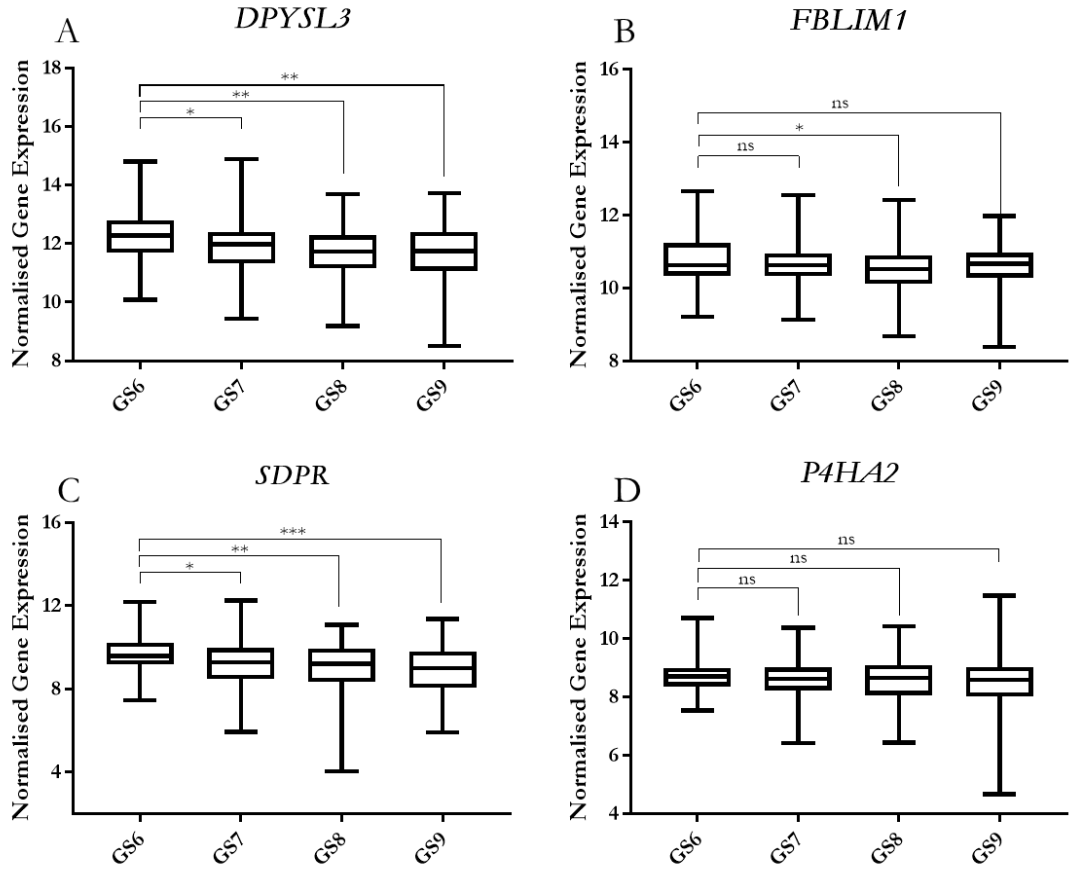
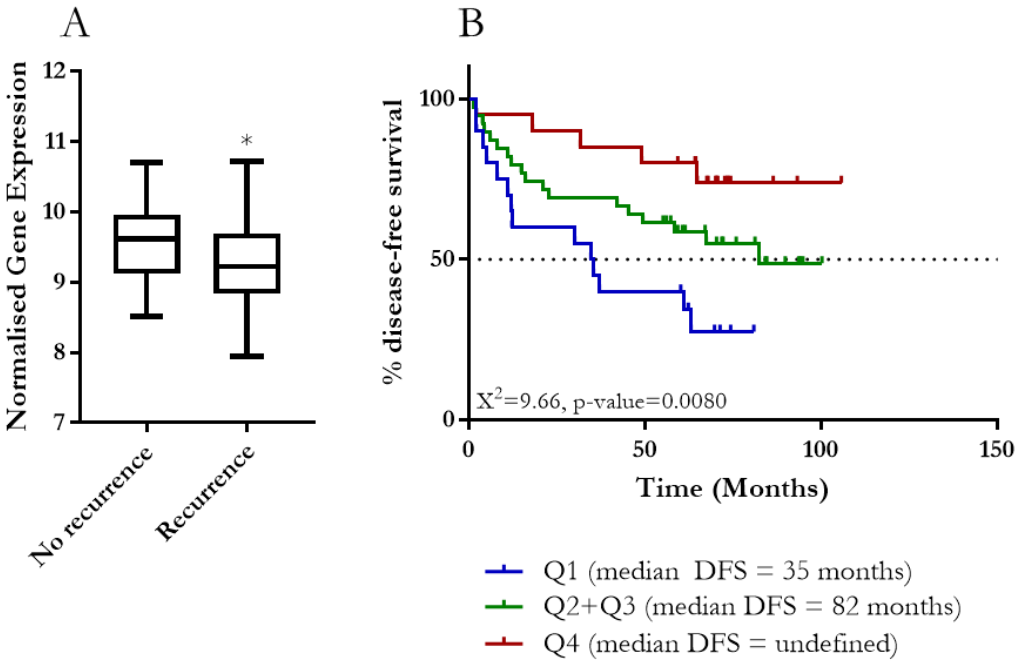


Figure 5.22: Gene expression of *DPYSL3* (A), *FBLIM1* (B), *SDPR* (C) and *P4HA2* (D) across four different Gleason scores generated from patient-derived whole transcriptome analyses (Abeshouse, Ahn et al. 2015). The data is publicly available from the TCGA data portal (https://portal.gdc.cancer.gov/) under the project number TCGA-PRAD. The profiles were generated using RNA-sequencing on a HiSeq2000 platform. The sample material was categorised into the Gleason scores GS6 (n =44), GS7 (n = 247), GS8 (n = 64) and GS9 (n = 137).

5.2.6.2.3 Impact of *DPYSL3* and *SDPR* expression on disease-free survival of prostate cancer patients

Previous analyses (Chapter 5.2.6.2.2) showed a significant association of *DPYSL3* and *SDPR* expression with the tissue-derived Gleason score (Fig. 5.23A+C) and were therefore subjected to further validation. For this, another publicly available dataset was selected, in which gene expression profiles as well as clinical information regarding the relapse-status were supplied (Glinsky, Glinskii et al. 2004). Overall, the dataset contained 79 patients, 37 without and 42 with disease-recurrence.

The expression of *DPYSL3* was shown to be significantly lower in patients with disease-recurrence compared to patients without (Fig. 5.23A). Additionally, the Kaplan-Meier analysis of quartiles, sorted by gene expression from low to high, showed visible differences (Fig. 5.23B). In *DPYSL3*, patients assigned to Q4 (presenting the highest expression) showed a significant longer disease-free survival length compared to the other quartiles, in particular Q1 (lowest gene expression). In addition, less than 50 % of patients assigned to Q4 suffered disease recurrence, therefore no median recurrence-free survival (RFS) was available. In a further comparison, the RFS of Q2+Q3 represents more than double the RFS of patients assigned to Q1. To further validate the predictive abilities of *DPYSL3*, the data was subjected to a univariate cox regression analysis, showing a significant association of *DPYSL3* expression with time to relapse (Fig. 5.23C). Overall, these findings further support the previously detected changes of *DPYSL3* in which the reduced expression was annotated with a poorer cancer phenotype (Fig. 5.22A and Fig. 5.21A).

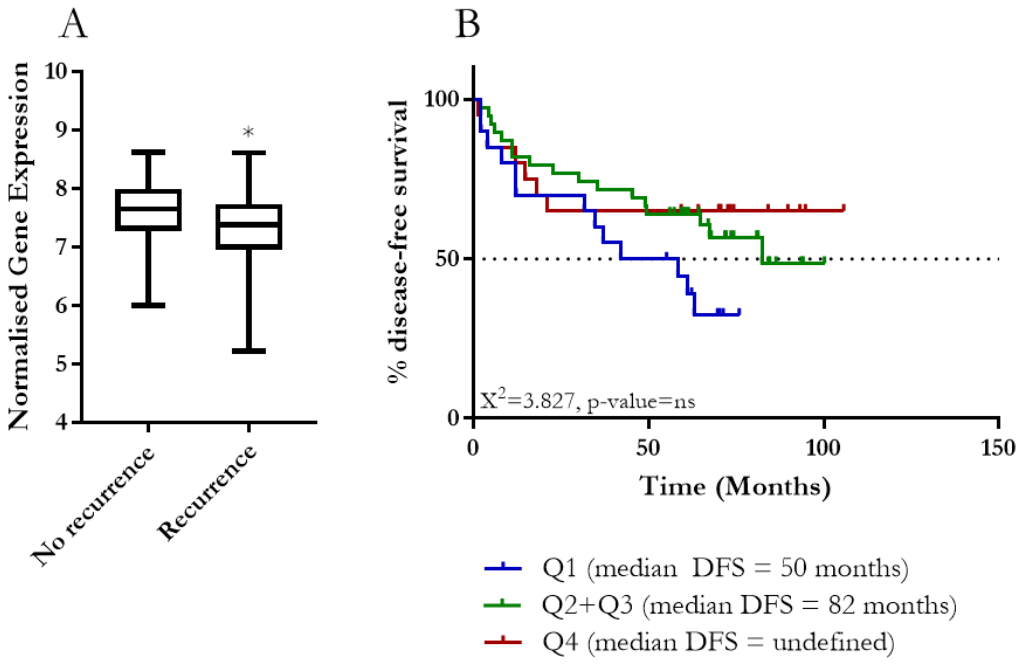


C

Variable	HR (95% CI)	p-value	β-value
DPYSL3	0.515 (0.295 to 0.898)	0.019	-0.663

Figure 5.23: Gene expression of *DPYSL3* in patient-derived whole transcriptome datasets of recurrent and non-recurrent PCa (Glinsky, Glinskii et al. 2004). The data was downloaded through the following website: http://web.bioinformatics.cicbiogune.es/CANCERTOOL/index.html (Cortazar, Ana R., Torrano et al. 2018). The profiles were generated using Affymetrix U95Av2. The sample material was categorised, dependent on comparison. A: No recurrence n = 37, recurrence n = 42; B: sorted by expression from lowest to highest and then separated into quartiles (Q1: n = 20, Q2+Q3: n = 39, Q4: n = 20); C: Univariate Cox regression analysis using *DPYSL3*, here the cases were not categorised, and all cases were used.

Using *SDPR*, comparison of patients without and with relapse showed a significant decrease in the overall expression of this gene (Fig. 5.24A), however the Kaplan-Meier analysis of the 3 quartile groups was unable to show a significant difference (Fig. 5.24B). Despite this, the comparison of median RFS times showed strong variations; patients with a lower *SDPR* expression show a RFS of 50 months, compared to 82 months of Q2 and Q3. Less than 50 % of patients assigned to Q4 experienced disease recurrence, therefore, no median time could be defined. Furthermore, the univariate cox regression analysis highlighted a significant association of *SDPR* expression with RFS (Fig. 5.24C). However, this association was less significant compared to *DPYSL3* (Fig. 5.23C).



C

Variable	HR (95% CI)	p-value	β-value
SDPR	0.620 (0.392 to 0.982)	0.042	-0.478

Figure 5.24: Gene expression of *SDPR* in patient-derived whole transcriptome datasets of recurrent and non-recurrent PCa (Glinsky, Glinskii et al. 2004). The data was downloaded through the following website: http://web.bioinformatics.cicbiogune.es/CANCERTOOL/index.html (Cortazar, Ana R., Torrano et al. 2018). The profiles were generated using Affymetrix U95Av2. The sample material was categorised, dependent on comparison. A: No recurrence n = 37, recurrence n = 42; B: sorted by expression from lowest to highest and then separated into quartiles (Q1: n = 20, Q2+Q3: n = 39, Q4: n = 20); C: Univariate Cox regression analysis using *SDPR*, here the cases were not categorised, and all cases were used.

5.3 Discussion

As previously mentioned, the generation of omic profiles commonly results in long lists of potential candidates. Here, the transcriptomic and proteomic profiles of both EMT models have resulted in large numbers of significant markers. The lists of genes and proteins have shown clear differences in their numbers (Tab. 5.1) and in total, a higher number of genes were significantly altered compared to proteins. This variation is mainly based on the detection and quantification limitations of proteomic approaches. Improvements in the technology and instrumentation over the last 5 years have enabled increases in the number of quantifiable proteins identified via mass spectrometry approaches (Shishkova, Hebert et al. 2016). A repeated analysis of the sample material generated in this PhD project using our more advanced mass spectrometry approaches, could most likely result in an increased number of quantified proteins.

In this study, the chosen approach for the identification of key markers was the generation of a highly confident core set of deregulated markers. This resulted in the identification of 13 conserved markers, consistently detected in both models (Tab. 5.2). Out of this list, 4 markers were selected for further studies, consisting of DPYL3, FBLI1, SDPR and P4HA2. The validation approaches for each potential novel biomarker were applied to cell line and patient-derived material.

The analysis of FBLI1 has shown a consistent upregulation across both cell lines and omic levels, showing slightly stronger induction in DU145 (Tab. 5.3). The increased expression upon stimulation was also shown through the analysis of MCF10A in both a stimulated and unstimulated state (Fig. 5.1), confirming the induction through TGF-β. Overall, in the analysis of cancer cell lines, *FBLIM1* showed the highest expression of all 4 analysed markers (Fig. 5.1), however it also demonstrated a lower expression in BCa cell lines compared to PCa cell lines (Fig. 5.1). A similar observation was shown in the screening of healthy tissue, where the expression of *FBLIM1* was the overall strongest compared to the other 4 markers. The *in silico* analysis of 5 additional EMT models have presented variable results in which 3 have shown an increased expression, whereas 2 have shown a reduced expression (Fig. 5.18). This inconsistency could potentially indicate a limited suitability of FBLI1 as a potential new biomarker for EMT. As mentioned previously, EMT is a highly conserved processed, and it would be expected that strongly associated markers present a strong consistency across multiple models. This might indicate that

FBLI1 is altered through EMT but is not directly associated to the activated pathway. These potential limitations were further supported through the comparison of the model expression with healthy prostate RNA, which indicated inverse results for both untreated cell lines, in which P5B3 was showing a higher expression, whereas DU145 showed a lower expression (Fig. 5.2). However, a significant increase, higher compared to healthy tissue, was shown in both cell lines upon stimulation (Fig. 5.2). Despite this increased expression in a healthy tissue RNA panel, the protein detected using IHC was very limited and the staining intensity ranged from low to not detected, which was also true for prostate cancer specimens (Fig. 5.10, Fig 5.14). Therefore, the analysis provided only limited information and no association with staining intensity and disease stage could be performed.

The *in silico* analysis of the expression of *FBLIM1* presented no significant difference between healthy and CRPC tissue, whereas primary PCa expressed a significantly lower expression compared to healthy and CRPC tissue (Fig. 5.21B). Also, the comparison of gene expression across the Gleason scores 6, 7, 8 and 9 showed limited significance (Fig. 5.22B). Previous studies on FBLI1, the Filamin-binding LIM protein 1, were also published using the name "Migfilin". It was shown that FBLI1 plays a role in cell adhesion, the actin cytoskeleton and as an integrin-activator (Das, M., Ithychanda et al. 2011, Ithychanda, Das et al. 2009). Research articles on the function and association of FBLI1 with cancer have shown information on an increase in disease malignancy through an increased FBLI1 expression (He, H., Ding et al. 2014, Toeda, Kasamatsu et al. 2018). Studies in oesophageal squamous cell carcinoma (ESCC) and oral squamous cell carcinoma (OSCC) have shown a regulation of cell migration and invasion through FBLI1 (He, H., Ding et al. 2014, Toeda, Kasamatsu et al. 2018). Also, the analysis of clinical specimens showed a significantly higher expression of FBLI1 in cancerous tissue compared to healthy specimens, confirming the here observed expression changes from primary PCa to CRPC (Fig. 5.21B), potentially highlighting an association of FBLI1 expression with disease outcome and overall survival length (Ou, Ma et al. 2012). In ESCC, a nuclear-cytoplasmic translocation from healthy to diseased tissue was observed, however this observation could not be confirmed in the analysed prostate TMAs of this study, based on the limited staining intensity across all samples (Fig. 5.14). The study on OSCC has furthermore identified an association of FBLI1 expression and its promotion of cellular migration, invasiveness and transendothelial migration (Toeda, Kasamatsu et

al. 2018). A potential explanation for the malignant function of FBLI1 was supplied by Seguin et al., who have stated that the increased expression of certain integrins can enhance the metastatic potential of tumours (Seguin, Desgrosellier et al. 2015). At the same time, Das et al., have shown that FBLI1 is enriched at cell-cell and cell-ECM sites, promoting integrin-activation through the displacement of filamin from integrins (Das, M., Ithychanda et al. 2011) and therefore categorising FBLI1 as an integrin activator. Furthermore, the silencing of *FBLIM1* resulted in a downregulation of *FN1*, a commonly known marker of EMT-induction (Das, M., Ithychanda et al. 2011).

The analysis of *P4HA2* (Prolyl 4-hydroxylase subunit alpha-2) identified a consistent upregulation of approximately 2.5-fold in both cell line models, with a slightly higher increase at the proteomic level (Tab. 5.3). The screening of various cell lines (Fig. 5.1) has highlighted an increase of *P4HA2* expression in MCF10A upon stimulation with TGF-β, also P4B6B and SAOS presented elevated levels of *P4HA2* expression. P4B6B (Harner-Foreman, Vadakekolathu et al. 2017) is a highly mesenchymal cell type, potentially supporting the induction of *P4HA2* through the development of a mesenchymal morphology upon stimulation. SAOS was shown to have a high expression of collagen IV (Pautke, Schieker et al. 2004), whereas P4HA2 is involved in the collagen synthesis. The *in silico* analysis of 5 cell line models (Fig. 5.20) has also presented a consistent increase of *P4HA2* expression in all 4 cancer cell line models, whereas a reduced expression was detected in the healthy tissue cell line. This mimics the expression pattern previously shown in the comparison of different prostate cancer stages (Fig. 5.21D). In the comparison of healthy tissue and both cell line models, a consistent increase from healthy to unstimulated to treated cells was observed (Fig. 5.2D+H). Furthermore, the expression of *P4HA2* in prostate tissue presented a lower expression compared to the majority of analysed tissue RNA (Fig. 5.6). The analysis of protein expression in healthy tissue could confirmed an elevated expression of *P4HA2* in the kidney and an overall very low to no expression in the remaining analysed tissues (Fig. 5.12). The protein expression observed in the prostate cancer TMA has also shown a very low expression overall, but the use of fluorescently-tagged secondary antibody has shown indications of a reduced expression in advanced prostate cancer (Fig. 5.16), however based on the limited number of patients no firm conclusions could be made.

An improved understanding on the impact of *P4HA2* on disease progression was possible through the *in silico* analysis of clinically-derived expression profiles. The expression of *P4HA2* was slightly decreased comparing healthy with primary PCa tissue, followed by a strong increase from primary PCa to CRPC (Fig. 5.21D). However, no differences in the expression intensity could be observed across the Gleason scores 6, 7, 8 and 9, highlighting limitations of *P4HA2* as a disease progression marker (Fig. 5.22D). A study by Xiang et al, has shown that silencing of P4HA2 decreases proliferation and invasiveness in 3D culture as well as impairment of collagen deposition (Xiong, Deng et al. 2014). Not only does the analysis of *in vitro* models highlight the association of *P4HA2* with cancer progression and survival, but also the analysis of patient material has shown that the expression of *P4HA2* is increased in BCa compared to healthy tissue, and in addition, is correlated with a poor prognosis (Gilkes, Chaturvedi et al. 2013, Xiong, Deng et al. 2014). Gilkes et al, suggested an association of P4HA2 with the organisation of collagen fibres of the ECM (Gilkes, Chaturvedi et al. 2013). The alignment of collagen surrounding the tumour can function as a disease prognosticator in BCa. Fibres that are aligned in a 90° angle to the tumour, so called perpendicular collagen, have been shown to be associated with a worse outcome (Conklin, Eickhoff et al. 2011). A study on the function of P4HA2 in breast cancer has shown that knockdown of P4HA2 in MDA-MB-231 cells results in an inhibition of tumour growth, as well as a reduction in tumour stiffness (Gilkes, Chaturvedi et al. 2013), which inhibits the migratory capabilities. Previous studies have shown that an elevated tumour stiffness can increase cell invasion and tumour metastasis (Reid, Kay et al. 2017). Such an increased expression of P4HA2 with a high stage PCa might represent a preparation of the primary tumour, through the increase of tumour stiffness, to spread in surrounding tissue and to develop metastasis.

The serum deprivation-response protein (SDPR), also known as caveolae-associated protein 2 (Cavin-2) was the only marker, out of the 4 analysed, that is downregulated upon stimulation with TGF-β; the downregulation was more intense in DU145 cells compared to P5B3 cells (Tab. 5.3). This downregulation was also documented in MCF10A cells upon stimulation with TGF-β (Fig. 5.1), as well as in the 5 *in silico* cell line models (Fig. 5.19), with all demonstrating a reduction in expression upon stimulation.

The comparison of *SDPR* expression in both cell line models with healthy prostate mRNA has shown a significant reduction in both models (Fig. 5.2C+G). Furthermore,

the validation using IHC has shown indications of a reduced expression with disease progression from stage II to stage IV PCa (Fig. 5.19), which was more apparent through the use of an IF-tagged secondary antibody on the same tissue sections (Fig. 5.19). The *in silico* analysis of patient-derived expression profiles further supported the *SDPR*-reduction associated with disease progression and EMT induction (Fig. 5.121C). In addition, a reduction of SDPR expression was correlated with increasing Gleason score (Fig. 5.22C), however *SDPR* did present limited capabilities for the prognosis of disease-recurrence (Fig. 5.24) based on a Kaplan-Meier analysis. Significantly lower expression levels of SDPR were measured in patients with disease-recurrence compared to patients without (Fig. 5.24B). Furthermore, the cox regression analysis has shown a significant association of *SDPR* expression with RFS (Fig. 5.24C). Previous studies have already proposed that a loss of SDPR could function as a marker for tumour progression in breast cancer (Ozturk, Papageorgis et al. 2016, Tian, Yu et al. 2016), and that SDPR is commonly silenced epigenetically by promotor DNA methylation (Tian, Yu et al. 2016). On the contrary, a depletion or SDPR loss was shown to enhance EMT induction and TGF-β pathway signalling activation (Tian, Yu et al. 2016). In general, the loss or reduction of SDPR was previously documented in various cancers, including bladder, colorectal, lung, pancreatic and ovarian cancers (Ozturk, Papageorgis et al. 2016), suggesting a conserved role across different tissue types in the inhibition of metastasis development through TGF-β signalling.

The analysis of both EMT models has shown a significant upregulation of DPYL3 on a gene and protein level, whereas the expression was visibly more strongly induced in P5B3 cells compared to DU145 cells (Tab. 5.2). This difference in the induction intensity could potentially be related to the nature of P5B3, being a single cell clone and showing a full response on the stimulation (Fig. 3.9), whereas DU145 is a heterogeneous cell line with a limited response to TGF-β, which is restricted to a subset of cells (Fig. 3.13). DU145 has shown responding and non-responding cells to the stimulation with TGF-β. Most likely, responding cells highlighted a strong deregulation of EMT-associated markers, whereas non-responding cells did not. The generated expression intensities of proteins and genes were therefore based on cells with a strong and a weak change of expression. The overall intensity must therefore be based on an averaged expression of a marker of interest. The mix of responding and non-responding cells might have resulted in a dilution of the mRNA/proteins and therefore resulted in a lower detected fold change.

The analysis of various cancer cell lines (Fig. 5.1) has not shown an association of *DPYSL3* with specific cancer aggressiveness or EMT state. This was for example shown in the lack of expression in P4B6. This model was previously described as spontaneous EMT and increased migratory potential was documented (Harner-Foreman, Vadakekolathu et al. 2017), and also the expression in P4B6B, a highly mesenchymal cell line, was shown to be very low, potentially representing a limited association of *DPYSL3* with a mesenchymal cell state. The additional *in silico* analysis of 5 independent models of EMT, induced through TGF-β alone or in combination with TNF-α (Fig. 5.17), has shown a consistent induction of *DPYSL3* upon stimulation, further supporting the measured expression changes of *DPYSL3* in the studied EMT models of P5B3 and DU145. Despite the significant induction across 7 independent EMT models, the analysis of patient-derived sample material highlighted an inverse directionality of *DPYSL3* expression with progressive disease.

Healthy tissue RNA has shown a higher expression when compared to the unstimulated cells of P5B3 (Fig. 5.2A+E), which was further supported by a strong protein expression measured in healthy prostate tissue (Fig. 5.8 and 5.9). Despite the measured upregulation of *DPSYL3* with EMT induction, the analysis of healthy prostate tissue specimens has shown a reduction of its expression from healthy tissue to adjacent prostate tissue, followed by stage II PCa. The lowest expression was shown in advanced PCa (Stage IV) (Fig. 5.13, Tab. 5.5). These initial observations were further confirmed through the *in silico* analysis of patient-derived gene expression profiles (Fig. 5.21A), highlighting a progressive reduction with disease state, showing the lowest expression in patients with CRPC. An expression reduction was also shown with increased Gleason scoring (Fig. 5.22A). Furthermore, *DPYSL3* showed predictive capabilities for disease-free survival (Fig. 5.23).

A literature review on the functional analysis of *DPYSL3* and its potential function as a cancer biomarker has also shown inconsistent results varying from cancer type to cancer type. The analysis of its function in hepatocellular carcinoma (HCC) has shown that DPYL3 suppresses cell proliferation and that knockdown of *DPYSL3* results in increased migratory capability of HCC cells. It was also observed that the mean expression of *DPYSL3* was reduced in HCC compared to healthy specimens (Oya, Kanda et al. 2015) and patients with a lower expression presented a significantly lower OS and RFS. An additional study on the methylation status of DPYL3 and its prognostic abilities for pelvic

lymph node metastasis in PCa has shown that DPYL3 promotor methylation of 15 % and above is highly predictive for lymph node metastasis (LNM) (Gao, X., Li et al. 2017). On the other hand, high expression in gastric cancer was associated with worse survival and a more malignant cancer phenotype. Furthermore, the expression positively correlates with a shorter recurrence-free survival (Kanda, Mitsuro, Nomoto et al. 2014). These results highlight the potential variation on the impact of DPYL3 on survival. A study by Matsunuma et al. on DPYL3 in claudin-low breast cancer has resulted in the hypothesis that DPYL3 functions as an EMT suppressor, which is activated by EMT regulators, resulting in a negative-feedback loop (Matsunuma, Chan et al. 2018). As previously discussed in the utility of biomarkers as therapeutic targets, markers need to be separated into "messengers" and "drivers" (Shen 2013). Messenger markers are changed as a consequence of activation, but do not cause effects such as tumour progression or metastasis. It might be the case that the upregulation of DPYL3 was driven as a response to EMT, which confirms the successful induction of the process in all analysed EMT models. On the other hand, the general function of DPYL3 would be the suppression of EMT, however this was not possible based on the continuous supply of stimulating and inducing cytokines, which overpowered the ability of DPYL3 to inhibit the process of EMT. In patients, the reduction of DPYL3 might result in a misbalance of EMT activity enabling a tumour to spread.

In conclusion, it can be said that all four markers can be reliably associated with the process of disease progression. Two markers, DPYL3 and SDPR, presented more consistent and conclusive results compared to P4HA2 and FBLI1. Despite this, the integrative approach of combining the transcriptomic and proteomic profiles of two independent EMT models has successfully identified a key collection of markers affiliated with the process of EMT and metastasis. However, it should be noted that this selection approach does not represent an ultimate solution and other methods for the integration of multi-omics datasets could have resulted in other, potentially better disease-associated biomarkers.

6. Chapter VI - Final discussion, conclusions and future work

6.1 General discussion

6.1.1 Introduction

Worldwide an estimated number of approximately 1.3 million men were diagnosed with prostate cancer in 2018, of whom about 500 000 were newly diagnosed in Europe (WHO, 2019). This identifies prostate cancer as the most common cancer in men in Europe. More than 95 % of these cases were diagnosed in men over 55, whereby the frequency of advanced disease is increasing with patient age (Scosyrev, Messing et al. 2012). Aside from the high incidence rate of PCa in men, the disease is also the 3rd most common cause of cancer in men in Europe with an estimated number of 110 000 deaths through PCa in 2018 alone (WHO, 2019a). Reduced chances of survival are correlated with the stage of prostate cancer at the time of diagnosis and increased mortality is mainly due to the development of metastasis (Chowdhury, Robinson et al. 2013). About 4 % of PCa patients will develop metastases, which reduces their 5-year survival rate to only 30 % (Thobe, Clark et al. 2011).

Since the majority of cancer-related deaths, not only in PCa, but overall, are related to the development of metastasis (Taketo 2011, Mehlen, Puisieux 2006), markers associated with this process are likely to be of high clinical utility in the surveillance and treatment of cancer patients. Markers for this process would enable the improved treatment decisions of potential systemic treatment after surgical removal of the primary tumour. It has been shown that the present uncertainty results in overtreatment, for example in BCa patients with lymph node negative diagnosis (Pantel, Brakenhoff 2004), where approximately 20 to 25 % of patients develop metastatic disease within 10 years, however 90 % of the patients within this category were subjected to chemotherapeutic treatment.

A key process commonly associated with the development of metastasis is "epithelial-mesenchymal transition". In general, EMT is an evolutionary highly conserved process (Lim, J., Thiery 2012), which is implicated during embryonal development, wound healing and fibrosis (Kalluri, Weinberg 2009). However, during cancer this process is activated, resulting in a metastasis-initiating mechanism. Here, polarised epithelial cells, which are

attached to a basement membrane and the neighbouring cells, undergo multifactorial changes to acquire mesenchymal cell properties. These changes result in altered gene and protein expression, which leads to increased motility through the degradation of intracellular contacts, increased invasiveness, migratory potential and resistance to apoptotic signals (Kalluri, Weinberg 2009). On a molecular level these changes are shown through a reduction of epithelial gene expression and an increase in mesenchymal associated genes. These changes are based on multiple molecular alterations such as the activation of EMT specific transcription factors (*SNAI1*, *SNAI2*, *TWIST1*, *TWIST2*, *ZEB1*) and an altered expression of additional proteins, including VIME, FINC, CADH1 and CADH2. This process is not unidirectional and it should be highlighted that the cells can reverse the process back into an epithelial morphology; this is called mesenchymal to epithelial transition (Lim, J., Thiery 2012).

The focus of this study was on the use of an integrated multi-omics approach for the discovery of novel disease-associated biomarkers in PCa and markers indicative for the process of EMT. Based on this, the study could be separated into 3 major milestones; (1) the development of well-characterised EMT models, (2) the generation and validation of omic profiles and the use of those to further characterise the derived EMT models, and (3) the discovery and validation of novel disease-associated biomarkers in PCa

6.1.2 TGF-β stimulation induces an EMT-like phenotype in the prostate cancer cell lines P5B3 and DU145 and alters EMT-associated signalling pathways

Many studies on the use of *in vitro* models of EMT achieved the induction of this process using cytokines, such as epidermal growth factor (EGF) (Grassi, de Souza Palma et al. 2017), tumour necrosis factor α (TNF-α) (Wang, H., Wang et al. 2013), as well as hepatocyte growth factor (HGF) (Liu, Fang, Song et al. 2017). In addition, many studies also supported the use of TGF-β for the activation of the EMT program in cell line models of liver (Lin, X., Liu et al. 2018), breast (Melzer, von der Ohe et al. 2017) and gastric cancer (Zhang, H., Liu et al. 2013). However, despite these findings, many of them have based their validation and further characterisation on the analysis of single EMT-associated genes and proteins (Zhau, Odero-Marah et al. 2008, Waldmeier, Meyer-Schaller et al. 2012, Liang, Fu et al. 2015), or the analysis of single omic levels. Furthermore, the generation of single omic data is often not used to its fullest potential and is mostly analysed for the identification of major deregulated markers (Katz, Dubois-Marshall et al.

2011, Lenferink, Cantin et al. 2010, Mikula, Rubel et al. 2011). Despite their infrequent use, the application of pathway analysis tools or systems biology approaches can harbour a more in-depth understanding of changes induced or present within the analysed sample cohort (Kim, Park et al. 2010, Kanda, M., Shimizu et al. 2016). In addition, the studies on EMT are commonly based on the selection of one single cell line. Such cell lines are commonly generated from the metastatic tumour site (Zhau, Odero-Marah et al. 2008, Lim, M., Chuong et al. 2011, Neal, Mckeithen et al. 2011).

In this study however, two cell lines (P5B3 and DU145), of which one was derived from a primary tumour (P5B3), were selected for the generation of inducible models of EMT and both were characterised using an integrated multi-omics approach, analysing gene and protein expression profiles from sample material generated at the same time point. This enabled the validation of the suitability of both models, not only on a wet-lab based approach but also through the analysis of omic profiles using an *in silico* pathway analysis based on pathway topology.

Both cell line models were treated with 10 ng/ml TGF-β over a period of 10 days, which has highlighted morphological changes indicative of a response to the stimulation with TGF-β, which presented itself with single, elongated cells of P5B3 (Fig. 3.9) and grouped, elongated cells of DU145 (Fig. 3.13). The analysis of molecular and proteomic EMT-markers have correlated the changes induced through the stimulation with an increased mesenchymal cell state. This was confirmed through an upregulation of EMT markers, such as VIME and FINC, as well as the downregulation of CADH1. The comparison of time point expression measurements of both models with published EMT state profiling indicated the association of both models to an intermediate mesenchymal phenotype (Huang, R. Y., Wong et al. 2013).

In addition, the migratory potential of both models was analysed using a scratch/wound healing assay. Stimulated P5B3 cells have shown a strong increase in their behaviour, enabling a complete wound closure after 24 hours, whereas untreated cells presented a closure of less than 10 % during the same time frame. This supports the findings of a successful induction of a mesenchymal cell state in P5B3 and its use as a model for EMT and potentially a proxy for metastasis. DU145, however, did not show any significant differences in its behaviour between untreated and treated cells. Nonetheless, it has been

shown that EMT and migration do not necessarily correlate and that sometimes pre-EMT cells present a higher migratory potential compared to post-EMT cells (Schaeffer, Somarelli et al. 2014).

These results support the use of both models for the generation of omic profiles and the profiles generated led to an EMT-phenotype and furthermore achieved a more in-depth understanding of changes induced through the stimulation with TGF-β. The analysis of both models using their omics profiles have highlighted the induction of EMT through the stimulation with TGF-β was enabled through the activation of both, SMAD-dependent and SMAD-independent, signalling pathways. This induction was shown, independent from the morphological changes, which were only limited in the cell line DU145. It also highlighted that a holistic approach, using multi-omic profiles, can explain the observed changes in cellular behaviour more accurately. Therefore, the use of these models highlights their potential for a better selection of novel biomarkers based on the targeted pathway. Overall, the generation of omic profiles have shown that both analysis methods and cell line models enable a characterisation of the desired and induced phenotype, whereas the proteomic analysis has shown an enrichment of cytoskeletal-related changes. The strong enrichment of cytoskeletal-associated proteins detected through the proteomic analysis were shown to be strongly associated the induction of EMT and changes in cell motility and adhesion.

6.1.3 The integration of transcriptomic and proteomic profiles can identify novel biomarkers associated with EMT and prostate cancer progression

The majority of omic studies for the discovery of novel biomarkers focus on the study of single omic levels (Kafetzopoulou, Boocock et al. 2013, Hou, Lou et al. 2015, Cheng, Lei, Yang et al. 2012), commonly the transcriptome or genome, which are also frequently analysed together (Wang, L., Xiao et al. 2014). The proteome however, is commonly discussed as harbouring a great potential for biomarker discovery (Borrebaeck 2017, Jacobs, Adkins et al. 2005, McDonald, Yates 2002), but developed only more recently into a routinely analysed omics level for the discovery of novel disease-associated biomarker (Hou, Lou et al. 2015, Øverbye, Skotland et al. 2015, Beretov, Wasinger et al. 2015). A large proportion of multi-omic studies have used sample material that was mostly generated separately at different time points or is derived from publicly available sources (Gupta, Jayaram et al. 2015, Li, L., Wei et al. 2014, Wagner, Ball et al. 2018).

In this study, matching transcriptomic (P5B3 n=10, DU145 n=9) and proteomic (P5B3 n=10, DU145U n=9, DU145T n=8) profiles were generated from the same samples under the same conditions. In addition, the sample material was collected within 1 hour to reduce protein degradation. This number of replicates presents, based on current knowledge, one of the largest matching cancer cell-line derived datasets based on the proteome and transcriptome of two cell line models. The improved quality of this integrated time-correlated approach was highlighted through an improved association of genes and proteins in both cell line models. However, it needs to be noted that this information is based on a reduced number of markers due to the limited amount of confidently identified proteins (P5B3 n=84, DU145 n=38). This limited number of proteins can be attributed to the limitations of technologies at the time of sample generation. A potential repeat for the sample analysis would most likely result in an increased identification and quantitation of proteins of up to 5000 proteins (Hülsmann, Kravic et al. 2018, Shishkova, Hebert et al. 2016).

The integration of both models and omic profiles enabled the identification of a core marker set of 13 genes and proteins, which were highly associated with the induced morphological and phenotypic changes. Four of these markers (DPYL3, FBLI1, SDPR and P4HA2) were subjected to further wet-lab and *in silico* validation approaches. The standard approach for the validation of a potential novel biomarker is the analysis of tissue microarrays (TMA) (Bubendorf, Lukas, Nocito et al. 2001). TMAs are glass slides spotted with small sections of tumour tissue of multiple patients. They commonly represent an easy route to obtain patient material for validation purposes. (Hassan, Ferrario et al. 2008). Despite this, as it was also the case here, the validation is not always successful and shows only limited differences between desired clinical parameters (De Matos, Trufelli et al. 2010, O'Hurley, Sjöstedt et al. 2014), especially in biomarkers that are proposed to be specific to certain cells. It has been suggested that the use of TMAs is more suitable for homogeneously distributed biomarkers (Merseburger, Kuczyk et al. 2003), whereas the process of EMT is most likely focussed on a subpopulation of cells, which are potentially not represented on this particular tumour section. Certainly, the possibility that a marker is not suitable always exists and is commonly the reason for a lack of validation. However, many other factors can influence the validation process, such as antibody specificity and the tumour sections present on the TMA themselves, since these sections only represent a snapshot of the tumour (Quagliata, Schlageter et al. 2014).

Based on the above-mentioned limitations of TMAs, an alternative method using patient derived transcriptomic profiles, which are publicly available, can enable a more in-depth study and characterisation of potentially novel biomarkers across larger patient pools. Here, all four markers were analysed in previously developed EMT models of independent studies (Sun, Yuting, Daemen et al. 2014, Maupin, Sinha et al. 2010, Takahashi, Nagano et al. 2010), which has highlighted the significant change of all markers through the induction of EMT. In addition to the cell-line based EMT models, the samples were validated in patient-derived transcriptomic profiles, in regards to their association with disease stage (benign, primary PCa or CRPC) and Gleason score. This analysis has shown a consistent and significant correlation of *DPYSL3* and *SDPR*. For this reason, these two markers were additionally analysed for their impact on disease-recurrence and disease-free survival. Despite the lack of detection or identified correlations with clinical information based on the TMAs used, the results have highlighted that all markers are strongly associated with EMT and PCa. The most significant results were achieved for *DPYSL3* and *SDPR*, of which their loss was shown to be highly associated with disease progression and recurrence in patient-derived data of PCa patients.

Studies have shown a significant association of a reduced *DPYSL3* expression with metastasis development, disease progression and migration. This was presented in studies of lung (Yang, Jiang et al. 2018), prostate (Gao, X., Li et al. 2017, Li, B., Li 2017) and liver cancer (Oya, Kanda et al. 2015). Overall, a large proportion of publications have identified *DPYSL3* as a metastasis-inhibitor and that a reduced expression has a negative impact on clinical outcome. The work of Gao et al has highlighted a potential link between the changed *DPYSL3* expression and its promotor methylation (Gao, X., Li et al. 2017). In this study, the expression of *DPYSL3* was shown to be increased in both cell line models, which could be explained by the work of Matsunuma and colleagues, which have proposed that *DPYSL3* functions as EMT suppressor regulating the EMT activation through a negative feedback loop (Matsunuma, Chan et al. 2018).

The tumour suppressor gene *SDPR* was shown to be reduced in this study. Previous studies have shown concordant results in a study on BCa, which has suggested that *SDPR* could be of potential use as clinical biomarker in BCa (Ozturk, Papageorgis et al. 2016). Its clinical applicability for the prognosis of disease progression was further supported

through a study in hepatocellular carcinoma, where the expression of *SDPR* was significantly associated with tumour differentiation and TNM stage. Furthermore, a lower expression of *SDPR* was associated with poorer survival (Jing, Luo et al. 2016).

In conclusion, it can be said that the integration of multi-omic profiles, derived from two independent cell line models, has enabled the identification of potential novel disease-associated biomarkers in PCa, which was supported through previously conducted studies highlighting the suitability of DPYL3 and SDPR in a clinical setting.

6.2 Conclusion

This study has generated two inducible models of EMT and successfully applied these to a novel pipeline describing a process from model to biomarker. This approach resulted in the identification of SDPR and DPYL3 as potential novel biomarkers for disease-progression in PCa. In addition to this, the generation of the matching omic datasets of two independent cell lines was able to contribute to the understanding of gene and protein expression correlation, highlighting the improvements in the correlation that can be made through connected sample collection with minimal time difference. In addition to the novel findings and discoveries made in this study, using the generated data, the potential of this dataset is not yet exhausted and can be used for future studies, for example in a more in-depth study on changes upon stimulation and EMT.

However, despite the successful use of this data, potential limitations need to be highlighted, such as the use of cell line models for the discovery of disease-associated biomarkers. Here, in this case, the discovered biomarkers could be successfully validated, however, it is crucial to select the model of choice carefully. Cell lines are highly artificial systems and it is of crucial importance to generate meaningful output, whereas their artificial nature can be overcome partially through the use of multiple models.

6.3 Future work

This study has highlighted the development of two inducible EMT models and their successful application in a multi-omics approach for the discovery of markers associated with the process of EMT and the progression of prostate cancer. This study has identified 13 markers of particular interest, of which 4 were characterised through the use of *in vitro* experiments and further validated using wet-lab approaches on clinically-derived specimens and *in silico* analyses. This process has shown an association of all 4 markers with EMT and disease progression, whereas SDPR and DPYL3 have presented a stronger potential in their function as novel candidates in the prediction of disease progression and recurrence in prostate cancer (see Chapter V). Additional work should be focussed on multiple aspects of this study, regarding technical advances, model characterisation and the function of selected markers, and furthermore the evaluation of both markers for the use as routine biomarkers.

To further understand the biological association of induced changes and identified markers, a more in-depth characterisation of both developed EMT models is necessary. Studies have shown that cells that underwent EMT commonly present an increased resistance to therapy (Shibue, Weinberg 2017) in association with a reduced proliferation rate (Tsai, Yang 2013). For this reason, assays to define therapy resistance of both models to standard care therapeutics, such as Docetaxel or Dabazitaxel, as well as proliferation rate, are important for the development of a more in-depth understanding. In this study, scratch assays were performed and gave first insights in changes of migratory behaviour of both cell line models, however for a better understanding a more advanced and realistic approach should be chosen. Such an approach could be the use of a migration assay based on Transwell plates, which characterises the migration of cells through their capability to move from an upper layer through a permeable membrane.

As mentioned previously, advances in mass spectrometry analysis and data processing enables the routine identification and quantitation of a higher number of proteins (up to 5000) within 90 min, compared to the number identified in this study (~2000). For this reason, a repeated analysis of the sample material for the generation of an improved library could deliver important information on potential newly identified proteins associated with EMT, as well as information on the correlation of gene and protein expression.

Aside from the use of advances in the instrumentation, the study of the biological function and potential associated interactions of DPYL3 and SDPR should be performed. As mentioned in Chapter V, published studies have postulated an association of changes in the methylation of the promoter region of SDPR (Tian, Yu et al. 2016) and DPYL3 (Gao, X., Li et al. 2017) with the induced changes in gene expression. Based on these results, future work should be focussed on the study of the methylation status of SDPR and DPYL3 in the generated cell line models and downstream in clinical specimens. The identification of altered methylation intensities of promoter regions associated with *DPYLS3* and *SDPR* and their function in disease progression could support the treatment decision of clinicians for the use of demethylating agents as alternative treatment options (Howell, Liu et al. 2010).

The generation and integration of the transcriptomic and proteomic datasets has highlighted two potential new biomarkers for the use in a clinical setting, mainly based on the *in silico* analysis of clinically-derived transcriptomic profiles. However, the suitability of both markers for the prediction of disease-progression in PCa has to be further evaluated. A biomarker predictive for disease progression should present certain capabilities, such as the detection in easily obtained and mini-invasive sample material. Most-routinely used sources are blood and urine samples.

For this reason, an initial evaluation step could be the analysis of gene and protein expression of both markers in the secretomes of both cell line models. A successful detection of either of the two, or both markers, in the secretomes of the cell line model, could be followed by the analysis of urine and blood samples of healthy, early stage and advanced prostate cancer. The screening of these different tissues would help collecting information on the presence/absence of these markers and also the potential variation in the expression across different disease stages.

Web references

Abcam, 2019. Protocol for previously generated cytoplasmic and nuclear fractions [Online]. Available at: https://www.abcam.com/protocols/subcellular-fractionation-protocol. (Accessed 22nd January 2019)

ATCC, 2018a. Characteristics of prostate cancer cell line DU145 [Online]. Available at: https://www.lgcstandards-atcc.org/products/all/HTB-81.aspx?geo_country=gb#characteristics. (Accessed 10th July 2018)

Broad Institute, 2018a. Example for biomarker discovery pipeline [Online]. Available at: https://www.broadinstitute.org/cptac/team-approach. (Accessed 21st May 2018)

Cancer Research UK, 2017a. Lifetime risk for development of cancer [Online]. Available at:http://www.cancerresearchuk.org/health-professional/cancer-statistics/risk/lifetime-risk (Accessed 23rd November 2017)

Cancer Research UK, 2017b. Cancer mortality rates [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/common-cancers-compared#heading-Zero (Accessed 19th November 2017).

Cancer Research UK, 2018a. Cancer incidence rates [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared#heading-Zero (Accessed 25th April 2018).

Cancer Research UK, 2018b. Types of cancer by originating cell [Online]. Available at: http://www.cancerresearchuk.org/what-is-cancer/how-cancer-starts/types-of-cancer#carcinomas (Accessed 21st May 2018)

Cancer Research UK, 2018c. Prostate cancer incidence rates [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence#collapseOne. (Accessed 21st May 2018).

Cancer Research UK, 2018d. Prostate cancer mortality rates [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/mortality#heading-One (Accessed 21st May 2018)

Cancer Research UK, 2019a. HPV and cancer [Online]. Available at: https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/infections-eg-hpv-and-cancer/hpv-and-cancer (Accessed 5th February 2019).

DeepSeq 2019 (https://www.nottingham.ac.uk/deepseq/) Accessed 11/01/19

Human Proteome Map, 2018. Human Proteome Map [Online]. Available at: http://www.humanproteomemap.org. (Accessed 21st May 2018).

Illumina, 2018. TruSeq® Stranded Total RNA Sample Preparation Guide [Online]. Available at: http://emea.support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqstrandedtotalrna/truseq-stranded-total-rna-sample-prep-guide-15031048-e.pdf. (Accessed 4th July 2018)

National Cancer Institute, 2018a. Definition of tumour marker [Online]. Available at: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/tumor-marker. (Accessed 21st May 2018).

PCEC, 2019. Histopathological representation of PCa tissue representing the 5 Gleason scores [Online]. Available at: https://www.prostateconditions.org/about-prostate-conditions/prostate-cancer/newly-diagnosed/gleason-score (Accessed 19th January 2019).

Prostate Cancer UK, 2017a. Risk of developing prostate cancer [Online]. Available at: https://prostatecanceruk.org/prostate-information/are-you-at-risk/infographic-what-is-my-risk (Accessed 31st October 2017)

Sciex, 2018a. Variable Window SWATH Acquisition [Online]. Available at: https://sciex.com/Documents/tech %20notes/Variable_Window_SWATH_Acquisition_TripleTOF_RUO-MKT-02-2879-B.pdf. (Accessed 21st May 2018).

Sciex, 2018. Improvements in data-independent acquisition in mass spectrometry [Online]. Available at: https://sciex.com/community/blogs/blogs/data-independent-acquisition-mass-spectrometry-with-the-power-of-swath. (Accessed 24th July 2018)

Waters, 2018. DIA acquisition through MS^E [Online]. Available at: http://www.waters.com/waters/promotionDetail.htm?id=10205700&alias=Alias_mse_INS&locale=en_GB. (Accessed 20th December 2018)

WHO, 2017a. Cancer statistics worldwide [Online]. Available at: http://www.who.int/cancer/en/ (Accessed 13th November 2017)

WHO, 2019a. Worldwide estimated deaths through prostate cancer [Online]. Available at: https://gco.iarc.fr/. (Accessed 16th January 2019)

WHO, 2019b. Impact of cancer development and HPV infection [Online]. Available at: https://www.who.int/news-room/fact-sheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer (Accessed 23rd January 2019)

Bibliography

AARON, L., FRANCO, O.E. and HAYWARD, S.W., 2016. Review of Prostate Anatomy and Embryology and the Etiology of Benign Prostatic Hyperplasia. *The Urologic clinics of North America,* **43**(3), pp. 279-288.

ABESHOUSE, A., AHN, J., AKBANI, R., ALLY, A., AMIN, S., ANDRY, C.D., ANNALA, M., APRIKIAN, A., ARMENIA, J. and ARORA, A., 2015. The molecular taxonomy of primary prostate cancer. *Cell,* **163**(4), pp. 1011-1025.

ACLOQUE, H., ADAMS, M.S., FISHWICK, K., BRONNER-FRASER, M. and NIETO, M.A., 2009. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *The Journal of clinical investigation,* **119**(6), pp. 1438-1449.

ADHYAM, M. and GUPTA, A.K., 2012. A review on the clinical utility of PSA in cancer prostate. *Indian journal of surgical oncology,* **3**(2), pp. 120-129.

AEBERSOLD, R. and MANN, M., 2016. Mass-spectrometric exploration of proteome structure and function. *Nature,* **537**(7620), pp. 347.

ANAND, P., KUNNUMAKARA, A.B., SUNDARAM, C., HARIKUMAR, K.B., THARAKAN, S.T., LAI, O.S., SUNG, B. and AGGARWAL, B.B., 2008. Cancer is a preventable disease that requires major lifestyle changes. *Pharmaceutical research,* **25**(9), pp. 2097-2116.

ANDERSON, N.L. and ANDERSON, N.G., 2002a. The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP,* **1**(11), pp. 845-867.

ANDERSON, N.L. and ANDERSON, N.G., 2002b. The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP,* **1**(11), pp. 845-867.

ARCHAKOV, A., LISITSA, A., PONOMARENKO, E. and ZGODA, V., 2015. *Recent advances in proteomic profiling of human blood: clinical scope,* .

ARGAST, G.M., KRUEGER, J.S., THOMSON, S., SUJKA-KWOK, I., CAREY, K., SILVA, S., O'CONNOR, M., MERCADO, P., MULFORD, I.J. and YOUNG, G.D., 2011. Inducible expression of TGFβ, snail and Zeb1 recapitulates EMT in vitro and in vivo in a NSCLC model. *Clinical & experimental metastasis,* **28**(7), pp. 593.

ARMSTRONG, A.J., HEALY, P., HALABI, S., VOLLMER, R., LARK, A., KEMENY, G., WARE, K. and FREEDLAND, S.J., 2016. Evaluation of an epithelial plasticity biomarker panel in men with localized prostate cancer. *Prostate cancer and prostatic diseases,* **19**(1), pp. 40-45.

ASSEFNIA, S., DAKSHANAMURTHY, S., GUIDRY AUVIL, J.M., HAMPEL, C., ANASTASIADIS, P.Z., KALLAKURY, B., UREN, A., FOLEY, D.W., BROWN,

M.L., SHAPIRO, L., BRENNER, M., HAIGH, D. and BYERS, S.W., 2014. Cadherin-11 in poor prognosis malignancies and rheumatoid arthritis: common target, common therapies. *Oncotarget,* **5**(6), pp. 1458-1474.

ATTISANO, L. and WRANA, J.L., 2002. Signal Transduction by the TGF-β Superfamily. *Science,* **296**(5573), pp. 1646-1647.

BAINBRIDGE, M.N., WARREN, R.L., HIRST, M., ROMANUIK, T., ZENG, T., GO, A., DELANEY, A., GRIFFITH, M., HICKENBOTHAM, M. and MAGRINI, V., 2006. Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach. *BMC genomics,* **7**(1), pp. 246.

BALBIN, O.A., PRENSNER, J.R., SAHU, A., YOCUM, A., SHANKAR, S., MALIK, R., FERMIN, D., DHANASEKARAN, S.M., CHANDLER, B. and THOMAS, D., 2013. Reconstructing targetable pathways in lung cancer by integrating diverse omics data. *Nature communications,* **4**, pp. 2617.

BALDI, P. and HATFIELD, G.W., 2011. *DNA microarrays and gene expression: from experiments to data analysis and modeling.* Cambridge university press.

BALK, S.P., KO, Y. and BUBLEY, G.J., 2003. Biology of prostate-specific antigen. *Journal of Clinical Oncology,* **21**(2), pp. 383-391.

BECK, M., SCHMIDT, A., MALMSTROEM, J., CLAASSEN, M., ORI, A., SZYMBORSKA, A., HERZOG, F., RINNER, O., ELLENBERG, J. and AEBERSOLD, R., 2011. The quantitative proteome of a human cell line. *Molecular systems biology,* **7**, pp. 549.

BENSALAH, K., MONTORSI, F. and SHARIAT, S.F., 2007. Challenges of cancer biomarker profiling. *European urology,* **52**(6), pp. 1601-1609.

BENTLEY, D.R., BALASUBRAMANIAN, S., SWERDLOW, H.P., SMITH, G.P., MILTON, J., BROWN, C.G., HALL, K.P., EVERS, D.J., BARNES, C.L. and BIGNELL, H.R., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature,* **456**(7218), pp. 53.

BERETOV, J., WASINGER, V.C., MILLAR, E.K., SCHWARTZ, P., GRAHAM, P.H. and LI, Y., 2015. Proteomic analysis of urine to identify breast cancer biomarker candidates using a label-free LC-MS/MS approach. *PloS one,* **10**(11), pp. e0141876.

BERTRAM, J.S., 2000. The molecular biology of cancer. *Molecular aspects of medicine,* **21**(6), pp. 167-223.

BIZZARRO, V., BELVEDERE, R., MIGLIARO, V., ROMANO, E., PARENTE, L. and PETRELLA, A., 2017. Hypoxia regulates ANXA1 expression to support prostate cancer cell invasion and aggressiveness. *Cell adhesion & migration,* **11**(3), pp. 247-260.

BLACK, D.L., 2000. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell,* **103**(3), pp. 367-370.

BONNOMET, A., SYNE, L., BRYSSE, A., FEYEREISEN, E., THOMPSON, E., NOËL, A., FOIDART, J., BIREMBAUT, P., POLETTE, M. and GILLES, C., 2012. A dynamic in vivo model of epithelial-to-mesenchymal transitions in circulating tumor cells and metastases of breast cancer. *Oncogene,* **31**(33), pp. 3741.

BORRÀS, E. and SABIDÓ, E., 2017. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. *Proteomics,* **17**(17-18), pp. 1700180.

BORREBAECK, C.A., 2017. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nature Reviews Cancer,* **17**(3), pp. 199.

BOUVARD, V., LOOMIS, D., GUYTON, K.Z., GROSSE, Y., EL GHISSASSI, F., BENBRAHIM-TALLAA, L., GUHA, N., MATTOCK, H. and STRAIF, K., 2015. Carcinogenicity of consumption of red and processed meat. *The Lancet Oncology,* **16**(16), pp. 1599-1600.

BRENNAN, E.P., MORINE, M.J., WALSH, D.W., ROXBURGH, S.A., LINDENMEYER, M.T., BRAZIL, D.P., GAORA, P.Ó, ROCHE, H.M., SADLIER, D.M. and COHEN, C.D., 2012. Next-generation sequencing identifies TGF-β1-associated gene expression profiles in renal epithelial cells reiterated in human diabetic nephropathy. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* **1822**(4), pp. 589-599.

BRUDERER, R., BERNHARDT, O.M., GANDHI, T., MILADINOVIC, S.M., CHENG, L.Y., MESSNER, S., EHRENBERGER, T., ZANOTELLI, V., BUTSCHEID, Y., ESCHER, C., VITEK, O., RINNER, O. and REITER, L., 2015. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Molecular & cellular proteomics : MCP,* **14**(5), pp. 1400-1410.

BUBENDORF, L., NOCITO, A., MOCH, H. and SAUTER, G., 2001. Tissue microarray (TMA) technology: miniaturized pathology archives for high-throughput in situ studies. *The Journal of pathology,* **195**(1), pp. 72-79.

BUBENDORF, L., SCHOPFER, A., WAGNER, U., SAUTER, G., MOCH, H., WILLI, N., GASSER, T.C. and MIHATSCH, M.J., 2000. Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Human pathology,* **31**(5), pp. 578-583.

BUERMANS, H. and DEN DUNNEN, J., 2014. Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* **1842**(10), pp. 1932-1941.

BURD, E.M., 2003. Human papillomavirus and cervical cancer. *Clinical microbiology reviews,* **16**(1), pp. 1-17.

BURKHOLDER, B., HUANG, R., BURGESS, R., LUO, S., JONES, V.S., ZHANG, W., LV, Z., GAO, C., WANG, B. and ZHANG, Y., 2014. Tumor-induced perturbations of cytokines and immune cell networks. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* **1845**(2), pp. 182-201.

BUSSEMAKERS, M.J., VAN BOKHOVEN, A., VERHAEGH, G.W., SMIT, F.P., KARTHAUS, H.F., SCHALKEN, J.A., DEBRUYNE, F.M., RU, N. and ISAACS, W.B., 1999. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer research,* **59**(23), pp. 5975-5979.

CALLE, E.E. and KAAKS, R., 2004. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nature Reviews Cancer,* **4**(8), pp. 579.

CARTER, H.B., PARTIN, A.W., WALSH, P.C., TROCK, B.J., VELTRI, R.W., NELSON, W.G., COFFEY, D.S., SINGER, E.A. and EPSTEIN, J.I., 2012. Gleason score 6 adenocarcinoma: should it be labeled as cancer? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* **30**(35), pp. 4294-4296.

CATALONA, W.J., SMITH, D.S., RATLIFF, T.L. and BASLER, J.W., 1993. Detection of organ-confined prostate cancer is increased through prostate-specific antigen—based screening. *Jama,* **270**(8), pp. 948-954.

CATALONA, W.J., PARTIN, A.W., SANDA, M.G., WEI, J.T., KLEE, G.G., BANGMA, C.H., SLAWIN, K.M., MARKS, L.S., LOEB, S., BROYLES, D.L., SHIN, S.S., CRUZ, A.B., CHAN, D.W., SOKOLL, L.J., ROBERTS, W.L., VAN SCHAIK, R.H. and MIZRAHI, I.A., 2011. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. *The Journal of urology,* **185**(5), pp. 1650-1655.

CHAFFER, C.L. and WEINBERG, R.A., 2011. A perspective on cancer cell metastasis. *Science (New York, N.Y.),* **331**(6024), pp. 1559-1564.

CHALLA, A.A. and STEFANOVIC, B., 2011. A novel role of vimentin filaments: binding and stabilization of collagen mRNAs. *Molecular and cellular biology,* **31**(18), pp. 3773-3789.

CHEN, X., LIAO, J., LU, Y., DUAN, X. and SUN, W., 2011. Activation of the PI3K/Akt pathway mediates bone morphogenetic protein 2-induced invasion of pancreatic cancer cells Panc-1. *Pathology & Oncology Research,* **17**(2), pp. 257-261.

CHENG, L., YANG, S., YANG, Y., ZHANG, W., XIAO, H., GAO, H., DENG, X. and ZHANG, Q., 2012. Global gene expression and functional network analysis of gastric cancer identify extended pathway maps and GPRC5A as a potential biomarker. *Cancer letters,* **326**(1), pp. 105-113.

CHENG, L., MONTIRONI, R., BOSTWICK, D.G., LOPEZ-BELTRAN, A. and BERNEY, D.M., 2012. Staging of prostate cancer. *Histopathology,* **60**(1), pp. 87-117.

CHENG, L., MONTIRONI, R., BOSTWICK, D.G., LOPEZ-BELTRAN, A. and BERNEY, D.M., 2012. Staging of prostate cancer. *Histopathology,* **60**(1), pp. 87-117.

CHIAM, K., RICCIARDELLI, C. and BIANCO-MIOTTO, T., 2014. Epigenetic biomarkers in prostate cancer: Current and future uses. *Cancer letters,* **342**(2), pp. 248-256.

CHOWDHURY, S., ROBINSON, D., CAHILL, D., RODRIGUEZ-VIDA, A., HOLMBERG, L. and MØLLER, H., 2013. Causes of death in men with prostate cancer: an analysis of 50 000 men from the Thames Cancer Registry. *BJU international,* **112**(2), pp. 182-189.

CHOYKE, P.L. and LOEB, S., 2017. Active Surveillance of Prostate Cancer. *Oncology (Williston Park, N.Y.),* **31**(1), pp. 67-70.

CHUNTHAPONG, J., SEFTOR, E.A., KHALKHALI-ELLIS, Z., SEFTOR, R.E., AMIR, S., LUBAROFF, D.M., HEIDGER, P.M. and HENDRIX, M.J., 2004. Dual roles of E-cadherin in prostate cancer invasion. *Journal of cellular biochemistry,* **91**(4), pp. 649-661.

COCK, P.J., FIELDS, C.J., GOTO, N., HEUER, M.L. and RICE, P.M., 2009. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic acids research,* **38**(6), pp. 1767-1771.

CONKLIN, M.W., EICKHOFF, J.C., RICHING, K.M., PEHLKE, C.A., ELICEIRI, K.W., PROVENZANO, P.P., FRIEDL, A. and KEELY, P.J., 2011. Aligned collagen is a prognostic signature for survival in human breast carcinoma. *The American journal of pathology,* **178**(3), pp. 1221-1232.

COOPER, C.R., CHAY, C.H. and PIENTA, K.J., 2002. The role of alpha(v)beta(3) in prostate cancer progression. *Neoplasia (New York, N.Y.),* **4**(3), pp. 191-194.

COOPER, J.P. and YOULE, R.J., 2012. Balancing cell growth and death. *Current opinion in cell biology,* **24**(6), pp. 802-803.

CORTAZAR, A.R., TORRANO, V., MARTÍN-MARTÍN, N., CARO-MALDONADO, A., CAMACHO, L., HERMANOVA, I., GURUCEAGA, E., LORENZO-MARTÍN, L.F., CALOTO, R. and GOMIS, R.R., 2018. CANCERTOOL: a visualization and representation interface to exploit cancer datasets. *Cancer research,* **78**(21), pp. 6320-6328.

CORTAZAR, A.R., TORRANO, V., MARTIN-MARTIN, N., CARO-MALDONADO, A., CAMACHO, L., HERMANOVA, I., GURUCEAGA, E., LORENZO-MARTIN, L.F., CALOTO, R., GOMIS, R.R., APAOLAZA, I., QUESADA, V., TRKA, J., GOMEZ-MUNOZ, A., VINCENT, S., BUSTELO, X.R.,

PLANES, F.J., ARANSAY, A.M. and CARRACEDO, A., 2018. CANCERTOOL: A Visualization and Representation Interface to Exploit Cancer Datasets. *Cancer research,* **78**(21), pp. 6320-6328.

DAI, J., ZHANG, N., WANG, J., CHEN, M. and CHEN, J., 2014. Gastrokine-2 is downregulated in gastric cancer and its restoration suppresses gastric tumorigenesis and cancer metastasis. *Tumor Biology,* **35**(5), pp. 4199-4207.

D'AMICO, A.V., WHITTINGTON, R., MALKOWICZ, S.B., SCHULTZ, D., BLANK, K., BRODERICK, G.A., TOMASZEWSKI, J.E., RENSHAW, A.A., KAPLAN, I. and BEARD, C.J., 1998. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *Jama,* **280**(11), pp. 969-974.

DAS, M., ITHYCHANDA, S.S., QIN, J. and PLOW, E.F., 2011. Migfilin and filamin as regulators of integrin activation in endothelial cells and neutrophils. *PloS one,* **6**(10), pp. e26355.

DAS, R., GREGORY, P.A., HOLLIER, B.G., TILLEY, W.D. and SELTH, L.A., 2014. Epithelial plasticity in prostate cancer: principles and clinical perspectives. *Trends in molecular medicine,* **20**(11), pp. 643-651.

DE MATOS, L.L., TRUFELLI, D.C., DE MATOS, MARIA GRACIELA LUONGO and DA SILVA PINHAL, MARIA APARECIDA, 2010. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomarker insights,* **5**, pp. BMI. S2185.

DERYNCK, R. and ZHANG, Y.E., 2003. Smad-dependent and Smad-independent pathways in TGF-β family signalling. *Nature,* **425**(6958), pp. 577.

DOMON, B. and AEBERSOLD, R., 2006. Mass spectrometry and protein analysis. *Science (New York, N.Y.),* **312**(5771), pp. 212-217.

DOUPLE, E.B., MABUCHI, K., CULLINGS, H.M., PRESTON, D.L., KODAMA, K., SHIMIZU, Y., FUJIWARA, S. and SHORE, R.E., 2011. Long-term radiation-related health effects in a unique human population: lessons learned from the atomic bomb survivors of Hiroshima and Nagasaki. *Disaster medicine and public health preparedness,* **5**(S1), pp. S122-S133.

DRABOVICH, A.P., MARTINEZ-MORILLO, E. and DIAMANDIS, E.P., 2015. Toward an integrated pipeline for protein biomarker development. *Biochimica et biophysica acta,* **1854**(6), pp. 677-686.

DRUCKER, E. and KRAPFENBAUER, K., 2013. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA journal,* **4**(1), pp. 7.

DU, R., LIU, B., ZHOU, L., WANG, D., HE, X., XU, X., ZHANG, L., NIU, C. and LIU, S., 2018. Downregulation of annexin A3 inhibits tumor metastasis and decreases drug resistance in breast cancer. *Cell death & disease,* **9**(2), pp. 126.

DUNN, M.W. and KAZER, M.W., 2011. Prostate cancer overview, *Seminars in oncology nursing* 2011, Elsevier, pp. 241-250.

EDMOND DE HOFFMANN, VINCENT STROOBANT, 2007. *Mass Spectrometry: Principles and Applications.* 3 edn. John Wiley & Sons Ltd.

EGGENER, S.E., BADANI, K., BAROCAS, D.A., BARRISFORD, G.W., CHENG, J.S., CHIN, A.I., CORCORAN, A., EPSTEIN, J.I., GEORGE, A.K., GUPTA, G.N., HAYN, M.H., KAUFFMAN, E.C., LANE, B., LISS, M.A., MIRZA, M., MORGAN, T.M., MOSES, K., NEPPLE, K.G., PRESTON, M.A., RAIS-BAHRAMI, S., RESNICK, M.J., SIDDIQUI, M.M., SILBERSTEIN, J., SINGER, E.A., SONN, G.A., SPRENKLE, P., STRATTON, K.L., TAYLOR, J., TOMASZEWSKI, J., TOLLEFSON, M., VICKERS, A., WHITE, W.M. and LOWRANCE, W.T., 2015. Gleason 6 Prostate Cancer: Translating Biology into Population Health. *The Journal of urology,* **194**(3), pp. 626-634.

ELLENBROEK, S.I. and COLLARD, J.G., 2007. Rho GTPases: functions and association with cancer. *Clinical & experimental metastasis,* **24**(8), pp. 657-672.

FAN, L., JENG, Y., LU, Y. and LIEN, H., 2016. SPOCK1 Is a Novel Transforming Growth Factor-β–Induced Myoepithelial Marker That Enhances Invasion and Correlates with Poor Prognosis in Breast Cancer. *PloS one,* **11**(9), pp. e0162933.

FANG, X., CAI, Y., LIU, J., WANG, Z., WU, Q., ZHANG, Z., YANG, C., YUAN, L. and OUYANG, G., 2011. Twist2 contributes to breast cancer progression by promoting an epithelial–mesenchymal transition and cancer stem-like cell self-renewal. *Oncogene,* **30**(47), pp. 4707.

FERREIRA, L.B., PALUMBO, A., DE MELLO, K.D., STERNBERG, C., CAETANO, M.S., DE OLIVEIRA, F.L., NEVES, A.F., NASCIUTTI, L.E., GOULART, L.R. and GIMBA, E.R., 2012. PCA3 noncoding RNA is involved in the control of prostate-cancer cell survival and modulates androgen receptor signaling. *BMC cancer,* **12**, pp. 507-2407-12-507.

FILELLA, X. and FOJ, L., 2016. Prostate cancer detection and prognosis: from Prostate Specific Antigen (PSA) to exosomal biomarkers. *International journal of molecular sciences,* **17**(11), pp. 1784.

FORONI, C., BROGGINI, M., GENERALI, D. and DAMIA, G., 2012. Epithelial-mesenchymal transition and breast cancer: role, molecular mechanisms and clinical impact. *Cancer treatment reviews,* **38**(6), pp. 689-697.

FOUAD, Y.A. and AANEI, C., 2017. Revisiting the hallmarks of cancer. *American journal of cancer research,* **7**(5), pp. 1016.

FREED-PASTOR, W.A. and PRIVES, C., 2012. Mutant p53: one name, many proteins. *Genes & development,* **26**(12), pp. 1268-1286.

FRISCH, S.M. and SCREATON, R.A., 2001. Anoikis mechanisms. *Current opinion in cell biology,* **13**(5), pp. 555-562.

FUSTAINO, V., PRESUTTI, D., COLOMBO, T., CARDINALI, B., PAPOFF, G., BRANDI, R., BERTOLAZZI, P., FELICI, G. and RUBERTI, G., 2017. Characterization of epithelial-mesenchymal transition intermediate/hybrid phenotypes associated to resistance to EGFR inhibitors in non-small cell lung cancer cell lines. *Oncotarget,* **8**(61), pp. 103340.

GAO, X., LI, L., RASSLER, J., PANG, J., CHEN, M., LIU, W., CHEN, Z., REN, S., ZHOU, F. and XIE, K., 2017. Prospective study of CRMP4 promoter methylation in prostate biopsies as a predictor for lymph node metastases. *JNCI: Journal of the National Cancer Institute,* **109**(6),.

GAO, D., VAHDAT, L.T., WONG, S., CHANG, J.C. and MITTAL, V., 2012. Microenvironmental regulation of epithelial-mesenchymal transitions in cancer. *Cancer research,* **72**(19), pp. 4883-4889.

GARCIA DE HERREROS, A., 2014. Epithelial to mesenchymal transition in tumor cells as consequence of phenotypic instability. *Frontiers in cell and developmental biology,* **2**, pp. 71.

GIANCOTTI, F.G., 2014. Deregulation of cell signaling in cancer. *FEBS letters,* **588**(16), pp. 2558-2570.

GILKES, D.M., CHATURVEDI, P., BAJPAI, S., WONG, C.C., WEI, H., PITCAIRN, S., HUBBI, M.E., WIRTZ, D. and SEMENZA, G.L., 2013. Collagen prolyl hydroxylases are essential for breast cancer metastasis. *Cancer research,* **73**(11), pp. 3285-3296.

GILLET, L.C., NAVARRO, P., TATE, S., ROST, H., SELEVSEK, N., REITER, L., BONNER, R. and AEBERSOLD, R., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & cellular proteomics : MCP,* **11**(6), pp. O111.016717.

GLEASON, D.F., 1966. Classification of prostatic carcinomas. *Cancer chemotherapy reports,* **50**(3), pp. 125-128.

GLINSKY, G.V., GLINSKII, A.B., STEPHENSON, A.J., HOFFMAN, R.M. and GERALD, W.L., 2004. Gene expression profiling predicts clinical outcome of prostate cancer. *The Journal of clinical investigation,* **113**(6), pp. 913-923.

GOOSSENS, N., NAKAGAWA, S., SUN, X. and HOSHIDA, Y., 2015. Cancer biomarker discovery and validation. *Translational cancer research,* **4**(3), pp. 256-269.

GRASSI, M.L., DE SOUZA PALMA, C., THOMÉ, C.H., LANFREDI, G.P., POERSCH, A. and FAÇA, V.M., 2017. Proteomic analysis of ovarian cancer cells during epithelial-mesenchymal transition (EMT) induced by epidermal growth factor (EGF) reveals mechanisms of cell cycle control. *Journal of proteomics,* **151**, pp. 2-11.

GRASSO, C.S., WU, Y., ROBINSON, D.R., CAO, X., DHANASEKARAN, S.M., KHAN, A.P., QUIST, M.J., JING, X., LONIGRO, R.J. and BRENNER, J.C., 2012. The mutational landscape of lethal castration-resistant prostate cancer. *Nature,* **487**(7406), pp. 239.

GREEN, D.R. and LLAMBI, F., 2015. Cell Death Signaling. *Cold Spring Harbor perspectives in biology,* **7**(12), pp. 10.1101/cshperspect.a006080.

GREENBURG, G. and HAY, E.D., 1982. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *The Journal of cell biology,* **95**(1), pp. 333-339.

GUO, Y., XIAO, P., LEI, S., DENG, F., XIAO, G.G., LIU, Y., CHEN, X., LI, L., WU, S. and CHEN, Y., 2008. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta biochimica et biophysica Sinica,* **40**(5), pp. 426-436.

GUPTA, M.K., JAYARAM, S., REDDY, D.N., POLISETTY, R.V. and SIRDESHMUKH, R., 2015. Transcriptomic and Proteomic Data Integration and Two-Dimensional Molecular Maps with Regulatory and Functional Linkages: Application to Cell Proliferation and Invasion Networks in Glioblastoma. *Journal of proteome research,* **14**(12), pp. 5017-5027.

GYGI, S.P., ROCHON, Y., FRANZA, B.R. and AEBERSOLD, R., 1999. Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology,* **19**(3), pp. 1720-1730.

HAGGLOF, C., HAMMARSTEN, P., STROMVALL, K., EGEVAD, L., JOSEFSSON, A., STATTIN, P., GRANFORS, T. and BERGH, A., 2014. TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. *PloS one,* **9**(2), pp. e86824.

HANAHAN, D. and WEINBERG, R.A., 2011. Hallmarks of cancer: the next generation. *Cell,* **144**(5), pp. 646-674.

HANAHAN, D. and WEINBERG, R.A., 2000. The hallmarks of cancer. *Cell,* **100**(1), pp. 57-70.

HANNA, S. and EL-SIBAI, M., 2013. Signaling networks of Rho GTPases in cell motility. *Cellular signalling,* **25**(10), pp. 1955-1961.

HARNER-FOREMAN, N., VADAKEKOLATHU, J., LAVERSIN, S.A., MATHIEU, M.G., REEDER, S., POCKLEY, A.G., REES, R.C. and BOOCOCK, D.J., 2017. A novel spontaneous model of epithelial-mesenchymal transition (EMT) using a primary prostate cancer derived cell line demonstrating distinct stem-like characteristics. *Scientific reports,* **7**, pp. 40633.

HASSAN, S., FERRARIO, C., MAMO, A. and BASIK, M., 2008. Tissue microarrays: emerging standard for biomarker validation. *Current opinion in biotechnology,* **19**(1), pp. 19-25.

HATAKEYAMA, S., YONEYAMA, T., TOBISAWA, Y. and OHYAMA, C., 2017. Recent progress and perspectives on prostate cancer biomarkers. *International journal of clinical oncology,* **22**(2), pp. 214-221.

HATTA, M., MIYAKE, Y., UCHIDA, K. and YAMAZAKI, J., 2018. Keratin 13 gene is epigenetically suppressed during transforming growth factor-β1-induced epithelial-mesenchymal transition in a human keratinocyte cell line. *Biochemical and biophysical research communications,* **496**(2), pp. 381-386.

HE, H., DING, F., LI, S., CHEN, H. and LIU, Z., 2014. Expression of migfilin is increased in esophageal cancer and represses the Akt-β-catenin activation. *American journal of cancer research,* **4**(3), pp. 270.

HE, M., GOU, M., QI, M., XIANG, W., JI, Z., WANG, W., ZHAO, S. and LIU, Y., 2018. Label free quantitative proteomics reveals the role of miR-200b in androgen-independent prostate cancer cells. *Clinical Proteomics,* **15**(1), pp. 8.

HECHT, S.S., 2006. Cigarette smoking: cancer risks, carcinogens, and mechanisms. *Langenbeck's Archives of Surgery,* **391**(6), pp. 603-613.

HEERBOTH, S., HOUSMAN, G., LEARY, M., LONGACRE, M., BYLER, S., LAPINSKA, K., WILLBANKS, A. and SARKAR, S., 2015. EMT and tumor metastasis. *Clinical and translational medicine,* **4**, pp. 6-015-0048-3. eCollection 2015.

HEIDENREICH, A., BASTIAN, P.J., BELLMUNT, J., BOLLA, M., JONIAU, S., VAN DER KWAST, T., MASON, M., MATVEEV, V., WIEGEL, T., ZATTONI, F., MOTTET, N. and EUROPEAN ASSOCIATION OF UROLOGY, 2014a. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *European urology,* **65**(1), pp. 124-137.

HEIDENREICH, A., BASTIAN, P.J., BELLMUNT, J., BOLLA, M., JONIAU, S., VAN DER KWAST, T., MASON, M., MATVEEV, V., WIEGEL, T., ZATTONI, F., MOTTET, N. and EUROPEAN ASSOCIATION OF UROLOGY, 2014b. EAU guidelines on prostate cancer. Part II: Treatment of advanced, relapsing, and castration-resistant prostate cancer. *European urology,* **65**(2), pp. 467-479.

HIEW, M.S.Y., CHENG, H.P., HUANG, C., CHONG, K.Y., CHEONG, S.K., CHOO, K.B. and KAMARUL, T., 2018. Incomplete cellular reprogramming of colorectal cancer cells elicits an epithelial/mesenchymal hybrid phenotype. *Journal of Biomedical Science,* **25**(1), pp. 57.

HINCK, A.P., 2012. Structural studies of the TGF-βs and their receptors–insights into evolution of the TGF-β superfamily. *FEBS letters,* **586**(14), pp. 1860-1870.

HOU, G., LOU, X., SUN, Y., XU, S., ZI, J., WANG, Q., ZHOU, B., HAN, B., WU, L. and ZHAO, X., 2015. Biomarker discovery and verification of esophageal squamous cell carcinoma using integration of SWATH/MRM. *Journal of proteome research,* **14**(9), pp. 3793-3803.

HOWELL, P.M., LIU, Z. and KHONG, H.T., 2010. Demethylating agents in the treatment of cancer. *Pharmaceuticals,* **3**(7), pp. 2022-2044.

HU, A., NOBLE, W.S. and WOLF-YADLIN, A., 2016. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research,* **5**, pp. 10.12688/f1000research.7042.1. eCollection 2016.

HUANG, B., JOLLY, M.K., LU, M., TSARFATY, I., BEN-JACOB, E. and JOSE'N, O., 2015. Modeling the transitions between collective and solitary migration phenotypes in cancer metastasis. *Scientific reports,* **5**, pp. 17379.

HUANG, R.Y., WONG, M., TAN, T., KUAY, K., NG, A., CHUNG, V., CHU, Y., MATSUMURA, N., LAI, H. and LEE, Y., 2013. An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell death & disease,* **4**(11), pp. e915.

HUANG, S., CHAUDHARY, K. and GARMIRE, L.X., 2017. More is better: recent progress in multi-omics data integration methods. *Frontiers in genetics,* **8**, pp. 84.

HUGO, C., 2003. The thrombospondin 1–TGF-β axis in fibrotic renal disease. *Nephrology Dialysis Transplantation,* **18**(7), pp. 1241-1245.

HÜLSMANN, J., KRAVIC, B., WEITH, M., GSTAIGER, M., AEBERSOLD, R.H., COLLINS, B.C. and MEYER, H., 2018. AP-SWATH reveals direct involvement of VCP/p97 in integrated stress response signaling through facilitating CReP/PPP1R15B degradation. *Molecular & Cellular Proteomics,* , pp. mcp. RA117. 000471.

HUMPHREY, P.A., 2004. Gleason grading and prognostic factors in carcinoma of the prostate. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc,* **17**(3), pp. 292-306.

IKUSHIMA, H. and MIYAZONO, K., 2010. TGFbeta signalling: a complex web in cancer progression. *Nature reviews.Cancer,* **10**(6), pp. 415-424.

ILYIN, S.E., BELKOWSKI, S.M. and PLATA-SALAMAN, C.R., 2004. Biomarker discovery and validation: technologies and integrative approaches. *Trends in biotechnology,* **22**(8), pp. 411-416.

IMANI, S., HOSSEINIFARD, H., CHENG, J., WEI, C. and FU, J., 2016. Prognostic value of EMT-inducing transcription factors (EMT-TFs) in metastatic breast cancer: a systematic review and meta-analysis. *Scientific reports,* **6**, pp. 28587.

INGOLIA, N.T., BRAR, G.A., ROUSKIN, S., MCGEACHY, A.M. and WEISSMAN, J.S., 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature protocols,* **7**(8), pp. 1534-1550.

IORDANSKAIA, T. and NAWSHAD, A., 2011. Mechanisms of transforming growth factor beta induced cell cycle arrest in palate development. *Journal of cellular physiology,* **226**(5), pp. 1415-1424.

ITHYCHANDA, S.S., DAS, M., MA, Y.Q., DING, K., WANG, X., GUPTA, S., WU, C., PLOW, E.F. and QIN, J., 2009. Migfilin, a molecular switch in regulation of integrin activation. *The Journal of biological chemistry,* **284**(7), pp. 4713-4722.

JACKSON, S.E. and CHESTER, J.D., 2015. Personalised cancer medicine. *International journal of cancer,* **137**(2), pp. 262-266.

JACOBS, J.M., ADKINS, J.N., QIAN, W., LIU, T., SHEN, Y., CAMP, D.G. and SMITH, R.D., 2005. Utilizing human blood plasma for proteomic biomarker discovery. *Journal of proteome research,* **4**(4), pp. 1073-1085.

JANSEN, S., GOSENS, R., WIELAND, T. and SCHMIDT, M., 2017. Paving the Rho in cancer metastasis: Rho GTPases and beyond. *Pharmacology & therapeutics,* .

JIANG, N., ZHU, S., CHEN, J., NIU, Y. and ZHOU, L., 2013. A-methylacyl-CoA racemase (AMACR) and prostate-cancer risk: a meta-analysis of 4,385 participants. *PLoS One,* **8**(10), pp. e74386.

JING, W., LUO, P., ZHU, M., AI, Q., CHAI, H. and TU, J., 2016. Prognostic and Diagnostic Significance of SDPR-Cavin-2 in Hepatocellular Carcinoma. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology,* **39**(3), pp. 950-960.

JONES, R.G. and THOMPSON, C.B., 2009. Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes & development,* **23**(5), pp. 537-548.

KAFETZOPOULOU, L.E., BOOCOCK, D.J., DHONDALAY, G.K.R., POWE, D.G. and BALL, G.R., 2013. Biomarker identification in breast cancer: beta-adrenergic receptor signaling and pathways to therapeutic response. *Computational and structural biotechnology journal,* **6**(7), pp. e201303003.

KALLURI, R., 2009. EMT: when epithelial cells decide to become mesenchymal-like cells. *The Journal of clinical investigation,* **119**(6), pp. 1417-1419.

KALLURI, R. and WEINBERG, R.A., 2009. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation,* **119**(6), pp. 1420-1428.

KANDA, M., NOMOTO, S., OYA, H., SHIMIZU, D., TAKAMI, H., HIBINO, S., HASHIMOTO, R., KOBAYASHI, D., TANAKA, C. and YAMADA, S., 2014. Dihydropyrimidinase-like 3 facilitates malignant behavior of gastric cancer. *Journal of Experimental & Clinical Cancer Research,* **33**(1), pp. 66.

KANDA, M., SHIMIZU, D., TANAKA, H., SHIBATA, M., IWATA, N., HAYASHI, M., KOBAYASHI, D., TANAKA, C., YAMADA, S., FUJII, T., NAKAYAMA, G., SUGIMOTO, H., KOIKE, M., FUJIWARA, M. and KODERA, Y., 2016. Metastatic pathway-specific transcriptome analysis identifies MFSD4 as a

putative tumor suppressor and biomarker for hepatic metastasis in patients with gastric cancer. *Oncotarget,* **7**(12), pp. 13667-13679.

KARANTANOS, T., CORN, P.G. and THOMPSON, T.C., 2013. Prostate cancer progression after androgen deprivation therapy: mechanisms of castrate resistance and novel therapeutic approaches. *Oncogene,* **32**(49), pp. 5501.

KATZ, E., DUBOIS-MARSHALL, S., SIMS, A.H., GAUTIER, P., CALDWELL, H., MEEHAN, R.R. and HARRISON, D.J., 2011. An in vitro model that recapitulates the epithelial to mesenchymal transition (EMT) in human breast cancer. *PLoS One,* **6**(2), pp. e17083.

KHATRI, P., SIROTA, M. and BUTTE, A.J., 2012. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS computational biology,* **8**(2), pp. e1002375.

KHURANA, E., FU, Y., CHAKRAVARTY, D., DEMICHELIS, F., RUBIN, M.A. and GERSTEIN, M., 2016. Role of non-coding sequence variants in cancer. *Nature reviews.Genetics,* **17**(2), pp. 93-108.

KIM, S.M., PARK, Y., PARK, E.S., CHO, J.Y., IZZO, J.G., ZHANG, D., KIM, S., LEE, J.H., BHUTANI, M.S. and SWISHER, S.G., 2010. Prognostic biomarkers for esophageal adenocarcinoma identified by analysis of tumor transcriptome. *PloS one,* **5**(11), pp. e15074.

KING, M.C., MARKS, J.H., MANDELL, J.B. and NEW YORK BREAST CANCER STUDY GROUP, 2003. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science (New York, N.Y.),* **302**(5645), pp. 643-646.

KUKURBA, K.R. and MONTGOMERY, S.B., 2015. RNA Sequencing and Analysis. *Cold Spring Harbor protocols,* **2015**(11), pp. 951-969.

KULASINGAM, V., PAVLOU, M.P. and DIAMANDIS, E.P., 2010. Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. *Nature Reviews Cancer,* **10**(5), pp. 371.

KULASINGAM, V. and DIAMANDIS, E.P., 2008a. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nature clinical practice.Oncology,* **5**(10), pp. 588-599.

KULASINGAM, V. and DIAMANDIS, E.P., 2008b. Tissue culture-based breast cancer biomarker discovery platform. *International journal of cancer,* **123**(9), pp. 2007-2012.

KURPINSKI, K., CHU, J., WANG, D. and LI, S., 2009. Proteomic profiling of mesenchymal stem cell responses to mechanical strain and TGF-β1. *Cellular and molecular bioengineering,* **2**(4), pp. 606.

LAMBERT, J., IVOSEV, G., COUZENS, A.L., LARSEN, B., TAIPALE, M., LIN, Z., ZHONG, Q., LINDQUIST, S., VIDAL, M. and AEBERSOLD, R., 2013. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature methods,* **10**(12), pp. 1239.

LAMOUILLE, S., XU, J. and DERYNCK, R., 2014. Molecular mechanisms of epithelial–mesenchymal transition. *Nature reviews Molecular cell biology,* **15**(3), pp. 178-196.

LANDRY, Y. and GIES, J., 2008. Drugs and their molecular targets: an updated overview. *Fundamental & clinical pharmacology,* **22**(1), pp. 1-18.

LANGMEAD, B. and SALZBERG, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods,* **9**(4), pp. 357.

LAPOINTE, J., LI, C., HIGGINS, J.P., VAN DE RIJN, M., BAIR, E., MONTGOMERY, K., FERRARI, M., EGEVAD, L., RAYFORD, W., BERGERHEIM, U., EKMAN, P., DEMARZO, A.M., TIBSHIRANI, R., BOTSTEIN, D., BROWN, P.O., BROOKS, J.D. and POLLACK, J.R., 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America,* **101**(3), pp. 811-816.

LEBRUN, J.J., 2012. The Dual Role of TGFbeta in Human Cancer: From Tumor Suppression to Cancer Metastasis. *ISRN molecular biology,* **2012**, pp. 381428.

LEE, C.H., AKIN-OLUGBADE, O. and KIRSCHENBAUM, A., 2011. Overview of prostate anatomy, histology, and pathology. *Endocrinology and metabolism clinics of North America,* **40**(3), pp. 565-575.

LENFERINK, A.E.G., CANTIN, C., NANTEL, A., WANG, E., DUROCHER, Y., BANVILLE, M., PAUL-ROC, B., MARCIL, A., WILSON, M.R. and O'CONNOR-MCCOURT, M., 2010. Transcriptome profiling of a TGF-β-induced epithelial-to-mesenchymal transition reveals extracellular clusterin as a target for therapeutic antibodies. *Oncogene,* **29**(6), pp. 831.

LI, B. and LI, C., 2017. Suppression of Prostate Cancer Metastasis by DPYSL3-Targeted saRNA. *RNA Activation.* Springer, pp. 207-216.

LI, E., LIU, L., LI, F., LUO, L., ZHAO, S., WANG, J., KANG, R., LUO, J. and ZHAO, Z., 2017. PSCA promotes prostate cancer proliferation and cell-cycle progression by up-regulating c-Myc. *The Prostate,* **77**(16), pp. 1563-1572.

LI, L., WEI, Y., TO, C., ZHU, C., TONG, J., PHAM, N., TAYLOR, P., IGNATCHENKO, V., IGNATCHENKO, A. and ZHANG, W., 2014. Integrated omic analysis of lung cancer reveals metabolism proteome signatures with prognostic impact. *Nature communications,* **5**, pp. 5469.

LIANG, W.C., FU, W.M., WONG, C.W., WANG, Y., WANG, W.M., HU, G.X., ZHANG, L., XIAO, L.J., WAN, D.C., ZHANG, J.F. and WAYE, M.M., 2015. The

lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. *Oncotarget,* **6**(26), pp. 22513-22525.

LIM, M., CHUONG, C. and ROY-BURMAN, P., 2011. PI3K, Erk signaling in BMP7-induced epithelial-mesenchymal transition (EMT) of PC-3 prostate cancer cells in 2-and 3-dimensional cultures. *Hormones and Cancer,* **2**(5), pp. 298.

LIM, J. and THIERY, J.P., 2012. Epithelial-mesenchymal transitions: insights from development. *Development (Cambridge, England),* **139**(19), pp. 3471-3486.

LIN, K., CROSWELL, J.M., KOENIG, H., LAM, C. and MALTZ, A., 2011. Prostate-specific antigen-based screening for prostate cancer: an evidence update for the US Preventive Services Task Force. *Evidence Synthesis,* (90),.

LIN, X., LIU, M., LIU, Y., HU, H., PAN, Y., ZOU, W., FAN, X. and HU, X., 2018. Transforming growth factor β1 promotes migration and invasion in HepG2 cells: Epithelial-to-mesenchymal transition via JAK/STAT3 signaling. *International journal of molecular medicine,* **41**(1), pp. 129-136.

LIN, B., FERGUSON, C., WHITE, J.T., WANG, S., VESSELLA, R., TRUE, L.D., HOOD, L. and NELSON, P.S., 1999. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer research,* **59**(17), pp. 4180-4184.

LIU, F., SONG, S., YI, Z., ZHANG, M., LI, J., YANG, F., YIN, H., YU, X., GUAN, C. and LIU, Y., 2017. HGF induces EMT in non-small-cell lung cancer through the hBVR pathway. *European journal of pharmacology,* **811**, pp. 180-190.

LIU, Y., HUETTENHAIN, R., COLLINS, B. and AEBERSOLD, R., 2013. Mass spectrometric protein maps for biomarker discovery and clinical research. *Expert review of molecular diagnostics,* **13**(8), pp. 811-825.

LIU, F., GU, L.N., SHAN, B.E., GENG, C.Z. and SANG, M.X., 2016. Biomarkers for EMT and MET in breast cancer: An update. *Oncology letters,* **12**(6), pp. 4869-4876.

LONG, Q., XU, J., OSUNKOYA, A.O., SANNIGRAHI, S., JOHNSON, B.A., ZHOU, W., GILLESPIE, T., PARK, J.Y., NAM, R.K., SUGAR, L., STANIMIROVIC, A., SETH, A.K., PETROS, J.A. and MORENO, C.S., 2014. Global transcriptome analysis of formalin-fixed prostate cancer specimens identifies biomarkers of disease recurrence. *Cancer research,* **74**(12), pp. 3228-3237.

LOUGHRAN, E.A., LEONARD, A.K., HILLIARD, T.S., PHAN, R.C., YEMC, M.G., HARPER, E., SHEEDY, E., KLYMENKO, Y., ASEM, M. and LIU, Y., 2018. Aging Increases Susceptibility to Ovarian Cancer Metastasis in Murine Allograft Models and Alters Immune Composition of Peritoneal Adipose Tissue. *Neoplasia,* **20**(6), pp. 621-631.

LOZITO, T.P. and TUAN, R.S., 2011. Mesenchymal stem cells inhibit both endogenous and exogenous MMPs via secreted TIMPs. *Journal of cellular physiology,* **226**(2), pp. 385-396.

LUDWIG, C., GILLET, L., ROSENBERGER, G., AMON, S., COLLINS, B.C. and AEBERSOLD, R., 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Molecular systems biology,* **14**(8), pp. e8126.

LUNDGREN, K., NORDENSKJÖLD, B. and LANDBERG, G., 2009. Hypoxia, Snail and incomplete epithelial–mesenchymal transition in breast cancer. *British journal of cancer,* **101**(10), pp. 1769.

MAIER, T., GUELL, M. and SERRANO, L., 2009. Correlation of mRNA and protein in complex biological samples. *FEBS letters,* **583**(24), pp. 3966-3973.

MALATI, T., 2007. Tumour markers: An overview. *Indian journal of clinical biochemistry : IJCB,* **22**(2), pp. 17-31.

MAO, Y., XU, J., LI, Z., ZHANG, N., YIN, H. and LIU, Z., 2013. The role of nuclear β-catenin accumulation in the Twist2-induced ovarian cancer EMT. *PLoS One,* **8**(11), pp. e78200.

MARESH, E.L., MAH, V., ALAVI, M., HORVATH, S., BAGRYANOVA, L., LIEBESKIND, E.S., KNUTZEN, L.A., ZHOU, Y., CHIA, D. and LIU, A.Y., 2010. Differential expression of anterior gradient gene AGR2 in prostate cancer. *BMC cancer,* **10**(1), pp. 680.

MARKS, L.S., FRADET, Y., DERAS, I.L., BLASE, A., MATHIS, J., AUBIN, S.M., CANCIO, A.T., DESAULNIERS, M., ELLIS, W.J. and RITTENHOUSE, H., 2007. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. *Urology,* **69**(3), pp. 532-535.

MATSUNUMA, R., CHAN, D.W., KIM, B.J., SINGH, P., HAN, A., SALTZMAN, A.B., CHENG, C., LEI, J.T., WANG, J., ROBERTO DA SILVA, L., SAHIN, E., LENG, M., FAN, C., PEROU, C.M., MALOVANNAYA, A. and ELLIS, M.J., 2018. DPYSL3 modulates mitosis, migration, and epithelial-to-mesenchymal transition in claudin-low breast cancer. *Proceedings of the National Academy of Sciences of the United States of America,* .

MATTICK, J.S. and MAKUNIN, I.V., 2006. Non-coding RNA. *Human molecular genetics,* **15 Spec No 1**, pp. R17-29.

MAUPIN, K.A., SINHA, A., EUGSTER, E., MILLER, J., ROSS, J., PAULINO, V., KESHAMOUNI, V.G., TRAN, N., BERENS, M. and WEBB, C., 2010. Glycogene expression alterations associated with pancreatic cancer epithelial-mesenchymal transition in complementary model systems. *PloS one,* **5**(9), pp. e13002.

MCDONALD, W.H. and YATES, J.R., 2002. Shotgun proteomics and biomarker discovery. *Disease markers,* **18**(2), pp. 99-105.

MCGRATH, S., CHRISTIDIS, D., PERERA, M., HONG, S.K., MANNING, T., VELA, I. and LAWRENTSCHUK, N., 2016. Prostate cancer biomarkers: are we hitting the mark? *Prostate international,* **4**(4), pp. 130-135.

MEHLEN, P. and PUISIEUX, A., 2006. Metastasis: a question of life or death. *Nature reviews cancer,* **6**(6), pp. 449.

MELZER, C., VON DER OHE, J., HASS, R. and UNGEFROREN, H., 2017. TGF-β-Dependent Growth Arrest and Cell Migration in Benign and Malignant Breast Epithelial Cells Are Antagonistically Controlled by Rac1 and Rac1b. *International journal of molecular sciences,* **18**(7), pp. 1574.

MERSEBURGER, A.S., KUCZYK, M.A., SERTH, J., BOKEMEYER, C., YOUNG, D.Y., SUN, L., CONNELLY, R.R., MCLEOD, D.G., MOSTOFI, F.K. and SRIVASTAVA, S.K., 2003. Limitations of tissue microarrays in the evaluation of focal alterations of bcl-2 and p53 in whole mount derived prostate tissues. *Oncology reports,* **10**(1), pp. 223-228.

MICALIZZI, D.S., FARABAUGH, S.M. and FORD, H.L., 2010. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *Journal of mammary gland biology and neoplasia,* **15**(2), pp. 117-134.

MIKI, Y., SWENSEN, J., SHATTUCK-EIDENS, D., FUTREAL, P.A., HARSHMAN, K., TAVTIGIAN, S., LIU, Q., COCHRAN, C., BENNETT, L.M. and DING, W., 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (New York, N.Y.),* **266**(5182), pp. 66-71.

MIKULA, M., RUBEL, T., KARCZMARSKI, J., GORYCA, K., DADLEZ, M. and OSTROWSKI, J., 2011. Integrating proteomic and transcriptomic high-throughput surveys for search of new biomarkers of colon tumors. *Functional & integrative genomics,* **11**(2), pp. 215-224.

MITARAI, H., WADA, N., HASEGAWA, D., YOSHIDA, S., SONODA, M., TOMOKIYO, A., HAMANO, S., SERITA, S., MIZUMACHI, H. and MAEDA, H., 2017. Transgelin mediates transforming growth factor-β1-induced proliferation of human periodontal ligament cells. *Journal of periodontal research,* **52**(6), pp. 984-993.

MORISHITA, Y. and KUSANO, E., 2011. Cellular and Molecular Basis of Epithelial-Mesenchymal Transition in Renal Fibrosis. *Journal of Nephrology and Therapeutics,* , pp. 1-4.

MULLER, P.A. and VOUSDEN, K.H., 2013. p53 mutations in cancer. *Nature cell biology,* **15**(1), pp. 2-8.

MURPHY, K., MURPHY, B.T., BOYCE, S., FLYNN, L., GILGUNN, S., O'ROURKE, C.J., ROONEY, C., STÖCKMANN, H., WALSH, A.L. and FINN, S., 2018. Integrating biomarkers across omic platforms: an approach to improve stratification of patients with indolent and aggressive prostate cancer. *Molecular oncology,* **12**(9), pp. 1513-1525.

NALLURI, S.M., O'CONNOR, J.W. and GOMEZ, E.W., 2015. Cytoskeletal signaling in TGF β-induced epithelial–mesenchymal transition. *Cytoskeleton,* **72**(11), pp. 557-569.

NARAYANAN, D.L., SALADI, R.N. and FOX, J.L., 2010. Ultraviolet radiation and skin cancer. *International journal of dermatology,* **49**(9), pp. 978-986.

NAUSEEF, J.T. and HENRY, M.D., 2011. Epithelial-to-mesenchymal transition in prostate cancer: paradigm or puzzle? *Nature reviews Urology,* **8**(8), pp. 428.

NEAL, C.L., MCKEITHEN, D. and ODERO-MARAH, V.A., 2011. Snail negatively regulates cell adhesion to extracellular matrix and integrin expression via the MAPK pathway in prostate cancer cells. *Cell adhesion & migration,* **5**(3), pp. 249-257.

NEUZILLET, C., TIJERAS-RABALLAND, A., COHEN, R., CROS, J., FAIVRE, S., RAYMOND, E. and DE GRAMONT, A., 2015. Targeting the TGFβ pathway for cancer therapy. *Pharmacology & therapeutics,* **147**, pp. 22-31.

NIETO, M.A., 2013. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science (New York, N.Y.),* **342**(6159), pp. 1234850.

NOOKAEW, I., PAPINI, M., PORNPUTTAPONG, N., SCALCINATI, G., FAGERBERG, L., UHLÉN, M. and NIELSEN, J., 2012. A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae. *Nucleic acids research,* **40**(20), pp. 10084-10097.

NOZATO, M., KANEKO, S., NAKAGAWARA, A. and KOMURO, H., 2013. Epithelial-mesenchymal transition-related gene expression as a new prognostic marker for neuroblastoma. *International journal of oncology,* **42**(1), pp. 134-140.

O'HURLEY, G., SJÖSTEDT, E., RAHMAN, A., LI, B., KAMPF, C., PONTÉN, F., GALLAGHER, W.M. and LINDSKOG, C., 2014. Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. *Molecular oncology,* **8**(4), pp. 783-798.

OU, Y., MA, L., DONG, L., MA, L., ZHAO, Z., MA, L., ZHOU, W., FAN, J., WU, C., YU, C., ZHAN, Q. and SONG, Y., 2012. Migfilin protein promotes migration and invasion in human glioma through epidermal growth factor receptor-mediated phospholipase C-gamma and STAT3 protein signaling pathways. *The Journal of biological chemistry,* **287**(39), pp. 32394-32405.

ØVERBYE, A., SKOTLAND, T., KOEHLER, C.J., THIEDE, B., SEIERSTAD, T., BERGE, V., SANDVIG, K. and LLORENTE, A., 2015. Identification of prostate cancer biomarkers in urinary exosomes. *Oncotarget,* **6**(30), pp. 30357.

OYA, H., KANDA, M., SUGIMOTO, H., SHIMIZU, D., TAKAMI, H., HIBINO, S., HASHIMOTO, R., OKAMURA, Y., YAMADA, S. and FUJII, T., 2015. Dihydropyrimidinase-like 3 is a putative hepatocellular carcinoma tumor suppressor. *Journal of gastroenterology,* **50**(5), pp. 590-600.

OZTURK, S., PAPAGEORGIS, P., WONG, C.K., LAMBERT, A.W., ABDOLMALEKY, H.M., THIAGALINGAM, A., COHEN, H.T. and THIAGALINGAM, S., 2016. SDPR functions as a metastasis suppressor in breast cancer by promoting apoptosis. *Proceedings of the National Academy of Sciences of the United States of America,* **113**(3), pp. 638-643.

PADUA, D. and MASSAGUÉ, J., 2009. Roles of TGFβ in metastasis. *Cell research,* **19**(1), pp. 89-102.

PANKOV, R. and YAMADA, K.M., 2002. Fibronectin at a glance. *Journal of cell science,* **115**(Pt 20), pp. 3861-3863.

PANTEL, K. and BRAKENHOFF, R.H., 2004. Dissecting the metastatic cascade. *Nature reviews cancer,* **4**(6), pp. 448.

PARK, A., LEE, J., MUN, S., KIM, D.J., CHA, B.H., MOON, K.T., YOO, T.K. and KANG, H., 2017. Identification of Transcription Factor YY1 as a Regulator of a Prostate Cancer-Specific Pathway Using Proteomic Analysis. *Journal of Cancer,* **8**(12), pp. 2303.

PARK, Y., KIM, J.W., KIM, D.S., KIM, E.B., PARK, S.J., PARK, J.Y., CHOI, W.S., SONG, J.G., SEO, H.Y., OH, S.C., KIM, B.S., PARK, J.J., KIM, Y.H. and KIM, J.S., 2008. The Bone Morphogenesis Protein-2 (BMP-2) is associated with progression to metastatic disease in gastric cancer. *Cancer research and treatment : official journal of Korean Cancer Association,* **40**(3), pp. 127-132.

PAUTKE, C., SCHIEKER, M., TISCHER, T., KOLK, A., NETH, P., MUTSCHLER, W. and MILZ, S., 2004. Characterization of osteosarcoma cell lines MG-63, Saos-2 and U-2 OS in comparison to human osteoblasts. *Anticancer Research,* **24**(6), pp. 3743-3748.

PEĆINA-ŠLAUS, N., 2003. Tumor suppressor gene E-cadherin and its role in normal and malignant cells. *Cancer cell international,* **3**(1), pp. 17.

PECORINO, L., 2012. *Molecular biology of cancer: mechanisms, targets, and therapeutics.* Oxford university press.

PELTIER, A., AOUN, F., LEMORT, M., KWIZERA, F., PAESMANS, M. and VAN VELTHOVEN, R., 2015. MRI-targeted biopsies versus systematic transrectal ultrasound guided biopsies for the diagnosis of localized prostate cancer in biopsy naive men. *BioMed research international,* **2015**.

PHAN, L.M., YEUNG, S.C. and LEE, M.H., 2014. Cancer metabolic reprogramming: importance, main features, and potentials for precise targeted anti-cancer therapies. *Cancer biology & medicine,* **11**(1), pp. 1-19.

PICKUP, M., NOVITSKIY, S. and MOSES, H.L., 2013. The roles of TGFbeta in the tumour microenvironment. *Nature reviews.Cancer,* **13**(11), pp. 788-799.

PLUMB, R.S., JOHNSON, K.A., RAINVILLE, P., SMITH, B.W., WILSON, I.D., CASTRO-PEREZ, J.M. and NICHOLSON, J.K., 2006. UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry,* **20**(13), pp. 1989-1994.

PONOMARENKO, E.A., POVERENNAYA, E.V., ILGISONIS, E.V., PYATNITSKIY, M.A., KOPYLOV, A.T., ZGODA, V.G., LISITSA, A.V. and ARCHAKOV, A.I., 2016. The size of the human proteome: the width and depth. *International journal of analytical chemistry,* **2016**.

POSTE, G., 2011. Bring on the biomarkers. *Nature,* **469**(7329), pp. 156.

PRENSNER, J.R., RUBIN, M.A., WEI, J.T. and CHINNAIYAN, A.M., 2012. Beyond PSA: the next generation of prostate cancer biomarkers. *Science translational medicine,* **4**(127), pp. 127rv3.

PURVINE, S., EPPEL*, J., YI, E.C. and GOODLETT, D.R., 2003. Shotgun collision-induced dissociation of peptides using a time of flight mass analyzer. *Proteomics,* **3**(6), pp. 847-850.

QIN, W., PAN, Y., ZHENG, X., LI, D., BU, J., XU, C., TANG, J., CUI, R., LIN, P. and YU, X., 2014. MicroRNA-124 regulates TGF-α-induced epithelial-mesenchymal transition in human prostate cancer cells. *International journal of oncology,* **45**(3), pp. 1225-1231.

QUAGLIATA, L., SCHLAGETER, M., QUINTAVALLE, C., TORNILLO, L. and TERRACCIANO, L.M., 2014. Identification of new players in hepatocarcinogenesis: limits and opportunities of using tissue microarray (TMA). *Microarrays,* **3**(2), pp. 91-102.

RAGHAVAN, C.T., SMUDA, M., SMITH, A.J., HOWELL, S., SMITH, D.G., SINGH, A., GUPTA, P., GLOMB, M.A., WORMSTONE, I.M. and NAGARAJ, R.H., 2016. AGEs in human lens capsule promote the TGFbeta2-mediated EMT of lens epithelial cells: implications for age-associated fibrosis. *Aging cell,* **15**(3), pp. 465-476.

RAMSKOLD, D., LUO, S., WANG, Y.C., LI, R., DENG, Q., FARIDANI, O.R., DANIELS, G.A., KHREBTUKOVA, I., LORING, J.F., LAURENT, L.C., SCHROTH, G.P. and SANDBERG, R., 2012. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature biotechnology,* **30**(8), pp. 777-782.

REID, S.E., KAY, E.J., NEILSON, L.J., HENZE, A.T., SERNEELS, J., MCGHEE, E.J., DHAYADE, S., NIXON, C., MACKEY, J.B., SANTI, A., SWAMINATHAN, K., ATHINEOS, D., PAPALAZAROU, V., PATELLA, F., ROMAN-FERNANDEZ, A., ELMAGHLOOB, Y., HERNANDEZ-FERNAUD, J.R., ADAMS, R.H., ISMAIL, S., BRYANT, D.M., SALMERON-SANCHEZ, M., MACHESKY, L.M., CARLIN, L.M., BLYTH, K., MAZZONE, M. and ZANIVAN, S., 2017. Tumor

matrix stiffness promotes metastatic cancer cell interaction with the endothelium. *The EMBO journal,* **36**(16), pp. 2373-2389.

RIDLEY, A.J., 2015. Rho GTPase signalling in cell migration. *Current opinion in cell biology,* **36**, pp. 103-112.

ROCK, K.L. and KONO, H., 2008. The inflammatory response to cell death. *Annu.Rev.pathmechdis.Mech.Dis.,* **3**, pp. 99-126.

ROMERO OTERO, J., GARCIA GOMEZ, B., CAMPOS JUANATEY, F. and TOUIJER, K.A., 2014. Prostate cancer biomarkers: an update. *Urologic oncology,* **32**(3), pp. 252-260.

ROOBOL, M.J., HAESE, A. and BJARTELL, A., 2011. Tumour markers in prostate cancer III: biomarkers in urine. *Acta Oncologica (Stockholm, Sweden),* **50 Suppl 1**, pp. 85-89.

ROSS, H.M., KRYVENKO, O.N., COWAN, J.E., SIMKO, J.P., WHEELER, T.M. and EPSTEIN, J.I., 2012. Do adenocarcinomas of the prostate with Gleason score (GS) *The American Journal of Surgical Pathology,* **36**(9), pp. 1346-1352.

SAHA, S., CHOI, H., KIM, B., DAYEM, A., YANG, G., KIM, K., YIN, Y. and CHO, S., 2017. KRT19 directly interacts with β-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. *Oncogene,* **36**(3), pp. 332.

SAINI, S., 2016a. PSA and beyond: alternative prostate cancer biomarkers. *Cellular Oncology,* **39**(2), pp. 97-106.

SAINI, S., 2016b. PSA and beyond: alternative prostate cancer biomarkers. *Cellular Oncology,* **39**(2), pp. 97-106.

SAJIC, T., LIU, Y. and AEBERSOLD, R., 2015. Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *PROTEOMICS-Clinical Applications,* **9**(3-4), pp. 307-321.

SANGER, F. and COULSON, A.R., 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology,* **94**(3), pp. 441-448.

SCHAEFFER, D., SOMARELLI, J.A., HANNA, G., PALMER, G.M. and GARCIA-BLANCO, M.A., 2014. Cellular migration and invasion uncoupled: increased migration is not an inexorable consequence of epithelial-to-mesenchymal transition. *Molecular and cellular biology,* **34**(18), pp. 3486-3499.

SCHMITTGEN, T.D. and LIVAK, K.J., 2008. Analyzing real-time PCR data by the comparative C T method. *Nature protocols,* **3**(6), pp. 1101.

SCHUBERT, O.T., GILLET, L.C., COLLINS, B.C., NAVARRO, P., ROSENBERGER, G., WOLSKI, W.E., LAM, H., AMODEI, D., MALLICK, P. and

MACLEAN, B., 2015. Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nature protocols,* **10**(3), pp. 426.

SCHULZE, A. and DOWNWARD, J., 2001. Navigating gene expression using microarrays—a technology review. *Nature cell biology,* **3**(8), pp. E190.

SCHWANHÄUSSER, B., BUSSE, D., LI, N., DITTMAR, G., SCHUCHHARDT, J., WOLF, J., CHEN, W. and SELBACH, M., 2011. Global quantification of mammalian gene expression control. *Nature,* **473**(7347), pp. 337.

SCOSYREV, E., MESSING, E.M., MOHILE, S., GOLIJANIN, D. and WU, G., 2012. Prostate cancer in the elderly: frequency of advanced disease at presentation and disease-specific mortality. *Cancer,* **118**(12), pp. 3062-3070.

SEGUIN, L., DESGROSELLIER, J.S., WEIS, S.M. and CHERESH, D.A., 2015. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. *Trends in cell biology,* **25**(4), pp. 234-240.

SERRANO, I., MCDONALD, P.C., LOCK, F.E. and DEDHAR, S., 2013. Role of the integrin-linked kinase (ILK)/Rictor complex in TGFbeta-1-induced epithelial-mesenchymal transition (EMT). *Oncogene,* **32**(1), pp. 50-60.

SERRANO-GOMEZ, S.J., MAZIVEYI, M. and ALAHARI, S.K., 2016. Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Molecular cancer,* **15**(1), pp. 18.

SEYHAN, A.A., 2010. Biomarkers in drug discovery and development. *Eur.Biopharm.Rev.,* , pp. 19-25.

SHABALINA, S.A. and SPIRIDONOV, N.A., 2004. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome biology,* **5**(4), pp. 105.

SHANKAR, J. and NABI, I.R., 2015. Actin cytoskeleton regulation of epithelial mesenchymal transition in metastatic cancer cells. *PloS one,* **10**(3), pp. e0119954.

SHARIAT, S.F., SCHERR, D.S., GUPTA, A., BIANCO, F.J.,Jr, KARAKIEWICZ, P.I., ZELTSER, I.S., SAMADI, D.B. and AKHAVAN, A., 2011. Emerging biomarkers for prostate cancer diagnosis, staging, and prognosis. *Archivos Espanoles de Urologia,* **64**(8), pp. 681-694.

SHARMA, S., KELLY, T.K. and JONES, P.A., 2010. Epigenetics in cancer. *Carcinogenesis,* **31**(1), pp. 27-36.

SHEN, Z., 2013. Cancer biomarkers and targeted therapies. *Cell & bioscience,* **3**(1), pp. 6.

SHI, L., WANG, H., WANG, F., FENG, M., WANG, M. and GUAN, W., 2014. Effects of gastrokine-2 expression on gastric cancer cell apoptosis by activation of extrinsic apoptotic pathways. *Molecular medicine reports,* **10**(6), pp. 2898-2904.

SHIBUE, T. and WEINBERG, R.A., 2017. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nature reviews Clinical oncology,* **14**(10), pp. 611.

SHISHKOVA, E., HEBERT, A.S. and COON, J.J., 2016. Now, more than ever, proteomics needs better chromatography. *Cell systems,* **3**(4), pp. 321-324.

SHUKLA, S., BHARTI, A.C., MAHATA, S., HUSSAIN, S., KUMAR, R., HEDAU, S. and DAS, B.C., 2009. Infection of human papillomaviruses in cancers of different human organ sites. *Indian Journal of Medical Research,* **130**(3), pp. 222.

SHUKLA, S., ZHANG, X., NIKNAFS, Y.S., XIAO, L., MEHRA, R., CIEŚLIK, M., ROSS, A., SCHAEFFER, E., MALIK, B. and GUO, S., 2016. Identification and validation of PCAT14 as prognostic biomarker in prostate cancer. *Neoplasia,* **18**(8), pp. 489-499.

SIDOLI, S., LIN, S., XIONG, L., BHANU, N.V., KARCH, K.R., JOHANSEN, E., HUNTER, C., MOLLAH, S. and GARCIA, B.A., 2015. Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) Analysis for Characterization and Quantification of Histone Post-translational Modifications. *Molecular & cellular proteomics : MCP,* **14**(9), pp. 2420-2428.

SISTIGU, A., DI MODUGNO, F., MANIC, G. and NISTICÒ, P., 2017. Deciphering the loop of epithelial-mesenchymal transition, inflammatory cytokines and cancer immunoediting. *Cytokine & growth factor reviews,* **36**, pp. 67-77.

SLATKOFF, S., GAMBOA, S., ZOLOTOR, A.J., MOUNSEY, A.L. and JONES, K., 2011. PURLs: PSA testing: when it's useful, when it's not. *The Journal of family practice,* **60**(6), pp. 357-360.

SOBIN, L. and GOSPODAROWICZ, M., 2009. Wittekind Ch. Eds. TNM Classification of Malignant Tumors.

SPORN, M.B., 1996. The war on cancer. *The lancet,* **347**(9012), pp. 1377-1381.

STARK, J.R., PERNER, S., STAMPFER, M.J., SINNOTT, J.A., FINN, S., EISENSTEIN, A.S., MA, J., FIORENTINO, M., KURTH, T. and LODA, M., 2009. Gleason score and lethal prostate cancer: does 3 4= 4 3? *Journal of Clinical Oncology,* **27**(21), pp. 3459-3464.

STONE, K.R., MICKEY, D.D., WUNDERLI, H., MICKEY, G.H. and PAULSON, D.F., 1978. Isolation of a human prostate carcinoma cell line (DU 145). *International journal of cancer,* **21**(3), pp. 274-281.

STONE, R.C., PASTAR, I., OJEH, N., CHEN, V., LIU, S., GARZON, K.I. and TOMIC-CANIC, M., 2016. Epithelial-mesenchymal transition in tissue repair and fibrosis. *Cell and tissue research,* **365**(3), pp. 495-506.

STRAUB, B.K., BODA, J., KUHN, C., SCHNOELZER, M., KORF, U., KEMPF, T., SPRING, H., HATZFELD, M. and FRANKE, W.W., 2003. A novel cell-cell

junction system: the cortex adhaerens mosaic of lens fiber cells. *Journal of cell science*, **116**(Pt 24), pp. 4985-4995.

SUDO, T., IWAYA, T., NISHIDA, N., SAWADA, G., TAKAHASHI, Y., ISHIBASHI, M., SHIBATA, K., FUJITA, H., SHIROUZU, K. and MORI, M., 2013. Expression of mesenchymal markers vimentin and fibronectin: the clinical significance in esophageal squamous cell carcinoma. *Annals of surgical oncology*, **20**(3), pp. 324-335.

SUN, B., FANG, Y., LI, Z., CHEN, Z. and XIANG, J., 2015. Role of cellular cytoskeleton in epithelial-mesenchymal transition process during cancer progression. *Biomedical reports*, **3**(5), pp. 603-610.

SUN, Y., DAEMEN, A., HATZIVASSILIOU, G., ARNOTT, D., WILSON, C., ZHUANG, G., GAO, M., LIU, P., BOUDREAU, A. and JOHNSON, L., 2014. Metabolic and transcriptional profiling reveals pyruvate dehydrogenase kinase 4 as a mediator of epithelial-mesenchymal transition and drug resistance in tumor cells. *Cancer & metabolism*, **2**(1), pp. 20.

SUN, Y., WANG, B.E., LEONG, K.G., YUE, P., LI, L., JHUNJHUNWALA, S., CHEN, D., SEO, K., MODRUSAN, Z., GAO, W.Q., SETTLEMAN, J. and JOHNSON, L., 2012. Androgen deprivation causes epithelial-mesenchymal transition in the prostate: implications for androgen-deprivation therapy. *Cancer research*, **72**(2), pp. 527-536.

SUZUKI, M., YOKOBORI, T., GOMBODORJ, N., YASHIRO, M., TURTOI, A., HANDA, T., OGATA, K., OYAMA, T., SHIRABE, K. and KUWANO, H., 2018. High stromal transforming growth factor β–induced expression is a novel marker of progression and poor prognosis in gastric cancer. *Journal of surgical oncology*, **118**(6), pp. 966-974.

SWEETWYNE, M.T. and MURPHY-ULLRICH, J.E., 2012. Thrombospondin1 in tissue repair and fibrosis: TGF-β-dependent and independent mechanisms. *Matrix Biology*, **31**(3), pp. 178-186.

SZÁSZ, A.M., LÁNCZKY, A., NAGY, Á, FÖRSTER, S., HARK, K., GREEN, J.E., BOUSSIOUTAS, A., BUSUTTIL, R., SZABÓ, A. and GYŐRFFY, B., 2016. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget*, **7**(31), pp. 49322.

TAKAHASHI, E., NAGANO, O., ISHIMOTO, T., YAE, T., SUZUKI, Y., SHINODA, T., NAKAMURA, S., NIWA, S., IKEDA, S., KOGA, H., TANIHARA, H. and SAYA, H., 2010. Tumor necrosis factor-alpha regulates transforming growth factor-beta-dependent epithelial-mesenchymal transition by promoting hyaluronan-CD44-moesin interaction. *The Journal of biological chemistry*, **285**(6), pp. 4060-4073.

TAKAISHI, S. and WANG, T.C., 2007. Gene expression profiling in a mouse model of Helicobacter-induced gastric cancer. *Cancer science*, **98**(3), pp. 284-293.

TAKETO, M.M., 2011. Reflections on the spread of metastasis to cancer prevention. *Cancer prevention research (Philadelphia, Pa.),* **4**(3), pp. 324-328.

TAKEUCHI, Y. and NISHIKAWA, H., 2016. Roles of regulatory T cells in cancer immunity. *International immunology,* **28**(8), pp. 401-409.

TALMADGE, J.E. and FIDLER, I.J., 2010. AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer research,* **70**(14), pp. 5649-5669.

TANGUAY, S., BEGIN, L., ELHILALI, M., BEHLOULI, H., KARAKIEWICZ, P. and APRIKIAN, A., 2002. Comparative evaluation of total PSA, free/total PSA, and complexed PSA in prostate cancer detection. *Urology,* **59**(2), pp. 261-265.

TENNAKOON, A., IZAWA, T., KUWAMURA, M. and YAMATE, J., 2015. Pathogenesis of type 2 epithelial to mesenchymal transition (EMT) in renal and hepatic fibrosis. *Journal of clinical medicine,* **5**(1), pp. 4.

TEPLY, B.A. and HAUKE, R.J., 2016. Chemotherapy options in castration-resistant prostate cancer. *Indian journal of urology : IJU : journal of the Urological Society of India,* **32**(4), pp. 262-270.

THIERY, J.P. and SLEEMAN, J.P., 2006. Complex networks orchestrate epithelial–mesenchymal transitions. *Nature reviews Molecular cell biology,* **7**(2), pp. 131-142.

THOBE, M.N., CLARK, R.J., BAINER, R.O., PRASAD, S.M. and RINKER-SCHAEFFER, C.W., 2011. From prostate to bone: key players in prostate cancer bone metastasis. *Cancers,* **3**(1), pp. 478-493.

THOMAS, P.A., KIRSCHMANN, D.A., CERHAN, J.R., FOLBERG, R., SEFTOR, E.A., SELLERS, T.A. and HENDRIX, M.J., 1999. Association between keratin and vimentin expression, malignant phenotype, and survival in postmenopausal breast cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research,* **5**(10), pp. 2698-2703.

TIAN, Y., YU, Y., HOU, L., CHI, J., MAO, J., XIA, L., WANG, X., WANG, P. and CAO, X., 2016. Serum deprivation response inhibits breast cancer progression by blocking transforming growth factor-β signaling. *Cancer science,* **107**(3), pp. 274-280.

TOEDA, Y., KASAMATSU, A., KOIKE, K., ENDO-SAKAMOTO, Y., FUSHIMI, K., KASAMA, H., YAMANO, Y., SHIIBA, M., TANZAWA, H. and UZAWA, K., 2018. FBLIM1 enhances oral cancer malignancy via modulation of the epidermal growth factor receptor pathway. *Molecular carcinogenesis,* **57**(12), pp. 1690-1697.

TOMCZAK, K., CZERWINSKA, P. and WIZNEROWICZ, M., 2015. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology (Poznan, Poland),* **19**(1A), pp. A68-77.

TONG, M., FUNG, T., LUK, S.T., NG, K., LEE, T.K., LIN, C., YAM, J.W., CHAN, K.W., NG, F. and ZHENG, B., 2015. ANXA3/JNK signaling promotes self-renewal

and tumor growth, and its blockade provides a therapeutic target for hepatocellular carcinoma. *Stem cell reports,* **5**(1), pp. 45-59.

TRAPNELL, C., PACHTER, L. and SALZBERG, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics,* **25**(9), pp. 1105-1111.

TRAPNELL, C., ROBERTS, A., GOFF, L., PERTEA, G., KIM, D., KELLEY, D.R., PIMENTEL, H., SALZBERG, S.L., RINN, J.L. and PACHTER, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols,* **7**(3), pp. 562.

TRAPNELL, C., WILLIAMS, B.A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M.J., SALZBERG, S.L., WOLD, B.J. and PACHTER, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology,* **28**(5), pp. 511-515.

TSAI, J.H. and YANG, J., 2013. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes & development,* **27**(20), pp. 2192-2206.

UCHIKADO, Y., INOUE, H., HARAGUCHI, N., MIMORI, K., NATSUGOE, S., OKUMURA, H., AIKOU, T. and MORI, M., 2006. Gene expression profiling of lymph node metastasis by oligomicroarray analysis using laser microdissection in esophageal squamous cell carcinoma. *International journal of oncology,* **29**(6), pp. 1337-1347.

UNGEFROREN, H., WITTE, D. and LEHNERT, H., 2018. The role of small GTPases of the Rho/Rac family in TGF-β-induced EMT and cell motility in cancer. *Developmental Dynamics,* **247**(3), pp. 451-461.

VALASTYAN, S. and WEINBERG, R.A., 2011. Tumor metastasis: molecular insights and evolving paradigms. *Cell,* **147**(2), pp. 275-292.

VALKENBURG, K.C. and WILLIAMS, B.O., 2011. Mouse models of prostate cancer. *Prostate cancer,* **2011**, pp. 895238.

VAN ZIJL, F., KRUPITZA, G. and MIKULITS, W., 2011. Initial steps of metastasis: cell invasion and endothelial transmigration. *Mutation Research/Reviews in Mutation Research,* **728**(1), pp. 23-34.

VENTER, J.C., ADAMS, M.D., MYERS, E.W., LI, P.W., MURAL, R.J., SUTTON, G.G., SMITH, H.O., YANDELL, M., EVANS, C.A., HOLT, R.A., GOCAYNE, J.D., AMANATIDES, P., BALLEW, R.M., HUSON, D.H., WORTMAN, J.R., ZHANG, Q., KODIRA, C.D., ZHENG, X.H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P.D., ZHANG, J., GABOR MIKLOS, G.L., NELSON, C., BRODER, S., CLARK, A.G., NADEAU, J., MCKUSICK, V.A., ZINDER, N., LEVINE, A.J., ROBERTS, R.J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-

THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., DI FRANCESCO, V., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A.E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T.J., HIGGINS, M.E., JI, R.R., KE, Z., KETCHUM, K.A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G.V., MILSHINA, N., MOORE, H.M., NAIK, A.K., NARAYAN, V.A., NEELAM, B., NUSSKERN, D., RUSCH, D.B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z., WANG, A., WANG, X., WANG, J., WEI, M., WIDES, R., XIAO, C., YAN, C., YAO, A., YE, J., ZHAN, M., ZHANG, W., ZHANG, H., ZHAO, Q., ZHENG, L., ZHONG, F., ZHONG, W., ZHU, S., ZHAO, S., GILBERT, D., BAUMHUETER, S., SPIER, G., CARTER, C., CRAVCHIK, A., WOODAGE, T., ALI, F., AN, H., AWE, A., BALDWIN, D., BADEN, H., BARNSTEAD, M., BARROW, I., BEESON, K., BUSAM, D., CARVER, A., CENTER, A., CHENG, M.L., CURRY, L., DANAHER, S., DAVENPORT, L., DESILETS, R., DIETZ, S., DODSON, K., DOUP, L., FERRIERA, S., GARG, N., GLUECKSMANN, A., HART, B., HAYNES, J., HAYNES, C., HEINER, C., HLADUN, S., HOSTIN, D., HOUCK, J., HOWLAND, T., IBEGWAM, C., JOHNSON, J., KALUSH, F., KLINE, L., KODURU, S., LOVE, A., MANN, F., MAY, D., MCCAWLEY, S., MCINTOSH, T., MCMULLEN, I., MOY, M., MOY, L., MURPHY, B., NELSON, K., PFANNKOCH, C., PRATTS, E., PURI, V., QURESHI, H., REARDON, M., RODRIGUEZ, R., ROGERS, Y.H., ROMBLAD, D., RUHFEL, B., SCOTT, R., SITTER, C., SMALLWOOD, M., STEWART, E., STRONG, R., SUH, E., THOMAS, R., TINT, N.N., TSE, S., VECH, C., WANG, G., WETTER, J., WILLIAMS, S., WILLIAMS, M., WINDSOR, S., WINN-DEEN, E., WOLFE, K., ZAVERI, J., ZAVERI, K., ABRIL, J.F., GUIGO, R., CAMPBELL, M.J., SJOLANDER, K.V., KARLAK, B., KEJARIWAL, A., MI, H., LAZAREVA, B., HATTON, T., NARECHANIA, A., DIEMER, K., MURUGANUJAN, A., GUO, N., SATO, S., BAFNA, V., ISTRAIL, S., LIPPERT, R., SCHWARTZ, R., WALENZ, B., YOOSEPH, S., ALLEN, D., BASU, A., BAXENDALE, J., BLICK, L., CAMINHA, M., CARNES-STINE, J., CAULK, P., CHIANG, Y.H., COYNE, M., DAHLKE, C., MAYS, A., DOMBROSKI, M., DONNELLY, M., ELY, D., ESPARHAM, S., FOSLER, C., GIRE, H., GLANOWSKI, S., GLASSER, K., GLODEK, A., GOROKHOV, M., GRAHAM, K., GROPMAN, B., HARRIS, M., HEIL, J., HENDERSON, S., HOOVER, J., JENNINGS, D., JORDAN, C., JORDAN, J., KASHA, J., KAGAN, L., KRAFT, C., LEVITSKY, A., LEWIS, M., LIU, X., LOPEZ, J., MA, D., MAJOROS, W., MCDANIEL, J., MURPHY, S., NEWMAN, M., NGUYEN, T., NGUYEN, N., NODELL, M., PAN, S., PECK, J., PETERSON, M., ROWE, W., SANDERS, R., SCOTT, J., SIMPSON, M., SMITH, T., SPRAGUE, A., STOCKWELL, T., TURNER, R., VENTER, E., WANG, M., WEN, M., WU, D., WU, M., XIA, A., ZANDIEH, A. and ZHU, X., 2001. The sequence of the human genome. *Science (New York, N.Y.),* **291**(5507), pp. 1304-1351.

VINAY, D.S., RYAN, E.P., PAWELEC, G., TALIB, W.H., STAGG, J., ELKORD, E., LICHTOR, T., DECKER, W.K., WHELAN, R.L. and KUMARA, H.S., 2015. Immune evasion in cancer: Mechanistic basis and therapeutic strategies, *Seminars in cancer biology* 2015, Elsevier, pp. S185-S198.

VOELKERDING, K.V., DAMES, S.A. and DURTSCHI, J.D., 2009. Next-generation sequencing: from basic research to diagnostics. *Clinical chemistry,* **55**(4), pp. 641-658.

VOGEL, C. and MARCOTTE, E.M., 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics,* **13**(4), pp. 227.

VOGELSTEIN, B. and KINZLER, K.W., 2004. Cancer genes and the pathways they control. *Nature medicine,* **10**(8), pp. 789.

WAGNER, S., BALL, G.R., POCKLEY, A.G. and MILES, A.K., 2018. Application of omic technologies in cancer research. *Translational Medicine Reports,* **2**(1),.

WAISBERG, J., DE SOUZA VIANA, L., AFFONSO JUNIOR, R.J., SILVA, S.R., DENADAI, M.V., MARGEOTTO, F.B., DE SOUZA, C.S. and MATOS, D., 2014. Overexpression of the ITGAV gene is associated with progression and spread of colorectal cancer. *Anticancer Research,* **34**(10), pp. 5599-5607.

WALDMEIER, L., MEYER-SCHALLER, N., DIEPENBRUCK, M. and CHRISTOFORI, G., 2012. Py2T murine breast cancer cells, a versatile model of TGFβ-induced EMT in vitro and in vivo. *PLoS One,* **7**(11), pp. e48651.

WANG, F., YANG, J., YU, K., XU, M., XU, Y., CHEN, L., LU, Y., FANG, H., WANG, X. and HU, Z., 2015. Activation of the NF-κB pathway as a mechanism of alcohol enhanced progression and metastasis of human hepatocellular carcinoma. *Molecular cancer,* **14**(1), pp. 10.

WANG, H., WANG, H., ZHOU, B., LI, C., ZHANG, F., WANG, X., ZHANG, G., BU, X., CAI, S. and DU, J., 2013. Epithelial–mesenchymal transition (EMT) induced by TNF-α requires AKT/GSK-3β-mediated stabilization of snail in colorectal cancer. *PloS one,* **8**(2), pp. e56664.

WANG, J., ZHANG, B., WU, H., CAI, J., SUI, X., WANG, Y., LI, H., QIU, Y., WANG, T. and CHEN, Z., 2017. CD51 correlates with the TGF-beta pathway and is a functional marker for colorectal cancer stem cells. *Oncogene,* **36**(10), pp. 1351.

WANG, L., XIAO, Y., PING, Y., LI, J., ZHAO, H., LI, F., HU, J., ZHANG, H., DENG, Y. and TIAN, J., 2014. Integrating multi-omics for uncovering the architecture of cross-talking pathways in breast cancer. *PloS one,* **9**(8), pp. e104282.

WANG, T., LI, Y., WANG, W., TUERHANJIANG, A., WU, Z., YANG, R., YUAN, M., MA, D., WANG, W. and WANG, S., 2014. Twist2, the key Twist isoform related to prognosis, promotes invasion of cervical cancer by inducing epithelial-mesenchymal transition and blocking senescence. *Human pathology,* **45**(9), pp. 1839-1846.

WANG, W., WANG, L., MIZOKAMI, A., SHI, J., ZOU, C., DAI, J., KELLER, E.T., LU, Y. and ZHANG, J., 2017. Down-regulation of E-cadherin enhances

prostate cancer chemoresistance via Notch signaling. *Chinese journal of cancer,* **36**(1), pp. 35.

WANG, Y., LIU, X.J. and YAO, X.D., 2014. Function of PCA3 in prostate tissue and clinical research progress on developing a PCA3 score. *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu,* **26**(4), pp. 493-500.

WANG, Z., GERSTEIN, M. and SNYDER, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews.Genetics,* **10**(1), pp. 57-63.

WARD, A.M., CATTO, J. and HAMDY, F., 2001. Prostate specific antigen: biology, biochemistry and available commercial assays. *Annals of Clinical Biochemistry,* **38**(6), pp. 633-651.

WIKSTRÖM, P., STATTIN, P., FRANCK-LISSBRANT, I., DAMBER, J. and BERGH, A., 1998. Transforming growth factor ?1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer. *The Prostate,* **37**(1), pp. 19-29.

WOOLF, MD, MPH, STEVEN H and ROTHEMICH, M., Stephen F, 1999. Screening for prostate cancer: the roles of science, policy, and opinion in determining what is best for patients. *Annual Review of Medicine,* **50**(1), pp. 207-221.

WOOSTER, R., BIGNELL, G., LANCASTER, J., SWIFT, S., SEAL, S., MANGION, J., COLLINS, N., GREGORY, S., GUMBS, C. and MICKLEM, G., 1995. Identification of the breast cancer susceptibility gene BRCA2. *Nature,* **378**(6559), pp. 789-792.

WRIGHT, J.L., SALINAS, C.A., LIN, D.W., KOLB, S., KOOPMEINERS, J., FENG, Z. and STANFORD, J.L., 2009. Prostate cancer specific mortality and Gleason 7 disease differences in prostate cancer outcomes between cases with Gleason 4 3 and Gleason 3 4 tumors in a population based cohort. *The Journal of urology,* **182**(6), pp. 2702-2707.

WU, Q., GUO, M., LU, Z., LI, T., QIAO, H. and KE, Y., 2003. Detection of human papillomavirus-16 in ovarian malignancy. *British journal of cancer,* **89**(4), pp. 672.

XIA, Y., SHEN, S. and VERMA, I.M., 2014. NF-kappaB, an active player in human cancers. *Cancer immunology research,* **2**(9), pp. 823-830.

XIONG, G., DENG, L., ZHU, J., RYCHAHOU, P.G. and XU, R., 2014. Prolyl-4-hydroxylase α subunit 2 promotes breast cancer progression and metastasis by regulating collagen deposition. *BMC cancer,* **14**(1), pp. 1.

XU, J., NEALE, A.V., DAILEY, R.K., EGGLY, S. and SCHWARTZ, K.L., 2012. Patient perspective on watchful waiting/active surveillance for localized prostate cancer. *Journal of the American Board of Family Medicine : JABFM,* **25**(6), pp. 763-770.

YAMASAKI, S. and ANDERSON, P., 2008. Reprogramming mRNA translation during stress. *Current opinion in cell biology,* **20**(2), pp. 222-226.

YANG, Y., JIANG, Y., XIE, D., LIU, M., SONG, N., ZHU, J., FAN, J. and ZHU, C., 2018. Inhibition of cell-adhesion protein DPYSL3 promotes metastasis of lung cancer. *Respiratory research,* **19**(1), pp. 41.

YU, H., XU, Q., LIU, F., YE, X., WANG, J. and MENG, X., 2015. Identification and Validation of Long Noncoding RNA Biomarkers in human non–small-cell lung carcinomas. *Journal of thoracic oncology,* **10**(4), pp. 645-654.

ZEISBERG, M. and NEILSON, E.G., 2009. Biomarkers for epithelial-mesenchymal transitions. *The Journal of clinical investigation,* **119**(6), pp. 1429-1437.

ZHANG, H., LIU, L., WANG, Y., ZHAO, G., XIE, R., LIU, C., XIAO, X., WU, K., NIE, Y. and ZHANG, H., 2013. KLF8 involves in TGF-beta-induced EMT and promotes invasion and migration in gastric cancer cells. *Journal of cancer research and clinical oncology,* **139**(6), pp. 1033-1042.

ZHANG, Y., BILBAO, A., BRUDERER, T., LUBAN, J., STRAMBIO-DE-CASTILLIA, C., LISACEK, F., HOPFGARTNER, G. and VARESIO, E., 2015. The use of variable Q1 isolation windows improves selectivity in LC–SWATH–MS acquisition. *Journal of proteome research,* **14**(10), pp. 4359-4371.

ZHAO, P., GUAN, H., DAI, Z., MA, Y., LIU, X. and WANG, X., 2016. Knockdown of SPOCK1 inhibits the proliferation and invasion in colorectal cancer cells by suppressing the PI3K/Akt pathway. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics,* **24**(6), pp. 437-445.

ZHAO, S., FUNG-LEUNG, W., BITTNER, A., NGO, K. and LIU, X., 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PloS one,* **9**(1), pp. e78644.

ZHAU, H.E., ODERO-MARAH, V., LUE, H., NOMURA, T., WANG, R., CHU, G., LIU, Z., ZHOU, B.P., HUANG, W. and CHUNG, L.W., 2008. Epithelial to mesenchymal transition (EMT) in human prostate cancer: lessons learned from ARCaP model. *Clinical & experimental metastasis,* **25**(6), pp. 601.

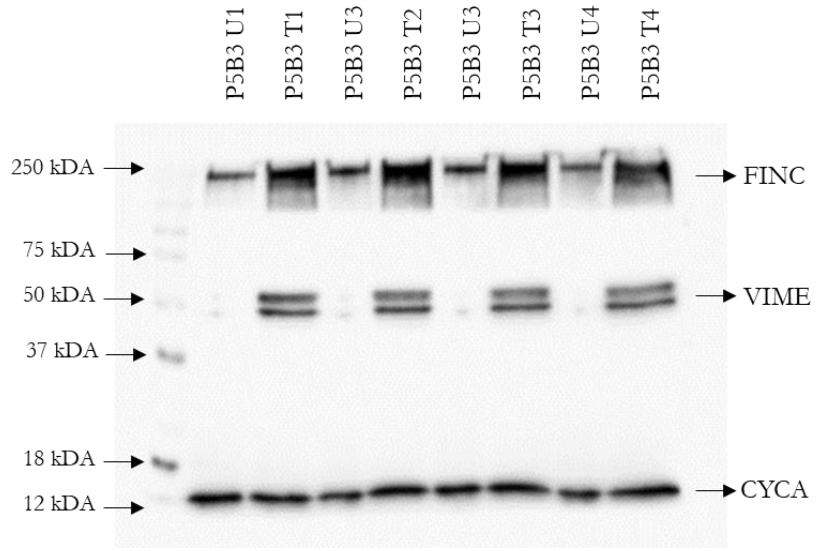
ZHOU, H., FANG, Y., WEINBERGER, P.M., DING, L., COWELL, J.K., HUDSON, F.Z., REN, M., LEE, J.R., CHEN, Q. and SU, H., 2016. Transgelin increases metastatic potential of colorectal cancer cells in vivo and alters expression of genes involved in cell motility. *BMC cancer,* **16**(1), pp. 55.

ZUBAREV, R.A., 2013. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics,* **13**(5), pp. 723-726.

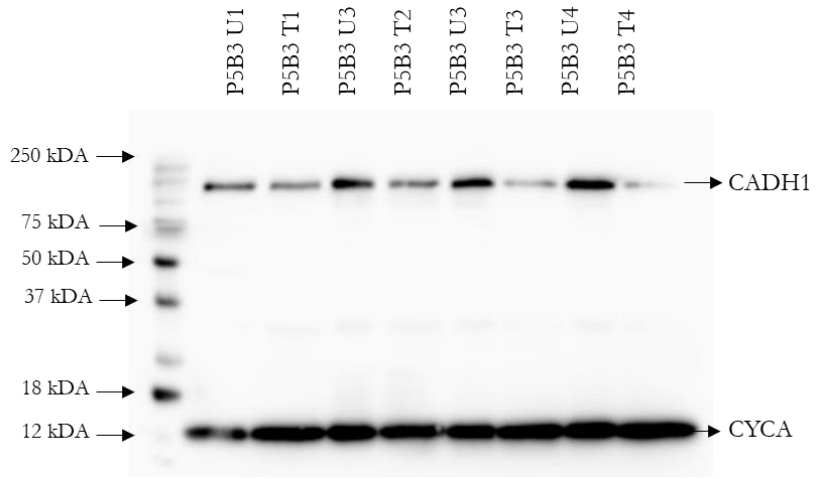
Appendix

A1 Unmodified images of Western blot analyses used in Chapter III

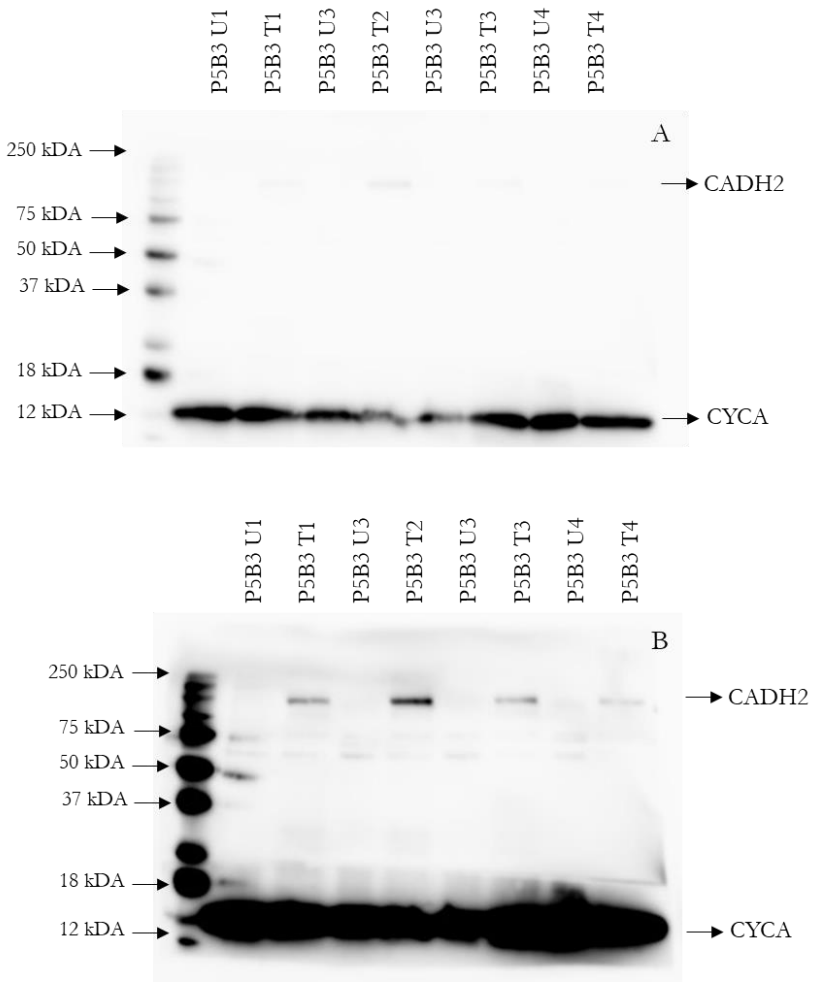
A1.1 Western blot analysis of fibronectin, vimentin and CYCA (Loading Control) in P5B3



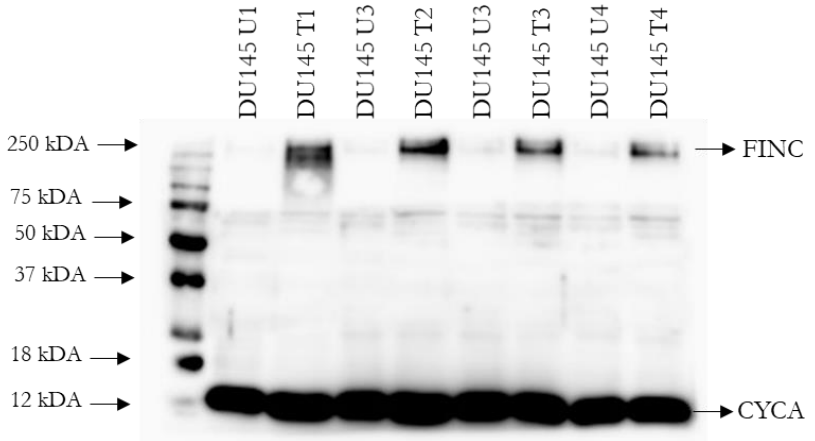
A1.2 Western blot analysis of E-cadherin and CYCA (Loading Control) in P5B3



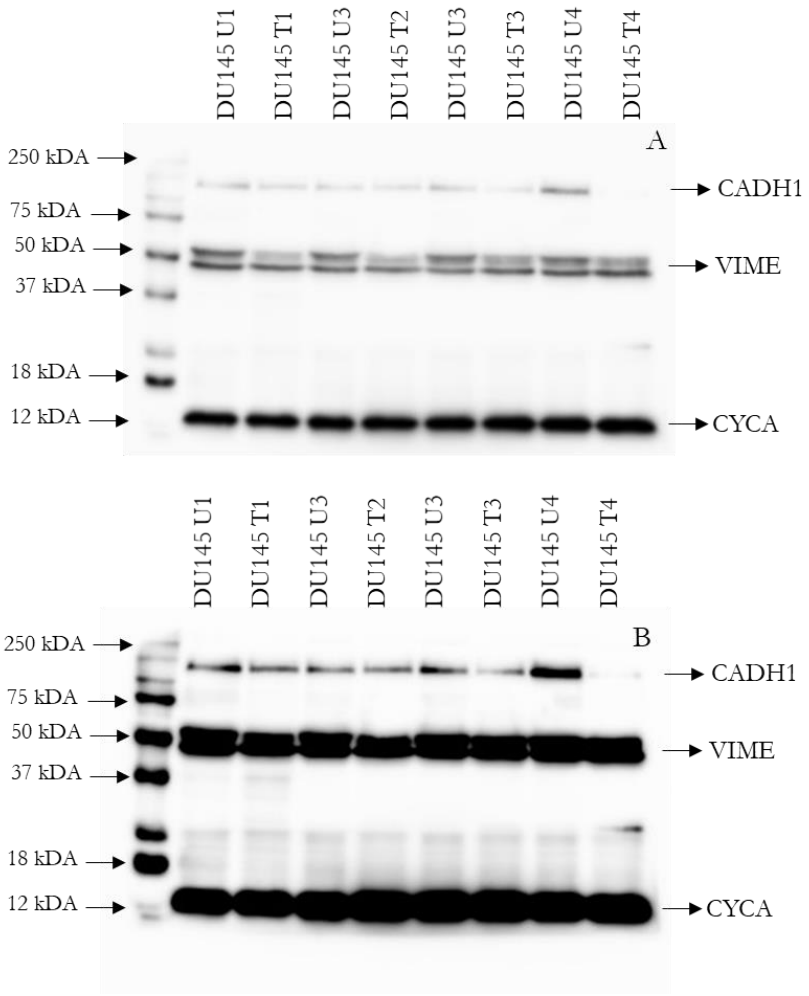
A1.3 Western blot analysis of N-cadherin and CYCA (Loading Control) in P5B3. A = optimised exposure for loading control, B = Extended exposure for detection of N-cadherin



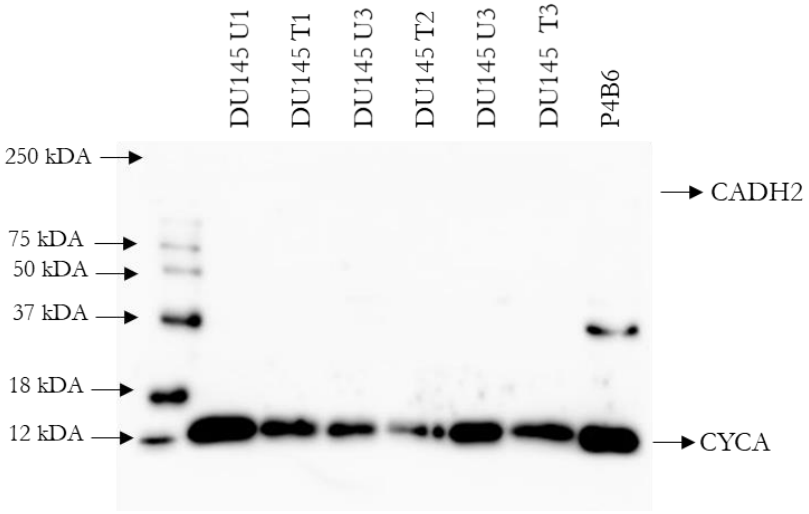
A1.4 Western blot analysis of fibronectin and CYCA (Loading Control) in DU145



A1.5 Western blot analysis of E-cadherin, vimentin and CYCA (Loading Control) in DU145, A = optimised exposure for loading control and vimentin, B = increased exposure for E-cadherin.

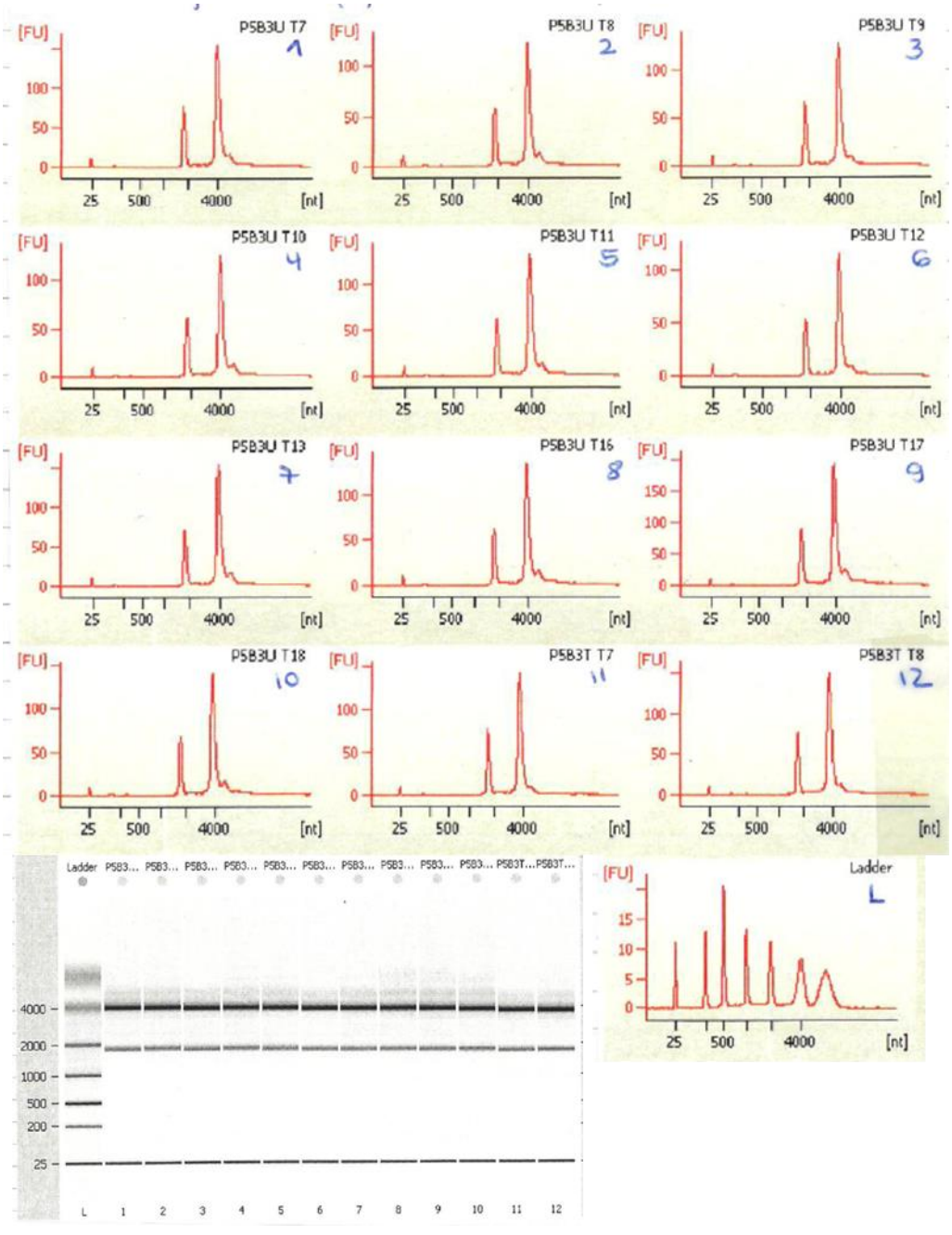


A1.6 Western blot analysis of N-cadherin and CYCA (Loading Control) in DU145

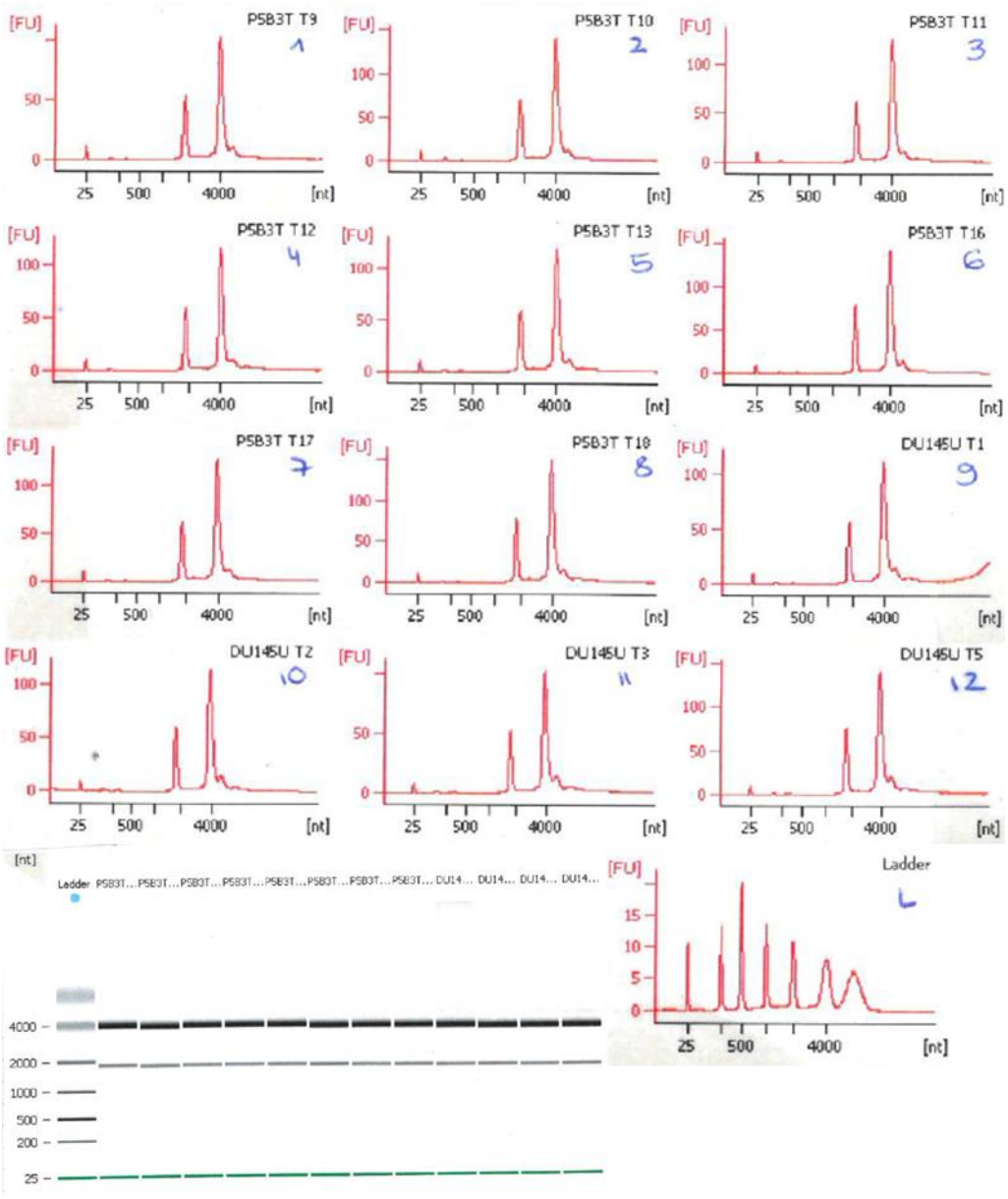


A2 Electropherogram of sample material used in RNA-sequencing experiment generated in Chapter IV

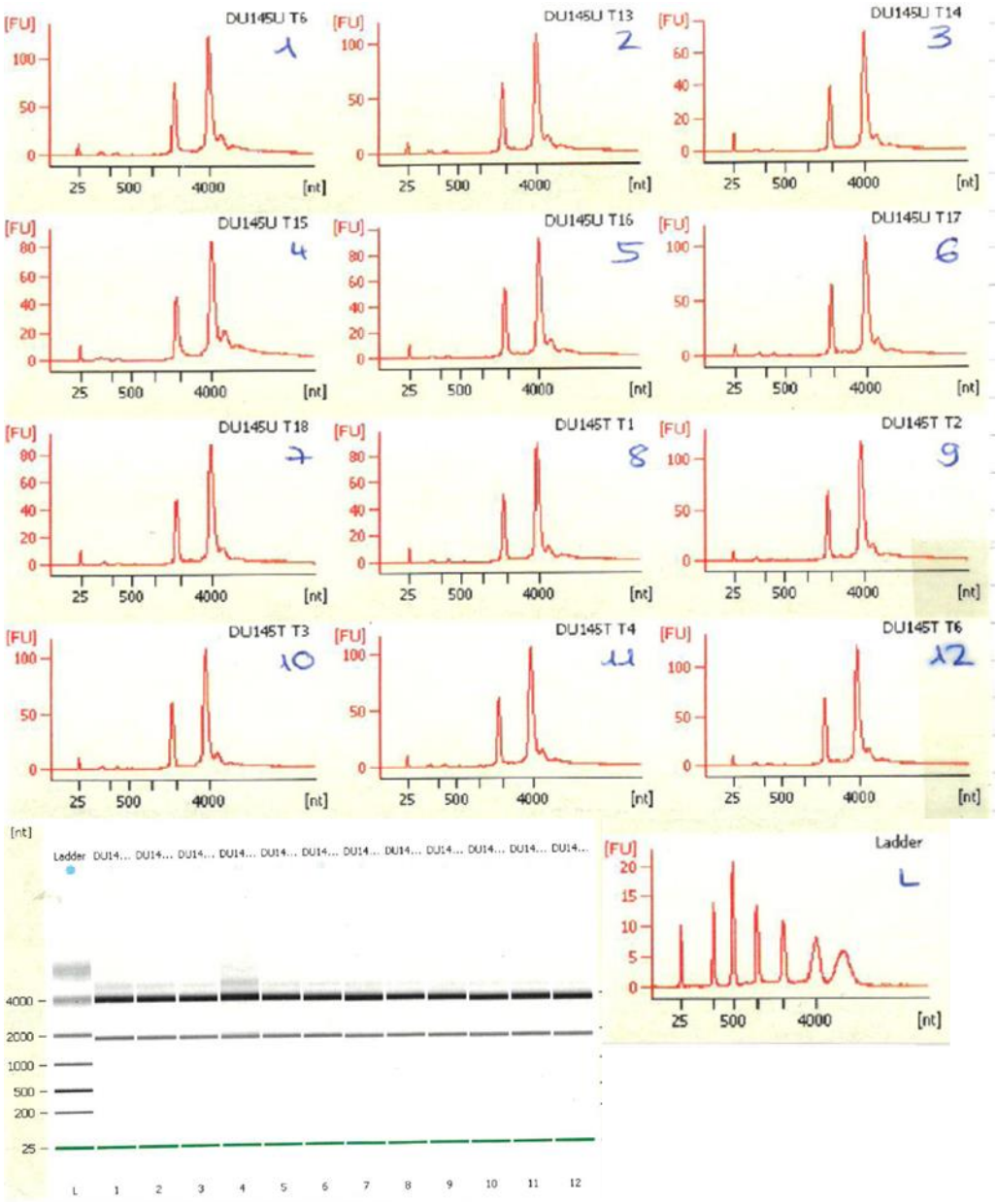
A2.1 Electropherogram of RNA extracted from untreated and treated P5B3



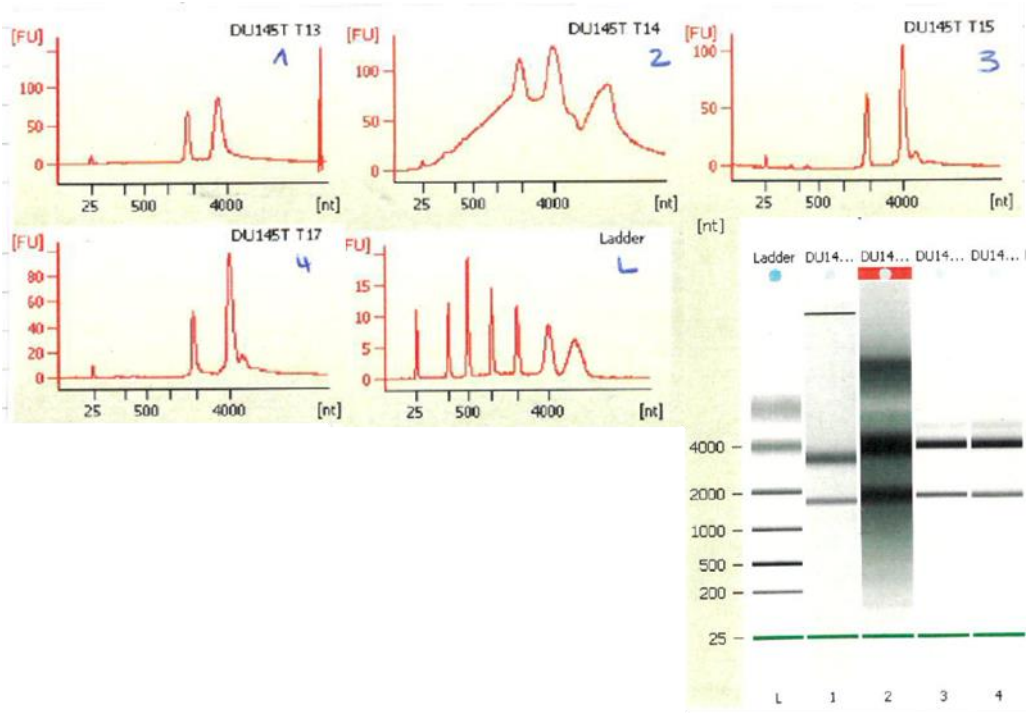
A2.2 Electropherogram of RNA extracted from treated P5B3 and untreated DU145

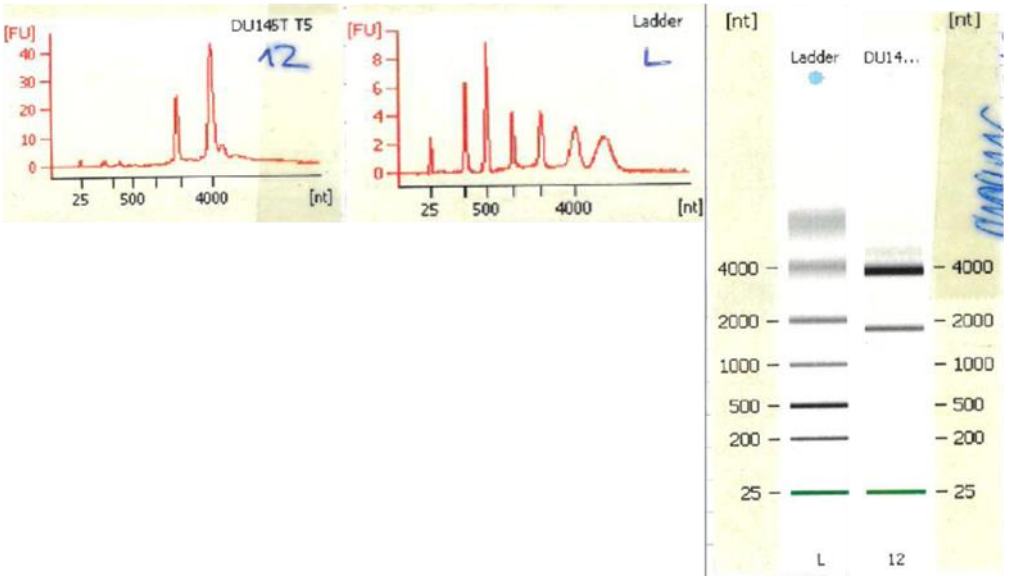


A2.3 Electropherogram of RNA extracted from untreated and treated DU145



A2.4 Electropherogram of RNA extracted from treated DU145





A3 Significant altered markers

A3.1 Genes not detected in P5B3 untreated or P5B3 treated after stimulation with TGF-β

Rank	Gene	Bonferroni	FC
1	***LOC102724279***	0.00000	Not detected in P5B3U
2	MSC-AS1	0.00000	Not detected in P5B3U
3	LINC01583	0.00000	Not detected in P5B3U
4	CADPS	0.00001	Not detected in P5B3U
5	FBXL7	0.00001	Not detected in P5B3U
6	CLEC18B	0.00001	Not detected in P5B3U
7	CACNG7	0.00002	Not detected in P5B3U
8	KLHDC8A	0.00007	Not detected in P5B3U
9	NTRK3	0.00008	Not detected in P5B3U
10	JAKMIP2-AS1	0.00012	Not detected in P5B3U
11	RTL1	0.00040	Not detected in P5B3U
12	COL22A1	0.00085	Not detected in P5B3U
13	COL5A3	0.00179	Not detected in P5B3U
14	IGFBP5	0.00548	Not detected in P5B3U
15	MYCT1	0.00604	Not detected in P5B3U
16	JAKMIP2	0.00631	Not detected in P5B3U
17	DEC1	0.01190	Not detected in P5B3U
18	MEG9	0.01798	Not detected in P5B3U
19	NAP1L3	0.02233	Not detected in P5B3U
20	SHANK1	0.03151	Not detected in P5B3U
21	BPIFB1	0.00167	Not detected in P5B3T

A3.2 Significantly deregulated genes in P5B3 upon stimulation with TGF-β

Rank	Gene	Corrected p-value	FC	Rank	Gene	Corrected p-value	FC
22	***MSC***	0.00000	1494.53	78	***SERPINE1***	0.00000	60.27
23	DIO3	0.00000	594.44	79	ZEB1	0.00085	60.04
24	***MEG3***	0.00001	483.72	80	KCNJ8	0.00002	59.97
25	***GDF6***	0.00002	456.82	81	TMEM119	0.00000	59.69
26	DGKI	0.00000	438.41	82	HMCN1	0.00000	59.26
27	***GPC4***	0.00000	433.11	83	CPED1	0.00040	58.23
28	***NKAIN4***	0.00858	418.27	84	***RBP1***	0.00000	55.13
29	TRPM2	0.00000	402.27	85	***TLL1***	0.00000	53.02
30	ITGB3	0.00000	375.01	86	AGTR1	0.04355	52.60
31	***HS3ST3A1***	0.00000	328.71	87	TNS1	0.00000	51.89
32	CDH11	0.00000	303.39	88	EPHA5-AS1	0.02757	51.03
33	PPAPDC1A	0.00004	264.20	89	***SIRPA***	0.00000	50.58
34	SPARC	0.00000	258.11	90	CRTAC1	0.00012	47.39
35	***BGN***	0.00470	248.56	91	ZC4H2	0.00000	45.50
36	POSTN	0.00001	235.44	92	NLRP1	0.00000	44.74
37	***VCAN***	0.00000	214.21	93	***CBLN2***	0.00000	44.28
38	***CNN1***	0.00000	203.77	94	SIRPB1	0.02057	43.77
39	***LOC728392***	0.00000	192.20	95	IGLON5	0.00005	41.72
40	***TWIST2***	0.00000	178.19	96	BIRC7	0.00124	41.55
41	MARCH4	0.00000	175.12	97	***MMP9***	0.00000	41.32
42	SLC28A3	0.00084	157.07	98	FGF1	0.00000	40.55
43	MEDAG	0.00991	156.78	99	FOXS1	0.00000	39.87
44	CTTNBP2	0.00001	150.16	100	LZTS1	0.00080	39.80
45	***VIM***	0.00000	141.81	101	TSHZ3	0.01452	39.76
46	***ADAM12***	0.00000	135.63	102	MMP10	0.00000	39.28
47	PTPRN	0.00000	126.54	103	***CCIN***	0.00000	38.53
48	***CASC15***	0.00002	122.86	104	***CDH2***	0.00000	37.98
49	GLI2	0.00000	122.85	105	***EPHA3***	0.00001	37.07
50	HHIP	0.00000	121.59	106	LDLRAD4	0.00001	36.13
51	SNHG24	0.00000	116.76	107	***NCF2***	0.00000	35.97
52	DCHS1	0.00000	102.49	108	FGF5	0.00000	34.82
53	***CREB3L1***	0.00000	93.77	109	***IL11***	0.00000	34.58
54	GBP5	0.00002	87.98	110	PRRX1	0.00002	34.16
55	ADAMTS12	0.00000	85.07	111	SCN8A	0.00000	34.14
56	***APCDD1L-AS1***	0.00000	84.98	112	CGB5	0.00552	33.83
57	CHST10	0.00000	82.44	113	RNF182	0.00000	33.27
58	IGF2	0.00000	80.08	114	***MMP1***	0.00141	32.92
59	SUSD4	0.00000	80.00	115	***MAP1B***	0.00002	32.90
60	***RASSF10***	0.00000	79.84	116	LOC101448202	0.02821	32.77
61	SLC16A2	0.00145	78.80	117	CSDC2	0.00001	32.57
62	***MMP2***	0.00000	76.91	118	ELN	0.00025	32.45
63	FBLL1	0.00062	74.57	119	MOV10L1	0.00000	31.88
64	CDH12	0.00001	73.38	120	MYH16	0.00000	31.77
65	NLRP3	0.00001	73.15	121	APCDD1L	0.00000	31.76
66	PDPN	0.00717	72.20	122	GFRA1	0.01958	31.36
67	***FBXL21***	0.00000	70.49	123	ADAMTS10	0.00004	31.31
68	***ANXA6***	0.00000	70.48	124	***LTB***	0.00000	30.73
69	CCDC85A	0.00028	70.19	125	COL3A1	0.01797	30.54
70	***CSF2***	0.00799	68.37	126	PGBD5	0.00002	30.44
71	DKK2	0.00129	66.22	127	FXYD6	0.00006	30.14
72	MEX3B	0.00000	62.92	128	RBMS3	0.00000	29.63
73	***CGB8***	0.00000	62.75	129	***GPR68***	0.00000	29.43
74	KCNH1	0.00000	62.39	130	***CCNJL***	0.00000	29.10
75	***DACT1***	0.00000	61.99	131	SHC2	0.00002	29.06
76	VIM-AS1	0.00004	61.77	132	SERPINA1	0.00511	28.76
77	MGC12916	0.00073	60.67	133	***KCNG1***	0.00000	28.02

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
134	***COL8A1***	0.00000	27.95	190	KRTAP2-3	0.00086	18.97
135	***HAPLN3***	0.00000	27.83	191	KIF1A	0.00000	18.92
136	***COL1A1***	0.00001	27.48	192	***TNFRSF19***	0.00000	18.85
137	***PCDH10***	0.00000	27.42	193	SDK2	0.03831	18.80
138	FRMPD4	0.00000	27.20	194	LAPTM5	0.00000	18.78
139	***C5orf46***	0.00059	27.18	195	***FAM101A***	0.00000	18.77
140	***C4orf26***	0.00000	27.09	196	***DIXDC1***	0.00000	18.43
141	RGS7	0.00005	26.17	197	ZP4	0.04097	18.33
142	***ADAM19***	0.00000	26.16	198	***HHIP-AS1***	0.00000	18.09
143	MYOZ1	0.00000	26.05	199	ADAMTS9	0.00053	18.07
144	EPHA5	0.00000	26.02	200	GOLGA7B	0.00000	17.79
145	LOC79160	0.00000	25.83	201	ADD2	0.00101	17.73
146	***XYLT1***	0.00000	25.82	202	EGOT	0.00585	17.68
147	ACTC1	0.00000	25.70	203	***S1PR1***	0.00000	17.65
148	SDK1	0.00000	25.25	204	***TMSB15A***	0.00000	17.60
149	DLX2	0.00001	25.18	205	***EDIL3***	0.00000	17.47
150	ZDHHC22	0.01213	25.10	206	SCN2A	0.00000	17.45
151	SLIT3	0.00000	25.05	207	SYT11	0.00000	17.31
152	IGF2BP1	0.03813	24.68	208	***DPYSL4***	0.00000	17.26
153	SENCR	0.00022	24.51	209	KCNK3	0.00006	17.24
154	***FLRT2***	0.00000	24.26	210	***COL27A1***	0.00000	17.23
155	PMP22	0.00060	24.17	211	***IGFBP7***	0.00000	17.22
156	***COL5A1***	0.00000	23.96	212	***SRPX***	0.00000	17.20
157	***GLIPR2***	0.00000	23.95	213	NAV3	0.00008	17.14
158	ROBO3	0.00000	23.93	214	F2R	0.00000	17.09
159	PALM2	0.00013	23.78	215	***NSG1***	0.00000	16.97
160	ZCCHC12	0.00170	23.77	216	CLDN14	0.00000	16.91
161	BEST3	0.00016	23.67	217	NCALD	0.00000	16.91
162	***FILIP1L***	0.00000	23.61	218	IGFBP7-AS1	0.00000	16.89
163	***RCN3***	0.00000	23.33	219	FMN2	0.00405	16.83
164	C6orf15	0.00017	23.01	220	TOX	0.00000	16.83
165	LOC101928370	0.04371	22.95	221	KCNMA1	0.00035	16.82
166	***RFTN1***	0.00000	22.91	222	TNFRSF9	0.00001	16.82
167	***ARL4C***	0.00000	22.86	223	CABP7	0.00006	16.80
168	ADRA1D	0.00000	22.77	224	***ADAMTS6***	0.00000	16.79
169	***TUBA1A***	0.00000	22.66	225	THBS2	0.00000	16.55
170	MYOM3	0.00000	21.76	226	GUCY1A2	0.00000	16.55
171	***DPYSL3***	0.00000	21.51	227	RGS4	0.00227	16.43
172	COL6A3	0.00000	21.21	228	***G0S2***	0.00007	16.42
173	COL6A2	0.00000	21.14	229	NTRK1	0.00332	16.37
174	DLX1	0.00001	21.14	230	ITGA11	0.00332	16.14
175	FIBIN	0.00002	20.88	231	NFATC2	0.00000	16.04
176	CACNA1H	0.03038	20.78	232	C1QTNF2	0.00000	16.02
177	ELFN1	0.00000	20.43	233	CHI3L2	0.02396	16.00
178	FAM172BP	0.00003	20.42	234	***MYH15***	0.00057	15.97
179	PALM2-AKAP2	0.00000	20.39	235	DACT3	0.00589	15.92
180	***FAP***	0.00000	20.33	236	LOC541472	0.00000	15.88
181	OSCAR	0.00001	20.32	237	***LOC100507431***	0.00000	15.87
182	***P2RY6***	0.00000	20.25	238	SOGA3	0.00033	15.76
183	AQP1	0.00000	20.12	239	DTNA	0.00666	15.69
184	***STEAP1B***	0.00001	19.92	240	PAK3	0.03491	15.65
185	POU3F1	0.00804	19.69	241	***LTBP2***	0.00000	15.64
186	C14orf37	0.00169	19.45	242	***SUN3***	0.00000	15.63
187	***GPR176***	0.00000	19.33	243	ACTBL2	0.00012	15.62
188	FLI1	0.00046	19.26	244	***TGFBI***	0.00000	15.60
189	***MFGE8***	0.00000	19.15	245	***UCN2***	0.00000	15.55
190	KRTAP2-3	0.00086	18.97	246	STK32A	0.00136	15.49

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
247	***PRR5L***	0.00000	15.26	304	***IL32***	0.00000	11.85
248	CNTNAP2	0.00762	15.25	305	MRAS	0.00000	11.75
249	***BICC1***	0.00000	15.10	306	HRASLS	0.00062	11.66
250	NEGR1	0.00856	15.07	307	GNG2	0.00000	11.57
251	GAL3ST3	0.00003	15.05	308	***NUAK1***	0.00000	11.54
252	LOC101929532	0.01484	14.98	309	KCNN1	0.04568	11.54
253	RAI2	0.00000	14.78	310	OLFML3	0.01467	11.37
254	***PLAU***	0.00000	14.76	311	JAM2	0.00000	11.31
255	***CYP24A1***	0.00000	14.65	312	CWH43	0.00001	11.31
256	BARX1	0.00000	14.56	313	***IL6***	0.00101	11.26
257	PRR16	0.00013	14.56	314	MIA	0.00017	11.14
258	HS3ST3B1	0.00493	14.50	315	ROS1	0.00106	11.13
259	***ANTXR2***	0.00000	14.48	316	***PSD***	0.00000	10.98
260	DIRAS1	0.00815	14.37	317	DCLK2	0.00000	10.95
261	PCDHGC5	0.00014	14.32	318	SAMD14	0.00002	10.91
262	KCNE4	0.00559	14.23	319	PDCD1LG2	0.00004	10.86
263	ZNF474	0.00006	14.19	320	KALRN	0.00000	10.73
264	LOC284581	0.00049	14.16	321	***NPR3***	0.00001	10.72
265	FAM13C	0.00315	14.08	322	GUCY1A3	0.00007	10.68
266	CAMK1D	0.00003	14.01	323	LINC00607	0.03900	10.62
267	***MFAP2***	0.00000	14.00	324	***LOX***	0.00000	10.61
268	***MRC2***	0.00000	13.94	325	***LCK***	0.00000	10.60
269	***PDE4C***	0.00000	13.91	326	PRKAR2B	0.00000	10.59
270	GBP6	0.00000	13.89	327	***FRMD5***	0.00000	10.56
271	***IL23A***	0.00001	13.86	328	FBLN5	0.00000	10.56
272	***GJA1***	0.00000	13.66	329	KCNA7	0.00218	10.54
273	MSRB3	0.00001	13.62	330	***KIF12***	0.00000	10.51
274	***WNT7A***	0.00000	13.60	331	GHR	0.00344	10.39
275	ELMOD1	0.00016	13.50	332	IKZF3	0.00000	10.39
276	TENM4	0.00000	13.50	333	FZD10-AS1	0.00000	10.32
277	USP2	0.00000	13.36	334	ZDHHC8P1	0.00004	10.31
278	***TMEM98***	0.00000	13.35	335	IRX4	0.00998	10.24
279	KCND1	0.00499	13.25	336	EBI3	0.00000	10.21
280	NTNG1	0.00000	13.18	337	***NPBWR1***	0.00000	10.20
281	***SNAI1***	0.00075	13.05	338	LOC101928718	0.00052	10.16
282	SH3GL3	0.00000	13.04	339	HOXB9	0.01307	10.13
283	***AEBP1***	0.00000	13.03	340	PLA2G4C	0.00025	10.10
284	EFNB3	0.00000	12.95	341	***NIPAL4***	0.00000	10.09
285	MYO3B	0.00000	12.95	342	***NUAK2***	0.00002	10.06
286	***GAS6-AS2***	0.00000	12.90	343	FAM110B	0.01726	10.02
287	ADRA2C	0.00408	12.89	344	DPYSL5	0.00000	9.92
288	CISH	0.00017	12.87	345	LRRTM3	0.01597	9.91
289	***ATP8B2***	0.00000	12.77	346	PIK3AP1	0.00046	9.83
290	SYT1	0.00000	12.76	347	***DYSF***	0.00000	9.83
291	ZNF469	0.00000	12.68	348	CPQ	0.00007	9.82
292	***PLEKHO1***	0.00000	12.66	349	***TNC***	0.00000	9.66
293	MSX1	0.00786	12.61	350	MDGA1	0.00000	9.64
294	ASGR1	0.00001	12.56	351	***KIAA1549L***	0.00000	9.63
295	***FN1***	0.00000	12.51	352	***RAI14***	0.00000	9.62
296	PNMA2	0.00776	12.39	353	TMCC2	0.00000	9.62
297	***PLAT***	0.00000	12.36	354	LOC100506178	0.00000	9.59
298	***ADAMTS15***	0.00000	12.32	355	ALPL	0.00276	9.51
299	HTR7	0.00000	12.21	356	***HAS2***	0.00001	9.50
300	KCNJ12	0.01067	12.14	357	***EHD3***	0.00000	9.48
301	***SERPINE2***	0.00000	12.09	358	NXPH2	0.00868	9.45
302	***TAGLN***	0.00000	12.04	359	***CD74***	0.00000	9.39
303	PCDHB5	0.00002	11.85	360	GRASP	0.03258	9.38

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
361	***SLC22A17***	0.00001	9.35	418	***PIK3CD***	0.00000	8.02
362	ACTG2	0.00019	9.34	419	***NREP***	0.00000	8.01
363	***NACAD***	0.00000	9.32	420	KBTBD11	0.04797	7.98
364	MAGEH1	0.00001	9.29	421	***LOC101928710***	0.00000	7.96
365	***CD83***	0.00000	9.29	422	***DHRS2***	0.00000	7.88
366	RPSAP52	0.00003	9.28	423	***CYR61***	0.00000	7.86
367	***SLC2A12***	0.00000	9.27	424	XKR5	0.03171	7.83
368	GABRB3	0.00107	9.23	425	***IL1B***	0.00001	7.83
369	HAGLROS	0.00697	9.18	426	APBA2	0.00123	7.82
370	ARHGAP22	0.00000	9.16	427	L1CAM	0.00000	7.80
371	***SMAD7***	0.00000	9.10	428	KRT33B	0.02168	7.78
372	***KLF2***	0.00000	9.07	429	IL4I1	0.00001	7.74
373	***PCSK1N***	0.00000	9.06	430	PAPPA	0.00000	7.73
374	PLEKHG1	0.00000	9.00	431	C10orf55	0.00000	7.69
375	***SEMA7A***	0.00000	8.98	432	RASSF9	0.00066	7.67
376	NOX5	0.00001	8.94	433	***RAB3B***	0.00029	7.63
377	PARP11	0.00053	8.90	434	SOBP	0.01386	7.62
378	SCARF1	0.00926	8.87	435	ANGPTL2	0.00003	7.60
379	CHRNB4	0.00002	8.86	436	MAP6	0.00088	7.58
380	***LBH***	0.00000	8.85	437	***JAG1***	0.00000	7.57
381	***THBS1***	0.00000	8.84	438	TRPC4	0.00000	7.56
382	BATF3	0.00000	8.83	439	PODNL1	0.00021	7.56
383	***CXCL1***	0.00001	8.81	440	***APOE***	0.00000	7.52
384	P4HA3	0.00000	8.81	441	***CHST11***	0.00000	7.52
385	***CTGF***	0.00000	8.80	442	***LOC729683***	0.00000	7.49
386	***TM4SF19***	0.00000	8.76	443	DYRK3	0.00001	7.46
387	***C15orf48***	0.00000	8.65	444	CNGB1	0.00000	7.45
388	CSMD3	0.00005	8.65	445	***FAM171A1***	0.00000	7.40
389	***SLC29A4***	0.00000	8.64	446	BEX4	0.03512	7.38
390	SHOX2	0.00056	8.62	447	***IFI6***	0.00598	7.37
391	***TSPAN5***	0.00000	8.60	448	TRAF1	0.00000	7.36
392	SNCB	0.00000	8.60	449	KIAA1644	0.00002	7.36
393	MIR100HG	0.00359	8.55	450	***AMIGO2***	0.00000	7.34
394	LYL1	0.00000	8.54	451	CHRNA3	0.00001	7.32
395	***PROC***	0.00000	8.50	452	CCDC184	0.00510	7.31
396	MN1	0.00000	8.50	453	***NPTX1***	0.00000	7.31
397	ATP10A	0.01549	8.47	454	***SOX4***	0.00000	7.29
398	CCDC69	0.00000	8.38	455	SARDH	0.00974	7.28
399	PNMAL1	0.00082	8.37	456	***CD274***	0.00095	7.27
400	***OAS2***	0.00001	8.32	457	***CNTN1***	0.00371	7.22
401	ZFP57	0.02673	8.28	458	TG	0.00031	7.22
402	TAGLN3	0.00004	8.26	459	***COL7A1***	0.00000	7.20
403	***NKILA***	0.00000	8.25	460	ISM1	0.02118	7.15
404	***LHB***	0.00002	8.22	461	***IRAK2***	0.00000	7.12
405	FAM171A2	0.00000	8.21	462	POU2F2	0.03129	7.11
406	CRMP1	0.00000	8.21	463	***MICAL2***	0.00000	7.11
407	SLC27A6	0.00000	8.20	464	***NKX3-1***	0.00000	7.10
408	RNF150	0.00000	8.19	465	***TNFAIP3***	0.00000	7.09
409	RND1	0.00144	8.16	466	IFFO1	0.00000	7.06
410	***PODXL***	0.00000	8.15	467	***FLNC***	0.00000	7.04
411	***FBXO32***	0.00000	8.13	468	SNCAIP	0.00005	6.99
412	***FRMD6***	0.00000	8.12	469	TMEFF1	0.00051	6.98
413	IL31RA	0.00000	8.12	470	***PMEPA1***	0.00000	6.98
414	***DKK3***	0.00000	8.11	471	***PID1***	0.00000	6.92
415	RAB39B	0.00000	8.09	472	LINC00623	0.00000	6.85
416	***SH2D2A***	0.00000	8.06	473	KRT34	0.00035	6.76
417	CCDC136	0.00000	8.03	474	ARMCX2	0.00000	6.75

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
475	PRICKLE2	0.00000	6.75	532	***TPM1***	0.00000	5.98
476	PDZD4	0.00000	6.74	533	***TMEM92***	0.00000	5.98
477	***ZNF697***	0.00000	6.74	534	RASGRP3	0.00001	5.97
478	***FNDC4***	0.00000	6.73	535	FAM131B	0.00008	5.94
479	***GPRC5B***	0.00000	6.72	536	KRT6B	0.00000	5.93
480	***LAMC2***	0.00000	6.66	537	***GREM1***	0.00018	5.93
481	CCDC8	0.00000	6.66	538	***PLCB4***	0.00145	5.93
482	***GPSM3***	0.00000	6.65	539	TRPC3	0.04051	5.92
483	***LIPG***	0.00000	6.63	540	***LIMD2***	0.00000	5.91
484	PRRT4	0.00005	6.62	541	LOC102724094	0.03915	5.90
485	LRRC71	0.00446	6.59	542	***CTHRC1***	0.00000	5.87
486	***LOC101059948***	0.00000	6.58	543	***MLLT11***	0.00000	5.87
487	***CAPN5***	0.00000	6.55	544	***LINC00704***	0.00000	5.87
488	GEM	0.00000	6.54	545	LOC100129940	0.02743	5.85
489	PRCD	0.00000	6.51	546	PAPPA-AS1	0.00208	5.84
490	CYTH4	0.00535	6.49	547	***KRT81***	0.00000	5.81
491	***FERMT1***	0.00000	6.47	548	***PDGFB***	0.00000	5.81
492	CNIH3	0.00539	6.45	549	MEX3A	0.00000	5.80
493	COL6A1	0.00000	6.45	550	***LPAR3***	0.00000	5.79
494	LOC389332	0.03911	6.44	551	SLC4A8	0.00000	5.78
495	ATP2A3	0.00000	6.42	552	PGF	0.00000	5.77
496	STARD4-AS1	0.00000	6.39	553	DUSP8	0.00000	5.75
497	***CYGB***	0.00000	6.38	554	DISC1	0.00004	5.74
498	***TM4SF19-AS1***	0.00055	6.38	555	TMEM178A	0.00017	5.73
499	EDNRA	0.00145	6.37	556	KDR	0.03055	5.71
500	LARGE	0.00033	6.34	557	GAB3	0.00004	5.71
501	***FOXD1***	0.00000	6.33	558	***MMP24***	0.00000	5.70
502	***SOCS2-AS1***	0.00000	6.33	559	SPRED3	0.00000	5.69
503	***MFAP5***	0.00001	6.32	560	CHRM4	0.00000	5.68
504	CNTNAP1	0.00000	6.31	561	***C1QTNF5***	0.00000	5.68
505	EFR3B	0.00013	6.30	562	SNAI3-AS1	0.00004	5.68
506	***RHOB***	0.00000	6.30	563	***CAMK4***	0.00019	5.67
507	SYNE1	0.00001	6.29	564	***FBN2***	0.00000	5.66
508	PIWIL2	0.03397	6.29	565	COL25A1	0.03135	5.63
509	NCR3LG1	0.00000	6.28	566	LRRN4	0.00663	5.63
510	***COL4A1***	0.00000	6.28	567	***ARHGEF40***	0.00000	5.63
511	***GRB10***	0.00000	6.26	568	***TGM2***	0.00000	5.63
512	***GLIPR1***	0.00000	6.26	569	CACNA1G	0.00011	5.61
513	CREB5	0.03005	6.26	570	ARHGAP31	0.00000	5.59
514	***HMGA2***	0.00238	6.25	571	***SALL4***	0.00001	5.59
515	***SERPINB2***	0.04672	6.22	572	KCNE5	0.00001	5.59
516	***RELB***	0.00000	6.21	573	***PKIA***	0.00000	5.58
517	***TLN2***	0.00000	6.18	574	***ALOX5AP***	0.00000	5.57
518	***SLC46A3***	0.00223	6.16	575	UGT3A2	0.00186	5.56
519	ADAMTS1	0.00000	6.16	576	***DNAJB5***	0.00000	5.56
520	***SEPT5***	0.00000	6.16	577	TMEM74B	0.00001	5.55
521	***SPHK1***	0.00000	6.15	578	PDLIM3	0.00028	5.55
522	PHYHIPL	0.02815	6.14	579	NEXN	0.00199	5.54
523	LOC100130476	0.01275	6.12	580	***FAM20C***	0.00000	5.52
524	***SOCS2***	0.00000	6.12	581	***IFI27***	0.00002	5.52
525	LINC00941	0.00000	6.10	582	SHROOM4	0.00000	5.51
526	SCNN1D	0.01256	6.10	583	***LRP1***	0.00024	5.49
527	***IRS1***	0.00000	6.08	584	***TUBB2B***	0.00000	5.46
528	PLXNA4	0.00000	6.06	585	ATP6V0A4	0.00101	5.46
529	C1S	0.00230	6.02	586	***EMP3***	0.00000	5.45
530	***SHC3***	0.00000	6.02	587	FAXC	0.00000	5.45
531	MEIS3	0.00001	5.99	588	GIPC3	0.03571	5.43

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
589	***DSE***	0.00000	5.41	646	LINC00960	0.00087	4.84
590	GRIP1	0.00011	5.41	647	DLL1	0.00000	4.83
591	***CALD1***	0.00010	5.41	648	***RAMP1***	0.00000	4.82
592	***PIK3IP1***	0.00000	5.41	649	TRIM36	0.00318	4.81
593	***PLEK2***	0.00000	5.40	650	IL1RAPL2	0.00004	4.79
594	***NGF***	0.00000	5.39	651	***KCNJ15***	0.00000	4.79
595	***INHBA***	0.00000	5.38	652	AZIN2	0.00004	4.79
596	CACHD1	0.00000	5.36	653	CRYAB	0.00026	4.79
597	ADRB1	0.02587	5.36	654	TRHDE-AS1	0.00000	4.79
598	***CDK5R1***	0.00000	5.35	655	***SH3PXD2A***	0.00000	4.77
599	REEP2	0.00000	5.33	656	***BAMBI***	0.00000	4.77
600	***CCL5***	0.00000	5.32	657	PRDM8	0.00001	4.76
601	ZCCHC24	0.00000	5.32	658	FAM231D	0.01606	4.76
602	RIMS3	0.00000	5.31	659	***FUT8***	0.00000	4.75
603	***ITGA5***	0.00000	5.29	660	TMEM86A	0.00008	4.75
604	HTRA3	0.00028	5.28	661	KIAA1324	0.00000	4.73
605	***TUBB3***	0.00000	5.28	662	PRR29	0.00947	4.72
606	PLXDC2	0.00000	5.27	663	***IFI44L***	0.00000	4.72
607	***PDGFA***	0.00000	5.27	664	DCLK1	0.00001	4.71
608	***DCBLD1***	0.00000	5.25	665	***SH3KBP1***	0.00000	4.70
609	ZFPM2	0.00007	5.25	666	***NT5E***	0.00004	4.68
610	GPNMB	0.00001	5.24	667	OVCH2	0.00073	4.68
611	EGF	0.04375	5.20	668	***SPON2***	0.01278	4.67
612	***NEDD9***	0.00000	5.16	669	RAET1K	0.00016	4.67
613	***SLITRK6***	0.00001	5.16	670	CERS4	0.03456	4.67
614	MURC	0.04706	5.16	671	***NRIP3***	0.00000	4.67
615	***RBM24***	0.00000	5.15	672	ARID3B	0.00000	4.66
616	PCDHGA6	0.04675	5.12	673	EML1	0.00008	4.66
617	***PALLD***	0.00000	5.12	674	GPR173	0.00084	4.62
618	SV2A	0.00000	5.09	675	***POPDC3***	0.00000	4.61
619	***PTHLH***	0.00000	5.08	676	***CAPRIN2***	0.00000	4.60
620	PITX3	0.00412	5.08	677	***LMBR1L***	0.00000	4.59
621	***FZD10***	0.00000	5.07	678	SAP30L-AS1	0.00000	4.58
622	***APLP1***	0.00000	5.04	679	SLC6A17	0.03089	4.57
623	***AXL***	0.00000	5.04	680	ZNF135	0.00016	4.57
624	***CYP1A1***	0.00000	5.03	681	***TCF4***	0.00000	4.55
625	PCDH18	0.00000	5.03	682	PARD6G	0.00000	4.52
626	***DPYSL2***	0.00000	5.02	683	***TBX3***	0.00000	4.52
627	***SLC8A1***	0.00000	5.01	684	SOX6	0.00000	4.51
628	HUNK	0.04879	5.01	685	KSR1	0.00000	4.49
629	***SPANXD***	0.00001	5.00	686	***MYL9***	0.00000	4.48
630	***MYO10***	0.00000	5.00	687	KIF5C	0.00008	4.48
631	***WNT5B***	0.00000	5.00	688	HIC1	0.00000	4.47
632	***HOXB2***	0.00000	5.00	689	NR5A2	0.00000	4.46
633	MKX	0.00604	5.00	690	***COL4A2***	0.00000	4.45
634	***RNF122***	0.00000	4.99	691	CYP26B1	0.00000	4.44
635	***HTR1D***	0.00000	4.98	692	***SCARF2***	0.00000	4.43
636	***ETS1***	0.00000	4.98	693	***MX2***	0.00838	4.43
637	***COL18A1***	0.00000	4.93	694	SLC22A1	0.00006	4.42
638	NES	0.00003	4.93	695	SPATA4	0.00211	4.41
639	BOC	0.00000	4.93	696	***TRHDE***	0.00000	4.40
640	***MSANTD3-TMEFF1***	0.00066	4.93	697	***LINC01137***	0.00000	4.40
641	***CHN1***	0.00000	4.92	698	CNKSR2	0.01203	4.39
642	PCDHB14	0.00052	4.89	699	***SMIM3***	0.00000	4.37
643	CXCL11	0.00096	4.88	700	SP5	0.00000	4.37
644	KIAA0226L	0.00002	4.86	701	***MATN3***	0.00000	4.37
645	DCHS2	0.00000	4.85	702	SPRR1B	0.00054	4.36

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
703	BACH2	0.00000	4.36	760	POU6F1	0.00000	3.91
704	TM6SF2	0.00015	4.36	761	***ADGRL2***	0.00000	3.91
705	UNC5C	0.00016	4.33	762	***LINC01605***	0.00000	3.90
706	***LRRC75B***	0.00000	4.32	763	***ENC1***	0.00000	3.90
707	***CXCL3***	0.00484	4.32	764	MAGEE1	0.00136	3.90
708	***ST6GALNAC5***	0.00003	4.31	765	KIAA1614	0.00125	3.89
709	SRSF12	0.00082	4.29	766	***MSANTD3***	0.00000	3.89
710	SHF	0.00000	4.26	767	***FHL1***	0.00001	3.88
711	DOK1	0.00000	4.25	768	BATF2	0.00965	3.87
712	***KRT17***	0.00016	4.24	769	TXK	0.00196	3.87
713	***BDNF***	0.00007	4.22	770	FKBP7	0.00000	3.86
714	LOC642366	0.00158	4.21	771	***MX1***	0.00184	3.85
715	PCDHGB2	0.01376	4.19	772	CDON	0.00000	3.84
716	***FAM101B***	0.00000	4.19	773	***MCAM***	0.00000	3.84
717	***GALNT10***	0.00000	4.18	774	***OCIAD2***	0.00000	3.84
718	OMP	0.00283	4.17	775	***MTSS1***	0.00000	3.84
719	***BVES***	0.00000	4.16	776	MCOLN3	0.00001	3.84
720	RUNDC3A-AS1	0.00003	4.16	777	PEAR1	0.00003	3.83
721	DOCK4	0.00024	4.16	778	GADD45G	0.00004	3.81
722	SCN1B	0.00116	4.15	779	***MIR31HG***	0.00000	3.81
723	***CLDN6***	0.00000	4.14	780	***PCNX***	0.00000	3.81
724	GUCY1B3	0.00000	4.13	781	SOCS1	0.00000	3.81
725	***KPNA7***	0.00000	4.12	782	***GAS6***	0.00000	3.80
726	***EFEMP2***	0.00000	4.12	783	LRRC3	0.00000	3.79
727	***RBAKDN***	0.01702	4.11	784	SNPH	0.00000	3.79
728	***UCHL1***	0.00000	4.11	785	STARD9	0.00003	3.79
729	APLN	0.00060	4.11	786	***MB21D2***	0.00000	3.79
730	***ITGA2***	0.00003	4.10	787	***LMCD1***	0.00000	3.77
731	***B4GALNT1***	0.00000	4.10	788	FAM225A	0.00854	3.77
732	***HHAT***	0.00000	4.10	789	***ASPHD2***	0.00000	3.77
733	***RNF130***	0.00000	4.09	790	HHIPL1	0.00242	3.77
734	TRIM46	0.00000	4.09	791	***CPA4***	0.00000	3.76
735	TRPV3	0.00701	4.08	792	***HES2***	0.00000	3.76
736	***NFKBIE***	0.00000	4.08	793	HOXD13	0.01691	3.74
737	GNAZ	0.02798	4.04	794	MGAT3	0.00105	3.72
738	***MME***	0.00000	4.04	795	SCUBE3	0.00001	3.70
739	DOCK10	0.00075	4.03	796	***GPR161***	0.00000	3.70
740	***TMEM121***	0.00000	4.01	797	OPRL1	0.00266	3.69
741	***IKBKE***	0.00000	4.00	798	CUBN	0.03728	3.69
742	***C16orf45***	0.00000	3.98	799	LINC01436	0.01126	3.68
743	GPRASP2	0.00001	3.97	800	TEK	0.00000	3.67
744	PRDM11	0.01883	3.96	801	***TEAD2***	0.00000	3.67
745	CSF1	0.00000	3.96	802	***MARCKSL1***	0.00000	3.66
746	SERPING1	0.00364	3.96	803	***CEP170***	0.00089	3.66
747	***CLSTN3***	0.00000	3.95	804	CNTNAP3P2	0.00001	3.66
748	***XAF1***	0.00028	3.95	805	***ABR***	0.00000	3.66
749	***EPHB2***	0.00000	3.94	806	***CLIP2***	0.00000	3.65
750	***GADD45A***	0.00001	3.94	807	PCDH9	0.00020	3.65
751	LOC103091866	0.00012	3.94	808	***PXDC1***	0.00000	3.64
752	***GBP1***	0.00160	3.93	809	***LARP6***	0.00000	3.64
753	***KCTD11***	0.00000	3.93	810	***MIR181A2HG***	0.00035	3.64
754	H2AFY2	0.00012	3.93	811	PARM1	0.00005	3.64
755	NANOS3	0.00824	3.92	812	***NFIX***	0.00000	3.63
756	ULBP3	0.00002	3.92	813	***S1PR5***	0.00000	3.63
757	LRRN2	0.00005	3.92	814	MMP11	0.02510	3.62
758	***CARD11***	0.00000	3.92	815	DIO2	0.00802	3.62
759	HSPA12A	0.00009	3.92	816	***CSRP2***	0.00000	3.62

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
817	***CST6***	0.00001	3.61	874	LOC101060542	0.00000	3.36
818	***ADTRP***	0.00000	3.61	875	TEPP	0.04241	3.35
819	***HDAC9***	0.00003	3.60	876	***LAMA3***	0.00000	3.35
820	DZIP1L	0.00000	3.59	877	***PDLIM7***	0.00000	3.35
821	***SPOCK1***	0.00000	3.58	878	NLGN2	0.00003	3.34
822	RASL11B	0.02042	3.57	879	HOXB4	0.00000	3.34
823	NXPH3	0.00010	3.56	880	CGNL1	0.00001	3.33
824	RASGEF1A	0.00026	3.55	881	***CRYGC***	0.00368	3.33
825	FAT4	0.00294	3.55	882	***PHLDB1***	0.00000	3.32
826	***CACNB1***	0.00000	3.54	883	***DNAJB2***	0.00000	3.32
827	***LOXL1***	0.00000	3.54	884	***TMEM132A***	0.00000	3.32
828	***ISL1***	0.00000	3.54	885	CDO1	0.04652	3.32
829	***CYTH1***	0.00000	3.54	886	***SLC6A15***	0.00045	3.32
830	***LAYN***	0.00000	3.54	887	***KCTD15***	0.00000	3.32
831	C17orf51	0.00000	3.54	888	***RWDD2A***	0.00000	3.31
832	ARHGEF25	0.00123	3.53	889	***KIF3C***	0.00000	3.30
833	***CCDC71L***	0.00000	3.53	890	FOXL1	0.00110	3.30
834	***SYNPO***	0.00000	3.52	891	***TMEM2***	0.00000	3.29
835	CYFIP2	0.00075	3.52	892	TET1	0.00000	3.29
836	***S1PR2***	0.00001	3.52	893	BNC2	0.00000	3.28
837	FAM196B	0.00025	3.51	894	***CDKN2B***	0.00000	3.28
838	***LPCAT2***	0.00000	3.51	895	***SMURF2***	0.01083	3.27
839	FRAS1	0.00012	3.51	896	***FAM214B***	0.00000	3.26
840	SOX9-AS1	0.00085	3.51	897	***TRAF5***	0.00000	3.26
841	MUM1L1	0.00778	3.51	898	TMEM171	0.00129	3.26
842	CXCL6	0.00509	3.50	899	***TMCC1***	0.00000	3.25
843	***BASP1***	0.00000	3.50	900	***CDC42EP2***	0.00000	3.25
844	***ATP2C2***	0.00000	3.49	901	***FBN1***	0.00004	3.25
845	ALDH1A2	0.00000	3.49	902	MFSD2A	0.00000	3.25
846	***EVC***	0.00000	3.49	903	***FMNL3***	0.00000	3.24
847	DOCK2	0.01809	3.48	904	***TP53I3***	0.00000	3.24
848	***STMN3***	0.00000	3.48	905	***SIPA1L1***	0.00000	3.24
849	***CBX2***	0.00000	3.47	906	***P3H3***	0.00000	3.23
850	MYO7B	0.00000	3.47	907	***FOSL1***	0.00000	3.23
851	***PTPN21***	0.00000	3.46	908	***FSTL1***	0.00000	3.23
852	DYNC1I1	0.01215	3.46	909	***LOC654342***	0.00000	3.23
853	***PDGFC***	0.00000	3.46	910	***ELL2***	0.00002	3.23
854	SMIM10L2B	0.00132	3.45	911	***IFI44***	0.00002	3.22
855	LHX5	0.00505	3.45	912	SPRY4	0.00000	3.22
856	***FZD2***	0.00000	3.45	913	***GPR153***	0.00000	3.21
857	***CCBE1***	0.00000	3.44	914	TOX2	0.00000	3.21
858	***ETV5***	0.00000	3.43	915	APLF	0.00008	3.20
859	***CTIF***	0.00000	3.42	916	***OTUB2***	0.00000	3.20
860	HAS2-AS1	0.00147	3.42	917	***IER3***	0.00000	3.20
861	SATB2	0.00000	3.42	918	***SYDE1***	0.00000	3.20
862	***DAB2***	0.00000	3.39	919	TBXA2R	0.00168	3.19
863	***MT1M***	0.00000	3.39	920	***TGFB2***	0.00000	3.17
864	PACERR	0.01694	3.38	921	***GPR143***	0.00000	3.16
865	***SFXN3***	0.00000	3.38	922	KCNQ5	0.02279	3.16
866	PDGFRL	0.00001	3.37	923	***UBA6-AS1***	0.00000	3.16
867	***RBP7***	0.00000	3.37	924	***LHX6***	0.00000	3.16
868	***MTCL1***	0.00000	3.37	925	A1BG-AS1	0.02642	3.16
869	***HERC3***	0.00000	3.37	926	***FOXN3***	0.00000	3.15
870	WIPF1	0.00000	3.36	927	VASH1	0.00912	3.15
871	C10orf25	0.00001	3.36	928	***CAP2***	0.00059	3.14
872	***PHTF1***	0.00000	3.36	929	***MICAL1***	0.00000	3.14
873	***TNNT1***	0.00000	3.36	930	***ACTG1***	0.00000	3.14

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
931	***SERPINB8***	0.00000	3.14	988	RAB30	0.00051	2.97
932	***PTPRK***	0.00000	3.13	989	RNF152	0.00001	2.97
933	MEGF9	0.00000	3.13	990	MSI1	0.00000	2.97
934	HCN2	0.00000	3.13	991	RECK	0.00210	2.96
935	ISM2	0.00152	3.12	992	***BCAS4***	0.00000	2.96
936	***C2CD4C***	0.00000	3.12	993	***HSF2BP***	0.00000	2.96
937	***P3H1***	0.00000	3.12	994	DPY19L2P1	0.00034	2.96
938	MTMR9LP	0.00024	3.12	995	***DCBLD2***	0.00054	2.96
939	***LOC100506844***	0.00000	3.12	996	STX1B	0.00000	2.96
940	HBEGF	0.00455	3.12	997	***USB1***	0.00000	2.96
941	***TBC1D2B***	0.00000	3.11	998	***AFAP1L2***	0.00000	2.96
942	***PEA15***	0.00000	3.11	999	ADGRA2	0.00000	2.96
943	***CD44***	0.00000	3.10	1000	***MAP1LC3A***	0.00000	2.96
944	NCF1C	0.00195	3.10	1001	***SMARCD3***	0.00001	2.96
945	NGFR	0.00474	3.10	1002	ZFP69B	0.00669	2.95
946	***SDC3***	0.00000	3.10	1003	ACHE	0.00003	2.94
947	***IL7R***	0.00005	3.10	1004	DNAH17	0.00030	2.94
948	***PLSCR3***	0.00000	3.09	1005	***ABL2***	0.00000	2.94
949	LINC00883	0.00126	3.09	1006	***ARSI***	0.00000	2.94
950	***SKIL***	0.00000	3.08	1007	SALL2	0.00004	2.93
951	***RAMP2***	0.00057	3.08	1008	***FZD1***	0.00000	2.93
952	***GADD45B***	0.00000	3.08	1009	***KCTD17***	0.00000	2.93
953	***TTPAL***	0.00000	3.08	1010	***LIX1L***	0.00000	2.92
954	***JARID2***	0.00000	3.07	1011	NLGN4X	0.04480	2.92
955	***SETBP1***	0.00000	3.07	1012	***HYI***	0.00000	2.92
956	SIRT4	0.00007	3.07	1013	SERPINB9	0.00023	2.91
957	WDR66	0.00005	3.07	1014	RASSF4	0.00000	2.91
958	RGL1	0.00000	3.07	1015	SCAMP5	0.00197	2.91
959	TCHH	0.02729	3.05	1016	***IL27RA***	0.00000	2.91
960	***ANGPTL4***	0.00001	3.05	1017	***TIAM2***	0.00000	2.91
961	MAP3K7CL	0.00007	3.05	1018	ZNF503-AS2	0.00015	2.89
962	***PDZD2***	0.00000	3.04	1019	***DNAJC18***	0.00000	2.89
963	ZNF528-AS1	0.00231	3.04	1020	***THEMIS2***	0.00637	2.89
964	***MMP14***	0.00000	3.03	1021	***ARHGEF28***	0.00000	2.89
965	***MMD***	0.00003	3.03	1022	***TSPAN9***	0.00000	2.89
966	LCP1	0.00183	3.02	1023	***CDK6***	0.01343	2.89
967	***SP110***	0.00000	3.02	1024	FOXO3B	0.00000	2.89
968	***HCP5***	0.00212	3.02	1025	FZD7	0.00000	2.88
969	GPR132	0.00002	3.02	1026	***DDIT4L***	0.00594	2.88
970	RASD2	0.00239	3.02	1027	***SLC26A2***	0.00024	2.88
971	ACAP1	0.00048	3.02	1028	***KLHL25***	0.00000	2.88
972	MICALCL	0.00067	3.01	1029	***CCDC102A***	0.00000	2.88
973	***NAGK***	0.00000	3.01	1030	***KLF7***	0.00000	2.87
974	***TPST1***	0.00000	3.01	1031	***PLCG1***	0.00000	2.87
975	***SPANXB1***	0.00000	3.01	1032	ANKRD44	0.00036	2.87
976	***MVB12B***	0.00000	3.00	1033	***GDF11***	0.00000	2.87
977	***TNFSF15***	0.00000	3.00	1034	ARL15	0.00250	2.87
978	***DPF1***	0.00001	3.00	1035	***PCDHGB5***	0.00000	2.86
979	***IER5L***	0.00000	3.00	1036	***ADGRG1***	0.00000	2.86
980	CNTNAP3B	0.00000	2.99	1037	LOC644554	0.00001	2.86
981	AASS	0.00004	2.99	1038	***NXPE3***	0.00001	2.86
982	***BPGM***	0.00000	2.98	1039	***ITGB6***	0.00000	2.86
983	***EVA1A***	0.00000	2.98	1040	***IRF9***	0.00000	2.86
984	***DSEL***	0.00007	2.98	1041	***BRSK1***	0.00000	2.86
985	MAPK8IP2	0.00000	2.98	1042	***PDE2A***	0.00000	2.85
986	***AKAP2***	0.00005	2.97	1043	***FHOD3***	0.00000	2.85
987	***DBN1***	0.00000	2.97	1044	***GRK5***	0.00000	2.85

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1159	***LYPD1***	0.00000	2.61	1216	***NECAP2***	0.00000	2.48
1160	FUT8-AS1	0.00018	2.61	1217	CAMK2N2	0.00522	2.48
1161	LOC283299	0.02383	2.61	1218	***WDR91***	0.00000	2.48
1162	***SHROOM2***	0.00000	2.61	1219	LOXL1-AS1	0.00019	2.48
1163	FAM43B	0.00329	2.61	1220	***AKR1B1***	0.00000	2.48
1164	***DLC1***	0.00000	2.60	1221	***ZSCAN2***	0.00000	2.48
1165	***PPP2R5B***	0.00000	2.60	1222	***EPSTI1***	0.00000	2.48
1166	***SERTAD4-AS1***	0.03021	2.60	1223	***USPL1***	0.00002	2.48
1167	COL9A2	0.00053	2.59	1224	***SLC45A3***	0.00000	2.48
1168	***LDOC1***	0.00000	2.58	1225	***BTN2A2***	0.00000	2.47
1169	***MAP3K12***	0.00000	2.57	1226	***TTYH3***	0.00000	2.47
1170	DOCK11	0.00129	2.57	1227	***CIB2***	0.00000	2.47
1171	MYO7A	0.00990	2.57	1228	***GPRIN1***	0.00000	2.46
1172	***KANK4***	0.00000	2.57	1229	CACNA2D1	0.01294	2.46
1173	***ACVR1***	0.00000	2.56	1230	***ZNF365***	0.00000	2.46
1174	***KRT5***	0.00008	2.56	1231	***DDAH2***	0.00000	2.46
1175	***SMURF1***	0.00000	2.56	1232	***CSRNP2***	0.00146	2.45
1176	***RUNX1***	0.00000	2.55	1233	***ISG15***	0.00233	2.45
1177	***C1orf106***	0.00000	2.55	1234	LIFR	0.00062	2.45
1178	LINC01124	0.00025	2.55	1235	GHET1	0.03193	2.45
1179	***MT1F***	0.00084	2.55	1236	***PBXIP1***	0.00004	2.45
1180	***AGPAT4***	0.00000	2.55	1237	***NRGN***	0.00000	2.45
1181	***ORMDL3***	0.00000	2.54	1238	SHANK3	0.00000	2.44
1182	***ANKRD37***	0.00010	2.54	1239	***ZNF532***	0.00000	2.44
1183	***PPP1R18***	0.00000	2.54	1240	***PLA2R1***	0.00011	2.44
1184	SPRN	0.00008	2.54	1241	***NBPF1***	0.00002	2.44
1185	***XDH***	0.00639	2.54	1242	***SLC2A6***	0.00000	2.44
1186	***DAPK3***	0.00000	2.53	1243	PCSK6	0.01310	2.44
1187	***SACS***	0.03267	2.53	1244	***FYN***	0.00000	2.43
1188	***PACS1***	0.00000	2.53	1245	***HERC5***	0.00000	2.43
1189	***DLG4***	0.00000	2.52	1246	***PSMD2***	0.00000	2.43
1190	***TPST2***	0.00000	2.52	1247	***CCDC85B***	0.00002	2.43
1191	GABBR1	0.00852	2.52	1248	AKR1C3	0.02756	2.43
1192	C8orf46	0.00057	2.52	1249	***GLS***	0.00346	2.43
1193	***CLDN11***	0.00004	2.52	1250	***LRCH3***	0.00000	2.43
1194	CROCC	0.00000	2.52	1251	***RIPK2***	0.00001	2.43
1195	***SLC39A13***	0.00000	2.52	1252	***SH3BGRL3***	0.00000	2.43
1196	***CERKL***	0.00031	2.52	1253	***LINC01138***	0.00079	2.43
1197	***SRC***	0.00000	2.52	1254	***NXN***	0.00000	2.43
1198	***USP18***	0.02214	2.51	1255	NLRC5	0.00298	2.43
1199	***KDELR3***	0.00000	2.51	1256	***SYNGR3***	0.00000	2.42
1200	***GAREML***	0.00001	2.51	1257	***SAA1***	0.00076	2.41
1201	***LPAR5***	0.00001	2.51	1258	EID2B	0.01450	2.41
1202	C3orf18	0.00003	2.50	1259	***LOC101929128***	0.00702	2.41
1203	***ZC3H12A***	0.00000	2.50	1260	***RASA3***	0.00000	2.41
1204	***CHST15***	0.00000	2.50	1261	***PCDHGC3***	0.00003	2.41
1205	***COL16A1***	0.00000	2.50	1262	***KATNAL1***	0.00187	2.41
1206	GXYLT2	0.00111	2.50	1263	KCTD12	0.00003	2.40
1207	***PANX1***	0.00000	2.49	1264	***MYADM***	0.00000	2.40
1208	***NFKB2***	0.00000	2.49	1265	***IGFBP4***	0.00000	2.40
1209	ABCA13	0.00170	2.49	1266	***DLGAP4***	0.00000	2.40
1210	SLC4A4	0.00702	2.49	1267	***MT2A***	0.00000	2.40
1211	MYB	0.00380	2.49	1268	***FJX1***	0.00000	2.40
1212	ZNF385B	0.00045	2.49	1269	***ZNF618***	0.00000	2.40
1213	***OAS3***	0.00057	2.49	1270	***ICAM1***	0.00005	2.40
1214	MEF2C	0.00031	2.48	1271	***EVC2***	0.00000	2.40
1215	FAM95C	0.00087	2.48	1272	***P4HA2***	0.00000	2.39

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1273	SNCA	0.00050	2.39	1330	***SP100***	0.00000	2.31
1274	***GPX7***	0.00000	2.39	1331	***SLC22A3***	0.00000	2.31
1275	***CCDC74A***	0.00001	2.39	1332	***TNIP1***	0.00000	2.31
1276	ZNF329	0.00000	2.39	1333	***HSPB1***	0.00002	2.31
1277	***FAM89B***	0.00000	2.39	1334	HIC2	0.01228	2.31
1278	***IDS***	0.00000	2.39	1335	***MYH9***	0.00000	2.30
1279	***KIAA0930***	0.00000	2.39	1336	***SMO***	0.00000	2.30
1280	***RGS20***	0.00062	2.38	1337	***ARHGEF17***	0.00000	2.30
1281	MRGPRF	0.00007	2.38	1338	***MSN***	0.00000	2.30
1282	***APBB1***	0.00000	2.38	1339	***FBXO2***	0.00000	2.30
1283	***SCARA3***	0.00000	2.38	1340	***SORCS2***	0.00059	2.30
1284	***SLC25A37***	0.00000	2.38	1341	FAM65C	0.03115	2.29
1285	***ISG20***	0.00000	2.38	1342	***FKBP1A***	0.00000	2.29
1286	***MMP17***	0.00000	2.38	1343	***ZSCAN31***	0.00218	2.29
1287	***TAX1BP3***	0.00000	2.38	1344	***SHISA4***	0.00005	2.29
1288	FRY	0.00491	2.38	1345	***DCUN1D3***	0.00001	2.29
1289	CPAMD8	0.00156	2.37	1346	ZEB2	0.00211	2.29
1290	***LOC642852***	0.00001	2.37	1347	***CYLD***	0.00097	2.29
1291	***ABCA1***	0.00000	2.37	1348	***ZNF821***	0.00007	2.28
1292	NFATC1	0.00541	2.37	1349	MRAP2	0.00023	2.28
1293	***ARFGAP1***	0.00000	2.37	1350	***VEGFC***	0.00000	2.28
1294	***ENDOD1***	0.00000	2.37	1351	***IGSF3***	0.00000	2.28
1295	***VOPP1***	0.00000	2.37	1352	***ULBP2***	0.00000	2.28
1296	***SLC9A7***	0.00000	2.36	1353	***CORO1A***	0.00000	2.28
1297	***FGF2***	0.02218	2.36	1354	***BBC3***	0.00001	2.28
1298	APBA1	0.00000	2.36	1355	FAM13A-AS1	0.02023	2.28
1299	***SELM***	0.00000	2.36	1356	PRKD1	0.00799	2.28
1300	***CLSTN1***	0.00000	2.36	1357	***GUSBP4***	0.00000	2.27
1301	***PARVA***	0.00000	2.36	1358	***SLC30A4***	0.00022	2.27
1302	***IVL***	0.00014	2.36	1359	***FARP1***	0.00000	2.27
1303	***EPG5***	0.00001	2.36	1360	***HELZ2***	0.00376	2.27
1304	***ZNF699***	0.00004	2.35	1361	***MARVELD1***	0.00000	2.27
1305	***BMP2***	0.00000	2.35	1362	NXPH4	0.04986	2.27
1306	***TCF12***	0.00120	2.35	1363	***ETV4***	0.00014	2.26
1307	LOC100507053	0.00001	2.35	1364	***ACTRT3***	0.00000	2.26
1308	***PTPRB***	0.00000	2.35	1365	***NBPF3***	0.00000	2.26
1309	***KIRREL***	0.00000	2.34	1366	***STEAP3***	0.00000	2.25
1310	***TMEM255B***	0.00001	2.34	1367	***IFI27L2***	0.00000	2.25
1311	***LETM2***	0.00000	2.34	1368	ARHGAP44	0.00048	2.25
1312	***FHL3***	0.00000	2.33	1369	***WIPI1***	0.00000	2.25
1313	***GOPC***	0.00747	2.33	1370	***TJP1***	0.00002	2.25
1314	***FBLIM1***	0.00000	2.33	1371	***CCDC93***	0.00028	2.25
1315	***ZNF319***	0.00000	2.33	1372	***NAV1***	0.00025	2.24
1316	***OLFML2A***	0.00352	2.32	1373	STEAP3-AS1	0.00000	2.24
1317	LOC100499489	0.00023	2.32	1374	ZNF154	0.00072	2.24
1318	ZNF713	0.00171	2.32	1375	***TRIM8***	0.00000	2.24
1319	***CNN2***	0.00000	2.32	1376	***C10orf35***	0.00010	2.24
1320	***GATA6***	0.00000	2.32	1377	***ATP9A***	0.00000	2.24
1321	***EXTL2***	0.00001	2.32	1378	UBE2Q2P1	0.00030	2.23
1322	RNF144A	0.00001	2.32	1379	***LUZP1***	0.00000	2.23
1323	***PGBD1***	0.00048	2.32	1380	***BMP4***	0.00003	2.23
1324	***ZNF561-AS1***	0.00000	2.32	1381	C17orf97	0.00590	2.23
1325	BFSP1	0.00799	2.32	1382	***RHOBTB1***	0.00000	2.23
1326	***ABL1***	0.00000	2.32	1383	***INPP5F***	0.00012	2.23
1327	***TMSB10***	0.00000	2.31	1384	KIF21B	0.00005	2.23
1328	PDZD7	0.00025	2.31	1385	***MSRA***	0.00005	2.23
1329	***ARHGEF10***	0.00003	2.31	1386	***PLEKHG5***	0.00000	2.22

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1387	***COL17A1***	0.00000	2.22	1444	***SNURF***	0.00000	2.15
1388	***MAPK11***	0.00000	2.22	1445	FOXL2NB	0.00290	2.15
1389	ZNF582-AS1	0.00238	2.22	1446	***TGFB1***	0.00000	2.15
1390	***TNFRSF12A***	0.00000	2.22	1447	***SMC5-AS1***	0.00016	2.15
1391	***CNN3***	0.00000	2.22	1448	***MAP4K4***	0.00021	2.15
1392	***BMPR2***	0.00002	2.22	1449	***FRMD6-AS1***	0.00000	2.15
1393	***C3orf52***	0.00000	2.22	1450	TGFB3	0.00329	2.15
1394	***RBMS2***	0.00000	2.21	1451	***ENTPD7***	0.00000	2.14
1395	***NLK***	0.00001	2.21	1452	***NTAN1***	0.00000	2.14
1396	***MTMR11***	0.00003	2.21	1453	***ADPRH***	0.00000	2.14
1397	***ARL10***	0.01396	2.21	1454	***GFOD1***	0.00005	2.14
1398	***S100A3***	0.00831	2.21	1455	LUZP2	0.01721	2.14
1399	***ZFP90***	0.00021	2.21	1456	***SEPT6***	0.00000	2.14
1400	***TUBB2A***	0.00001	2.21	1457	PCDHGB6	0.00039	2.14
1401	***TLE1***	0.00000	2.21	1458	***ARHGEF18***	0.00000	2.13
1402	***PPP4R1L***	0.00003	2.21	1459	***LINC01224***	0.00000	2.13
1403	***VDR***	0.00000	2.20	1460	***STK10***	0.00000	2.13
1404	PI4KAP1	0.00735	2.20	1461	***NCKAP5L***	0.00000	2.13
1405	***TMEM198***	0.00000	2.20	1462	***TCF7***	0.00000	2.13
1406	***ZFAND2A***	0.00000	2.20	1463	PALD1	0.00160	2.13
1407	***VANGL2***	0.00000	2.20	1464	***JUN***	0.03680	2.13
1408	PCDHGA9	0.02415	2.20	1465	***FGF11***	0.00000	2.13
1409	PTCHD4	0.00020	2.19	1466	PHF21A	0.00003	2.13
1410	***CDK17***	0.00093	2.19	1467	PAX8	0.00000	2.13
1411	***KIFC3***	0.00000	2.19	1468	***SMYD2***	0.00195	2.12
1412	HVCN1	0.03339	2.19	1469	***C19orf66***	0.02386	2.12
1413	BHLHB9	0.03061	2.19	1470	PAPL	0.03552	2.12
1414	***CNTNAP3***	0.00000	2.18	1471	***PRKAB2***	0.00000	2.12
1415	***THBS3***	0.00006	2.18	1472	***KIAA1161***	0.00001	2.12
1416	***RNF121***	0.00000	2.18	1473	ATXN7L2	0.01283	2.12
1417	***ZNF853***	0.00000	2.18	1474	HDX	0.02216	2.12
1418	***LRRC49***	0.00112	2.18	1475	***RAB43***	0.00000	2.12
1419	***ANKLE2***	0.00000	2.18	1476	***TICAM1***	0.00000	2.12
1420	***SPRED1***	0.01215	2.17	1477	SRRM3	0.00027	2.12
1421	GPR137B	0.00037	2.17	1478	***TRIM3***	0.00000	2.12
1422	***ARTN***	0.00007	2.17	1479	***L3MBTL3***	0.00000	2.12
1423	***ANXA8***	0.00000	2.17	1480	***TNFRSF25***	0.00074	2.11
1424	***OPTN***	0.00329	2.17	1481	***EVL***	0.00000	2.11
1425	TTLL7	0.00765	2.17	1482	***P4HA2-AS1***	0.00685	2.11
1426	***IVNS1ABP***	0.00000	2.16	1483	***CDKN1A***	0.00000	2.11
1427	***ARHGAP32***	0.00000	2.16	1484	***STK38L***	0.03873	2.11
1428	***SERPINB7***	0.00000	2.16	1485	***ITGAV***	0.00306	2.11
1429	***ITGA6***	0.01002	2.16	1486	***GSTM3***	0.00000	2.11
1430	***SEMA3A***	0.00038	2.16	1487	ANKRD65	0.00027	2.11
1431	***SCD5***	0.00033	2.16	1488	***GNAI2***	0.00000	2.11
1432	***C8orf48***	0.00246	2.16	1489	***A1BG***	0.00003	2.11
1433	SUPT3H	0.00040	2.16	1490	***ZBED2***	0.00000	2.10
1434	***ATP1B1***	0.00000	2.16	1491	***LIMS1***	0.00442	2.10
1435	***TMEM156***	0.00365	2.16	1492	***IFFO2***	0.00001	2.10
1436	ABCC9	0.00244	2.16	1493	FUT4	0.00000	2.10
1437	***SNX25***	0.00000	2.16	1494	***ATP10D***	0.00001	2.10
1438	CYP27C1	0.00009	2.16	1495	PI4KAP2	0.00138	2.10
1439	CDC14A	0.00012	2.16	1496	FOXL2	0.00586	2.10
1440	CELSR3	0.00001	2.15	1497	***MIR22HG***	0.00038	2.09
1441	***CD3EAP***	0.00000	2.15	1498	***YPEL5***	0.00000	2.09
1442	SUSD5	0.00007	2.15	1499	***MXRA7***	0.00000	2.09
1443	***AGAP2-AS1***	0.00000	2.15	1500	***ZSCAN26***	0.02564	2.09

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1501	***LCAT***	0.00336	2.09	1558	***ZNF281***	0.00533	2.01
1502	***RTN4R***	0.00000	2.09	1559	***TBC1D16***	0.00004	2.01
1503	***IL11RA***	0.00010	2.09	1560	***STX2***	0.00187	2.01
1504	PUS10	0.00155	2.09	1561	THSD1	0.01510	2.01
1505	***GAS1***	0.00023	2.09	1562	NBPF12	0.00637	2.01
1506	***TTC7B***	0.00001	2.09	1563	NATD1	0.00000	2.01
1507	***MYO1D***	0.00000	2.09	1564	LONRF1	0.02422	2.01
1508	***KLC1***	0.00000	2.09	1565	***TNFRSF6B***	0.00002	2.01
1509	***VCL***	0.00000	2.09	1566	***SEL1L3***	0.00000	2.01
1510	CDSN	0.00040	2.09	1567	***TTTY14***	0.00197	2.00
1511	FBXO10	0.00005	2.09	1568	***AVEN***	0.00000	2.00
1512	***MIR205HG***	0.00322	2.08	1569	***SH3PXD2B***	0.00000	2.00
1513	***TPM2***	0.00000	2.08	1570	***NBPF15***	0.00000	2.00
1514	***MAP7D1***	0.00000	2.08	1571	***ZNF669***	0.00000	2.00
1515	***EXT2***	0.00000	2.07	1572	***SLC16A4***	0.00501	2.00
1516	***MOXD1***	0.00000	2.07	1573	***PARP8***	0.00000	2.00
1517	CCDC92	0.00032	2.07	1574	***CSRNP1***	0.00062	2.00
1518	PRRX2	0.03290	2.07	1575	ACTB	0.00000	2.00
1519	***ZHX3***	0.00001	2.07	1576	TCONS_00029157	0.00167	2.00
1520	***MDFI***	0.00000	2.07	1577	SLC38A7	0.00000	2.00
1521	***CCND1***	0.00723	2.06	1578	PPFIBP1	0.00410	2.00
1522	***GJB2***	0.00000	2.06	1579	CDC42EP5	0.00000	2.00
1523	***RAPGEF2***	0.00220	2.06	1580	E2F7	0.00026	2.00
1524	***LYPD6***	0.01757	2.06	1581	MARK1	0.00210	1.99
1525	***DDX58***	0.00005	2.06	1582	ZNF827	0.00000	1.99
1526	***TLE4***	0.00000	2.06	1583	GOLT1A	0.00584	1.99
1527	ZNF850	0.00713	2.05	1584	VLDLR	0.00000	1.99
1528	***AADAT***	0.00022	2.05	1585	DFNA5	0.00000	1.99
1529	FAXDC2	0.02958	2.05	1586	HN1	0.00000	1.99
1530	***NUMBL***	0.00001	2.05	1587	CDH4	0.01237	1.99
1531	DNMBP-AS1	0.00151	2.05	1588	IFI16	0.00054	1.99
1532	***GAD1***	0.00000	2.05	1589	ADAMTSL5	0.02783	1.99
1533	ACTA2	0.00640	2.05	1590	FCHSD1	0.00000	1.99
1534	NPR2	0.01520	2.05	1591	IFIT1	0.00144	1.99
1535	***TRAF3***	0.00000	2.05	1592	SPOCD1	0.00035	1.98
1536	***PIK3CD-AS2***	0.04945	2.05	1593	MAPK8IP1	0.00000	1.98
1537	***PBX3***	0.00003	2.05	1594	BTBD11	0.00000	1.98
1538	***INAFM2***	0.00002	2.05	1595	AREL1	0.00000	1.98
1539	***CDC42SE1***	0.00000	2.04	1596	PLEKHO2	0.00009	1.98
1540	***ANXA5***	0.00000	2.04	1597	HERC6	0.00000	1.98
1541	***MSANTD2***	0.00045	2.04	1598	GLCE	0.00007	1.98
1542	BCL11B	0.02837	2.04	1599	MAPKAPK2	0.00000	1.98
1543	***NF2***	0.00000	2.04	1600	ZFP64	0.00000	1.98
1544	CYP2U1	0.03162	2.04	1601	NOCT	0.00000	1.98
1545	***NOL4L***	0.00001	2.04	1602	C4orf48	0.00085	1.97
1546	***CEACAM19***	0.00058	2.04	1603	KIFC2	0.00001	1.97
1547	***KIAA0754***	0.00009	2.04	1604	C17orf67	0.00052	1.97
1548	***ZSCAN9***	0.00019	2.03	1605	LOC101927027	0.00093	1.97
1549	SPNS2	0.00067	2.03	1606	FURIN	0.00000	1.97
1550	***ZFP36L1***	0.00607	2.03	1607	GRINA	0.00000	1.97
1551	***EVA1B***	0.00001	2.03	1608	FTL	0.00000	1.97
1552	***ADGRB2***	0.00000	2.03	1609	HIP1	0.00000	1.97
1553	PRKY	0.00001	2.03	1610	AMZ1	0.01094	1.97
1554	***CTXN1***	0.00002	2.02	1611	MDK	0.00002	1.97
1555	HTRA1	0.00225	2.02	1612	FADS3	0.00003	1.97
1556	***TRIM16L***	0.00000	2.02	1613	LIMK2	0.00000	1.96
1557	***ZBTB38***	0.00250	2.02	1614	WTIP	0.00012	1.96

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1615	PKP1	0.00000	1.96	1672	CCDC28B	0.00000	1.90
1616	PCNXL2	0.00000	1.96	1673	MSL3	0.00000	1.90
1617	LAT2	0.02892	1.96	1674	VASN	0.00067	1.90
1618	DUSP22	0.00000	1.96	1675	FAM127C	0.00000	1.90
1619	ZIK1	0.00001	1.96	1676	KCNS3	0.00012	1.90
1620	GABARAPL1	0.00000	1.96	1677	PFN4	0.00033	1.90
1621	ZNF641	0.00002	1.96	1678	ZNF542P	0.00257	1.90
1622	RUNX2	0.00026	1.96	1679	AP5Z1	0.00000	1.90
1623	NET1	0.00175	1.95	1680	DENND5A	0.00000	1.90
1624	PRKCA	0.00001	1.95	1681	NME7	0.00520	1.89
1625	UAP1L1	0.00001	1.95	1682	TANGO6	0.00000	1.89
1626	MAPK8IP3	0.00002	1.95	1683	KIAA1211	0.00062	1.89
1627	C20orf194	0.00000	1.95	1684	SLFN5	0.00007	1.89
1628	FLJ23867	0.00000	1.95	1685	STARD4	0.02780	1.89
1629	PLK2	0.00013	1.94	1686	ZFYVE1	0.00045	1.89
1630	TGIF2	0.00000	1.94	1687	ZNF462	0.00028	1.89
1631	C1orf74	0.00000	1.94	1688	NAB2	0.00000	1.89
1632	ZBTB46	0.00009	1.94	1689	PIP4K2C	0.00000	1.89
1633	FOXN3-AS1	0.04877	1.94	1690	TTTY15	0.00004	1.89
1634	SLMO1	0.02943	1.94	1691	TNFAIP1	0.00000	1.89
1635	MLXIP	0.00009	1.94	1692	CYP2J2	0.01915	1.89
1636	HOMER3	0.00000	1.94	1693	DHRSX	0.00000	1.89
1637	SUGCT	0.00000	1.94	1694	SKI	0.00002	1.88
1638	SSC4D	0.00400	1.94	1695	CAND2	0.00002	1.88
1639	PRAF2	0.00011	1.94	1696	AKAP12	0.00173	1.88
1640	ATP13A2	0.00000	1.93	1697	ANOS1	0.00012	1.88
1641	MMP24-AS1	0.00000	1.93	1698	KCTD7	0.00002	1.88
1642	PORCN	0.00002	1.93	1699	IGF1R	0.00012	1.88
1643	SNX29	0.00000	1.93	1700	PKIG	0.00000	1.88
1644	LINC00842	0.00063	1.93	1701	FAAP100	0.00000	1.88
1645	STX1A	0.00001	1.93	1702	FGD6	0.00000	1.88
1646	NRP1	0.00000	1.93	1703	DNMT3A	0.00000	1.88
1647	ZNF486	0.00005	1.92	1704	HABP4	0.00007	1.87
1648	SERTAD4	0.00000	1.92	1705	MYO9B	0.00000	1.87
1649	SVIL-AS1	0.00001	1.92	1706	FAM65A	0.00293	1.87
1650	STAT5A	0.00000	1.92	1707	CORO1C	0.00001	1.87
1651	ARL16	0.00000	1.92	1708	MKL1	0.00000	1.87
1652	SLC35F2	0.00001	1.92	1709	TRNP1	0.00000	1.87
1653	INSIG1	0.00000	1.92	1710	SLC37A2	0.00001	1.87
1654	FST	0.00113	1.92	1711	FAM26F	0.00225	1.86
1655	PDLIM4	0.00000	1.92	1712	SPEG	0.00073	1.86
1656	CPE	0.00002	1.92	1713	SH2B3	0.00001	1.86
1657	GCNT1	0.00002	1.92	1714	TRIB2	0.00002	1.86
1658	ZFAND5	0.00000	1.92	1715	SEZ6L2	0.00002	1.86
1659	MAFK	0.00000	1.92	1716	ZDHHC17	0.04738	1.86
1660	TP53INP2	0.00825	1.91	1717	CHPF	0.00001	1.86
1661	TSPYL4	0.00000	1.91	1718	GALNT14	0.00001	1.86
1662	ZYX	0.00000	1.91	1719	DNAH5	0.00013	1.86
1663	MESDC1	0.00007	1.91	1720	RRBP1	0.00034	1.86
1664	PXDN	0.00001	1.91	1721	ITGB1	0.03724	1.86
1665	LEF1	0.03517	1.91	1722	ATXN1	0.00006	1.86
1666	FKRP	0.00002	1.91	1723	TATDN2	0.00000	1.85
1667	CHST3	0.00001	1.91	1724	CALCOCO1	0.00000	1.85
1668	CAMKK1	0.00002	1.91	1725	GOLIM4	0.00705	1.85
1669	WNT9A	0.00004	1.91	1726	RGS12	0.00000	1.85
1670	DNMBP	0.00000	1.90	1727	TVP23C	0.02451	1.85
1671	SLC31A2	0.00000	1.90	1728	STAT2	0.00020	1.85

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1729	SLC19A2	0.00002	1.85	1786	CDH24	0.00000	1.80
1730	SLITRK5	0.00304	1.85	1787	KLHL5	0.02780	1.80
1731	UPP1	0.00000	1.85	1788	LPIN2	0.00000	1.80
1732	NR2F1-AS1	0.00155	1.84	1789	KIAA1549	0.00882	1.80
1733	RNF44	0.00001	1.84	1790	FAM219A	0.00000	1.80
1734	SIK1	0.00025	1.84	1791	PTTG1IP	0.00000	1.80
1735	SLC36A1	0.00567	1.84	1792	GOSR2	0.00000	1.79
1736	NDNL2	0.00000	1.84	1793	MAGED1	0.00000	1.79
1737	DDX26B	0.00104	1.84	1794	STOML1	0.00000	1.79
1738	JAZF1	0.03078	1.84	1795	HOXA11-AS	0.00190	1.79
1739	PLEKHN1	0.00001	1.84	1796	DMTN	0.00005	1.79
1740	LRRC8A	0.00000	1.84	1797	ADAM23	0.01475	1.79
1741	DAP	0.00000	1.84	1798	ZNF530	0.00006	1.79
1742	ARG2	0.00001	1.84	1799	ZBED1	0.00000	1.79
1743	SDCCAG8	0.00201	1.84	1800	FLJ32255	0.00031	1.79
1744	RGS10	0.00002	1.84	1801	SQLE	0.00519	1.79
1745	HMGCS1	0.00058	1.84	1802	SLC22A4	0.00010	1.79
1746	SPATS2	0.00000	1.84	1803	ZBTB17	0.00001	1.79
1747	C11orf68	0.00000	1.84	1804	OPN3	0.00105	1.79
1748	HOXA3	0.00500	1.83	1805	MACF1	0.01911	1.79
1749	RHBDF2	0.00014	1.83	1806	C17orf85	0.00015	1.79
1750	PVRL3	0.01553	1.83	1807	SLC22A15	0.00180	1.79
1751	GSN-AS1	0.00087	1.83	1808	SEC24D	0.00317	1.79
1752	RAB32	0.00000	1.83	1809	ZNF625	0.00002	1.79
1753	VAV2	0.00000	1.82	1810	FAM229B	0.00002	1.78
1754	GYG2	0.01817	1.82	1811	CBR3	0.00001	1.78
1755	GJC1	0.00486	1.82	1812	RP9	0.00000	1.78
1756	RNF24	0.00000	1.82	1813	SLC27A1	0.00090	1.78
1757	C1orf122	0.00000	1.82	1814	YWHAH	0.00000	1.78
1758	RAPH1	0.01856	1.82	1815	SOX12	0.00001	1.78
1759	MCOLN1	0.00002	1.82	1816	MYL12A	0.00000	1.78
1760	FAM210B	0.00004	1.82	1817	MFSD12	0.00001	1.78
1761	FBXL18	0.00001	1.82	1818	NFYA	0.02313	1.78
1762	CTSB	0.00000	1.82	1819	ZNF211	0.00046	1.78
1763	IFNAR2	0.00001	1.81	1820	TMEM44	0.00000	1.78
1764	SLC20A1	0.00000	1.81	1821	PLEKHG3	0.00000	1.78
1765	RIMS2	0.04536	1.81	1822	GSN	0.00000	1.78
1766	TDRD7	0.00000	1.81	1823	COPZ2	0.00000	1.78
1767	C2orf16	0.02427	1.81	1824	RBM38	0.00000	1.77
1768	HPS4	0.00000	1.81	1825	FNIP2	0.00000	1.77
1769	STAT1	0.00000	1.81	1826	RGS19	0.00008	1.77
1770	CKAP4	0.00000	1.81	1827	CLTCL1	0.00003	1.77
1771	ADAMTS16	0.00000	1.81	1828	ITGAE	0.00386	1.77
1772	TSPAN3	0.00000	1.81	1829	FCGRT	0.00000	1.77
1773	LBX2-AS1	0.01675	1.81	1830	TNFSF12	0.00384	1.77
1774	ARMCX6	0.00000	1.81	1831	SHANK2	0.00009	1.77
1775	LINC00865	0.00398	1.81	1832	LINC01572	0.00070	1.77
1776	AKT3	0.00037	1.81	1833	LOC90768	0.00039	1.77
1777	EPOR	0.00162	1.81	1834	CERCAM	0.00000	1.77
1778	SPECC1	0.00000	1.80	1835	PIP4K2A	0.00005	1.77
1779	PRKD3	0.00861	1.80	1836	DOCK9	0.00173	1.77
1780	KLF10	0.00000	1.80	1837	MAPRE2	0.00001	1.77
1781	CIDECP	0.00007	1.80	1838	ZNF627	0.00000	1.77
1782	IL1RAP	0.00005	1.80	1839	PIP4K2B	0.00000	1.77
1783	SMPD1	0.00000	1.80	1840	PLXNA1	0.00002	1.77
1784	TGFBR2	0.00002	1.80	1841	RNF216P1	0.00000	1.77
1785	GSTM4	0.00010	1.80	1842	MAPKBP1	0.00001	1.77

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1843	ZNF71	0.00011	1.77	1900	FERMT2	0.03631	1.73
1844	MOB3B	0.00000	1.77	1901	PPP1R15A	0.00628	1.73
1845	KCTD10	0.00000	1.77	1902	FAM57A	0.00001	1.73
1846	PVR	0.00000	1.77	1903	ZNF232	0.00140	1.73
1847	BCL9	0.00712	1.77	1904	MICALL2	0.00017	1.73
1848	TPM4	0.00015	1.76	1905	AP1M1	0.00000	1.73
1849	NPC2	0.00000	1.76	1906	HLA-B	0.00437	1.73
1850	PDE4A	0.01142	1.76	1907	CTTNBP2NL	0.00442	1.73
1851	RNF25	0.00000	1.76	1908	TINAGL1	0.00000	1.73
1852	MARCH3	0.00034	1.76	1909	YIPF5	0.01952	1.72
1853	AMPD2	0.00000	1.76	1910	FCHSD2	0.00048	1.72
1854	ZNF134	0.00000	1.76	1911	ZNF35	0.00070	1.72
1855	TRIM32	0.00000	1.76	1912	SH3RF3	0.00000	1.72
1856	EEPD1	0.00009	1.76	1913	ASAP1	0.00001	1.72
1857	CCM2	0.00000	1.76	1914	SNTA1	0.00000	1.72
1858	RGMB	0.01129	1.76	1915	CCDC80	0.00410	1.72
1859	SSBP3	0.00010	1.76	1916	TFE3	0.00000	1.72
1860	N4BP3	0.00000	1.76	1917	ADH5	0.00000	1.72
1861	MAP3K9	0.00003	1.76	1918	MYL6B	0.00000	1.72
1862	ZNF880	0.01435	1.76	1919	KRBA2	0.00007	1.72
1863	CSRP1	0.00000	1.76	1920	CHFR	0.00000	1.72
1864	ZNF497	0.00272	1.76	1921	HARS	0.00000	1.72
1865	EPHB4	0.00000	1.75	1922	CCDC9	0.00001	1.72
1866	LYN	0.00002	1.75	1923	STMN1	0.00000	1.71
1867	IPO5P1	0.00000	1.75	1924	SGCB	0.03513	1.71
1868	IRGQ	0.00020	1.75	1925	HDAC7	0.00000	1.71
1869	ABTB2	0.00005	1.75	1926	FANK1	0.00014	1.71
1870	PTRF	0.00000	1.75	1927	MGAT5B	0.01760	1.71
1871	ECM1	0.02373	1.75	1928	PLK3	0.00312	1.71
1872	RIC8A	0.00000	1.75	1929	TMEM8A	0.00000	1.71
1873	IGFBP6	0.00006	1.75	1930	ZNF408	0.00009	1.71
1874	TMEM40	0.00000	1.75	1931	PTGFRN	0.00000	1.71
1875	PHC2	0.00003	1.74	1932	LINC00265	0.02879	1.71
1876	APOBEC3B	0.00002	1.74	1933	LDB1	0.00000	1.71
1877	ZNF428	0.00000	1.74	1934	CC2D1B	0.00000	1.71
1878	TLN1	0.00008	1.74	1935	CNOT4	0.00081	1.71
1879	IFI27L1	0.00043	1.74	1936	RAP1GAP2	0.00000	1.71
1880	PARP3	0.00000	1.74	1937	CTNNBIP1	0.00000	1.71
1881	MFHAS1	0.00000	1.74	1938	FAM109A	0.00001	1.71
1882	NBPF9	0.02918	1.74	1939	ULK4	0.00002	1.71
1883	PLEKHG2	0.00010	1.74	1940	VAT1	0.00000	1.71
1884	CD2BP2	0.00000	1.74	1941	USP11	0.00000	1.70
1885	TTL	0.00022	1.74	1942	KLHL18	0.00036	1.70
1886	IFNLR1	0.00051	1.74	1943	EXOG	0.01999	1.70
1887	C1orf216	0.00000	1.74	1944	PLOD1	0.00016	1.70
1888	CADM4	0.00016	1.74	1945	ACOT9	0.00000	1.70
1889	DTX3	0.00251	1.74	1946	GNA12	0.00000	1.70
1890	ITPKB	0.00011	1.74	1947	C14orf159	0.00138	1.70
1891	EXT1	0.00000	1.74	1948	THOC5	0.00000	1.70
1892	MGAT5	0.00000	1.74	1949	C9orf91	0.00000	1.70
1893	TRAF6	0.00000	1.74	1950	ARF4	0.00002	1.69
1894	SLC12A4	0.00000	1.74	1951	TULP4	0.00252	1.69
1895	ATF5	0.00001	1.74	1952	SEMA4F	0.00000	1.69
1896	TMEM265	0.00002	1.73	1953	ZNF3	0.00002	1.69
1897	MICAL3	0.00001	1.73	1954	CACNB3	0.00005	1.69
1898	SPSB1	0.00006	1.73	1955	PTK7	0.00019	1.69
1899	IFNGR2	0.00000	1.73	1956	RRAS	0.00000	1.69

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1957	RELT	0.00017	1.69	2014	KLHL21	0.00000	1.65
1958	EIF2AK4	0.01048	1.69	2015	ZSWIM8	0.00000	1.65
1959	ULK1	0.00138	1.69	2016	MEAF6	0.01901	1.65
1960	GLIS2	0.00839	1.69	2017	ZNF516	0.00000	1.65
1961	TIMP4	0.00002	1.69	2018	LOC389831	0.00255	1.65
1962	DGKD	0.00000	1.69	2019	CSGALNACT2	0.01355	1.65
1963	MANBA	0.00002	1.69	2020	KANK2	0.00000	1.65
1964	ARMC5	0.00008	1.68	2021	TBCB	0.00000	1.65
1965	ZNRF3	0.00154	1.68	2022	RIN2	0.00039	1.64
1966	CTSC	0.00000	1.68	2023	TCEA2	0.00004	1.64
1967	RPS6KC1	0.00012	1.68	2024	GDAP1	0.03738	1.64
1968	ATL1	0.00567	1.68	2025	SETMAR	0.00037	1.64
1969	TRAPPC10	0.00000	1.68	2026	SPTB	0.00005	1.64
1970	HIVEP1	0.00208	1.68	2027	ABHD4	0.00749	1.64
1971	POMT2	0.00001	1.68	2028	SNAPC4	0.00004	1.64
1972	MAP1S	0.00003	1.68	2029	ZNF140	0.00054	1.64
1973	HNRNPA1P33	0.00299	1.68	2030	PROCR	0.00000	1.64
1974	SLC35B4	0.00079	1.68	2031	SRF	0.00090	1.64
1975	VPS18	0.00031	1.68	2032	CAP1	0.00000	1.64
1976	MEX3D	0.00004	1.67	2033	ACTR1A	0.00000	1.64
1977	CDH3	0.00003	1.67	2034	PSTPIP2	0.00218	1.63
1978	PHLDA2	0.00005	1.67	2035	ELMO2	0.00000	1.63
1979	COMMD9	0.00000	1.67	2036	TMEM206	0.00020	1.63
1980	TMEM110	0.00025	1.67	2037	ADCY7	0.00024	1.63
1981	DRAP1	0.00000	1.67	2038	TMEM25	0.00000	1.63
1982	CD59	0.00000	1.67	2039	LIMK1	0.00000	1.63
1983	S100A2	0.00077	1.67	2040	CFL1	0.00000	1.63
1984	CYP27B1	0.03986	1.67	2041	CERS5	0.00030	1.63
1985	LTBP1	0.00012	1.67	2042	CERS6	0.00163	1.63
1986	SEC61A1	0.00000	1.67	2043	SESN2	0.00870	1.63
1987	CBX1	0.00013	1.66	2044	TFIP11	0.00000	1.63
1988	CRCP	0.00000	1.66	2045	TCF3	0.00001	1.62
1989	YKT6	0.00000	1.66	2046	IFIT3	0.03285	1.62
1990	ATP6V1B2	0.00000	1.66	2047	TTBK2	0.00890	1.62
1991	KRT16	0.00034	1.66	2048	SFN	0.00000	1.62
1992	SEC31A	0.00000	1.66	2049	BAG3	0.00002	1.62
1993	ATOX1	0.00000	1.66	2050	GNG4	0.00030	1.62
1994	MAPK7	0.00178	1.66	2051	TJP2	0.00404	1.62
1995	KDM5B	0.00061	1.66	2052	CD276	0.00001	1.62
1996	ZNF668	0.00000	1.66	2053	RCAN1	0.00038	1.62
1997	BHLHE40	0.00008	1.66	2054	ZC3H7B	0.00000	1.62
1998	KIF13A	0.02979	1.66	2055	CLIP4	0.00062	1.62
1999	HILPDA	0.00001	1.66	2056	RABAC1	0.00005	1.61
2000	CIC	0.00002	1.66	2057	AP4M1	0.00000	1.61
2001	PTBP2	0.01924	1.66	2058	MYL6	0.00000	1.61
2002	CDK14	0.00527	1.66	2059	NBPF8	0.00719	1.61
2003	CLTB	0.00002	1.65	2060	ZBTB47	0.00016	1.61
2004	ORAI2	0.00008	1.65	2061	SVIL	0.00029	1.61
2005	POFUT2	0.00003	1.65	2062	CEP170B	0.00000	1.61
2006	NT5DC2	0.00001	1.65	2063	UBE2F	0.00000	1.61
2007	CMTM3	0.00000	1.65	2064	TP53INP1	0.04134	1.61
2008	HOXA11	0.00003	1.65	2065	SNN	0.00002	1.61
2009	MT1E	0.00000	1.65	2066	ZPR1	0.00000	1.61
2010	TP63	0.00606	1.65	2067	UBTD1	0.00000	1.61
2011	OBSL1	0.00035	1.65	2068	PAK1	0.00000	1.61
2012	PPP1R14B	0.00000	1.65	2069	RAC2	0.00000	1.61
2013	B3GNT9	0.00001	1.65	2070	SEC14L1	0.00000	1.61

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1957	RELT	0.00017	1.69	2014	KLHL21	0.00000	1.65
1958	EIF2AK4	0.01048	1.69	2015	ZSWIM8	0.00000	1.65
1959	ULK1	0.00138	1.69	2016	MEAF6	0.01901	1.65
1960	GLIS2	0.00839	1.69	2017	ZNF516	0.00000	1.65
1961	TIMP4	0.00002	1.69	2018	LOC389831	0.00255	1.65
1962	DGKD	0.00000	1.69	2019	CSGALNACT2	0.01355	1.65
1963	MANBA	0.00002	1.69	2020	KANK2	0.00000	1.65
1964	ARMC5	0.00008	1.68	2021	TBCB	0.00000	1.65
1965	ZNRF3	0.00154	1.68	2022	RIN2	0.00039	1.64
1966	CTSC	0.00000	1.68	2023	TCEA2	0.00004	1.64
1967	RPS6KC1	0.00012	1.68	2024	GDAP1	0.03738	1.64
1968	ATL1	0.00567	1.68	2025	SETMAR	0.00037	1.64
1969	TRAPPC10	0.00000	1.68	2026	SPTB	0.00005	1.64
1970	HIVEP1	0.00208	1.68	2027	ABHD4	0.00749	1.64
1971	POMT2	0.00001	1.68	2028	SNAPC4	0.00004	1.64
1972	MAP1S	0.00003	1.68	2029	ZNF140	0.00054	1.64
1973	HNRNPA1P33	0.00299	1.68	2030	PROCR	0.00000	1.64
1974	SLC35B4	0.00079	1.68	2031	SRF	0.00090	1.64
1975	VPS18	0.00031	1.68	2032	CAP1	0.00000	1.64
1976	MEX3D	0.00004	1.67	2033	ACTR1A	0.00000	1.64
1977	CDH3	0.00003	1.67	2034	PSTPIP2	0.00218	1.63
1978	PHLDA2	0.00005	1.67	2035	ELMO2	0.00000	1.63
1979	COMMD9	0.00000	1.67	2036	TMEM206	0.00020	1.63
1980	TMEM110	0.00025	1.67	2037	ADCY7	0.00024	1.63
1981	DRAP1	0.00000	1.67	2038	TMEM25	0.00000	1.63
1982	CD59	0.00000	1.67	2039	LIMK1	0.00000	1.63
1983	S100A2	0.00077	1.67	2040	CFL1	0.00000	1.63
1984	CYP27B1	0.03986	1.67	2041	CERS5	0.00030	1.63
1985	LTBP1	0.00012	1.67	2042	CERS6	0.00163	1.63
1986	SEC61A1	0.00000	1.67	2043	SESN2	0.00870	1.63
1987	CBX1	0.00013	1.66	2044	TFIP11	0.00000	1.63
1988	CRCP	0.00000	1.66	2045	TCF3	0.00001	1.62
1989	YKT6	0.00000	1.66	2046	IFIT3	0.03285	1.62
1990	ATP6V1B2	0.00000	1.66	2047	TTBK2	0.00890	1.62
1991	KRT16	0.00034	1.66	2048	SFN	0.00000	1.62
1992	SEC31A	0.00000	1.66	2049	BAG3	0.00002	1.62
1993	ATOX1	0.00000	1.66	2050	GNG4	0.00030	1.62
1994	MAPK7	0.00178	1.66	2051	TJP2	0.00404	1.62
1995	KDM5B	0.00061	1.66	2052	CD276	0.00001	1.62
1996	ZNF668	0.00000	1.66	2053	RCAN1	0.00038	1.62
1997	BHLHE40	0.00008	1.66	2054	ZC3H7B	0.00000	1.62
1998	KIF13A	0.02979	1.66	2055	CLIP4	0.00062	1.62
1999	HILPDA	0.00001	1.66	2056	RABAC1	0.00005	1.61
2000	CIC	0.00002	1.66	2057	AP4M1	0.00000	1.61
2001	PTBP2	0.01924	1.66	2058	MYL6	0.00000	1.61
2002	CDK14	0.00527	1.66	2059	NBPF8	0.00719	1.61
2003	CLTB	0.00002	1.65	2060	ZBTB47	0.00016	1.61
2004	ORAI2	0.00008	1.65	2061	SVIL	0.00029	1.61
2005	POFUT2	0.00003	1.65	2062	CEP170B	0.00000	1.61
2006	NT5DC2	0.00001	1.65	2063	UBE2F	0.00000	1.61
2007	CMTM3	0.00000	1.65	2064	TP53INP1	0.04134	1.61
2008	HOXA11	0.00003	1.65	2065	SNN	0.00002	1.61
2009	MT1E	0.00000	1.65	2066	ZPR1	0.00000	1.61
2010	TP63	0.00606	1.65	2067	UBTD1	0.00000	1.61
2011	OBSL1	0.00035	1.65	2068	PAK1	0.00000	1.61
2012	PPP1R14B	0.00000	1.65	2069	RAC2	0.00000	1.61
2013	B3GNT9	0.00001	1.65	2070	SEC14L1	0.00000	1.61

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1957	RELT	0.00017	1.69	2014	KLHL21	0.00000	1.65
1958	EIF2AK4	0.01048	1.69	2015	ZSWIM8	0.00000	1.65
1959	ULK1	0.00138	1.69	2016	MEAF6	0.01901	1.65
1960	GLIS2	0.00839	1.69	2017	ZNF516	0.00000	1.65
1961	TIMP4	0.00002	1.69	2018	LOC389831	0.00255	1.65
1962	DGKD	0.00000	1.69	2019	CSGALNACT2	0.01355	1.65
1963	MANBA	0.00002	1.69	2020	KANK2	0.00000	1.65
1964	ARMC5	0.00008	1.68	2021	TBCB	0.00000	1.65
1965	ZNRF3	0.00154	1.68	2022	RIN2	0.00039	1.64
1966	CTSC	0.00000	1.68	2023	TCEA2	0.00004	1.64
1967	RPS6KC1	0.00012	1.68	2024	GDAP1	0.03738	1.64
1968	ATL1	0.00567	1.68	2025	SETMAR	0.00037	1.64
1969	TRAPPC10	0.00000	1.68	2026	SPTB	0.00005	1.64
1970	HIVEP1	0.00208	1.68	2027	ABHD4	0.00749	1.64
1971	POMT2	0.00001	1.68	2028	SNAPC4	0.00004	1.64
1972	MAP1S	0.00003	1.68	2029	ZNF140	0.00054	1.64
1973	HNRNPA1P33	0.00299	1.68	2030	PROCR	0.00000	1.64
1974	SLC35B4	0.00079	1.68	2031	SRF	0.00090	1.64
1975	VPS18	0.00031	1.68	2032	CAP1	0.00000	1.64
1976	MEX3D	0.00004	1.67	2033	ACTR1A	0.00000	1.64
1977	CDH3	0.00003	1.67	2034	PSTPIP2	0.00218	1.63
1978	PHLDA2	0.00005	1.67	2035	ELMO2	0.00000	1.63
1979	COMMD9	0.00000	1.67	2036	TMEM206	0.00020	1.63
1980	TMEM110	0.00025	1.67	2037	ADCY7	0.00024	1.63
1981	DRAP1	0.00000	1.67	2038	TMEM25	0.00000	1.63
1982	CD59	0.00000	1.67	2039	LIMK1	0.00000	1.63
1983	S100A2	0.00077	1.67	2040	CFL1	0.00000	1.63
1984	CYP27B1	0.03986	1.67	2041	CERS5	0.00030	1.63
1985	LTBP1	0.00012	1.67	2042	CERS6	0.00163	1.63
1986	SEC61A1	0.00000	1.67	2043	SESN2	0.00870	1.63
1987	CBX1	0.00013	1.66	2044	TFIP11	0.00000	1.63
1988	CRCP	0.00000	1.66	2045	TCF3	0.00001	1.62
1989	YKT6	0.00000	1.66	2046	IFIT3	0.03285	1.62
1990	ATP6V1B2	0.00000	1.66	2047	TTBK2	0.00890	1.62
1991	KRT16	0.00034	1.66	2048	SFN	0.00000	1.62
1992	SEC31A	0.00000	1.66	2049	BAG3	0.00002	1.62
1993	ATOX1	0.00000	1.66	2050	GNG4	0.00030	1.62
1994	MAPK7	0.00178	1.66	2051	TJP2	0.00404	1.62
1995	KDM5B	0.00061	1.66	2052	CD276	0.00001	1.62
1996	ZNF668	0.00000	1.66	2053	RCAN1	0.00038	1.62
1997	BHLHE40	0.00008	1.66	2054	ZC3H7B	0.00000	1.62
1998	KIF13A	0.02979	1.66	2055	CLIP4	0.00062	1.62
1999	HILPDA	0.00001	1.66	2056	RABAC1	0.00005	1.61
2000	CIC	0.00002	1.66	2057	AP4M1	0.00000	1.61
2001	PTBP2	0.01924	1.66	2058	MYL6	0.00000	1.61
2002	CDK14	0.00527	1.66	2059	NBPF8	0.00719	1.61
2003	CLTB	0.00002	1.65	2060	ZBTB47	0.00016	1.61
2004	ORAI2	0.00008	1.65	2061	SVIL	0.00029	1.61
2005	POFUT2	0.00003	1.65	2062	CEP170B	0.00000	1.61
2006	NT5DC2	0.00001	1.65	2063	UBE2F	0.00000	1.61
2007	CMTM3	0.00000	1.65	2064	TP53INP1	0.04134	1.61
2008	HOXA11	0.00003	1.65	2065	SNN	0.00002	1.61
2009	MT1E	0.00000	1.65	2066	ZPR1	0.00000	1.61
2010	TP63	0.00606	1.65	2067	UBTD1	0.00000	1.61
2011	OBSL1	0.00035	1.65	2068	PAK1	0.00000	1.61
2012	PPP1R14B	0.00000	1.65	2069	RAC2	0.00000	1.61
2013	B3GNT9	0.00001	1.65	2070	SEC14L1	0.00000	1.61

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1957	RELT	0.00017	1.69	2014	KLHL21	0.00000	1.65
1958	EIF2AK4	0.01048	1.69	2015	ZSWIM8	0.00000	1.65
1959	ULK1	0.00138	1.69	2016	MEAF6	0.01901	1.65
1960	GLIS2	0.00839	1.69	2017	ZNF516	0.00000	1.65
1961	TIMP4	0.00002	1.69	2018	LOC389831	0.00255	1.65
1962	DGKD	0.00000	1.69	2019	CSGALNACT2	0.01355	1.65
1963	MANBA	0.00002	1.69	2020	KANK2	0.00000	1.65
1964	ARMC5	0.00008	1.68	2021	TBCB	0.00000	1.65
1965	ZNRF3	0.00154	1.68	2022	RIN2	0.00039	1.64
1966	CTSC	0.00000	1.68	2023	TCEA2	0.00004	1.64
1967	RPS6KC1	0.00012	1.68	2024	GDAP1	0.03738	1.64
1968	ATL1	0.00567	1.68	2025	SETMAR	0.00037	1.64
1969	TRAPPC10	0.00000	1.68	2026	SPTB	0.00005	1.64
1970	HIVEP1	0.00208	1.68	2027	ABHD4	0.00749	1.64
1971	POMT2	0.00001	1.68	2028	SNAPC4	0.00004	1.64
1972	MAP1S	0.00003	1.68	2029	ZNF140	0.00054	1.64
1973	HNRNPA1P33	0.00299	1.68	2030	PROCR	0.00000	1.64
1974	SLC35B4	0.00079	1.68	2031	SRF	0.00090	1.64
1975	VPS18	0.00031	1.68	2032	CAP1	0.00000	1.64
1976	MEX3D	0.00004	1.67	2033	ACTR1A	0.00000	1.64
1977	CDH3	0.00003	1.67	2034	PSTPIP2	0.00218	1.63
1978	PHLDA2	0.00005	1.67	2035	ELMO2	0.00000	1.63
1979	COMMD9	0.00000	1.67	2036	TMEM206	0.00020	1.63
1980	TMEM110	0.00025	1.67	2037	ADCY7	0.00024	1.63
1981	DRAP1	0.00000	1.67	2038	TMEM25	0.00000	1.63
1982	CD59	0.00000	1.67	2039	LIMK1	0.00000	1.63
1983	S100A2	0.00077	1.67	2040	CFL1	0.00000	1.63
1984	CYP27B1	0.03986	1.67	2041	CERS5	0.00030	1.63
1985	LTBP1	0.00012	1.67	2042	CERS6	0.00163	1.63
1986	SEC61A1	0.00000	1.67	2043	SESN2	0.00870	1.63
1987	CBX1	0.00013	1.66	2044	TFIP11	0.00000	1.63
1988	CRCP	0.00000	1.66	2045	TCF3	0.00001	1.62
1989	YKT6	0.00000	1.66	2046	IFIT3	0.03285	1.62
1990	ATP6V1B2	0.00000	1.66	2047	TTBK2	0.00890	1.62
1991	KRT16	0.00034	1.66	2048	SFN	0.00000	1.62
1992	SEC31A	0.00000	1.66	2049	BAG3	0.00002	1.62
1993	ATOX1	0.00000	1.66	2050	GNG4	0.00030	1.62
1994	MAPK7	0.00178	1.66	2051	TJP2	0.00404	1.62
1995	KDM5B	0.00061	1.66	2052	CD276	0.00001	1.62
1996	ZNF668	0.00000	1.66	2053	RCAN1	0.00038	1.62
1997	BHLHE40	0.00008	1.66	2054	ZC3H7B	0.00000	1.62
1998	KIF13A	0.02979	1.66	2055	CLIP4	0.00062	1.62
1999	HILPDA	0.00001	1.66	2056	RABAC1	0.00005	1.61
2000	CIC	0.00002	1.66	2057	AP4M1	0.00000	1.61
2001	PTBP2	0.01924	1.66	2058	MYL6	0.00000	1.61
2002	CDK14	0.00527	1.66	2059	NBPF8	0.00719	1.61
2003	CLTB	0.00002	1.65	2060	ZBTB47	0.00016	1.61
2004	ORAI2	0.00008	1.65	2061	SVIL	0.00029	1.61
2005	POFUT2	0.00003	1.65	2062	CEP170B	0.00000	1.61
2006	NT5DC2	0.00001	1.65	2063	UBE2F	0.00000	1.61
2007	CMTM3	0.00000	1.65	2064	TP53INP1	0.04134	1.61
2008	HOXA11	0.00003	1.65	2065	SNN	0.00002	1.61
2009	MT1E	0.00000	1.65	2066	ZPR1	0.00000	1.61
2010	TP63	0.00606	1.65	2067	UBTD1	0.00000	1.61
2011	OBSL1	0.00035	1.65	2068	PAK1	0.00000	1.61
2012	PPP1R14B	0.00000	1.65	2069	RAC2	0.00000	1.61
2013	B3GNT9	0.00001	1.65	2070	SEC14L1	0.00000	1.61

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1957	RELT	0.00017	1.69	2014	KLHL21	0.00000	1.65
1958	EIF2AK4	0.01048	1.69	2015	ZSWIM8	0.00000	1.65
1959	ULK1	0.00138	1.69	2016	MEAF6	0.01901	1.65
1960	GLIS2	0.00839	1.69	2017	ZNF516	0.00000	1.65
1961	TIMP4	0.00002	1.69	2018	LOC389831	0.00255	1.65
1962	DGKD	0.00000	1.69	2019	CSGALNACT2	0.01355	1.65
1963	MANBA	0.00002	1.69	2020	KANK2	0.00000	1.65
1964	ARMC5	0.00008	1.68	2021	TBCB	0.00000	1.65
1965	ZNRF3	0.00154	1.68	2022	RIN2	0.00039	1.64
1966	CTSC	0.00000	1.68	2023	TCEA2	0.00004	1.64
1967	RPS6KC1	0.00012	1.68	2024	GDAP1	0.03738	1.64
1968	ATL1	0.00567	1.68	2025	SETMAR	0.00037	1.64
1969	TRAPPC10	0.00000	1.68	2026	SPTB	0.00005	1.64
1970	HIVEP1	0.00208	1.68	2027	ABHD4	0.00749	1.64
1971	POMT2	0.00001	1.68	2028	SNAPC4	0.00004	1.64
1972	MAP1S	0.00003	1.68	2029	ZNF140	0.00054	1.64
1973	HNRNPA1P33	0.00299	1.68	2030	PROCR	0.00000	1.64
1974	SLC35B4	0.00079	1.68	2031	SRF	0.00090	1.64
1975	VPS18	0.00031	1.68	2032	CAP1	0.00000	1.64
1976	MEX3D	0.00004	1.67	2033	ACTR1A	0.00000	1.64
1977	CDH3	0.00003	1.67	2034	PSTPIP2	0.00218	1.63
1978	PHLDA2	0.00005	1.67	2035	ELMO2	0.00000	1.63
1979	COMMD9	0.00000	1.67	2036	TMEM206	0.00020	1.63
1980	TMEM110	0.00025	1.67	2037	ADCY7	0.00024	1.63
1981	DRAP1	0.00000	1.67	2038	TMEM25	0.00000	1.63
1982	CD59	0.00000	1.67	2039	LIMK1	0.00000	1.63
1983	S100A2	0.00077	1.67	2040	CFL1	0.00000	1.63
1984	CYP27B1	0.03986	1.67	2041	CERS5	0.00030	1.63
1985	LTBP1	0.00012	1.67	2042	CERS6	0.00163	1.63
1986	SEC61A1	0.00000	1.67	2043	SESN2	0.00870	1.63
1987	CBX1	0.00013	1.66	2044	TFIP11	0.00000	1.63
1988	CRCP	0.00000	1.66	2045	TCF3	0.00001	1.62
1989	YKT6	0.00000	1.66	2046	IFIT3	0.03285	1.62
1990	ATP6V1B2	0.00000	1.66	2047	TTBK2	0.00890	1.62
1991	KRT16	0.00034	1.66	2048	SFN	0.00000	1.62
1992	SEC31A	0.00000	1.66	2049	BAG3	0.00002	1.62
1993	ATOX1	0.00000	1.66	2050	GNG4	0.00030	1.62
1994	MAPK7	0.00178	1.66	2051	TJP2	0.00404	1.62
1995	KDM5B	0.00061	1.66	2052	CD276	0.00001	1.62
1996	ZNF668	0.00000	1.66	2053	RCAN1	0.00038	1.62
1997	BHLHE40	0.00008	1.66	2054	ZC3H7B	0.00000	1.62
1998	KIF13A	0.02979	1.66	2055	CLIP4	0.00062	1.62
1999	HILPDA	0.00001	1.66	2056	RABAC1	0.00005	1.61
2000	CIC	0.00002	1.66	2057	AP4M1	0.00000	1.61
2001	PTBP2	0.01924	1.66	2058	MYL6	0.00000	1.61
2002	CDK14	0.00527	1.66	2059	NBPF8	0.00719	1.61
2003	CLTB	0.00002	1.65	2060	ZBTB47	0.00016	1.61
2004	ORAI2	0.00008	1.65	2061	SVIL	0.00029	1.61
2005	POFUT2	0.00003	1.65	2062	CEP170B	0.00000	1.61
2006	NT5DC2	0.00001	1.65	2063	UBE2F	0.00000	1.61
2007	CMTM3	0.00000	1.65	2064	TP53INP1	0.04134	1.61
2008	HOXA11	0.00003	1.65	2065	SNN	0.00002	1.61
2009	MT1E	0.00000	1.65	2066	ZPR1	0.00000	1.61
2010	TP63	0.00606	1.65	2067	UBTD1	0.00000	1.61
2011	OBSL1	0.00035	1.65	2068	PAK1	0.00000	1.61
2012	PPP1R14B	0.00000	1.65	2069	RAC2	0.00000	1.61
2013	B3GNT9	0.00001	1.65	2070	SEC14L1	0.00000	1.61

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2071	FAM201A	0.02880	1.61	2128	CDKN2A	0.00261	1.57
2072	ZSWIM4	0.00623	1.61	2129	ADRA1B	0.00027	1.57
2073	SAMHD1	0.00391	1.61	2130	RAB38	0.01084	1.57
2074	RBFOX2	0.00003	1.61	2131	UBE2E3	0.00662	1.57
2075	MTURN	0.01928	1.61	2132	CKB	0.00002	1.57
2076	NINJ1	0.00066	1.61	2133	ZBTB9	0.00102	1.57
2077	KRT6A	0.03810	1.61	2134	TOMM34	0.00000	1.57
2078	TFPT	0.00000	1.61	2135	ILK	0.00000	1.57
2079	PNP	0.00000	1.60	2136	FAM83G	0.00000	1.57
2080	CACNG4	0.03064	1.60	2137	WDR45	0.00000	1.57
2081	CHPF2	0.00001	1.60	2138	NFIL3	0.00030	1.57
2082	GS1-124K5.11	0.00000	1.60	2139	GZF1	0.01916	1.57
2083	TSPAN17	0.00000	1.60	2140	KIAA0226	0.00007	1.57
2084	SNX21	0.00000	1.60	2141	NDEL1	0.00003	1.57
2085	GPR39	0.00005	1.60	2142	THRA	0.00042	1.57
2086	WDR81	0.00166	1.60	2143	BNIP3	0.04579	1.57
2087	SIPA1L3	0.00003	1.60	2144	C3orf38	0.00503	1.57
2088	TNFRSF10D	0.00001	1.60	2145	TOM1	0.00008	1.57
2089	HARS2	0.00000	1.60	2146	NME4	0.00000	1.57
2090	NISCH	0.00032	1.60	2147	RHOC	0.00000	1.56
2091	B4GALT7	0.00000	1.60	2148	FBXO41	0.04018	1.56
2092	RPL23AP82	0.03696	1.60	2149	RIPK1	0.00000	1.56
2093	HMGXB3	0.00002	1.60	2150	KIF1C	0.00000	1.56
2094	TOMM40L	0.00005	1.60	2151	RAB11FIP5	0.00001	1.56
2095	NSFL1C	0.00000	1.60	2152	COMMD3	0.00000	1.56
2096	PHC1	0.00062	1.60	2153	IGF2BP2	0.04858	1.56
2097	TMEM57	0.00018	1.60	2154	MEA1	0.00000	1.56
2098	CTNNB1	0.00011	1.59	2155	GAB2	0.00020	1.56
2099	GPR156	0.03661	1.59	2156	SSR3	0.04065	1.56
2100	KLHDC8B	0.00337	1.59	2157	HERPUD1	0.00026	1.56
2101	MPV17L2	0.00005	1.59	2158	ARL4A	0.00029	1.56
2102	ABRACL	0.00000	1.59	2159	FMNL1	0.00260	1.56
2103	SLC29A1	0.00029	1.59	2160	DIEXF	0.00006	1.56
2104	CHSY1	0.00368	1.59	2161	HOMER2	0.00080	1.56
2105	CLCN7	0.00027	1.59	2162	GATAD2A	0.00000	1.56
2106	RRAGC	0.00977	1.59	2163	CDR2	0.00067	1.55
2107	WBP1L	0.00002	1.59	2164	ARHGAP17	0.00000	1.55
2108	GDI1	0.00000	1.59	2165	CENPO	0.00020	1.55
2109	BRD9	0.00000	1.58	2166	KRBA1	0.00999	1.55
2110	CD40	0.00000	1.58	2167	MGAT1	0.00000	1.55
2111	RIMKLB	0.00775	1.58	2168	RAB42	0.00020	1.55
2112	NTRK2	0.01620	1.58	2169	EIF1AD	0.00000	1.55
2113	CDYL	0.00000	1.58	2170	FBXL19	0.00023	1.55
2114	TXLNA	0.00000	1.58	2171	FOXP4	0.00243	1.55
2115	TCP11L1	0.00448	1.58	2172	ASB1	0.00000	1.55
2116	CD151	0.00000	1.58	2173	CXorf40B	0.00085	1.55
2117	SGPL1	0.00000	1.58	2174	TBC1D1	0.00003	1.55
2118	HMOX2	0.00000	1.58	2175	METTL1	0.00002	1.55
2119	LRP3	0.01570	1.57	2176	MYO1C	0.00000	1.55
2120	TP73-AS1	0.00006	1.57	2177	FZD4	0.00655	1.55
2121	ZFYVE26	0.00066	1.57	2178	CNPY4	0.01638	1.55
2122	SUSD6	0.02107	1.57	2179	PQLC2	0.00086	1.55
2123	C8orf58	0.00034	1.57	2180	PHF23	0.00210	1.55
2124	LOC101929709	0.00022	1.57	2181	NYNRIN	0.01986	1.55
2125	AJUBA	0.00003	1.57	2182	FLCN	0.00034	1.55
2126	CCDC137	0.00004	1.57	2183	CNP	0.00005	1.55
2127	CARD10	0.00004	1.57	2184	ATP2C1	0.01134	1.55

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2185	ZNF460	0.03624	1.55	2242	LYPLA2	0.00020	1.52
2186	ARHGAP1	0.00016	1.55	2243	KDM3A	0.02738	1.52
2187	CYTH2	0.00191	1.55	2244	LPXN	0.00931	1.52
2188	RNF187	0.00002	1.55	2245	SLC9A3R2	0.00298	1.52
2189	FAM168A	0.00003	1.55	2246	CDYL2	0.00004	1.52
2190	ZNF358	0.00004	1.55	2247	DCAF5	0.00037	1.52
2191	PHF1	0.00000	1.55	2248	ENTHD2	0.01109	1.52
2192	INA	0.00270	1.54	2249	CLN5	0.00067	1.52
2193	ARPC1B	0.00000	1.54	2250	B3GLCT	0.00932	1.52
2194	GPAT4	0.00002	1.54	2251	SSH1	0.00014	1.51
2195	HN1L	0.00000	1.54	2252	CDK13	0.00000	1.51
2196	RBM15B	0.00000	1.54	2253	BBS4	0.00009	1.51
2197	ZNF561	0.00513	1.54	2254	VAMP3	0.00030	1.51
2198	MTMR3	0.00001	1.54	2255	ACAT2	0.00546	1.51
2199	ZNF587	0.00032	1.54	2256	DUOXA1	0.00035	1.51
2200	TUBB6	0.00000	1.54	2257	ZNF79	0.00216	1.51
2201	LOXL3	0.00054	1.54	2258	HGSNAT	0.00004	1.51
2202	SAT1	0.00001	1.54	2259	EXO5	0.00023	1.51
2203	ZNF608	0.00421	1.54	2260	VGLL4	0.00611	1.51
2204	TSC22D1	0.00109	1.53	2261	BRF2	0.00126	1.51
2205	VMP1	0.00004	1.53	2262	GNB1	0.00003	1.51
2206	TOR1B	0.00000	1.53	2263	HJURP	0.04337	1.51
2207	CERK	0.00047	1.53	2264	FZR1	0.00287	1.51
2208	BACE1	0.00080	1.53	2265	OAZ2	0.00000	1.51
2209	ST6GALNAC6	0.00286	1.53	2266	B4GALNT4	0.00790	1.51
2210	MAP2	0.04992	1.53	2267	SRGAP2	0.00035	1.50
2211	PPIF	0.00025	1.53	2268	SAP30L	0.02210	1.50
2212	FAM118B	0.00000	1.53	2269	TP53BP1	0.00486	1.50
2213	ZNF8	0.00013	1.53	2270	RRP36	0.00000	1.50
2214	VPS39	0.00000	1.53	2271	IL18BP	0.01033	1.50
2215	SQSTM1	0.00000	1.53	2272	TEX10	0.04067	1.50
2216	PDE6D	0.00001	1.53	2273	INO80C	0.01586	1.50
2217	TMEM102	0.00187	1.53	2274	TM2D2	0.00001	1.50
2218	FLYWCH1	0.00075	1.53	2275	NGFRAP1	0.00000	1.50
2219	ZFAND3	0.00101	1.53	2276	DNLZ	0.00588	1.50
2220	SNX33	0.00059	1.53	2277	QSOX1	0.00007	1.50
2221	TVP23C-CDRT4	0.01578	1.53	2278	TIMM22	0.00000	1.50
2222	POLR3A	0.00041	1.53	2279	NIPSNAP1	0.00002	1.50
2223	ETV6	0.00033	1.53	2280	UBE2Z	0.00000	1.50
2224	ANKS3	0.03136	1.53	2281	RAB11FIP3	0.00149	1.50
2225	PNMA1	0.00005	1.53	2282	ISY1-RAB43	0.00000	1.50
2226	ANXA2	0.00000	1.53	2283	ESYT1	0.00000	1.50
2227	PGRMC2	0.00025	1.53	2284	MLPH	0.00000	1.49
2228	ZBTB4	0.00001	1.53	2285	JUND	0.00330	1.49
2229	HGS	0.00000	1.53	2286	TIMM23B	0.00130	1.49
2230	RUSC2	0.00000	1.53	2287	MAP4	0.00037	1.49
2231	STX4	0.00447	1.53	2288	MGST3	0.00003	1.49
2232	IQCK	0.00003	1.52	2289	CHD3	0.02483	1.49
2233	KIAA0355	0.00068	1.52	2290	C1orf50	0.00043	1.49
2234	EHBP1L1	0.00015	1.52	2291	TRAM2	0.00045	1.49
2235	DPP9	0.00005	1.52	2292	DDX60L	0.00080	1.49
2236	DSTYK	0.00014	1.52	2293	DIABLO	0.00002	1.49
2237	ATAT1	0.00206	1.52	2294	ZSWIM6	0.00012	1.49
2238	TECPR1	0.00023	1.52	2295	YRDC	0.00151	1.49
2239	CASP2	0.00313	1.52	2296	FKBP1B	0.00390	1.49
2240	PLEKHM2	0.00000	1.52	2297	WDYHV1	0.00000	1.49
2241	LRIG3	0.01923	1.52	2298	DCAF4	0.00000	1.49

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2299	ELF4	0.00078	1.49	2356	TESK1	0.00571	1.47
2300	MANF	0.00147	1.49	2357	NRBP1	0.00000	1.46
2301	CRYBB2P1	0.00357	1.49	2358	RBM4B	0.03293	1.46
2302	DCTN5	0.00000	1.49	2359	DGCR8	0.00018	1.46
2303	REXO2	0.01009	1.49	2360	CENPB	0.00001	1.46
2304	POLR3D	0.02013	1.49	2361	SUSD1	0.00530	1.46
2305	NCOA5	0.00000	1.49	2362	ARPC1A	0.00000	1.46
2306	TMEM234	0.00060	1.49	2363	ANXA2P2	0.01215	1.46
2307	ARPC4	0.00002	1.49	2364	PIAS3	0.00039	1.46
2308	INPPL1	0.00001	1.49	2365	PINX1	0.00066	1.46
2309	HYAL2	0.02124	1.49	2366	FUNDC2	0.00004	1.46
2310	HRAS	0.00005	1.49	2367	RIT1	0.00141	1.46
2311	APBB3	0.02116	1.49	2368	FAM21A	0.00369	1.46
2312	SLC39A14	0.00001	1.49	2369	FAM50B	0.00034	1.46
2313	FUT11	0.00074	1.49	2370	TBC1D25	0.00025	1.46
2314	NFKB1	0.00002	1.49	2371	OST4	0.00000	1.45
2315	RANGAP1	0.00000	1.49	2372	SMARCD1	0.00004	1.45
2316	TBC1D17	0.00001	1.48	2373	IL17RA	0.01767	1.45
2317	CCDC109B	0.00011	1.48	2374	CRAMP1L	0.00002	1.45
2318	GNG5	0.00000	1.48	2375	CHST14	0.00053	1.45
2319	SNAP47	0.00001	1.48	2376	TMEM189	0.00014	1.45
2320	BRK1	0.00000	1.48	2377	C7orf49	0.00000	1.45
2321	CPNE2	0.00000	1.48	2378	BSDC1	0.00000	1.45
2322	TMEM242	0.01259	1.48	2379	PFN1	0.00000	1.45
2323	TRAPPC1	0.00000	1.48	2380	OSBPL3	0.02321	1.45
2324	ATXN1L	0.00063	1.48	2381	SZRD1	0.00000	1.45
2325	TMEM185A	0.00085	1.48	2382	ZNF512	0.01215	1.45
2326	SEPW1	0.00000	1.48	2383	GABARAP	0.00000	1.45
2327	LLGL1	0.00042	1.48	2384	MARK4	0.00012	1.45
2328	PTPRF	0.00250	1.48	2385	RAPGEF1	0.01381	1.45
2329	EIF2B2	0.00000	1.48	2386	ASL	0.00002	1.45
2330	KDM6A	0.00107	1.48	2387	TSPAN6	0.00000	1.45
2331	ATAD3B	0.04511	1.48	2388	CDC16	0.00119	1.45
2332	TPM3	0.00000	1.48	2389	CASP6	0.00000	1.45
2333	BTN2A1	0.00104	1.48	2390	LAMTOR1	0.00000	1.45
2334	SLC13A3	0.00042	1.48	2391	DCTN1	0.00003	1.45
2335	ATG13	0.00000	1.48	2392	JMJD4	0.00168	1.45
2336	C15orf57	0.00027	1.48	2393	TRIM27	0.00077	1.45
2337	PFKP	0.00000	1.48	2394	BUD31	0.00000	1.45
2338	SMYD3	0.00000	1.47	2395	LDOC1L	0.01798	1.45
2339	FTSJ1	0.00000	1.47	2396	AQP11	0.04264	1.44
2340	STX3	0.00989	1.47	2397	SUPT5H	0.00000	1.44
2341	EXOC7	0.00000	1.47	2398	ADD1	0.00000	1.44
2342	TRIM21	0.01113	1.47	2399	ZNF707	0.00075	1.44
2343	SRR	0.00534	1.47	2400	CLCN6	0.02617	1.44
2344	TAGLN2	0.00000	1.47	2401	C1orf52	0.03364	1.44
2345	PML	0.00569	1.47	2402	ARMC9	0.00068	1.44
2346	ST5	0.00534	1.47	2403	RELA	0.00000	1.44
2347	GAS2L1	0.03242	1.47	2404	LOC389906	0.01191	1.44
2348	KANSL3	0.00128	1.47	2405	B3GALT6	0.01658	1.44
2349	TAOK2	0.00279	1.47	2406	STRADA	0.00000	1.44
2350	CMTM7	0.00158	1.47	2407	PANK2	0.00083	1.44
2351	TUBB	0.00000	1.47	2408	NPLOC4	0.00014	1.44
2352	TEP1	0.02708	1.47	2409	TECPR2	0.00390	1.44
2353	CMTR1	0.00185	1.47	2410	MGEA5	0.01989	1.44
2354	BRPF3	0.03586	1.47	2411	BAX	0.00001	1.44
2355	PPM1M	0.00044	1.47	2412	GABPB1	0.00458	1.44

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2413	BTBD9	0.00077	1.44	2470	SYS1	0.00000	1.41
2414	GFOD2	0.01231	1.44	2471	PWP1	0.00000	1.41
2415	RNF40	0.00000	1.44	2472	FAM131A	0.02562	1.41
2416	TM2D3	0.04091	1.44	2473	FAM53C	0.00243	1.41
2417	LPPR2	0.00668	1.44	2474	SIPA1	0.00001	1.41
2418	VPS53	0.00315	1.44	2475	KNOP1	0.00000	1.41
2419	FADD	0.00033	1.44	2476	ARF3	0.00000	1.41
2420	KDM4A	0.00312	1.44	2477	PCYT1A	0.00091	1.41
2421	ATG101	0.00006	1.44	2478	TSR2	0.00002	1.41
2422	AGTRAP	0.00015	1.44	2479	ME3	0.00266	1.41
2423	TMEM51	0.01132	1.44	2480	PDXK	0.03422	1.41
2424	C7orf43	0.03540	1.44	2481	TRPC4AP	0.00000	1.41
2425	FHOD1	0.00020	1.44	2482	FAM50A	0.00001	1.41
2426	GSTO1	0.00003	1.44	2483	HSD17B7	0.04408	1.41
2427	GPAM	0.00309	1.44	2484	GPN2	0.00506	1.41
2428	ENTPD4	0.00003	1.44	2485	MSTO1	0.01229	1.41
2429	GOLGA3	0.00344	1.43	2486	AMMECR1L	0.02856	1.41
2430	ORMDL2	0.00001	1.43	2487	USP35	0.00004	1.41
2431	CCT2	0.00074	1.43	2488	TNFRSF10B	0.00085	1.41
2432	SSR2	0.00000	1.43	2489	MED20	0.00000	1.41
2433	PTPRJ	0.00035	1.43	2490	LPIN3	0.00524	1.41
2434	ERCC3	0.00000	1.43	2491	RASSF1	0.00194	1.41
2435	SH3BP5L	0.00284	1.43	2492	ZDHHC18	0.02147	1.41
2436	RNF215	0.01873	1.43	2493	CYTH3	0.00000	1.41
2437	IRF2BP2	0.00020	1.43	2494	CENPBD1	0.00662	1.41
2438	ITGA3	0.00015	1.43	2495	PDE4DIP	0.00123	1.41
2439	GNL1	0.00000	1.43	2496	SAP30BP	0.00010	1.40
2440	CUTA	0.00004	1.43	2497	MICA	0.04038	1.40
2441	PLD3	0.02924	1.43	2498	TOLLIP	0.01828	1.40
2442	ENG	0.01868	1.43	2499	RING1	0.02789	1.40
2443	FBXL12	0.00124	1.43	2500	RAD54L2	0.02557	1.40
2444	SLC37A3	0.00001	1.43	2501	SNF8	0.00008	1.40
2445	P3H4	0.00051	1.43	2502	ETNK2	0.00003	1.40
2446	SLC35E1	0.00078	1.43	2503	CDC34	0.00009	1.40
2447	WDR82	0.00000	1.43	2504	TARBP2	0.00485	1.40
2448	NOSIP	0.00032	1.43	2505	TWF2	0.00003	1.40
2449	TSTD2	0.01877	1.43	2506	COMMD5	0.00001	1.40
2450	MOAP1	0.00047	1.43	2507	TP53RK	0.00086	1.40
2451	MTCH1	0.00001	1.43	2508	PINK1	0.00032	1.40
2452	MAST2	0.00002	1.43	2509	EMC10	0.00015	1.40
2453	WDR1	0.00000	1.43	2510	MSRB2	0.00007	1.40
2454	DMTF1	0.03617	1.43	2511	LAMTOR2	0.00046	1.40
2455	C17orf49	0.00323	1.43	2512	CENPV	0.01545	1.40
2456	B4GALT2	0.00041	1.42	2513	SLC2A1	0.03237	1.40
2457	FGFRL1	0.00474	1.42	2514	ARPC2	0.00017	1.39
2458	NARF	0.00113	1.42	2515	FKBP9	0.00381	1.39
2459	CDCA7L	0.00008	1.42	2516	FBRS	0.00607	1.39
2460	UROS	0.00005	1.42	2517	ATXN7L3B	0.00008	1.39
2461	JUP	0.00180	1.42	2518	UBXN7	0.04919	1.39
2462	ACLY	0.00191	1.42	2519	TRMT12	0.00020	1.39
2463	SCRN1	0.04685	1.42	2520	POLM	0.00188	1.39
2464	SERPINH1	0.00298	1.42	2521	MAPRE1	0.00030	1.39
2465	STRIP1	0.00000	1.42	2522	ZNF282	0.00213	1.39
2466	USP22	0.00004	1.42	2523	TAF7	0.00062	1.39
2467	AKT1	0.00005	1.42	2524	NOP14-AS1	0.00583	1.39
2468	E2F3	0.02288	1.41	2525	FAM120B	0.01361	1.39
2469	AP2B1	0.00179	1.41	2526	DBNL	0.00001	1.39

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2527	NINL	0.01965	1.39	2584	CXXC1	0.00340	1.36
2528	DTD1	0.00177	1.39	2585	KANSL1	0.00023	1.36
2529	FAM32A	0.00000	1.39	2586	PPP1R12C	0.00411	1.36
2530	SERINC3	0.00760	1.39	2587	PHACTR4	0.00402	1.36
2531	PDZD11	0.00000	1.39	2588	FXYD5	0.00001	1.36
2532	EBP	0.00048	1.39	2589	PFDN1	0.00756	1.36
2533	SNAPC2	0.00007	1.38	2590	ATP6V0A2	0.01253	1.36
2534	CAPZB	0.00000	1.38	2591	TMEM14C	0.00000	1.36
2535	TOR1A	0.00001	1.38	2592	SHKBP1	0.00421	1.36
2536	PFN2	0.00000	1.38	2593	RTCB	0.00000	1.36
2537	MAP3K4	0.00231	1.38	2594	RPP30	0.00025	1.36
2538	WIPI2	0.00000	1.38	2595	GBF1	0.00467	1.36
2539	ERC1	0.00451	1.38	2596	APOPT1	0.00909	1.36
2540	CMIP	0.00187	1.38	2597	GMIP	0.01767	1.36
2541	DVL1	0.00176	1.38	2598	RFX5	0.00053	1.36
2542	TNKS1BP1	0.02269	1.38	2599	PRDM4	0.00700	1.36
2543	BRMS1	0.00008	1.38	2600	PPP2R2D	0.00075	1.36
2544	ERF	0.00594	1.38	2601	FBXL5	0.03263	1.36
2545	TBC1D13	0.00013	1.38	2602	HEATR6	0.00024	1.36
2546	TBCC	0.00214	1.37	2603	SLC15A4	0.00233	1.36
2547	TRIM25	0.00104	1.37	2604	ADAR	0.01523	1.36
2548	MIR4435-2HG	0.02124	1.37	2605	DCTN2	0.00000	1.35
2549	PPRC1	0.00125	1.37	2606	DNAJC9	0.01097	1.35
2550	GRAMD1A	0.00313	1.37	2607	NUP62	0.00070	1.35
2551	ARFRP1	0.00058	1.37	2608	PVRL2	0.00028	1.35
2552	STX8	0.00891	1.37	2609	TES	0.00866	1.35
2553	SLC35E2B	0.02874	1.37	2610	TMEM173	0.00149	1.35
2554	AEN	0.00018	1.37	2611	COX19	0.00925	1.35
2555	MAD2L2	0.00005	1.37	2612	BIN3	0.00023	1.35
2556	RTKN	0.00016	1.37	2613	KBTBD2	0.00195	1.35
2557	DEXI	0.00022	1.37	2614	PDCD11	0.03775	1.35
2558	PBX2	0.00088	1.37	2615	NOL9	0.01192	1.35
2559	OS9	0.03367	1.37	2616	GPN1	0.00000	1.35
2560	TADA3	0.00084	1.37	2617	PELO	0.00000	1.35
2561	ACOT7	0.00007	1.37	2618	GRIPAP1	0.01870	1.35
2562	FDX1L	0.01983	1.37	2619	SURF4	0.00000	1.35
2563	TMEM208	0.00003	1.37	2620	S100A11	0.00001	1.35
2564	CAMSAP1	0.00000	1.37	2621	NUDCD3	0.00001	1.35
2565	ANO10	0.00093	1.37	2622	PRMT2	0.00000	1.35
2566	SNX12	0.00000	1.37	2623	BPHL	0.00000	1.35
2567	SQRDL	0.00002	1.37	2624	SDHAF2	0.00282	1.35
2568	FBXO44	0.00448	1.37	2625	TRAF2	0.02894	1.35
2569	PTPN1	0.00000	1.37	2626	EXOSC6	0.00313	1.35
2570	ZNF513	0.03891	1.37	2627	ADAT1	0.00001	1.35
2571	METTL22	0.00034	1.37	2628	IKBKG	0.02768	1.35
2572	CSNK1G1	0.03209	1.37	2629	MBD3	0.00268	1.34
2573	C12orf43	0.00000	1.37	2630	STK39	0.02536	1.34
2574	DAD1	0.00000	1.37	2631	TIMM10B	0.00006	1.34
2575	UBE2R2	0.00000	1.37	2632	RRAGA	0.00000	1.34
2576	PI4K2A	0.00005	1.37	2633	CASC3	0.00079	1.34
2577	PGS1	0.00049	1.37	2634	TNIP2	0.01089	1.34
2578	TNFRSF21	0.00007	1.36	2635	NTMT1	0.01728	1.34
2579	ANKRD11	0.00327	1.36	2636	CHD4	0.04903	1.34
2580	PYCARD	0.00403	1.36	2637	DEDD	0.00104	1.34
2581	EIF4EBP1	0.03087	1.36	2638	FOCAD	0.00061	1.34
2582	RNF19B	0.00672	1.36	2639	DUSP3	0.00092	1.34
2583	HIVEP2	0.03897	1.36	2640	CTBP1-AS2	0.00225	1.34

Rank	Gene	Corrected p-value	FC	Rank	Gene	Corrected p-value	FC
2641	IP6K2	0.02825	1.34	2698	DDX24	0.00000	1.30
2642	MORF4L1	0.00649	1.34	2699	AP3S2	0.01270	1.30
2643	POMGNT2	0.00017	1.34	2700	SNIP1	0.04128	1.30
2644	PNPLA6	0.01182	1.34	2701	PARD3	0.00042	1.30
2645	STAG3L4	0.00233	1.34	2702	ZNF304	0.00430	1.30
2646	ARHGDIA	0.00311	1.34	2703	S100A13	0.00013	1.30
2647	AGAP3	0.00648	1.34	2704	GNAS	0.00012	1.30
2648	RNASEH2C	0.00001	1.33	2705	ASB6	0.00096	1.30
2649	FAM134A	0.00258	1.33	2706	ZNF496	0.01425	1.30
2650	HLA-E	0.00001	1.33	2707	CENPT	0.00455	1.30
2651	ZC3HC1	0.00036	1.33	2708	CALM2	0.03566	1.30
2652	EVI5L	0.00122	1.33	2709	ZHX2	0.00617	1.30
2653	BOK	0.04219	1.33	2710	NRSN2	0.00143	1.30
2654	ISCA2	0.01587	1.33	2711	REXO4	0.01895	1.30
2655	GUSBP1	0.03645	1.33	2712	SLC39A7	0.00054	1.30
2656	HCCS	0.00001	1.33	2713	RPL39L	0.00550	1.30
2657	BAIAP2L1	0.00001	1.33	2714	UROD	0.00012	1.30
2658	RDH11	0.00673	1.33	2715	HOXD11	0.00798	1.30
2659	RAB34	0.00281	1.33	2716	KLC2	0.00554	1.30
2660	CIAPIN1	0.00044	1.33	2717	ZNF622	0.00001	1.30
2661	CLASP1	0.02046	1.33	2718	GUK1	0.02456	1.29
2662	ELOVL1	0.00003	1.33	2719	MAX	0.00419	1.29
2663	MAP3K11	0.01277	1.33	2720	JOSD1	0.00004	1.29
2664	TDP1	0.00255	1.33	2721	WDR37	0.01752	1.29
2665	CDK2	0.00004	1.32	2722	DNAJC8	0.00144	1.29
2666	STARD3	0.00011	1.32	2723	PHPT1	0.00775	1.29
2667	SRSF8	0.00109	1.32	2724	TMEM230	0.01746	1.29
2668	SF3B4	0.01132	1.32	2725	PMM2	0.00001	1.29
2669	TMEM199	0.00166	1.32	2726	YIPF3	0.00048	1.29
2670	SRP14	0.00747	1.32	2727	SLC35B2	0.02642	1.29
2671	CREB3	0.00003	1.32	2728	TUSC2	0.00305	1.29
2672	DDA1	0.00000	1.32	2729	PTPN9	0.00017	1.29
2673	SMG9	0.03239	1.32	2730	MCM7	0.00153	1.29
2674	LINC00094	0.04701	1.32	2731	SWI5	0.04182	1.29
2675	EHD4	0.00000	1.32	2732	CLTA	0.00001	1.29
2676	PREP	0.00191	1.32	2733	URB1	0.00822	1.29
2677	B4GALT3	0.00277	1.32	2734	CDK9	0.00370	1.29
2678	RNF216	0.00053	1.32	2735	SDF2	0.04215	1.29
2679	PDRG1	0.00047	1.32	2736	DAXX	0.00006	1.28
2680	DDX31	0.00625	1.32	2737	IST1	0.00037	1.28
2681	SMG6	0.01898	1.32	2738	WHSC1	0.00198	1.28
2682	BMS1P6	0.02988	1.32	2739	B3GAT3	0.03861	1.28
2683	COPS7B	0.01686	1.31	2740	BMS1P20	0.03775	1.28
2684	CDCA4	0.01252	1.31	2741	MBD1	0.00361	1.28
2685	C19orf12	0.00156	1.31	2742	ARHGEF7	0.00019	1.28
2686	VPS25	0.00148	1.31	2743	YWHAQ	0.00000	1.28
2687	ATXN7L3	0.03109	1.31	2744	GNB5	0.00024	1.28
2688	DCTD	0.00009	1.31	2745	PPP4R1	0.00022	1.28
2689	MXD4	0.00599	1.31	2746	SEC24C	0.04527	1.28
2690	OCRL	0.03815	1.31	2747	CDA	0.01677	1.28
2691	DPH7	0.00020	1.31	2748	SLC2A4RG	0.00834	1.28
2692	SHC1	0.00003	1.31	2749	CCNI	0.00005	1.28
2693	SLC25A38	0.04340	1.31	2750	NUTF2	0.00395	1.28
2694	PRRC2B	0.04074	1.31	2751	ZMIZ2	0.04813	1.28
2695	TMEM120B	0.00169	1.31	2752	COPA	0.00435	1.28
2696	KDM2B	0.00149	1.31	2753	CCNE1	0.02809	1.28
2697	ELOF1	0.00383	1.30	2754	LSG1	0.00275	1.28

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2755	TMEM39B	0.02279	1.28	2812	UBL5	0.00239	1.24
2756	BOD1	0.00303	1.28	2813	SRPRB	0.00002	1.24
2757	TMEM256	0.01619	1.28	2814	PIP5K1A	0.00857	1.24
2758	NGRN	0.00007	1.28	2815	NUDT1	0.02515	1.24
2759	SGTA	0.02029	1.28	2816	GON4L	0.01220	1.24
2760	LCMT1	0.00070	1.28	2817	DGKA	0.04669	1.24
2761	BCL7B	0.01885	1.28	2818	TMED9	0.02639	1.23
2762	TUBA1B	0.00012	1.28	2819	ACAA2	0.00057	1.23
2763	ERCC1	0.03076	1.28	2820	WDR45B	0.00470	1.23
2764	VPS33A	0.00086	1.28	2821	DDX23	0.00760	1.23
2765	HYOU1	0.01993	1.27	2822	TSR3	0.03138	1.23
2766	YTHDF1	0.02896	1.27	2823	EIF4A1	0.00297	1.23
2767	ZNF263	0.01992	1.27	2824	CPSF7	0.04129	1.23
2768	RBM4	0.00091	1.27	2825	TP53	0.03237	1.23
2769	SERF2	0.00390	1.27	2826	ISCU	0.00365	1.23
2770	MTMR14	0.00877	1.27	2827	ABCF2	0.00959	1.23
2771	RBM19	0.01570	1.27	2828	EHD1	0.00945	1.23
2772	WDR54	0.02954	1.27	2829	HK1	0.01558	1.23
2773	DDX19A	0.00022	1.27	2830	PHF19	0.02929	1.23
2774	SORBS3	0.00760	1.27	2831	AK1	0.00291	1.23
2775	MTHFD1L	0.01195	1.27	2832	SLC35D2	0.03950	1.23
2776	TMSB4X	0.00000	1.27	2833	NEDD8	0.00323	1.23
2777	MYEOV2	0.00015	1.27	2834	NUP93	0.04568	1.22
2778	NEK6	0.00146	1.27	2835	PGAM1	0.04551	1.22
2779	RTFDC1	0.00000	1.27	2836	GTF3C4	0.00069	1.22
2780	KIAA0391	0.00002	1.27	2837	SF3A3	0.00000	1.22
2781	AP2M1	0.00039	1.27	2838	PRPF4	0.00073	1.22
2782	RAF1	0.00003	1.27	2839	IRF5	0.02769	1.22
2783	FEN1	0.00139	1.27	2840	ACIN1	0.00235	1.22
2784	CBR1	0.00021	1.26	2841	SDC4	0.01387	1.22
2785	PI4KB	0.00036	1.26	2842	FAM134C	0.01689	1.22
2786	IFT52	0.00030	1.26	2843	C14orf119	0.00427	1.22
2787	MRPL14	0.02210	1.26	2844	NONO	0.02998	1.22
2788	RALY	0.00339	1.26	2845	RABIF	0.03593	1.21
2789	MPZL1	0.01963	1.26	2846	EZR	0.00690	1.21
2790	ASH2L	0.00868	1.26	2847	UBE2Q1	0.01446	1.21
2791	CLP1	0.02369	1.26	2848	MAEA	0.03081	1.21
2792	IPO9	0.00341	1.26	2849	GTF2F1	0.02364	1.21
2793	TOX4	0.00000	1.26	2850	MTFR1L	0.02221	1.21
2794	POLE3	0.00004	1.26	2851	PPP3CC	0.03546	1.20
2795	KIAA1191	0.00013	1.26	2852	UBE2J2	0.01930	1.20
2796	POLR2C	0.00001	1.25	2853	TPD52L2	0.00033	1.20
2797	ATP6V0E1	0.00003	1.25	2854	SF3B2	0.00017	1.20
2798	MICB	0.01138	1.25	2855	HNRNPAB	0.00141	1.20
2799	E2F6	0.01059	1.25	2856	GINS3	0.00486	1.20
2800	COA3	0.00218	1.25	2857	SAE1	0.00060	1.20
2801	BFAR	0.00987	1.25	2858	S100A10	0.03389	1.19
2802	CTSZ	0.02507	1.25	2859	LSM14B	0.03112	1.19
2803	MED8	0.00738	1.25	2860	SUPT4H1	0.00938	1.19
2804	SUMO3	0.00001	1.25	2861	DHDDS	0.02384	1.19
2805	MAPKAP1	0.00189	1.25	2862	ARPC5L	0.00515	1.19
2806	TMEM185B	0.01243	1.25	2863	GUCD1	0.01693	1.19
2807	KXD1	0.03701	1.25	2864	TANGO2	0.00275	1.19
2808	BRD2	0.01048	1.25	2865	DNAJC5	0.04598	1.19
2809	FAF2	0.00183	1.25	2866	ILF2	0.01515	1.18
2810	IPPK	0.00152	1.25	2867	TIMELESS	0.02162	1.17
2811	ADIPOR2	0.01265	1.25	2868	CTNNBL1	0.01972	1.17

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2869	ECD	0.02571	1.17	2926	GOT1	0.00186	-1.25
2870	RAN	0.00315	1.16	2927	TXNDC5	0.01948	-1.25
2871	DDX47	0.00854	1.16	2928	PMPCB	0.00011	-1.25
2872	EWSR1	0.03637	1.16	2929	TMEM69	0.01468	-1.25
2873	INTS4	0.02745	1.16	2930	MRPS27	0.00003	-1.25
2874	BUB3	0.02053	1.15	2931	MPRIP	0.00583	-1.25
2875	DERL1	0.02916	1.14	2932	MRPL37	0.01910	-1.25
2876	TLDC1	0.04691	1.13	2933	RPS15	0.04967	-1.25
2877	RHOA	0.04571	1.12	2934	RPS14	0.00761	-1.25
2878	SEC13	0.02105	1.12	2935	TMEM106C	0.00836	-1.25
2879	FBXO7	0.03641	1.12	2936	RPL12	0.00094	-1.25
2880	NDUFB4	0.02287	-1.14	2937	PRCP	0.00078	-1.25
2881	RPL15	0.00020	-1.15	2938	NUDT5	0.00000	-1.25
2882	PEX19	0.03314	-1.16	2939	SNRPD3	0.01077	-1.25
2883	EIF3D	0.02430	-1.17	2940	CCT7	0.00139	-1.25
2884	RPN1	0.03674	-1.17	2941	CAPN1	0.04256	-1.26
2885	RFC5	0.00004	-1.17	2942	RPL10A	0.00002	-1.26
2886	FUCA2	0.02180	-1.17	2943	EIF2D	0.00047	-1.26
2887	AK2	0.00128	-1.18	2944	ALG9	0.01392	-1.26
2888	EIF3K	0.04902	-1.18	2945	AMFR	0.00028	-1.26
2889	MTCH2	0.00013	-1.18	2946	KIF22	0.02011	-1.26
2890	ATP5H	0.00064	-1.18	2947	EIF2B5	0.00007	-1.26
2891	RPL3	0.00220	-1.19	2948	RBFA	0.00570	-1.26
2892	ALAS1	0.03248	-1.19	2949	PITRM1	0.00175	-1.26
2893	RPRD1B	0.00122	-1.19	2950	ECH1	0.00285	-1.26
2894	NUP88	0.00305	-1.19	2951	RPS11	0.01145	-1.27
2895	NDUFV3	0.02876	-1.19	2952	RPS18	0.00368	-1.27
2896	EIF4A3	0.00039	-1.19	2953	ATRAID	0.01206	-1.27
2897	CRTAP	0.01952	-1.19	2954	BRE	0.00056	-1.27
2898	EEF1D	0.02280	-1.19	2955	MRPL34	0.00048	-1.27
2899	GART	0.00052	-1.19	2956	DYNLT1	0.00010	-1.27
2900	CDK5RAP1	0.04036	-1.20	2957	WIBG	0.00311	-1.27
2901	GRPEL1	0.01396	-1.20	2958	CKLF	0.00177	-1.27
2902	ATP5L	0.00208	-1.20	2959	SUCLG1	0.00052	-1.27
2903	PRPS1	0.03055	-1.20	2960	UQCRFS1	0.00000	-1.27
2904	ECHS1	0.03458	-1.20	2961	NDUFAB1	0.00817	-1.27
2905	GATB	0.00290	-1.21	2962	TUBG1	0.03827	-1.28
2906	NHP2	0.00334	-1.21	2963	TIAM1	0.00560	-1.28
2907	MRPL21	0.04137	-1.21	2964	GTF3A	0.02346	-1.28
2908	C1orf43	0.00239	-1.22	2965	RPL19	0.00003	-1.28
2909	NDUFA9	0.00427	-1.22	2966	SLC39A9	0.00222	-1.28
2910	RPL27A	0.00019	-1.22	2967	DAZAP1	0.00001	-1.28
2911	TOMM7	0.03421	-1.22	2968	LAPTM4B	0.00657	-1.28
2912	CHMP4A	0.03954	-1.22	2969	ADI1	0.00144	-1.28
2913	RPL39	0.00022	-1.22	2970	RPL35A	0.00000	-1.28
2914	PCID2	0.00037	-1.22	2971	ERGIC3	0.01969	-1.28
2915	VPS37B	0.04643	-1.23	2972	CCNB2	0.02446	-1.28
2916	NIF3L1	0.00269	-1.23	2973	RPL13A	0.00258	-1.28
2917	AHCY	0.00244	-1.24	2974	DNAJA3	0.00444	-1.28
2918	CCDC85C	0.02404	-1.24	2975	ICT1	0.00854	-1.29
2919	NAA38	0.00904	-1.24	2976	NADSYN1	0.00042	-1.29
2920	NDUFB9	0.00000	-1.24	2977	APOA1BP	0.00029	-1.29
2921	NELFA	0.04971	-1.24	2978	EEF1G	0.00010	-1.29
2922	TUBGCP4	0.04553	-1.25	2979	COQ5	0.01243	-1.30
2923	SPCS1	0.00450	-1.25	2980	LHFPL2	0.04754	-1.30
2924	IQCE	0.02128	-1.25	2981	MRPL38	0.03147	-1.30
2925	MRPL53	0.03021	-1.25	2982	OXA1L	0.02706	-1.30

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2983	UBE2T	0.00094	-1.30	3040	KDF1	0.00106	-1.34
2984	RPS19	0.00123	-1.30	3041	GHITM	0.01629	-1.34
2985	RHOF	0.01390	-1.30	3042	LINC00493	0.03249	-1.34
2986	HIGD2A	0.00032	-1.30	3043	NDUFA10	0.00000	-1.34
2987	FAM162A	0.00479	-1.30	3044	NDUFV1	0.00155	-1.34
2988	PTTG1	0.00112	-1.31	3045	AKR1A1	0.00054	-1.34
2989	RPL27	0.00000	-1.31	3046	MYBL2	0.02924	-1.34
2990	MFF	0.00263	-1.31	3047	DPCD	0.04815	-1.34
2991	ATP5A1	0.00334	-1.31	3048	ARHGAP8	0.01725	-1.34
2992	SBNO2	0.00542	-1.31	3049	SNX8	0.00001	-1.34
2993	PA2G4	0.00000	-1.31	3050	JAGN1	0.00273	-1.34
2994	RPL13	0.02022	-1.31	3051	PMS2P1	0.02128	-1.34
2995	DDB2	0.02726	-1.31	3052	PSMG2	0.02087	-1.35
2996	RFT1	0.00853	-1.32	3053	TMEM59	0.04022	-1.35
2997	NENF	0.00101	-1.32	3054	MKNK1	0.01186	-1.35
2998	BLCAP	0.00129	-1.32	3055	PXMP2	0.01565	-1.35
2999	SCMH1	0.00889	-1.32	3056	FAH	0.00005	-1.35
3000	SYNGR2	0.00592	-1.32	3057	SOX17	0.00969	-1.35
3001	GLE1	0.00004	-1.32	3058	STOML2	0.00434	-1.35
3002	PDCD6	0.00007	-1.32	3059	PROSER2	0.00000	-1.35
3003	S100A6	0.01711	-1.32	3060	RBKS	0.00011	-1.35
3004	SLC25A5	0.00000	-1.32	3061	RPS3	0.00002	-1.35
3005	COQ9	0.00066	-1.32	3062	COQ7	0.00810	-1.35
3006	PMS2	0.04016	-1.32	3063	NDUFS2	0.00026	-1.35
3007	MRPL22	0.00093	-1.32	3064	RPS4X	0.00000	-1.35
3008	RPS8	0.00000	-1.32	3065	DUS2	0.00000	-1.35
3009	PYCR2	0.00000	-1.32	3066	KAT8	0.01370	-1.35
3010	PGM1	0.00044	-1.32	3067	SP1	0.01335	-1.35
3011	RPS21	0.00036	-1.32	3068	RPL4	0.00001	-1.36
3012	PGAM5	0.00037	-1.32	3069	CUX1	0.02233	-1.36
3013	TMED4	0.00018	-1.32	3070	GSTK1	0.00506	-1.36
3014	VAMP8	0.00001	-1.32	3071	CLPP	0.03442	-1.36
3015	ACO2	0.00236	-1.32	3072	LOC81691	0.03995	-1.36
3016	PNKD	0.00240	-1.32	3073	MRPS18B	0.00000	-1.36
3017	RPS12	0.00003	-1.32	3074	TECR	0.02710	-1.36
3018	RPS6	0.00000	-1.32	3075	ATP5G3	0.00000	-1.36
3019	QSOX2	0.00348	-1.33	3076	CDC42BPG	0.00971	-1.36
3020	MED11	0.01589	-1.33	3077	MCU	0.00048	-1.36
3021	RPS9	0.00005	-1.33	3078	RPL32	0.00000	-1.36
3022	RPS4Y1	0.00000	-1.33	3079	VWDE	0.00079	-1.36
3023	BBS1	0.04241	-1.33	3080	PRDX1	0.00267	-1.36
3024	TPD52L1	0.00060	-1.33	3081	ADCK4	0.00309	-1.36
3025	ZNF395	0.01391	-1.33	3082	EED	0.00411	-1.36
3026	VDAC1	0.00003	-1.33	3083	DPAGT1	0.00139	-1.36
3027	RPLP0	0.00022	-1.33	3084	TMCO4	0.00072	-1.37
3028	CISD3	0.04148	-1.33	3085	FARSB	0.00500	-1.37
3029	NOP56	0.01126	-1.33	3086	NOA1	0.00111	-1.37
3030	UQCC3	0.01212	-1.33	3087	ITM2C	0.02919	-1.37
3031	MDH2	0.00000	-1.33	3088	RPSA	0.00003	-1.37
3032	PLEKHA2	0.00512	-1.33	3089	RPL18	0.00014	-1.37
3033	LTBR	0.00013	-1.33	3090	RPS2	0.00074	-1.37
3034	NDUFS8	0.01252	-1.33	3091	ATP5C1	0.00001	-1.37
3035	MID1	0.00252	-1.33	3092	NPM3	0.00377	-1.37
3036	LITAF	0.01429	-1.34	3093	CLDND1	0.00517	-1.37
3037	MPI	0.00092	-1.34	3094	LRPAP1	0.00008	-1.37
3038	RPS28	0.00025	-1.34	3095	RPS23	0.00000	-1.37
3039	KRTCAP2	0.01061	-1.34	3096	C6orf106	0.00000	-1.37

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3097	RPL35	0.00001	-1.37	3154	ERBB2	0.00067	-1.41
3098	RREB1	0.00322	-1.38	3155	EFEMP1	0.00004	-1.41
3099	RPL38	0.00047	-1.38	3156	GALK1	0.00126	-1.41
3100	QPCTL	0.01569	-1.38	3157	ATP5G2	0.00001	-1.41
3101	SLC43A3	0.00515	-1.38	3158	RPS27A	0.00198	-1.41
3102	CYC1	0.00030	-1.38	3159	TMEM231	0.00332	-1.41
3103	CSRP2BP	0.01244	-1.38	3160	ERGIC1	0.00003	-1.41
3104	SMARCD2	0.00005	-1.38	3161	CHCHD7	0.00174	-1.41
3105	NADK	0.00381	-1.38	3162	IGSF8	0.01447	-1.41
3106	GRHPR	0.00000	-1.38	3163	AGFG2	0.00582	-1.42
3107	CIB1	0.00001	-1.38	3164	PDHX	0.02303	-1.42
3108	TUFM	0.00001	-1.38	3165	SLC16A1	0.02190	-1.42
3109	SLC50A1	0.00018	-1.38	3166	CSTB	0.00000	-1.42
3110	MCAT	0.00117	-1.38	3167	PYGL	0.00080	-1.42
3111	MTHFD1	0.00159	-1.39	3168	GUSB	0.00152	-1.42
3112	COX5A	0.00000	-1.39	3169	ITPK1	0.00207	-1.42
3113	OBFC1	0.00005	-1.39	3170	RPL37A	0.00000	-1.42
3114	TMED3	0.00017	-1.39	3171	CDC25C	0.00005	-1.42
3115	PPIP5K1	0.00451	-1.39	3172	IFRD2	0.04105	-1.42
3116	ZDHHC12	0.01180	-1.39	3173	RMDN3	0.00021	-1.42
3117	MPC1	0.01047	-1.39	3174	MARVELD3	0.00166	-1.42
3118	HAX1	0.00341	-1.39	3175	ABHD11	0.00050	-1.42
3119	RPSAP58	0.00194	-1.39	3176	RPL37	0.00000	-1.42
3120	SCARNA12	0.00073	-1.39	3177	RTN4IP1	0.00004	-1.42
3121	SNRPF	0.00006	-1.39	3178	CTDSPL	0.00019	-1.42
3122	PDHA1	0.01424	-1.39	3179	CAPG	0.00013	-1.42
3123	TMEM186	0.01207	-1.39	3180	ECSIT	0.00409	-1.42
3124	PAICS	0.00558	-1.39	3181	RTN3	0.00354	-1.42
3125	ANAPC16	0.00153	-1.39	3182	MROH6	0.00705	-1.42
3126	TANC1	0.03461	-1.39	3183	MRPL45	0.00091	-1.42
3127	JTB	0.00010	-1.39	3184	SLC37A4	0.00621	-1.42
3128	PRMT7	0.00380	-1.40	3185	LDLRAD3	0.00284	-1.42
3129	EPS8L1	0.00065	-1.40	3186	ZADH2	0.00763	-1.42
3130	TSEN54	0.00947	-1.40	3187	ATP5G1	0.00001	-1.43
3131	PACSIN3	0.00006	-1.40	3188	PIGV	0.00162	-1.43
3132	RNPEPL1	0.02843	-1.40	3189	RPL14	0.00000	-1.43
3133	FAM96B	0.00004	-1.40	3190	TMTC2	0.00802	-1.43
3134	AP1M2	0.00079	-1.40	3191	ELMO3	0.00003	-1.43
3135	DARS2	0.00008	-1.40	3192	G6PD	0.02936	-1.43
3136	PRKAG2	0.01307	-1.40	3193	RUVBL2	0.00405	-1.43
3137	SLC25A15	0.00053	-1.40	3194	BCKDHA	0.00104	-1.43
3138	TBCD	0.00009	-1.40	3195	DENND3	0.00061	-1.43
3139	ENKD1	0.00400	-1.40	3196	UQCRC2	0.00000	-1.43
3140	RPL29	0.00000	-1.40	3197	CEP72	0.00001	-1.43
3141	TAF4	0.01454	-1.40	3198	NUP37	0.00022	-1.43
3142	TOE1	0.00111	-1.40	3199	TXN	0.00669	-1.43
3143	ENOSF1	0.00000	-1.40	3200	LTA4H	0.01456	-1.43
3144	CDKN2C	0.00418	-1.40	3201	LOC554223	0.00269	-1.43
3145	ACAA1	0.00001	-1.41	3202	ELL3	0.00054	-1.44
3146	MRPL35	0.00068	-1.41	3203	SOD2	0.02729	-1.44
3147	KDSR	0.00287	-1.41	3204	RNPEP	0.00004	-1.44
3148	SLC48A1	0.00167	-1.41	3205	TALDO1	0.00009	-1.44
3149	GCN1	0.03436	-1.41	3206	FKBP4	0.00000	-1.44
3150	SNX5	0.00612	-1.41	3207	AFG3L2	0.00000	-1.44
3151	ARHGEF16	0.00038	-1.41	3208	NUDT19	0.01285	-1.44
3152	MTFMT	0.00989	-1.41	3209	HSPE1	0.00416	-1.44
3153	CRISPLD1	0.01398	-1.41	3210	FAM83A	0.00002	-1.44

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3211	IQSEC2	0.01956	-1.44	3268	LPCAT3	0.01175	-1.48
3212	PIH1D1	0.00009	-1.44	3269	RPS13	0.00000	-1.48
3213	ASCC1	0.00001	-1.44	3270	CCHCR1	0.00440	-1.48
3214	GFM1	0.04715	-1.45	3271	PCCB	0.00017	-1.48
3215	TSC22D2	0.00120	-1.45	3272	COQ4	0.00023	-1.48
3216	PDLIM2	0.01558	-1.45	3273	TMEM143	0.03069	-1.48
3217	SDHB	0.00000	-1.45	3274	IL10RB	0.00004	-1.48
3218	RPL21	0.00004	-1.45	3275	H2AFJ	0.00004	-1.48
3219	GPX4	0.00271	-1.45	3276	TYSND1	0.00069	-1.49
3220	TMEM179B	0.00615	-1.45	3277	WBSCR22	0.00000	-1.49
3221	GBAS	0.00052	-1.45	3278	SPAG16	0.01056	-1.49
3222	CCDC47	0.04777	-1.45	3279	PWWP2B	0.04210	-1.49
3223	HAGHL	0.00072	-1.45	3280	COL9A3	0.00412	-1.49
3224	RPS3A	0.00000	-1.45	3281	CEBPZOS	0.01540	-1.49
3225	PNPO	0.00000	-1.45	3282	COG7	0.00002	-1.49
3226	BTBD2	0.00303	-1.45	3283	HIST1H4C	0.00919	-1.49
3227	ANXA11	0.00000	-1.45	3284	RPS16	0.00779	-1.49
3228	CALHM2	0.00274	-1.45	3285	SLC25A10	0.00260	-1.49
3229	CCDC134	0.01394	-1.46	3286	C1orf116	0.00004	-1.49
3230	NACA	0.00020	-1.46	3287	NDUFA7	0.00009	-1.49
3231	RIBC2	0.01823	-1.46	3288	RPS10	0.00000	-1.49
3232	RPL30	0.00000	-1.46	3289	MID2	0.00044	-1.49
3233	WDR19	0.00166	-1.46	3290	LPCAT4	0.00117	-1.49
3234	RPL36	0.00001	-1.46	3291	NGEF	0.00087	-1.50
3235	R3HDM2	0.00368	-1.46	3292	STK24	0.00001	-1.50
3236	ISYNA1	0.02051	-1.46	3293	SELO	0.01018	-1.50
3237	C1orf210	0.00037	-1.46	3294	NUBPL	0.02088	-1.50
3238	NDUFS1	0.00001	-1.46	3295	SLC3A2	0.00003	-1.50
3239	RPL7A	0.00009	-1.46	3296	SIL1	0.00149	-1.50
3240	SEH1L	0.00161	-1.46	3297	MADD	0.00031	-1.50
3241	FOXRED1	0.00042	-1.46	3298	PCBD1	0.00001	-1.50
3242	RPS5	0.00033	-1.46	3299	ARID5B	0.02975	-1.50
3243	CHID1	0.00036	-1.46	3300	D2HGDH	0.00542	-1.50
3244	MAPK3	0.00010	-1.46	3301	METTL5	0.02464	-1.50
3245	CRELD2	0.00035	-1.46	3302	TCIRG1	0.00249	-1.50
3246	ASF1B	0.00008	-1.47	3303	NECAB3	0.02523	-1.50
3247	DGAT1	0.00193	-1.47	3304	RPL22	0.03948	-1.51
3248	RAD23A	0.00000	-1.47	3305	HAGH	0.00231	-1.51
3249	RPL18A	0.00003	-1.47	3306	RPP25L	0.00848	-1.51
3250	PITPNM1	0.01229	-1.47	3307	PTPRR	0.00024	-1.51
3251	PIGO	0.00106	-1.47	3308	DEGS1	0.00096	-1.51
3252	RBPMS	0.04372	-1.47	3309	CDH1	0.00016	-1.51
3253	PLCE1	0.00470	-1.47	3310	LRRC45	0.00122	-1.51
3254	NANS	0.00001	-1.47	3311	TMEM170A	0.00200	-1.51
3255	ERMARD	0.00459	-1.47	3312	MRPL24	0.00000	-1.51
3256	RPS25	0.00017	-1.47	3313	RMND5B	0.00002	-1.51
3257	PAN2	0.00780	-1.47	3314	HSPA9	0.00009	-1.51
3258	NAAA	0.00005	-1.47	3315	EEF1B2	0.00000	-1.52
3259	RPL8	0.00001	-1.47	3316	RNF103	0.01439	-1.52
3260	CD109	0.02362	-1.47	3317	SCRN2	0.00131	-1.52
3261	PXMP4	0.00005	-1.47	3318	GRTP1	0.00137	-1.52
3262	MAPK13	0.00000	-1.48	3319	SORD	0.00000	-1.52
3263	TTC19	0.00022	-1.48	3320	MTSS1L	0.00044	-1.52
3264	SLC25A39	0.00420	-1.48	3321	COA6	0.03875	-1.52
3265	TMX2	0.00001	-1.48	3322	ST3GAL1	0.00000	-1.52
3266	AIFM1	0.00835	-1.48	3323	ANKRD13C	0.00195	-1.52
3267	RAB25	0.00000	-1.48	3324	ATP5B	0.00000	-1.52

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3325	NQO2	0.00108	-1.53	3382	ZNF552	0.00246	-1.57
3326	LDHB	0.03348	-1.53	3383	ILVBL	0.00003	-1.57
3327	HIST1H4A	0.02177	-1.53	3384	PCDH1	0.00007	-1.57
3328	TACC2	0.00003	-1.53	3385	CYSRT1	0.00347	-1.57
3329	CLDN7	0.00000	-1.53	3386	CENPP	0.00912	-1.57
3330	CMBL	0.00963	-1.53	3387	IMMP2L	0.00488	-1.58
3331	TSEN2	0.03560	-1.53	3388	VAPA	0.00257	-1.58
3332	COQ6	0.00004	-1.53	3389	BAIAP2	0.00001	-1.58
3333	ARSJ	0.00419	-1.53	3390	HSD17B12	0.00849	-1.58
3334	SLC27A3	0.00012	-1.53	3391	PECR	0.00186	-1.58
3335	VSIG10	0.00094	-1.53	3392	LONP1	0.00011	-1.58
3336	XPNPEP3	0.00001	-1.54	3393	SLC27A5	0.00433	-1.58
3337	LGR4	0.00432	-1.54	3394	EBPL	0.00001	-1.59
3338	RAB11FIP1	0.00009	-1.54	3395	DLEU1	0.00540	-1.59
3339	C9orf142	0.00334	-1.54	3396	GSTO2	0.00000	-1.59
3340	METTL10	0.00579	-1.54	3397	DHRS13	0.01080	-1.59
3341	COQ2	0.01855	-1.54	3398	CASC8	0.00181	-1.59
3342	CAPS	0.01371	-1.54	3399	ARHGEF10L	0.00000	-1.59
3343	ABCD1	0.00188	-1.54	3400	ESRP2	0.00000	-1.59
3344	LDLRAP1	0.00002	-1.54	3401	LYAR	0.01286	-1.59
3345	RPLP2	0.00003	-1.54	3402	SH3GLB2	0.00010	-1.59
3346	RPS29	0.00000	-1.54	3403	BDH1	0.00005	-1.59
3347	PTPRU	0.00218	-1.54	3404	RELL1	0.00677	-1.59
3348	TMEM184A	0.00082	-1.55	3405	NDUFV2	0.00000	-1.59
3349	PGAP2	0.00011	-1.55	3406	NCAPD3	0.00001	-1.59
3350	DSG2	0.02474	-1.55	3407	SKP2	0.00000	-1.59
3351	PPT2	0.00000	-1.55	3408	ADGRE5	0.01860	-1.59
3352	PDSS2	0.00953	-1.55	3409	ABHD15	0.00000	-1.60
3353	RPPH1	0.00002	-1.55	3410	CASP7	0.00351	-1.60
3354	COMT	0.00002	-1.55	3411	SIGMAR1	0.00264	-1.60
3355	GAN	0.00620	-1.55	3412	CD9	0.00000	-1.60
3356	AKAP13	0.00078	-1.55	3413	C1QBP	0.00000	-1.60
3357	NR2F6	0.00574	-1.55	3414	GNAL	0.00085	-1.60
3358	TMEM246	0.00000	-1.55	3415	CHEK2	0.00000	-1.60
3359	SPATS2L	0.00119	-1.55	3416	TKT	0.00006	-1.60
3360	TST	0.00026	-1.56	3417	CLUH	0.00020	-1.60
3361	RPL23A	0.00000	-1.56	3418	TTC30B	0.00524	-1.61
3362	GDE1	0.00000	-1.56	3419	SNORA70	0.00994	-1.61
3363	DANCR	0.00000	-1.56	3420	SMIM22	0.00002	-1.61
3364	SHTN1	0.02832	-1.56	3421	KRT8	0.00000	-1.61
3365	FAM53B	0.00002	-1.56	3422	FGFR3	0.02268	-1.61
3366	RCC1	0.00000	-1.56	3423	ZNF431	0.03106	-1.61
3367	ERVMER34-1	0.04671	-1.56	3424	MSH5	0.01909	-1.61
3368	PPP1R15B	0.00456	-1.56	3425	JAK3	0.03382	-1.61
3369	ZCCHC2	0.02275	-1.56	3426	FAM46B	0.00021	-1.61
3370	ATP1A1	0.00001	-1.56	3427	TMEM205	0.00000	-1.61
3371	CCDC51	0.00000	-1.56	3428	EXOSC5	0.00060	-1.61
3372	SLC1A5	0.02205	-1.56	3429	HIST1H2AB	0.02061	-1.61
3373	ATP5D	0.01226	-1.56	3430	SCARNA17	0.00069	-1.62
3374	EPHA1	0.00025	-1.56	3431	PAIP2B	0.01828	-1.62
3375	TCF25	0.00001	-1.56	3432	LRPPRC	0.01741	-1.62
3376	CRB3	0.03630	-1.57	3433	PHYH	0.00663	-1.62
3377	SLC37A1	0.00000	-1.57	3434	GPHN	0.00093	-1.62
3378	ESRRA	0.00000	-1.57	3435	SARS2	0.00007	-1.62
3379	PNPLA2	0.00141	-1.57	3436	MSLN	0.00044	-1.62
3380	NAB1	0.00000	-1.57	3437	DNAJC19	0.01239	-1.62
3381	GSTCD	0.01505	-1.57	3438	KLF3	0.00635	-1.62

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3439	HIST1H2BI	0.03207	-1.62	3496	PRDX6	0.00000	-1.67
3440	NDUFC1	0.00000	-1.62	3497	PSME1	0.00000	-1.67
3441	MAN2A2	0.00095	-1.63	3498	TMEM147	0.00000	-1.67
3442	DLAT	0.00691	-1.63	3499	DENND2D	0.00000	-1.68
3443	CXCL16	0.00013	-1.63	3500	ESYT2	0.00001	-1.68
3444	FUK	0.00512	-1.63	3501	PHF10	0.00140	-1.68
3445	HIBADH	0.00210	-1.63	3502	CSGALNACT1	0.00031	-1.68
3446	MEIS2	0.00784	-1.63	3503	AK4	0.00020	-1.68
3447	EFNA5	0.00767	-1.63	3504	MIF4GD	0.00016	-1.68
3448	SH3BP2	0.00003	-1.63	3505	LRP8	0.02610	-1.68
3449	FBXW9	0.00019	-1.63	3506	ABLIM3	0.00038	-1.69
3450	TMC6	0.00000	-1.63	3507	LOC100288181	0.00000	-1.69
3451	HDAC4	0.00018	-1.63	3508	SNORA24	0.00953	-1.69
3452	NAA25	0.00463	-1.64	3509	PXN-AS1	0.02838	-1.69
3453	COASY	0.00000	-1.64	3510	CAMKMT	0.00086	-1.69
3454	DAPK1	0.01066	-1.64	3511	DSC3	0.01399	-1.69
3455	CAMK2G	0.00000	-1.64	3512	ZFP36	0.00002	-1.69
3456	PVRL1	0.01481	-1.64	3513	RALGDS	0.00002	-1.70
3457	ADCY6	0.00118	-1.64	3514	ACSF2	0.00021	-1.70
3458	DNM2	0.00000	-1.64	3515	TOP2A	0.02859	-1.70
3459	CBR4	0.00884	-1.64	3516	ADRBK1	0.00001	-1.70
3460	IARS2	0.00211	-1.64	3517	ATXN10	0.00000	-1.70
3461	RAB15	0.00006	-1.64	3518	GALM	0.00427	-1.70
3462	USP24	0.03602	-1.64	3519	TLCD1	0.00206	-1.70
3463	C21orf59	0.00358	-1.64	3520	NBEAL2	0.04412	-1.70
3464	PARS2	0.04761	-1.65	3521	MUTYH	0.00106	-1.70
3465	SCARNA22	0.00247	-1.65	3522	GK5	0.03449	-1.70
3466	NDUFAF2	0.00017	-1.65	3523	LOC113230	0.04795	-1.70
3467	MRPL39	0.00004	-1.65	3524	TKFC	0.00017	-1.70
3468	ATE1	0.00447	-1.65	3525	RNY1	0.00419	-1.71
3469	SCARNA13	0.00010	-1.65	3526	PLEKHA6	0.00003	-1.71
3470	CELSR1	0.01012	-1.65	3527	FARP2	0.00005	-1.71
3471	ADAM15	0.00000	-1.65	3528	TMEM241	0.00000	-1.71
3472	HDDC3	0.00224	-1.65	3529	RFFL	0.00000	-1.71
3473	SCARNA6	0.00047	-1.65	3530	PPARG	0.00000	-1.71
3474	SCARNA2	0.00049	-1.65	3531	OXCT1	0.02872	-1.71
3475	ADAT2	0.00584	-1.66	3532	MYH14	0.00159	-1.71
3476	HSBP1L1	0.00000	-1.66	3533	UCA1	0.00000	-1.71
3477	LRRC1	0.00933	-1.66	3534	TPCN1	0.00000	-1.71
3478	TXNRD1	0.00198	-1.66	3535	BRI3BP	0.00011	-1.72
3479	ATP7B	0.00259	-1.66	3536	TMEM141	0.00000	-1.72
3480	TAOK3	0.00161	-1.66	3537	NMI	0.04951	-1.73
3481	HEXB	0.00001	-1.66	3538	ACSL3	0.00117	-1.73
3482	RPUSD3	0.00009	-1.66	3539	MAOA	0.02625	-1.73
3483	BSCL2	0.00002	-1.66	3540	SLC17A5	0.00125	-1.73
3484	TIGD2	0.01395	-1.67	3541	FZD5	0.00199	-1.73
3485	B9D2	0.03499	-1.67	3542	SFXN4	0.00001	-1.73
3486	VARS	0.01234	-1.67	3543	ETFB	0.00001	-1.73
3487	ZDHHC23	0.00181	-1.67	3544	HOXA1	0.00010	-1.74
3488	TMEM99	0.00002	-1.67	3545	NR6A1	0.01041	-1.74
3489	IQCH	0.00435	-1.67	3546	PHKA1	0.00000	-1.74
3490	SLC22A18	0.00004	-1.67	3547	PLD1	0.01000	-1.74
3491	L2HGDH	0.02069	-1.67	3548	PIAS1	0.00206	-1.74
3492	PIGP	0.00041	-1.67	3549	NDC1	0.00001	-1.74
3493	RIPK4	0.00000	-1.67	3550	LRP11	0.00887	-1.74
3494	RPL36A	0.00000	-1.67	3551	SEMA4B	0.00000	-1.74
3495	MFSD6	0.00006	-1.67	3552	COQ3	0.00000	-1.74

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3553	NDRG1	0.00890	-1.74	3610	DNAAF3	0.00497	-1.80
3554	IBTK	0.01112	-1.74	3611	HIST1H2BE	0.00053	-1.80
3555	CCDC6	0.00019	-1.75	3612	PIM3	0.00012	-1.80
3556	AIM1	0.00366	-1.75	3613	ARHGAP42	0.04018	-1.80
3557	YBX3	0.00000	-1.75	3614	ACAT1	0.00235	-1.80
3558	OIP5	0.00105	-1.75	3615	CHAC2	0.01137	-1.81
3559	SYPL1	0.01799	-1.75	3616	CCDC106	0.00066	-1.81
3560	NEK2	0.02386	-1.75	3617	TCAF2	0.00352	-1.81
3561	ELMSAN1	0.00009	-1.75	3618	HR	0.00049	-1.81
3562	ECHDC3	0.00000	-1.75	3619	BCKDHB	0.00000	-1.81
3563	LOC154761	0.00226	-1.75	3620	EGFR	0.00002	-1.81
3564	AKAP1	0.00000	-1.75	3621	FAM43A	0.03370	-1.81
3565	KLF9	0.00030	-1.75	3622	SFXN2	0.00002	-1.81
3566	ATP1B3	0.00000	-1.75	3623	KIAA1804	0.00078	-1.81
3567	KCTD3	0.01343	-1.75	3624	BLVRA	0.00000	-1.82
3568	DAG1	0.00346	-1.75	3625	POLD2	0.00000	-1.82
3569	PTPRS	0.00612	-1.76	3626	RABGGTA	0.00000	-1.82
3570	OSBPL5	0.00000	-1.76	3627	OSTF1	0.00002	-1.82
3571	TMEM182	0.01058	-1.76	3628	PPP1R3G	0.01687	-1.82
3572	APRT	0.00000	-1.76	3629	PIK3R1	0.00102	-1.82
3573	RETSAT	0.00000	-1.76	3630	CPNE7	0.00000	-1.82
3574	SYTL1	0.01973	-1.76	3631	TSPAN15	0.00000	-1.82
3575	NSUN7	0.03244	-1.76	3632	DPH6	0.00013	-1.82
3576	FRAT2	0.00000	-1.76	3633	NABP1	0.03161	-1.82
3577	FTH1	0.00000	-1.76	3634	MAP7	0.00005	-1.83
3578	ZMYND8	0.00000	-1.77	3635	RHPN1	0.00073	-1.83
3579	DLGAP1-AS1	0.00043	-1.77	3636	CBX7	0.00227	-1.83
3580	IRAK1BP1	0.01605	-1.77	3637	LAMA5	0.00325	-1.83
3581	DENND4C	0.04203	-1.77	3638	CHP1	0.00000	-1.83
3582	RPARP-AS1	0.00717	-1.77	3639	CHKA	0.00000	-1.83
3583	PLEKHF1	0.00027	-1.77	3640	C3	0.00025	-1.83
3584	PRSS21	0.00434	-1.77	3641	UBXN8	0.00011	-1.83
3585	NT5C3A	0.03836	-1.78	3642	IMPDH2	0.00000	-1.83
3586	RNF141	0.00373	-1.78	3643	HAUS4	0.00000	-1.83
3587	IPO4	0.00417	-1.78	3644	HIST1H4D	0.00019	-1.83
3588	PGAP3	0.01057	-1.78	3645	TCEAL1	0.00036	-1.83
3589	CDC42SE2	0.00038	-1.78	3646	AHNAK	0.00397	-1.84
3590	SNORA67	0.01817	-1.78	3647	TFRC	0.00002	-1.84
3591	DHTKD1	0.00000	-1.78	3648	ABCB6	0.00000	-1.84
3592	FDX1	0.01196	-1.78	3649	SNX2	0.04582	-1.84
3593	NUSAP1	0.00088	-1.78	3650	SEMA3F	0.00432	-1.84
3594	OSR2	0.00042	-1.78	3651	WDR34	0.00001	-1.84
3595	PPT1	0.00000	-1.78	3652	DTWD2	0.00005	-1.84
3596	TMEM165	0.00000	-1.78	3653	MPZL3	0.00000	-1.84
3597	CPT2	0.00000	-1.78	3654	HIST1H2BB	0.00505	-1.85
3598	MTL5	0.00000	-1.79	3655	NMU	0.00002	-1.85
3599	GNE	0.00089	-1.79	3656	CHMP2B	0.02563	-1.85
3600	ICK	0.01626	-1.79	3657	HOXA5	0.00028	-1.85
3601	MAP3K1	0.00051	-1.79	3658	C11orf71	0.00540	-1.85
3602	ECHDC2	0.01705	-1.79	3659	ACTR3C	0.00980	-1.85
3603	MYO1B	0.01510	-1.79	3660	FAM83H-AS1	0.00046	-1.85
3604	CRYZL1	0.00000	-1.79	3661	DTX4	0.00000	-1.85
3605	MCCC1	0.00005	-1.80	3662	SPA17	0.00002	-1.85
3606	TTLL12	0.00002	-1.80	3663	ROR1	0.00001	-1.85
3607	SLC22A5	0.00000	-1.80	3664	CYB5A	0.00000	-1.85
3608	GPI	0.00001	-1.80	3665	TRAPPC9	0.00000	-1.86
3609	CRACR2B	0.03747	-1.80	3666	HOXA9	0.00000	-1.86

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3667	LINC01588	0.00026	-1.86	3724	OMA1	0.00332	-1.91
3668	PTGR2	0.00184	-1.86	3725	ANXA4	0.00125	-1.92
3669	CALML4	0.00266	-1.86	3726	SSH3	0.00000	-1.92
3670	PAG1	0.01381	-1.86	3727	LOC101927811	0.02121	-1.92
3671	MYO18A	0.00001	-1.86	3728	WDR90	0.00024	-1.92
3672	SLC25A19	0.00003	-1.86	3729	SH3BP5-AS1	0.00146	-1.92
3673	FAM86C2P	0.02767	-1.86	3730	***MARCH1***	0.00000	-1.92
3674	APEH	0.00000	-1.86	3731	ACYP1	0.01743	-1.92
3675	NEIL1	0.03239	-1.86	3732	GCLC	0.00191	-1.92
3676	TRERF1	0.00000	-1.87	3733	C10orf54	0.00000	-1.92
3677	ELOVL6	0.00608	-1.87	3734	TERC	0.00001	-1.93
3678	GCHFR	0.00001	-1.87	3735	CBLC	0.00001	-1.93
3679	LEO1	0.01722	-1.87	3736	ST6GALNAC2	0.01679	-1.93
3680	UBXN11	0.00062	-1.87	3737	SMPD2	0.00004	-1.93
3681	ABHD12	0.00001	-1.87	3738	CEACAM1	0.00294	-1.93
3682	C9orf116	0.00000	-1.87	3739	CRABP2	0.00002	-1.93
3683	MND1	0.01690	-1.87	3740	SAMD12	0.01401	-1.93
3684	CA2	0.00000	-1.87	3741	GLUD1	0.00000	-1.93
3685	ACBD4	0.00592	-1.88	3742	PARD6A	0.00000	-1.93
3686	E2F8	0.00033	-1.88	3743	ETFDH	0.00013	-1.93
3687	ARHGEF4	0.00010	-1.88	3744	FAM98C	0.00034	-1.93
3688	MIER3	0.02783	-1.88	3745	SNORA43	0.00009	-1.93
3689	BHLHE41	0.00453	-1.88	3746	KLK6	0.01085	-1.93
3690	SFI1	0.00215	-1.88	3747	GPD2	0.00232	-1.94
3691	EMP2	0.00000	-1.88	3748	ADAMTSL3	0.00458	-1.94
3692	SEMA5A	0.03211	-1.88	3749	PPTC7	0.00246	-1.94
3693	LAMA4	0.00000	-1.88	3750	FER1L4	0.00059	-1.94
3694	SNORA17	0.01517	-1.88	3751	ZBTB7B	0.00017	-1.94
3695	ARRB1	0.00000	-1.89	3752	FAM63B	0.01381	-1.94
3696	SNORA84	0.00060	-1.89	3753	CREG1	0.00021	-1.95
3697	SDHA	0.00000	-1.89	3754	DHRS11	0.00000	-1.95
3698	HIBCH	0.04692	-1.89	3755	PPP1R3D	0.00000	-1.95
3699	NTHL1	0.00000	-1.89	3756	CARS2	0.00000	-1.95
3700	C15orf62	0.00114	-1.89	3757	MIPEP	0.00000	-1.95
3701	FAM49A	0.03975	-1.89	3758	CLDN4	0.00000	-1.96
3702	KAZN	0.01779	-1.89	3759	GATA3	0.00035	-1.96
3703	SLC52A3	0.00017	-1.89	3760	ACSS2	0.00000	-1.96
3704	PTPRH	0.00000	-1.89	3761	SYK	0.00000	-1.96
3705	ECI1	0.00000	-1.89	3762	GPSM2	0.00060	-1.96
3706	TMEM30B	0.00030	-1.89	3763	CAMK2N1	0.00000	-1.96
3707	SMCHD1	0.02803	-1.89	3764	NANOS1	0.00007	-1.97
3708	RNASET2	0.00000	-1.89	3765	LMTK2	0.00000	-1.97
3709	PLCD1	0.00067	-1.90	3766	ERICH5	0.04034	-1.97
3710	PSMB10	0.00000	-1.90	3767	SCARNA10	0.00001	-1.97
3711	HSDL2	0.00001	-1.90	3768	FAM213A	0.00002	-1.97
3712	ZNF341	0.00028	-1.90	3769	IL17RB	0.00065	-1.97
3713	GALNT13	0.00190	-1.90	3770	IL15RA	0.00000	-1.97
3714	WFDC2	0.00000	-1.90	3771	SOX13	0.00000	-1.97
3715	HCAR2	0.04179	-1.90	3772	PROSER2-AS1	0.00000	-1.97
3716	BCAT2	0.00001	-1.90	3773	SNORA71D	0.00002	-1.98
3717	CPD	0.00048	-1.90	3774	MOCOS	0.00001	-1.98
3718	PTPN13	0.00170	-1.91	3775	ASCL2	0.01032	-1.98
3719	PLEKHA7	0.00000	-1.91	3776	NOV	0.01626	-1.98
3720	PROM2	0.00001	-1.91	3777	PPP1R3C	0.00014	-1.98
3721	MCCC2	0.00001	-1.91	3778	SAPCD2	0.00061	-1.98
3722	TCTN1	0.00003	-1.91	3779	TSPAN1	0.00000	-1.98
3723	NUDT16P1	0.00013	-1.91	3780	HS6ST1	0.00000	-1.99

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3781	RDM1	0.00206	-1.99	3838	***ARRDC1***	0.00000	-2.07
3782	SLC35E4	0.00048	-1.99	3839	***MARVELD2***	0.00003	-2.08
3783	SLK	0.01402	-1.99	3840	***IL18***	0.00577	-2.08
3784	HIST1H2BM	0.01678	-1.99	3841	***GGCT***	0.00018	-2.08
3785	PAXIP1-AS1	0.02080	-2.00	3842	OGDHL	0.04495	-2.08
3786	PLEKHG6	0.00000	-2.00	3843	FOXD2	0.02476	-2.08
3787	SLC7A5	0.00000	-2.00	3844	***TMEM38B***	0.00111	-2.08
3788	***PIGN***	0.00004	-2.00	3845	***CCDC57***	0.00001	-2.08
3789	***IDH3A***	0.00000	-2.00	3846	***COMTD1***	0.00020	-2.08
3790	LOC101927181	0.00781	-2.00	3847	***C6orf132***	0.00000	-2.09
3791	***ASAH1***	0.00008	-2.00	3848	***F12***	0.00018	-2.09
3792	LYPD5	0.00002	-2.00	3849	***LY6K***	0.00000	-2.09
3793	***FAR1***	0.00519	-2.00	3850	***PDE9A***	0.00002	-2.09
3794	***CIPC***	0.00013	-2.01	3851	***SNORA10***	0.00253	-2.09
3795	COBLL1	0.00650	-2.01	3852	***SIAE***	0.00015	-2.09
3796	***SUCLG2***	0.00050	-2.01	3853	***NR3C1***	0.00001	-2.10
3797	***UPK3B***	0.01175	-2.01	3854	***PPA1***	0.00253	-2.10
3798	***LOC100505666***	0.00432	-2.01	3855	***LYRM7***	0.00078	-2.10
3799	***MKNK2***	0.00001	-2.01	3856	SPEF2	0.02247	-2.10
3800	***SNORA71B***	0.01057	-2.01	3857	***CRYBG3***	0.00446	-2.11
3801	***NAPRT***	0.00001	-2.02	3858	***ABHD17C***	0.00000	-2.11
3802	***PPP1R12B***	0.00000	-2.02	3859	***EPN3***	0.00004	-2.11
3803	***ACY1***	0.00001	-2.02	3860	***SNORA47***	0.01824	-2.11
3804	***RAC3***	0.00001	-2.02	3861	***PPP2R5A***	0.00017	-2.11
3805	***OPLAH***	0.00000	-2.02	3862	***PKN2***	0.00059	-2.11
3806	***EVPL***	0.00008	-2.02	3863	***TLE2***	0.00000	-2.12
3807	***FAM111B***	0.00254	-2.02	3864	***KLHDC4***	0.00000	-2.12
3808	***PARD6B***	0.00704	-2.03	3865	***LRRC16A***	0.00000	-2.12
3809	ZFYVE28	0.00532	-2.03	3866	***TMEM53***	0.00004	-2.12
3810	***TMEM106B***	0.01878	-2.03	3867	***HDHD3***	0.00001	-2.12
3811	RIMS4	0.00030	-2.03	3868	***RXRA***	0.00000	-2.12
3812	***GALNT7***	0.00004	-2.03	3869	***MARCH1***	0.00451	-2.12
3813	***HINT3***	0.02303	-2.03	3870	MIPOL1	0.00002	-2.12
3814	***ZNF488***	0.01148	-2.03	3871	***ASAP3***	0.00000	-2.12
3815	***CASP4***	0.00014	-2.03	3872	***FAM86B1***	0.00239	-2.12
3816	LOC646762	0.00017	-2.03	3873	***SNORA74B***	0.00115	-2.13
3817	***MACC1***	0.00576	-2.04	3874	***BZW2***	0.00000	-2.13
3818	***ALDH3A2***	0.00001	-2.04	3875	***NALCN***	0.00000	-2.13
3819	***LINC01550***	0.00173	-2.04	3876	***PITX1***	0.00001	-2.13
3820	***RBP4***	0.00000	-2.04	3877	***KCNK1***	0.00001	-2.13
3821	***BLVRB***	0.00000	-2.04	3878	***CFAP36***	0.00164	-2.13
3822	***LGALS3***	0.00000	-2.04	3879	***TRIM2***	0.03377	-2.14
3823	***MORN1***	0.00014	-2.04	3880	***CASKIN2***	0.00001	-2.14
3824	***ATP2B4***	0.00037	-2.04	3881	***EGFR-AS1***	0.00000	-2.14
3825	***RHOV***	0.00028	-2.04	3882	***HIST1H1A***	0.00083	-2.14
3826	BANK1	0.00682	-2.04	3883	***SNORA44***	0.00942	-2.14
3827	***FAM84B***	0.00068	-2.05	3884	***TOB1***	0.00360	-2.15
3828	***STAC***	0.00000	-2.05	3885	***GMDS***	0.00000	-2.15
3829	***DCPS***	0.00000	-2.05	3886	***IRX2***	0.00000	-2.15
3830	***UHRF1BP1***	0.00001	-2.06	3887	***TESK2***	0.00001	-2.15
3831	TRIML2	0.03072	-2.06	3888	***THRIL***	0.00000	-2.16
3832	***FGD3***	0.00001	-2.06	3889	***UNC93B1***	0.00000	-2.16
3833	***ACADS***	0.00004	-2.06	3890	***KLHL2***	0.00001	-2.16
3834	HES5	0.01396	-2.06	3891	***RNF149***	0.00000	-2.16
3835	***DBP***	0.00309	-2.06	3892	***MAP3K8***	0.00230	-2.16
3836	***LACTB2***	0.00007	-2.07	3893	***GMPR***	0.00000	-2.16
3837	***TMCO6***	0.00000	-2.07	3894	***PITPNC1***	0.00000	-2.16

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
3895	***IFI30***	0.00000	-2.17	3952	***B4GALT4***	0.00000	-2.26
3896	***ABCC3***	0.00000	-2.17	3953	GRHL3	0.00005	-2.27
3897	***NFIA***	0.00006	-2.17	3954	***SERINC5***	0.00120	-2.27
3898	***SLC45A4***	0.00000	-2.17	3955	IQCD	0.00006	-2.27
3899	***PXYLP1***	0.00000	-2.17	3956	ZSCAN12P1	0.00262	-2.28
3900	***AGMAT***	0.00000	-2.18	3957	***ZNRF2***	0.00000	-2.28
3901	***ABCB10***	0.00014	-2.18	3958	CPEB3	0.00001	-2.28
3902	***ANKRD33B***	0.00005	-2.18	3959	***TBX6***	0.00417	-2.29
3903	***C1GALT1***	0.00315	-2.18	3960	***RASAL1***	0.00000	-2.29
3904	***LYNX1***	0.00002	-2.18	3961	***AP1AR***	0.00928	-2.29
3905	NMNAT2	0.00956	-2.18	3962	***HPCAL1***	0.00000	-2.29
3906	***TTC39A***	0.00000	-2.18	3963	KCNQ1OT1	0.00461	-2.29
3907	***HADH***	0.00000	-2.18	3964	***SNORA71A***	0.00007	-2.29
3908	B4GALNT3	0.00191	-2.18	3965	***SNORA23***	0.00188	-2.29
3909	***UPK3BL***	0.00031	-2.18	3966	***RHPN2***	0.00000	-2.29
3910	***SNORD89***	0.02726	-2.19	3967	***PLAC8***	0.00000	-2.29
3911	***PER2***	0.00001	-2.19	3968	***TMEM135***	0.00091	-2.29
3912	***CNNM4***	0.00001	-2.19	3969	***DMPK***	0.00030	-2.30
3913	INPP5J	0.00779	-2.19	3970	***FAM151A***	0.00000	-2.30
3914	***RARG***	0.00000	-2.19	3971	***EGLN3***	0.00000	-2.30
3915	***STOM***	0.00001	-2.19	3972	BCL2L10	0.03545	-2.30
3916	PADI1	0.00114	-2.19	3973	WNT2B	0.00184	-2.30
3917	***SPTSSA***	0.01138	-2.19	3974	***PCGF5***	0.00001	-2.30
3918	MCMDC2	0.00115	-2.21	3975	***PBK***	0.00084	-2.31
3919	***LMTK3***	0.00000	-2.21	3976	***OLMALINC***	0.00003	-2.31
3920	***SNORA38***	0.00464	-2.21	3977	***ACOT11***	0.00000	-2.31
3921	GPR78	0.00063	-2.21	3978	***HID1***	0.00000	-2.31
3922	***GCSH***	0.00026	-2.21	3979	TTC6	0.00416	-2.32
3923	RNFT2	0.03856	-2.21	3980	***KTN1-AS1***	0.00058	-2.32
3924	***HSPB11***	0.00001	-2.22	3981	***LINC00707***	0.00001	-2.32
3925	***FAM8A1***	0.00410	-2.22	3982	CHN2	0.00013	-2.34
3926	ADGRV1	0.00018	-2.22	3983	***RMND5A***	0.00003	-2.34
3927	LRRC8B	0.00010	-2.22	3984	***VGLL1***	0.00000	-2.34
3928	NAALADL2	0.01713	-2.22	3985	***ARHGAP26***	0.00000	-2.34
3929	WDR31	0.00080	-2.22	3986	***SCARNA16***	0.02443	-2.34
3930	***CCNO***	0.00378	-2.22	3987	***ACADM***	0.01009	-2.35
3931	***PDIK1L***	0.00005	-2.22	3988	SEMA3B	0.00000	-2.36
3932	***SCARNA9L***	0.00006	-2.22	3989	***B3GNT7***	0.00000	-2.36
3933	***TFAP2C***	0.00000	-2.22	3990	***POR***	0.00000	-2.36
3934	***RARRES3***	0.00002	-2.23	3991	***SNORA71C***	0.00096	-2.37
3935	***SNORA68***	0.00042	-2.23	3992	SRGAP3	0.00133	-2.37
3936	***KRT19***	0.00000	-2.23	3993	ILDR1	0.00282	-2.37
3937	***TRAFD1***	0.00014	-2.24	3994	***NUDT8***	0.00000	-2.37
3938	TMEM61	0.03938	-2.24	3995	***IFIT2***	0.00611	-2.37
3939	***CDC42EP4***	0.00000	-2.25	3996	***UPK2***	0.01491	-2.37
3940	***SULT2B1***	0.00000	-2.25	3997	***TMEM144***	0.00000	-2.37
3941	***DSG3***	0.00787	-2.25	3998	***EEF2K***	0.00000	-2.38
3942	LOC100294362	0.00006	-2.25	3999	***TFPI***	0.00003	-2.39
3943	***SNORD3A***	0.00078	-2.25	4000	***CES2***	0.00000	-2.39
3944	***SREBF1***	0.00000	-2.25	4001	***PADI2***	0.00917	-2.39
3945	***RNF144B***	0.03785	-2.25	4002	HLA-DMA	0.00532	-2.39
3946	***SYNE2***	0.00003	-2.25	4003	***PDLIM1***	0.00000	-2.40
3947	STC1	0.01679	-2.25	4004	***PDE4D***	0.00068	-2.40
3948	***QTRT1***	0.00009	-2.26	4005	***ARHGDIB***	0.00000	-2.40
3949	***NFIB***	0.00040	-2.26	4006	***SNORA26***	0.00045	-2.40
3950	ATP7A	0.00045	-2.26	4007	***SLC29A2***	0.00000	-2.40
3951	***ACSL1***	0.00000	-2.26	4008	***MACROD1***	0.00008	-2.40

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
4009	***HPSE***	0.00000	-2.41	4066	***BTG2***	0.01781	-2.58
4010	***FAM102A***	0.00000	-2.41	4067	***FAM102B***	0.02333	-2.58
4011	***DPP7***	0.00000	-2.41	4068	***RHOBTB3***	0.00000	-2.59
4012	***HTATIP2***	0.00000	-2.41	4069	***B3GNT3***	0.00000	-2.59
4013	***AVPI1***	0.00001	-2.41	4070	***KRTCAP3***	0.00000	-2.59
4014	ENPP4	0.00310	-2.42	4071	***TCEA3***	0.00000	-2.59
4015	***TMEM123***	0.01493	-2.42	4072	SEMA3E	0.00003	-2.59
4016	***SKAP2***	0.01917	-2.42	4073	***LMO7***	0.00003	-2.59
4017	***PEPD***	0.00000	-2.42	4074	***TRAP1***	0.00000	-2.59
4018	***SNORA45A***	0.00095	-2.42	4075	***SLC1A6***	0.00000	-2.60
4019	***PRMT3***	0.00000	-2.42	4076	***PIK3C2B***	0.00000	-2.60
4020	***NTN4***	0.00011	-2.42	4077	***FAM195A***	0.00000	-2.61
4021	***SLC16A14***	0.00000	-2.43	4078	***NPW***	0.02834	-2.61
4022	***CPT1A***	0.00000	-2.43	4079	***LIPE***	0.00000	-2.61
4023	STS	0.00509	-2.43	4080	***IFITM2***	0.00000	-2.61
4024	***FOXA1***	0.00000	-2.45	4081	***SNORA64***	0.00098	-2.62
4025	SPON1	0.00401	-2.45	4082	***PITPNM3***	0.00000	-2.62
4026	***SPTLC3***	0.00054	-2.45	4083	***ITGB8***	0.00028	-2.63
4027	***RHOBTB2***	0.00000	-2.45	4084	***FRMD4B***	0.00139	-2.63
4028	CDKL1	0.00022	-2.45	4085	LOC101927954	0.00051	-2.64
4029	***LIPH***	0.00000	-2.46	4086	***DAPK2***	0.00018	-2.64
4030	***GRAMD4***	0.00000	-2.46	4087	***ABCC5***	0.00000	-2.64
4031	***ANO1***	0.00000	-2.46	4088	***FGFBP1***	0.00000	-2.65
4032	PRTG	0.02292	-2.47	4089	***RASEF***	0.00001	-2.65
4033	***SNORA52***	0.01994	-2.47	4090	***KIAA0040***	0.00002	-2.65
4034	STARD8	0.00012	-2.47	4091	***FOXC1***	0.00001	-2.66
4035	PPP1R9A	0.00672	-2.49	4092	***LRRC8D***	0.00001	-2.66
4036	STAT4	0.00272	-2.49	4093	***TNIK***	0.00000	-2.67
4037	***EHF***	0.00004	-2.49	4094	***KCNS1***	0.00000	-2.67
4038	***NRP2***	0.00002	-2.49	4095	MGAT4A	0.00050	-2.67
4039	***MATN2***	0.00004	-2.49	4096	***MEGF6***	0.00087	-2.67
4040	***EML2***	0.00000	-2.50	4097	***ANKRD22***	0.00008	-2.68
4041	***KLHDC9***	0.00001	-2.50	4098	***ARHGEF3***	0.00000	-2.68
4042	RFESD	0.01286	-2.50	4099	***RBM47***	0.00000	-2.69
4043	***ERP27***	0.00001	-2.50	4100	***C5orf38***	0.00001	-2.69
4044	***WDR89***	0.00006	-2.50	4101	***CMTM4***	0.00000	-2.70
4045	***HK2***	0.00000	-2.51	4102	***ID3***	0.00000	-2.71
4046	***GEMIN8P4***	0.00000	-2.51	4103	WNT3A	0.03229	-2.71
4047	***CEBPD***	0.00022	-2.52	4104	TEX15	0.00842	-2.71
4048	***ZNF652***	0.00117	-2.52	4105	***ATP8B1***	0.00007	-2.71
4049	***EXPH5***	0.00011	-2.52	4106	***APOL6***	0.00002	-2.71
4050	***FNBP1***	0.00000	-2.52	4107	***EAF2***	0.00447	-2.71
4051	KLHL29	0.00000	-2.52	4108	***HOOK1***	0.00124	-2.72
4052	RBPMS-AS1	0.01187	-2.53	4109	***LLGL2***	0.00000	-2.72
4053	PA2G4P4	0.00429	-2.53	4110	***KIZ***	0.00012	-2.73
4054	***DCXR***	0.00000	-2.54	4111	***VAV3***	0.00006	-2.73
4055	***AHR***	0.00169	-2.54	4112	DNAH11	0.00000	-2.73
4056	***TNFAIP2***	0.00000	-2.54	4113	***OAF***	0.00000	-2.73
4057	***RALGPS2***	0.00003	-2.55	4114	***RNLS***	0.00000	-2.73
4058	***LRP5***	0.00000	-2.55	4115	***EMP1***	0.00000	-2.74
4059	CEBPA	0.00523	-2.56	4116	***SMAD6***	0.00000	-2.74
4060	***CHPT1***	0.00000	-2.56	4117	***SNORA21***	0.01380	-2.75
4061	***KIF13B***	0.00000	-2.56	4118	***CTH***	0.02133	-2.76
4062	***PLCD3***	0.00000	-2.56	4119	***SPATA13***	0.00000	-2.77
4063	***TTC22***	0.00002	-2.57	4120	***SULT1A1***	0.00000	-2.77
4064	***PIR***	0.00001	-2.57	4121	***ERMP1***	0.00000	-2.77
4065	***PLIN2***	0.00000	-2.58	4122	***CLMN***	0.00000	-2.78

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
4123	***SH3YL1***	0.00004	-2.79	4180	***PLA2G16***	0.00000	-3.03
4124	GPT	0.01312	-2.79	4181	***RAPGEF5***	0.00000	-3.03
4125	CCSER1	0.00001	-2.80	4182	***ADORA2B***	0.00000	-3.03
4126	***PTGER4***	0.00460	-2.80	4183	***CSPG4***	0.00003	-3.03
4127	***DPYD***	0.00000	-2.80	4184	***SYNE4***	0.00000	-3.04
4128	***LYPD6B***	0.00000	-2.81	4185	***TMPRSS2***	0.00000	-3.04
4129	***REEP6***	0.00003	-2.81	4186	***CABLES1***	0.00000	-3.04
4130	***KIAA1958***	0.00000	-2.81	4187	***PBX1***	0.00000	-3.05
4131	***SHMT1***	0.00000	-2.82	4188	ITPR1	0.00002	-3.05
4132	WNK2	0.00028	-2.82	4189	***AP1S3***	0.00000	-3.06
4133	***ALDH1A3***	0.00000	-2.82	4190	***PRRG4***	0.00000	-3.06
4134	QRFPR	0.00447	-2.82	4191	***MGST1***	0.00000	-3.06
4135	***NUP210***	0.00001	-2.83	4192	***SCNN1A***	0.00000	-3.07
4136	***KAT2B***	0.00002	-2.83	4193	***LOC100130987***	0.00001	-3.08
4137	***FAM46A***	0.00000	-2.84	4194	***MBOAT1***	0.00000	-3.08
4138	***LPIN1***	0.00000	-2.84	4195	***DLK2***	0.00000	-3.09
4139	***DDIT4***	0.00000	-2.85	4196	***OCLN***	0.00000	-3.10
4140	CASQ2	0.02123	-2.85	4197	***TFPI2***	0.00000	-3.10
4141	***RHOU***	0.00543	-2.86	4198	FTCDNL1	0.02232	-3.10
4142	***A4GALT***	0.00000	-2.86	4199	***RAB20***	0.00000	-3.11
4143	***RAB27B***	0.00326	-2.86	4200	***RAB11FIP4***	0.00000	-3.11
4144	TMEM52	0.01123	-2.87	4201	***MAFB***	0.00002	-3.11
4145	***SLC25A45***	0.00000	-2.87	4202	***NR2F2***	0.00000	-3.12
4146	SH3TC2	0.00000	-2.89	4203	WBSCR27	0.00500	-3.13
4147	CA9	0.03209	-2.89	4204	***SASH1***	0.00001	-3.13
4148	***FAM134B***	0.00005	-2.89	4205	***NR1H3***	0.00000	-3.13
4149	***GPD1L***	0.00075	-2.89	4206	***UGT1A6***	0.00000	-3.15
4150	CFAP43	0.00537	-2.89	4207	***C4orf32***	0.00143	-3.16
4151	***PRICKLE4***	0.00000	-2.90	4208	***MAP3K5***	0.00000	-3.17
4152	LOC100506271	0.04977	-2.91	4209	***AREG***	0.01738	-3.18
4153	SGPP2	0.01129	-2.91	4210	TNFRSF11A	0.00000	-3.19
4154	***SH3TC1***	0.00000	-2.91	4211	***ESR1***	0.00000	-3.19
4155	***SECTM1***	0.00000	-2.91	4212	ABCA5	0.00013	-3.20
4156	SLC16A7	0.00008	-2.92	4213	LOC100128770	0.03572	-3.21
4157	***CFB***	0.00023	-2.92	4214	***SMPDL3A***	0.00003	-3.26
4158	***ST6GALNAC4***	0.00000	-2.92	4215	***SNORD14B***	0.02578	-3.27
4159	***S100A4***	0.00000	-2.92	4216	PTPRO	0.03234	-3.27
4160	***TC2N***	0.00025	-2.93	4217	***PLA2G10***	0.00012	-3.27
4161	***BAG1***	0.00000	-2.95	4218	***PPL***	0.00000	-3.27
4162	***DKFZP586I1420***	0.00000	-2.95	4219	***S100A9***	0.00001	-3.29
4163	EPB41L4A-AS2	0.00024	-2.95	4220	***ISPD***	0.00000	-3.30
4164	***PRR15***	0.00000	-2.96	4221	CYP3A7	0.00013	-3.31
4165	***PRKCD***	0.00000	-2.96	4222	***TH***	0.00000	-3.32
4166	LRRC32	0.00691	-2.98	4223	***CKMT1A***	0.00000	-3.33
4167	ACER2	0.00090	-2.98	4224	***CXADR***	0.00000	-3.34
4168	***ANK3***	0.00003	-2.98	4225	***CHCHD10***	0.00000	-3.34
4169	***BIRC3***	0.00065	-2.99	4226	NOS1AP	0.00000	-3.34
4170	***KLF5***	0.00002	-2.99	4227	***PLS1***	0.00073	-3.34
4171	ARNT2	0.00114	-2.99	4228	***GAREM***	0.00000	-3.38
4172	***TMEM238***	0.00000	-2.99	4229	***RASSF5***	0.00000	-3.38
4173	***FOXQ1***	0.00006	-2.99	4230	***RGCC***	0.00445	-3.40
4174	***HOXC13***	0.00000	-3.00	4231	***SNCG***	0.00000	-3.41
4175	***P4HTM***	0.00000	-3.00	4232	***CAMK2D***	0.00001	-3.43
4176	***TMEM91***	0.00000	-3.01	4233	***SORL1***	0.00000	-3.43
4177	***EPB41L1***	0.00000	-3.02	4234	***OVOL2***	0.00000	-3.44
4178	***PEX11A***	0.00000	-3.02	4235	***GRAMD1C***	0.01933	-3.44
4179	***BTBD3***	0.00000	-3.02	4236	***OLR1***	0.00000	-3.44

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
4237	***PDCD4***	0.00030	-3.45	4294	***CD55***	0.00003	-3.97
4238	C6orf165	0.00068	-3.48	4295	***GLUL***	0.00000	-4.04
4239	***SYT8***	0.00002	-3.49	4296	***GSTA4***	0.00000	-4.04
4240	PLCXD3	0.00002	-3.49	4297	***ABHD11-AS1***	0.00000	-4.04
4241	***BSPRY***	0.00000	-3.51	4298	***MUC20***	0.00003	-4.05
4242	***PRSS12***	0.00000	-3.52	4299	LOC102723373	0.00347	-4.07
4243	***MYZAP***	0.00000	-3.53	4300	***COBL***	0.00000	-4.07
4244	***EPAS1***	0.00000	-3.53	4301	IGSF11	0.00729	-4.09
4245	***RNF223***	0.00128	-3.54	4302	***IL1RN***	0.00000	-4.09
4246	ANKDD1A	0.00000	-3.54	4303	TGFBR3	0.00001	-4.10
4247	***RHBDL1***	0.02801	-3.55	4304	MRPL23-AS1	0.03742	-4.11
4248	***C11orf70***	0.00000	-3.55	4305	ARFGEF3	0.00000	-4.11
4249	***SUSD2***	0.00016	-3.55	4306	***PLLP***	0.00000	-4.14
4250	***FAAH***	0.00002	-3.55	4307	***RARRES2***	0.00019	-4.14
4251	***EPS8***	0.00000	-3.55	4308	***THEM6***	0.00000	-4.14
4252	C5orf66-AS1	0.00354	-3.55	4309	ERBB4	0.00005	-4.15
4253	***MPV17L***	0.00000	-3.56	4310	***RINL***	0.00000	-4.16
4254	PCDH20	0.01153	-3.57	4311	***GRHL1***	0.00000	-4.17
4255	***IMPA2***	0.00000	-3.57	4312	KLHL30	0.00000	-4.20
4256	***ID1***	0.00000	-3.58	4313	PLIN4	0.00003	-4.23
4257	ABCA12	0.00000	-3.58	4314	***MAL2***	0.00000	-4.23
4258	***CKMT1B***	0.00001	-3.58	4315	ACPP	0.00000	-4.23
4259	RGS11	0.00000	-3.59	4316	PSG4	0.00045	-4.23
4260	***CASP10***	0.00001	-3.61	4317	***SLC9A3R1***	0.00000	-4.23
4261	***TNNI2***	0.00008	-3.62	4318	FREM2	0.00333	-4.23
4262	SLC9A2	0.00032	-3.62	4319	***SULT1A2***	0.00000	-4.25
4263	LRRC4	0.00000	-3.62	4320	***FLVCR2***	0.00000	-4.25
4264	***CXXC5***	0.00001	-3.65	4321	***PTGES***	0.00000	-4.25
4265	***PHACTR3***	0.01293	-3.67	4322	***SRD5A3***	0.00000	-4.26
4266	EXOC3L4	0.00826	-3.68	4323	***KLK5***	0.00000	-4.32
4267	***MAP2K6***	0.00006	-3.69	4324	***LOC101927934***	0.00000	-4.33
4268	***GSR***	0.00000	-3.69	4325	***ECEL1P2***	0.00000	-4.34
4269	ARHGAP20	0.00001	-3.72	4326	***MITF***	0.00000	-4.35
4270	ID4	0.00045	-3.73	4327	***PC***	0.00000	-4.36
4271	***KIF26A***	0.00000	-3.73	4328	***PHLPP1***	0.00000	-4.37
4272	CMAHP	0.00000	-3.74	4329	COLCA2	0.00183	-4.38
4273	***FKBP5***	0.00000	-3.74	4330	***ALDH3B2***	0.00009	-4.38
4274	***TNFRSF18***	0.00000	-3.76	4331	***GAL***	0.00000	-4.43
4275	***MYO5C***	0.00000	-3.77	4332	FA2H	0.00000	-4.45
4276	***C1orf226***	0.00000	-3.77	4333	IL12A	0.00001	-4.45
4277	***NAGA***	0.00000	-3.77	4334	PDE8B	0.00001	-4.47
4278	TMPRSS3	0.00002	-3.78	4335	***CLDN3***	0.00354	-4.50
4279	MCF2L	0.00912	-3.79	4336	***TLR3***	0.00411	-4.52
4280	***PPM1H***	0.00000	-3.80	4337	***SGPP1***	0.00001	-4.53
4281	***TMPRSS13***	0.00001	-3.80	4338	***FRMD3***	0.00000	-4.53
4282	***ZG16B***	0.00000	-3.80	4339	***ST3GAL4-AS1***	0.00000	-4.54
4283	IL6R	0.00054	-3.83	4340	FAM3B	0.04917	-4.59
4284	***ADRB2***	0.00000	-3.83	4341	***GRAMD2***	0.00000	-4.65
4285	RGS9BP	0.00103	-3.84	4342	SLC25A25-AS1	0.00000	-4.66
4286	MXRA5	0.00000	-3.84	4343	***SLC27A2***	0.00000	-4.67
4287	***NRG4***	0.00002	-3.86	4344	***ASS1***	0.00000	-4.67
4288	***C4orf19***	0.00001	-3.88	4345	***HS6ST2***	0.00000	-4.67
4289	***E2F2***	0.00000	-3.88	4346	***ST3GAL4***	0.00000	-4.69
4290	***AIM1L***	0.00000	-3.89	4347	***NPNT***	0.00000	-4.69
4291	***XK***	0.00000	-3.90	4348	***TFCP2L1***	0.00000	-4.71
4292	***ADAP1***	0.00000	-3.93	4349	***ALDH3B1***	0.00000	-4.72
4293	***SYBU***	0.00000	-3.96	4350	***TMC5***	0.00000	-4.72

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
4351	PPM1L	0.00005	-4.76	4408	***BCAS1***	0.00004	-6.21
4352	***FOXO6***	0.00027	-4.77	4409	***HPDL***	0.00000	-6.24
4353	***MPZL2***	0.00000	-4.77	4410	***UNC5B-AS1***	0.00001	-6.24
4354	***LINC01133***	0.00017	-4.79	4411	***FAM174B***	0.00000	-6.28
4355	TCN1	0.04962	-4.79	4412	BEST2	0.01963	-6.29
4356	***TBC1D8***	0.00000	-4.80	4413	ADGRE2	0.00000	-6.31
4357	MAG	0.00000	-4.80	4414	***KLK8***	0.00000	-6.31
4358	***CREG2***	0.00000	-4.81	4415	***PTGS1***	0.00000	-6.40
4359	RSPH1	0.01389	-4.82	4416	***ACSL5***	0.00000	-6.40
4360	***ENPP1***	0.00000	-4.82	4417	GPX2	0.03780	-6.42
4361	***TNFSF10***	0.00196	-4.84	4418	***SCARB1***	0.00000	-6.45
4362	LOC100506834	0.03013	-4.89	4419	ABCB1	0.00102	-6.50
4363	***LIPA***	0.00000	-4.89	4420	CD22	0.00001	-6.56
4364	***BCL6***	0.00000	-4.89	4421	***RAPGEFL1***	0.00000	-6.58
4365	***ZNF114***	0.00000	-4.92	4422	***EPB41L4A***	0.00000	-6.58
4366	LOC101928738	0.03180	-4.94	4423	***MB***	0.00000	-6.69
4367	***IRX3***	0.00000	-4.95	4424	***SPOCK2***	0.00000	-6.74
4368	***SCEL***	0.00000	-4.97	4425	***LIMCH1***	0.00000	-6.76
4369	***ANXA9***	0.00001	-4.98	4426	***CDKN1C***	0.00003	-6.80
4370	ALPPL2	0.01005	-4.99	4427	***KRT15***	0.00000	-6.82
4371	***WDR86-AS1***	0.00000	-5.01	4428	ZNF503-AS1	0.00001	-6.83
4372	***SMIM5***	0.00000	-5.02	4429	***TINCR***	0.00000	-6.85
4373	***CFD***	0.00000	-5.03	4430	***EVA1C***	0.00000	-6.91
4374	LRG1	0.00000	-5.04	4431	***HRASLS2***	0.00000	-6.94
4375	EVPLL	0.00047	-5.05	4432	***RASL11A***	0.00000	-6.96
4376	KCNQ3	0.00000	-5.06	4433	DHRS9	0.00063	-6.98
4377	FUT9	0.00026	-5.10	4434	***TMPRSS11E***	0.00014	-7.07
4378	MMRN2	0.00000	-5.14	4435	DOCK8	0.00000	-7.10
4379	ATP8A1	0.00006	-5.15	4436	***HPGD***	0.02030	-7.18
4380	***CCDC64B***	0.00000	-5.20	4437	POU2F3	0.00118	-7.27
4381	***UPK1B***	0.00000	-5.26	4438	***ELF3***	0.00000	-7.35
4382	SPINK5	0.00005	-5.30	4439	***KLK10***	0.00000	-7.39
4383	***PNMT***	0.00000	-5.30	4440	***ANKRD2***	0.00000	-7.40
4384	SIDT1	0.00000	-5.34	4441	***KCNK5***	0.00000	-7.42
4385	***POU5F1***	0.00000	-5.35	4442	***FAM65B***	0.00000	-7.42
4386	***IKZF2***	0.00000	-5.39	4443	DLX3	0.00008	-7.44
4387	***IL20RA***	0.00000	-5.41	4444	***ADGRF1***	0.00000	-7.55
4388	***ANGPT1***	0.00015	-5.43	4445	***KCNK15***	0.00000	-7.56
4389	RTN4RL1	0.00000	-5.44	4446	***ADAMTSL4***	0.00000	-7.64
4390	***PSORS1C3***	0.00000	-5.45	4447	***EDN2***	0.00001	-7.71
4391	***SLC16A5***	0.00000	-5.46	4448	***GATA2***	0.00000	-7.72
4392	***S100A14***	0.00000	-5.49	4449	***LXN***	0.00001	-7.79
4393	PSG1	0.01603	-5.52	4450	***TRIM29***	0.00000	-7.80
4394	SPTSSB	0.00004	-5.57	4451	***MMP28***	0.00000	-7.86
4395	***GATA2-AS1***	0.00000	-5.68	4452	***CYP26A1***	0.00000	-7.97
4396	***SLC6A11***	0.00000	-5.73	4453	***TJP3***	0.00000	-7.97
4397	FFAR4	0.00901	-5.74	4454	KLK7	0.00000	-8.03
4398	LOC284344	0.00021	-5.81	4455	***MYPN***	0.00000	-8.07
4399	***EDAR***	0.00000	-5.82	4456	***ALDH3A1***	0.00000	-8.08
4400	***ID2***	0.00000	-5.97	4457	***ZBTB16***	0.00000	-8.10
4401	***SDPR***	0.00000	-5.99	4458	CLIC5	0.00000	-8.22
4402	CNR1	0.00002	-5.99	4459	***CD14***	0.00000	-8.26
4403	C5AR1	0.00006	-5.99	4460	***VEPH1***	0.00000	-8.26
4404	***SLC29A3***	0.00000	-6.00	4461	***SP6***	0.00000	-8.28
4405	LRRC26	0.00692	-6.08	4462	LINC01085	0.00063	-8.36
4406	***PTGDS***	0.00020	-6.16	4463	GCNT3	0.00599	-8.37
4407	PIK3C2G	0.00001	-6.17	4464	***HSD11B2***	0.00000	-8.41

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
4465	PDK4	0.00164	-8.62	4522	***TMEM45B***	0.00000	-15.42
4466	CALHM3	0.00000	-8.70	4523	***PKDCC***	0.00111	-15.70
4467	ANKRD20A5P	0.01155	-8.90	4524	***OSBP2***	0.00000	-15.79
4468	SLC25A21	0.00000	-8.94	4525	***ADIRF***	0.00000	-15.90
4469	MUC13	0.01139	-9.01	4526	***TMPRSS4***	0.00000	-16.00
4470	BCO1	0.00782	-9.01	4527	***VILL***	0.00000	-16.12
4471	***ALDH1L1***	0.00000	-9.03	4528	PADI4	0.00048	-16.16
4472	***TGM1***	0.00000	-9.06	4529	***ALOX5***	0.00000	-16.27
4473	***EPGN***	0.00000	-9.26	4530	NCCRP1	0.00000	-16.44
4474	TLE6	0.00035	-9.30	4531	SLURP1	0.00000	-16.75
4475	***SCNN1G***	0.00000	-9.55	4532	WFDC12	0.02774	-16.87
4476	PPARGC1B	0.00002	-9.70	4533	KLK11	0.00000	-16.88
4477	***CAPN8***	0.00000	-9.86	4534	B3GALT5	0.00015	-17.76
4478	DSG4	0.00013	-9.94	4535	***HRK***	0.00000	-17.95
4479	GRAMD1B	0.00118	-10.08	4536	CYP4X1	0.00000	-18.20
4480	***ABCG2***	0.00000	-10.16	4537	***CDH23***	0.00000	-18.65
4481	SMOC2	0.00000	-10.28	4538	***MUC1***	0.00000	-18.69
4482	S100A5	0.00986	-10.34	4539	***PRR15L***	0.00000	-18.95
4483	***LY6D***	0.00000	-10.42	4540	***TRIM31***	0.00000	-19.33
4484	CYP4B1	0.00003	-10.69	4541	SOX21-AS1	0.00000	-20.07
4485	***TSPAN8***	0.00000	-10.70	4542	***AGR2***	0.00000	-20.18
4486	PROM1	0.02362	-10.77	4543	***KCNQ1***	0.00000	-21.03
4487	***METTL7A***	0.00000	-10.95	4544	***PAX9***	0.00000	-21.35
4488	***AMN***	0.00000	-10.98	4545	***WISP2***	0.00000	-22.65
4489	ARHGEF38	0.02655	-10.99	4546	LDLRAD1	0.00914	-23.02
4490	***FOLR1***	0.00000	-10.99	4547	***SCNN1B***	0.00000	-23.07
4491	***S100P***	0.00000	-11.13	4548	***PI3***	0.00001	-24.97
4492	***UPK3A***	0.00005	-11.22	4549	BMP3	0.00001	-25.29
4493	SGSM1	0.00000	-11.35	4550	***LCN2***	0.00010	-25.65
4494	***SLCO4A1***	0.00000	-11.42	4551	SOX21	0.00000	-26.63
4495	SSTR1	0.00002	-11.67	4552	***UNC5B***	0.00000	-26.68
4496	SLCO4A1-AS1	0.00000	-11.70	4553	CLDN8	0.00005	-27.24
4497	***PCSK5***	0.00000	-11.82	4554	PPFIBP2	0.00027	-29.16
4498	***MMP7***	0.00000	-11.88	4555	PDE3B	0.00061	-29.45
4499	MGP	0.00011	-11.91	4556	CEACAM5	0.00000	-30.41
4500	UBXN10	0.00007	-12.06	4557	MSMB	0.02101	-30.90
4501	LINC01559	0.00000	-12.78	4558	***KRT13***	0.00000	-32.31
4502	SCARA5	0.00031	-12.96	4559	MGAM	0.00027	-36.26
4503	***RARRES1***	0.00000	-13.08	4560	***GKN1***	0.00002	-36.78
4504	TNXB	0.00000	-13.09	4561	C10orf105	0.00120	-37.62
4505	***FBP1***	0.00000	-13.22	4562	***CDH5***	0.00000	-37.64
4506	***GDPD3***	0.00000	-13.24	4563	***CEACAM6***	0.00000	-38.61
4507	SEMA6D	0.00000	-13.40	4564	***CYP4F12***	0.00000	-38.72
4508	***AOC1***	0.00002	-13.44	4565	***PPP1R16B***	0.00000	-39.34
4509	***ALPP***	0.00000	-13.78	4566	***CRISP3***	0.00001	-50.88
4510	***EREG***	0.00051	-14.02	4567	CYP4F29P	0.02164	-55.30
4511	***H19***	0.00000	-14.13	4568	***INHBB***	0.00000	-57.80
4512	***FXYD3***	0.00000	-14.15	4569	SLC12A3	0.00006	-65.71
4513	PPEF1	0.01523	-14.43	4570	***KRT4***	0.00000	-89.42
4514	***SYT12***	0.00000	-14.54	4571	LYPD2	0.00797	-99.94
4515	***CYP4F3***	0.00000	-14.70	4572	***RERG***	0.00000	-102.87
4516	***LOC729966***	0.00000	-14.86	4573	***PSCA***	0.00000	-103.70
4517	DEGS2	0.00002	-14.88	4574	***GKN2***	0.00000	-118.65
4518	GLB1L2	0.00000	-14.91	4575	***C11orf86***	0.00000	-120.44
4519	***AQP3***	0.00000	-15.18				
4520	LINC00974	0.00001	-15.22				
4521	***SLPI***	0.00000	-15.24				

A3.3 Significantly deregulated genes in DU145 upon stimulation with TGF-β

Rank	Gene	Corrected p-value	FC	Rank	Gene	Corrected p-value	FC
1	***BMP2***	0.00000	84.71	56	***TPM1***	0.00000	11.63
2	DOCK2	0.00002	77.30	57	***FRMD6***	0.00000	11.48
3	***SPOCK1***	0.00000	70.63	58	KIAA1614	0.00000	11.48
4	***C4orf26***	0.00000	67.91	59	***P4HA3***	0.00000	11.40
5	MYOCD	0.03454	63.42	60	***NIPAL4***	0.00000	11.21
6	LPAR5	0.00013	56.57	61	***ADAMTS6***	0.00000	11.16
7	COL20A1	0.00000	46.29	62	BCL11A	0.00002	11.07
8	R3HDML	0.00228	43.91	63	SORCS2	0.00042	10.54
9	GDF6	0.04188	41.02	64	MPP4	0.00197	10.54
10	***CRLF1***	0.00000	38.49	65	***PADI2***	0.00000	10.53
11	CCR1	0.00029	37.79	66	***TGFBI***	0.00000	10.50
12	***ITGA11***	0.00003	35.57	67	***TAGLN***	0.00000	10.48
13	ROS1	0.00003	31.85	68	***GPR87***	0.00000	10.45
14	***LBH***	0.00000	30.86	69	ADGRF4	0.00000	10.44
15	***COL4A1***	0.00015	30.53	70	CPED1	0.02424	10.42
16	SGCG	0.00644	29.28	71	FAM198B	0.03761	10.20
17	LOC100507431	0.00417	28.47	72	CASC15	0.00600	10.11
18	ADGRF2	0.03797	27.75	73	RASGRF2	0.00000	10.05
19	KLHDC8A	0.00000	27.52	74	DACT1	0.00083	9.81
20	***NRP2***	0.00000	27.23	75	***LCP1***	0.00000	9.64
21	CHRNA4	0.00009	26.35	76	***THBS1***	0.00004	9.61
22	FOXS1	0.00009	24.95	77	***NCF2***	0.00000	9.56
23	***ADAM19***	0.00000	24.73	78	***BMF***	0.00000	9.55
24	PLXNA4	0.00040	23.99	79	GRID1	0.00324	9.54
25	MC5R	0.01355	22.71	80	BAAT	0.00007	9.40
26	***SAP30L-AS1***	0.00002	21.75	81	***C1orf106***	0.00000	9.35
27	CD300C	0.00002	21.49	82	***COL1A1***	0.00015	9.07
28	***RASGRP3***	0.00001	21.37	83	***TP53I3***	0.00000	9.06
29	ALPK2	0.00000	20.88	84	***LINC00704***	0.00019	8.97
30	GNA14	0.00002	20.75	85	FLRT2	0.00203	8.91
31	***GALNT10***	0.00000	19.22	86	***MSC***	0.00000	8.61
32	***COL4A2***	0.00007	19.00	87	LDLRAD4	0.00143	8.47
33	GPR183	0.00002	18.76	88	JAM2	0.00000	8.44
34	***PMEPA1***	0.00000	18.71	89	LOC79160	0.00167	8.40
35	***RASGRP1***	0.00000	18.66	90	FAM26E	0.01653	8.38
36	NKAIN4	0.00000	17.33	91	***ITGAV***	0.00000	8.37
37	IGF2	0.00000	17.06	92	***LAMC2***	0.00004	8.21
38	MYO7B	0.01624	17.06	93	CLDN14	0.00000	8.18
39	***NKILA***	0.00000	16.48	94	***ISM2***	0.00000	8.15
40	SLAMF9	0.00285	15.74	95	ARHGAP31	0.00000	8.05
41	ACTBL2	0.00564	15.60	96	***CCDC80***	0.00000	7.79
42	***COL5A1***	0.00001	14.98	97	***WNT5B***	0.00000	7.71
43	***TMEM59L***	0.00000	14.54	98	***CTGF***	0.00004	7.58
44	CLEC19A	0.00181	14.16	99	***LTBP2***	0.00001	7.54
45	ESM1	0.04853	13.99	100	***ANXA8L1***	0.00002	7.52
46	***SERPINE1***	0.00061	13.98	101	GPR132	0.00001	7.51
47	***AMIGO2***	0.00000	13.79	102	***PPP1R14C***	0.00000	7.44
48	***PIK3IP1***	0.00000	13.59	103	***PCDH1***	0.00000	7.39
49	SEMA5B	0.00348	13.51	104	***MSC-AS1***	0.00000	7.37
50	***GLIPR1***	0.00001	12.99	105	***LMCD1***	0.00000	7.36
51	***ESAM***	0.00000	12.82	106	***CYP24A1***	0.00000	7.24
52	LINC01279	0.00170	12.71	107	***PKP1***	0.00000	7.09
53	***ACKR3***	0.00000	12.36	108	HS3ST3B1	0.00082	7.01
54	***NEDD9***	0.00000	11.98	109	AQP1	0.00009	6.90
55	***SERPINE2***	0.00000	11.65	110	RNF152	0.01339	6.90

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
111	***MBOAT2***	0.00000	6.77	168	***STK32A***	0.00000	4.96
112	LINC01561	0.01299	6.68	169	SCN2A	0.00001	4.96
113	***SYT11***	0.00000	6.61	170	***PHLDB1***	0.00000	4.94
114	LINC01537	0.03804	6.48	171	LMO1	0.00035	4.90
115	***AFAP1L2***	0.00000	6.40	172	***LTBP1***	0.00211	4.88
116	ZCCHC18	0.01586	6.37	173	***SMAD7***	0.00000	4.87
117	SYT13	0.00033	6.36	174	***DSE***	0.00009	4.85
118	***DCBLD1***	0.00000	6.33	175	IFFO1	0.00003	4.83
119	***KLF7***	0.00000	6.29	176	***TFPI2***	0.00001	4.83
120	COL7A1	0.00027	6.27	177	DHRS2	0.00001	4.77
121	***LINC00842***	0.00004	6.26	178	***TP53INP1***	0.00934	4.77
122	***PTPRK***	0.00000	6.26	179	***TSPAN2***	0.00016	4.75
123	***STK38L***	0.00003	6.17	180	***FSTL1***	0.00034	4.74
124	***VASN***	0.00000	6.14	181	***ADAMTS15***	0.00000	4.73
125	ANPEP	0.00035	6.10	182	LZTS1	0.00000	4.71
126	***LINC00623***	0.00000	6.06	183	FGF1	0.00821	4.71
127	***PGM2L1***	0.00165	6.05	184	***PTHLH***	0.00000	4.69
128	***PRUNE2***	0.00001	6.01	185	***FOXP1***	0.00000	4.65
129	***RAI14***	0.00005	6.00	186	***FBN1***	0.01403	4.64
130	CPNE4	0.02539	5.90	187	***ARL15***	0.00004	4.62
131	***PIK3AP1***	0.00170	5.89	188	C14orf37	0.00025	4.61
132	PGBD5	0.00015	5.88	189	***MEX3B***	0.00061	4.53
133	COL27A1	0.00000	5.80	190	***PKIA***	0.00006	4.51
134	***NRG1***	0.00003	5.77	191	***PPAPDC1A***	0.00000	4.50
135	QPCT	0.00026	5.76	192	***EVA1A***	0.00105	4.49
136	***LAMB1***	0.00002	5.76	193	***TARSL2***	0.00001	4.47
137	***WNT7A***	0.00015	5.74	194	***MFAP2***	0.00000	4.47
138	***LIMS2***	0.00006	5.73	195	***C18orf25***	0.00014	4.46
139	***INPP4B***	0.00000	5.66	196	***TCEAL1***	0.00009	4.45
140	DNAH7	0.04635	5.65	197	***MRC2***	0.00001	4.43
141	ENG	0.00001	5.64	198	***XPR1***	0.00001	4.42
142	LOC100128288	0.03848	5.61	199	***MATN3***	0.00002	4.42
143	***CDKN2B***	0.00059	5.57	200	***ACTC1***	0.00011	4.40
144	NLRP1	0.00000	5.55	201	***BEAN1***	0.00000	4.38
145	FKBP7	0.00043	5.54	202	PRR5L	0.00000	4.37
146	CACNA1G	0.00000	5.52	203	***FHOD3***	0.00000	4.37
147	***CHST11***	0.00000	5.52	204	***SLC35F3***	0.00047	4.36
148	***NAV1***	0.00080	5.52	205	***ITGA2***	0.00004	4.36
149	***DOCK10***	0.00012	5.51	206	***BAMBI***	0.00000	4.33
150	***PROC***	0.00002	5.49	207	MOB3B	0.00001	4.32
151	***LFNG***	0.00000	5.40	208	***ANGPTL4***	0.00000	4.32
152	ZNF365	0.00136	5.37	209	CFAP54	0.00002	4.27
153	***SKIL***	0.00053	5.31	210	***EPHB2***	0.00000	4.24
154	RCAN2	0.00006	5.27	211	***IL18BP***	0.00165	4.23
155	***SOX4***	0.00003	5.23	212	PITX2	0.00013	4.22
156	LOC729683	0.00073	5.23	213	ZNF697	0.00683	4.22
157	***LRRN2***	0.00000	5.18	214	SHC3	0.00526	4.20
158	***MACRH4***	0.00050	5.17	215	NR2F1-AS1	0.00017	4.20
159	***FSTL3***	0.00000	5.15	216	***MVB12B***	0.00000	4.20
160	TNFRSF19	0.00050	5.13	217	***NOG***	0.00000	4.20
161	***CACHD1***	0.00043	5.08	218	PTPRB	0.00482	4.19
162	***HNRNPA1P33***	0.00054	5.03	219	***TMEM45A***	0.00002	4.19
163	MAML2	0.00310	5.03	220	***SNAI2***	0.01168	4.16
164	DNAJC22	0.00069	5.03	221	***FERMT2***	0.00064	4.15
165	***FBXO32***	0.00000	5.02	222	***NEURL1B***	0.00000	4.15
166	***CORO2B***	0.00006	5.00	223	WIPF1	0.00011	4.12
167	KIAA1549L	0.00056	4.98	224	***DPY19L1***	0.00001	4.12

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
225	***TUBA4A***	0.00000	4.11	282	***STK33***	0.00801	3.61
226	***SGK1***	0.00000	4.11	283	***BVES***	0.04019	3.61
227	***PLS3***	0.00812	4.10	284	***IL11***	0.00123	3.60
228	***HTR1D***	0.00000	4.07	285	PLA2R1	0.00745	3.60
229	FST	0.01065	4.06	286	***TMEM2***	0.00187	3.58
230	***GALNT16***	0.00002	4.05	287	VSTM4	0.00507	3.57
231	GNG2	0.00228	4.04	288	***MYO10***	0.00011	3.57
232	***JARID2***	0.00000	4.03	289	***EFR3B***	0.00000	3.57
233	***LIMS1***	0.00140	4.03	290	***GPRC5B***	0.00000	3.57
234	***FZD2***	0.00000	4.03	291	***GLIPR2***	0.00000	3.56
235	MATK	0.00050	4.02	292	***CALD1***	0.01101	3.56
236	MN1	0.00159	3.99	293	WNT5A	0.00029	3.55
237	LOC643072	0.00009	3.98	294	***F2R***	0.00001	3.52
238	***DOCK4***	0.00000	3.98	295	***DPYSL3***	0.00000	3.52
239	GBP1	0.00004	3.96	296	CNTNAP2	0.01026	3.52
240	***MLLT11***	0.00000	3.95	297	***BST1***	0.00004	3.50
241	***LINC01138***	0.00000	3.95	298	***CDKN1C***	0.00000	3.50
242	***KCTD11***	0.00000	3.94	299	MME	0.02216	3.48
243	UNC5CL	0.00005	3.94	300	CLUL1	0.01930	3.48
244	***CADM1***	0.00001	3.94	301	***PGRMC2***	0.00003	3.46
245	***YPEL2***	0.00001	3.94	302	HLA-DPA1	0.00022	3.46
246	***EFNA2***	0.00000	3.94	303	***GOPC***	0.00245	3.46
247	***BPGM***	0.00000	3.92	304	***SLC19A2***	0.00051	3.44
248	***SUSD6***	0.00000	3.92	305	***EDN1***	0.00000	3.43
249	***MIR503HG***	0.00281	3.90	306	***CDK14***	0.00006	3.42
250	ACTG2	0.01424	3.90	307	GPC4	0.02720	3.42
251	***FAM228B***	0.00003	3.89	308	***RAB3B***	0.03706	3.42
252	***DNAJB2***	0.00000	3.89	309	***ST5***	0.00000	3.39
253	***HSD17B6***	0.01463	3.88	310	***ECM1***	0.00000	3.38
254	***IGFBP3***	0.01741	3.88	311	***MARCKSL1***	0.00000	3.38
255	ZCCHC12	0.00274	3.86	312	***FRMD6-AS1***	0.00000	3.38
256	***COL4A4***	0.00106	3.85	313	***TSPAN18***	0.00000	3.38
257	KALRN	0.00096	3.83	314	ABAT	0.00000	3.38
258	***FZD1***	0.00001	3.83	315	***PFN4***	0.00012	3.37
259	***EEPD1***	0.00000	3.82	316	***FBLIM1***	0.00000	3.37
260	ANOS1	0.00002	3.82	317	***ARHGEF40***	0.00000	3.36
261	***S1PR3***	0.00714	3.81	318	GLI1	0.00474	3.36
262	***NTN1***	0.00001	3.81	319	***SUSD4***	0.00007	3.36
263	PCDHB15	0.00101	3.81	320	***ELK3***	0.00214	3.35
264	LRCH2	0.01425	3.81	321	DYNC1I1	0.03052	3.34
265	MAMDC2	0.00057	3.79	322	PBX1	0.00253	3.33
266	***MMP14***	0.00000	3.79	323	MEX3A	0.02578	3.31
267	***SLC22A3***	0.00001	3.78	324	***OCIAD2***	0.00000	3.31
268	RAPGEF2	0.00506	3.77	325	***NREP***	0.00000	3.31
269	***TAPT1***	0.00000	3.76	326	ZNF112	0.02019	3.29
270	TLN2	0.00006	3.76	327	GABRQ	0.00010	3.29
271	ST8SIA6	0.00359	3.75	328	***TMEM92***	0.00122	3.27
272	***LINC00869***	0.00000	3.74	329	***TMEM65***	0.00550	3.25
273	***COL4A3***	0.00796	3.74	330	***RAP2A***	0.00421	3.25
274	***LRRC8C***	0.00005	3.73	331	SUSD5	0.00006	3.24
275	***PXDN***	0.00307	3.73	332	HLX	0.00002	3.24
276	***NPC2***	0.00000	3.73	333	***RASSF2***	0.00002	3.24
277	***TNFAIP8***	0.00051	3.70	334	LPCAT2	0.00001	3.24
278	***EML1***	0.00000	3.69	335	***TGFB1I1***	0.00000	3.23
279	***PDGFC***	0.00059	3.68	336	APCDD1	0.00043	3.21
280	KDM7A	0.00259	3.67	337	***HBEGF***	0.00002	3.21
281	LRP4	0.00018	3.66	338	BMPR1B	0.03023	3.20

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
339	***SEMA7A***	0.00000	3.19	396	***PLEKHG4B***	0.00000	2.96
340	***FAM105A***	0.00224	3.19	397	***ARHGAP32***	0.00026	2.96
341	C4orf19	0.00008	3.19	398	ADAMTSL3	0.00013	2.96
342	LINC00941	0.01609	3.18	399	***TSPAN12***	0.01099	2.94
343	***ARNTL2***	0.00004	3.18	400	***EPB41L2***	0.00206	2.94
344	RGL1	0.00131	3.18	401	***PTPN21***	0.00006	2.93
345	***FERMT1***	0.00033	3.16	402	***ARHGEF18***	0.00000	2.92
346	***ZNF185***	0.00000	3.16	403	***EMB***	0.00290	2.92
347	***SERTAD4***	0.00000	3.15	404	***FNDC3B***	0.00001	2.91
348	***C14orf132***	0.00009	3.14	405	IGF2BP1	0.01044	2.90
349	***SLC26A2***	0.00260	3.14	406	***LIPG***	0.00241	2.89
350	MYB	0.03136	3.13	407	***PCDHB2***	0.00003	2.89
351	***RFTN1***	0.00030	3.13	408	SH3PXD2A	0.00015	2.89
352	***PID1***	0.00000	3.13	409	***ZNF532***	0.00015	2.88
353	NR2F1	0.00604	3.12	410	***SMIM3***	0.00004	2.87
354	***RNF121***	0.00000	3.11	411	***PPP1R12A***	0.00465	2.87
355	***LOC730101***	0.00001	3.11	412	***ETS1***	0.00090	2.87
356	***BMPR2***	0.00026	3.11	413	***RECK***	0.00014	2.86
357	CDYL2	0.00101	3.11	414	***MRAS***	0.00002	2.85
358	***SLC2A10***	0.00005	3.11	415	***ULK1***	0.00000	2.85
359	PDZD2	0.00085	3.10	416	***SLAMF7***	0.00005	2.85
360	***FUT8***	0.00009	3.08	417	***SERPINB5***	0.00012	2.84
361	***FAM214B***	0.00000	3.08	418	***RNF182***	0.00155	2.84
362	PCDHB9	0.04642	3.07	419	***PLEK2***	0.00000	2.84
363	***P4HA1***	0.01728	3.07	420	FAXDC2	0.00000	2.83
364	BEST3	0.00066	3.07	421	***ZNF827***	0.00000	2.82
365	***TUFT1***	0.00000	3.06	422	***RBMS3***	0.00046	2.81
366	***LINC00673***	0.00000	3.06	423	***KRBA2***	0.00289	2.80
367	***CEP170***	0.00171	3.05	424	***MAP7***	0.00045	2.80
368	***MAP3K2***	0.03094	3.05	425	CCM2L	0.03419	2.79
369	***FNBP1L***	0.01295	3.04	426	***EFEMP2***	0.01569	2.79
370	***SPOCD1***	0.00000	3.04	427	***SPDL1***	0.03838	2.78
371	***PALLD***	0.00004	3.03	428	NFASC	0.00123	2.78
372	***TMCC1***	0.00001	3.03	429	***TSPAN14***	0.00000	2.78
373	***OLFM2***	0.00000	3.03	430	***NEBL***	0.00198	2.77
374	***SEMA3C***	0.00138	3.02	431	***LOC103091866***	0.00003	2.76
375	ATP7A	0.00489	3.02	432	***ELL2***	0.04939	2.75
376	***KCNH1***	0.00000	3.01	433	TGFB3	0.00009	2.74
377	***CD44***	0.00001	3.01	434	PEAR1	0.03533	2.74
378	***PDGFA***	0.00061	3.01	435	***PIK3CD***	0.00000	2.73
379	***LRP12***	0.00270	3.01	436	TNFRSF25	0.04894	2.73
380	***SERTAD4-AS1***	0.00001	3.00	437	***BLOC1S2***	0.00212	2.72
381	***ACSL4***	0.03111	3.00	438	***RHOB***	0.00001	2.72
382	***PPP3CA***	0.00041	3.00	439	PALM2	0.01814	2.71
383	***SYT1***	0.00042	2.99	440	***CYP26B1***	0.00010	2.71
384	CFH	0.04510	2.99	441	***SYCE1L***	0.00114	2.71
385	***ATP13A2***	0.00000	2.99	442	***FAM189A2***	0.00458	2.71
386	***LINC01137***	0.00000	2.99	443	***SGCB***	0.00004	2.71
387	***IGF2BP3***	0.01190	2.98	444	***SORT1***	0.00288	2.71
388	IL17RD	0.01330	2.98	445	***FAM168A***	0.00006	2.71
389	***PAX6***	0.00021	2.98	446	***PDIA3P1***	0.00002	2.70
390	CHSY3	0.01427	2.98	447	NXPH3	0.00034	2.70
391	***MYL9***	0.00000	2.98	448	MAFA	0.01426	2.70
392	***GLI2***	0.00012	2.97	449	***ZDHHC17***	0.01347	2.70
393	***ATP10D***	0.00065	2.97	450	***GADD45B***	0.00001	2.70
394	SBK1	0.04527	2.97	451	***PRKAB2***	0.00029	2.70
395	***PLOD2***	0.00064	2.96	452	***WWP1***	0.01856	2.70

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
453	PLEKHA8P1	0.01444	2.69	510	***ARMCX2***	0.00003	2.56
454	***ITGB1***	0.00438	2.69	511	***CPEB2***	0.00581	2.56
455	ZNF260	0.03529	2.69	512	***MAP1B***	0.02637	2.55
456	PCDHB13	0.00254	2.67	513	RNF165	0.01702	2.55
457	***HLA-DQB1***	0.00602	2.67	514	***LUZP1***	0.00012	2.55
458	***CBX2***	0.00016	2.67	515	CARD11	0.00467	2.55
459	***TPM4***	0.00013	2.67	516	***BCAT1***	0.01322	2.55
460	GPNMB	0.00010	2.67	517	FAM231D	0.02900	2.55
461	***CTSV***	0.00003	2.66	518	***CLSTN1***	0.00004	2.55
462	***DSC2***	0.01842	2.66	519	***AKT3***	0.01203	2.55
463	DFNB31	0.00000	2.66	520	RAB30	0.00016	2.55
464	***GPR161***	0.00542	2.65	521	***ZFP36L1***	0.00096	2.55
465	***TSPAN13***	0.00055	2.65	522	***ACVR1***	0.00024	2.54
466	***TRAM1***	0.02573	2.65	523	***MIR22HG***	0.00000	2.54
467	PLAC1	0.00395	2.65	524	***SLC4A7***	0.00079	2.54
468	***NDST1***	0.00039	2.65	525	***SLC16A2***	0.00475	2.54
469	***EFR3A***	0.00288	2.64	526	***GALNT1***	0.04169	2.54
470	***PACS1***	0.00200	2.64	527	***MFSD7***	0.00010	2.53
471	***GAS6-AS2***	0.00004	2.64	528	***FXYD6***	0.00083	2.53
472	***LOC728392***	0.00583	2.62	529	ABCA1	0.00035	2.53
473	***MOXD1***	0.00581	2.61	530	***DLC1***	0.00438	2.52
474	***CTHRC1***	0.00132	2.61	531	***P2RY2***	0.00013	2.52
475	***LIMA1***	0.00391	2.61	532	***HIVEP1***	0.01008	2.52
476	***F2RL1***	0.00000	2.61	533	***IL13RA1***	0.00440	2.51
477	***TRIB1***	0.00049	2.61	534	LGR6	0.00091	2.51
478	KIAA1211	0.00194	2.61	535	***CMTM3***	0.00000	2.51
479	***PBX3***	0.00002	2.60	536	***SIK1***	0.00009	2.51
480	***TANC2***	0.02026	2.60	537	***KIDINS220***	0.01440	2.51
481	PTGES3L	0.02293	2.60	538	***TTC3***	0.00398	2.51
482	***APLP1***	0.00171	2.60	539	***UBTD2***	0.00066	2.50
483	AGPAT4	0.00073	2.60	540	***MAP3K7CL***	0.00022	2.50
484	***ARNTL***	0.00001	2.60	541	***EXTL2***	0.01960	2.50
485	***KDELC1***	0.00021	2.60	542	***ITGB3***	0.00024	2.49
486	***STXBP5***	0.00004	2.60	543	***MORF4L2***	0.00002	2.49
487	***SDC2***	0.02157	2.60	544	***FRS2***	0.04053	2.49
488	***SMURF2***	0.00012	2.60	545	PCDHB6	0.00783	2.49
489	***ANO6***	0.00035	2.59	546	HSD17B8	0.01602	2.49
490	***APAF1***	0.00797	2.59	547	***CALU***	0.00016	2.48
491	***RBPJ***	0.03847	2.59	548	TNRC6C	0.02639	2.48
492	***TRAM2***	0.00001	2.59	549	***GOLIM4***	0.00333	2.48
493	BMP6	0.00109	2.58	550	***SEC24D***	0.00349	2.48
494	TRPS1	0.00866	2.58	551	***EFEMP1***	0.00143	2.48
495	***SLC45A3***	0.00007	2.58	552	***APBB2***	0.00000	2.48
496	HTRA3	0.00040	2.58	553	***ZC4H2***	0.00775	2.48
497	***NKX3-1***	0.00018	2.58	554	***ZBED2***	0.00002	2.47
498	***CSRP2***	0.00002	2.58	555	***YPEL5***	0.00030	2.47
499	***FAM114A1***	0.00060	2.57	556	***HLA-DRB1***	0.00011	2.47
500	***SNAP23***	0.00109	2.57	557	***PCDHB5***	0.00095	2.47
501	***SARAF***	0.00214	2.57	558	***CD59***	0.00000	2.47
502	***MLLT3***	0.00089	2.57	559	***EDIL3***	0.00056	2.46
503	***CTTNBP2NL***	0.01569	2.57	560	***YIPF5***	0.00008	2.46
504	***USP2***	0.00013	2.57	561	***RBP1***	0.00330	2.46
505	***ANKLE2***	0.00000	2.56	562	***JUNB***	0.00000	2.46
506	***TFEB***	0.00007	2.56	563	***MATN2***	0.02941	2.46
507	VANGL2	0.03451	2.56	564	***SOWAHC***	0.02071	2.46
508	***BHLHE40***	0.00003	2.56	565	HECW2	0.00796	2.45
509	KIAA1211L	0.01900	2.56	566	***PPP1R13L***	0.00308	2.45

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
567	PLCXD2	0.00009	2.45	624	***PRKD3***	0.00057	2.34
568	***SUSD1***	0.00017	2.45	625	C2CD4C	0.00035	2.34
569	***SWAP70***	0.01451	2.45	626	PLAGL1	0.02586	2.34
570	CACNA1H	0.00071	2.44	627	***CDK17***	0.00181	2.33
571	***CDKN1A***	0.00001	2.44	628	***ITGB5***	0.00043	2.33
572	***TIPARP***	0.00588	2.44	629	***KIAA1161***	0.00049	2.33
573	***ITPRIPL2***	0.01589	2.44	630	SLC2A12	0.00135	2.33
574	***CLTCL1***	0.00000	2.44	631	IPW	0.02061	2.33
575	***RNFT1***	0.02088	2.44	632	***ERAP2***	0.00894	2.33
576	***CAPRIN2***	0.00266	2.43	633	***STT3B***	0.02399	2.33
577	***RCAN1***	0.00398	2.43	634	PRKAA2	0.03740	2.32
578	***MIR181A2HG***	0.02415	2.43	635	***PDGFB***	0.00103	2.32
579	***ATP1B1***	0.00151	2.43	636	***PGRMC1***	0.00205	2.32
580	***SLC29A1***	0.00000	2.43	637	***PHTF1***	0.00740	2.32
581	***MKL1***	0.00000	2.43	638	***LOC100507487***	0.00243	2.32
582	***CCPG1***	0.00626	2.42	639	***STARD13***	0.00068	2.32
583	***MIR31HG***	0.00020	2.42	640	***DNAJC3***	0.00302	2.32
584	NRCAM	0.00277	2.42	641	***CHRNB1***	0.00002	2.32
585	GPR137C	0.02394	2.41	642	***B4GALT1***	0.00537	2.32
586	***HSP90B1***	0.02699	2.41	643	***TAB2***	0.00421	2.31
587	***KLHL24***	0.01856	2.41	644	CECR2	0.00217	2.31
588	IFI16	0.02494	2.41	645	***FGF11***	0.00036	2.31
589	***HHAT***	0.00043	2.41	646	***AOX1***	0.00055	2.31
590	***FNDC3A***	0.01635	2.41	647	***ADAM9***	0.00759	2.31
591	***ST6GAL1***	0.00002	2.41	648	***ATL1***	0.02570	2.31
592	***EIF2AK3***	0.01567	2.41	649	***HS2ST1***	0.03729	2.31
593	***SLC35D1***	0.00176	2.41	650	***ELF1***	0.00026	2.30
594	***FAM101B***	0.00235	2.40	651	***IL31RA***	0.00037	2.30
595	***PLEKHO1***	0.00000	2.40	652	***CACNG4***	0.00000	2.30
596	ADTRP	0.01074	2.40	653	***SH2D4A***	0.00000	2.30
597	***TES***	0.00053	2.40	654	***AREL1***	0.00005	2.30
598	***ETS2***	0.00048	2.40	655	***KCNN4***	0.00002	2.30
599	***WWC2***	0.00299	2.40	656	***ENTPD7***	0.00166	2.30
600	***TBC1D19***	0.00607	2.39	657	ERICH5	0.01185	2.30
601	***BDH2***	0.01107	2.39	658	***IGF1R***	0.00951	2.29
602	***AHNAK2***	0.01819	2.39	659	***CAP2***	0.00007	2.29
603	***ASPHD2***	0.00007	2.39	660	***TMEM57***	0.00015	2.29
604	***CYTH1***	0.00000	2.38	661	***UBA6-AS1***	0.00003	2.28
605	***EPHA4***	0.00711	2.38	662	***CCDC74B***	0.00077	2.28
606	***SLFN12***	0.00373	2.37	663	***P4HA2-AS1***	0.04429	2.28
607	***FAM177A1***	0.00033	2.37	664	***KLHL26***	0.00624	2.28
608	***EHBP1***	0.00068	2.36	665	***CGN***	0.00323	2.28
609	***P4HA2***	0.00066	2.36	666	***RAB5A***	0.02058	2.28
610	***CCSER2***	0.03227	2.36	667	HOXC8	0.02648	2.27
611	***TPD52***	0.00443	2.36	668	***RBM27***	0.04440	2.27
612	***PFKFB3***	0.00024	2.36	669	***KLHL25***	0.00000	2.27
613	***PTK7***	0.00004	2.36	670	KCNQ5	0.04276	2.27
614	***LHFP***	0.00691	2.36	671	***ITGA5***	0.00000	2.27
615	***STC1***	0.02228	2.35	672	***LAMP2***	0.00230	2.27
616	***FEZ2***	0.02747	2.35	673	***FURIN***	0.00004	2.26
617	KIAA0922	0.00142	2.35	674	PADI1	0.00230	2.26
618	***GPX8***	0.00003	2.35	675	***ERV3-1***	0.00006	2.26
619	***TSC22D3***	0.00004	2.35	676	***MAGT1***	0.00478	2.26
620	WNT3	0.00122	2.35	677	BEND4	0.01835	2.26
621	***CAMSAP2***	0.02082	2.35	678	TMEM117	0.00275	2.26
622	***KDM5B***	0.00031	2.35	679	***CNPY4***	0.00002	2.26
623	***PELI1***	0.01402	2.35	680	***ARFGAP1***	0.00000	2.25

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
681	***CITED4***	0.00117	2.25	738	***SH3BGRL***	0.01027	2.16
682	ZNF618	0.01376	2.25	739	***CHST7***	0.00003	2.16
683	***LMO4***	0.00342	2.25	740	***TNFSF9***	0.00000	2.16
684	***LRP11***	0.00791	2.24	741	***IGFBP4***	0.00003	2.16
685	***DOCK9***	0.00335	2.24	742	***LAMC1***	0.02064	2.16
686	***KDELC2***	0.00259	2.24	743	***ZNF627***	0.00155	2.15
687	TNFAIP8L3	0.00640	2.24	744	***MPZL3***	0.03492	2.15
688	***GADD45G***	0.00002	2.24	745	***WBP5***	0.00316	2.15
689	***PICALM***	0.00054	2.23	746	ZNF570	0.04719	2.15
690	ZNF853	0.01391	2.23	747	***PDLIM7***	0.00000	2.15
691	DMRTA2	0.02313	2.23	748	***KIRREL***	0.00020	2.14
692	***KIF3C***	0.00001	2.23	749	NEXN	0.02497	2.14
693	***PRKACB***	0.03055	2.23	750	***RTN4***	0.00092	2.14
694	***C3orf52***	0.00537	2.22	751	***INPPL1***	0.00102	2.14
695	***PROS1***	0.01801	2.22	752	***SOCS6***	0.02504	2.14
696	***EXT1***	0.00038	2.22	753	***PBXIP1***	0.00142	2.14
697	***MFAP3***	0.00673	2.22	754	***PLD1***	0.00083	2.14
698	***LOXL4***	0.00200	2.22	755	***C5orf24***	0.00211	2.14
699	***KLF10***	0.00658	2.22	756	***SDCCAG8***	0.00082	2.13
700	***BTG1***	0.00225	2.21	757	***DDX6***	0.00718	2.13
701	***FOCAD***	0.00000	2.21	758	DAB2	0.00985	2.13
702	ZNF93	0.02161	2.21	759	***SLC16A1***	0.01064	2.13
703	***RSPRY1***	0.00184	2.21	760	***BMPR1A***	0.00652	2.12
704	***SHC2***	0.00065	2.21	761	***TUBA1A***	0.00022	2.12
705	***SIPA1L1***	0.00266	2.20	762	***PXDC1***	0.00116	2.12
706	***AFF4***	0.04205	2.20	763	GALNT6	0.01060	2.12
707	***KCNQ2***	0.00388	2.20	764	***ITSN2***	0.01712	2.12
708	***SNAI1***	0.00060	2.19	765	***SCARB2***	0.00152	2.12
709	***SPATS2***	0.00003	2.19	766	***APLF***	0.03075	2.11
710	***UBASH3B***	0.00010	2.19	767	***FUT8-AS1***	0.00137	2.11
711	***SLC35F2***	0.00020	2.19	768	***KIF16B***	0.00813	2.11
712	***TMSB4X***	0.00000	2.19	769	***CCDC50***	0.00088	2.11
713	***MT1L***	0.00000	2.18	770	***TTYH3***	0.00010	2.11
714	***SLC22A17***	0.00237	2.18	771	***ACTR2***	0.00891	2.11
715	***VWA1***	0.00004	2.18	772	***PDIA3***	0.01953	2.11
716	***PVRL3***	0.00248	2.18	773	***GABARAPL1***	0.00502	2.11
717	***PLAUR***	0.00000	2.18	774	***IER3***	0.00000	2.11
718	CGNL1	0.03804	2.18	775	***SSR3***	0.00460	2.11
719	***PRR15***	0.00050	2.18	776	***SLC41A2***	0.00025	2.10
720	ITGA1	0.01093	2.18	777	***AAMDC***	0.00006	2.10
721	***TMEM237***	0.01845	2.18	778	***PAWR***	0.00375	2.10
722	ULBP1	0.00299	2.17	779	***ARHGEF28***	0.00340	2.10
723	***JADE3***	0.01122	2.17	780	NKD1	0.00330	2.09
724	***WTIP***	0.01745	2.17	781	***TMCC2***	0.00003	2.09
725	ANO4	0.00864	2.17	782	***MFHAS1***	0.04354	2.09
726	***YTHDF1***	0.00005	2.17	783	OPRL1	0.00931	2.09
727	***VDR***	0.00003	2.17	784	ADAM23	0.01802	2.09
728	***PWWP2A***	0.00092	2.17	785	KANSL1L	0.00621	2.09
729	***ECE1***	0.00067	2.17	786	***RNF170***	0.01349	2.09
730	ITPR2	0.04546	2.16	787	***HSF2BP***	0.00003	2.09
731	MAML3	0.00364	2.16	788	***SLC17A5***	0.00545	2.09
732	***KHDRBS3***	0.00002	2.16	789	***SERINC3***	0.00795	2.08
733	***CPT1C***	0.00015	2.16	790	***CCNJL***	0.00002	2.08
734	***TSPYL4***	0.00004	2.16	791	***PRRC1***	0.03556	2.08
735	***TMEM165***	0.00000	2.16	792	***NRBF2***	0.03001	2.08
736	***PEA15***	0.00000	2.16	793	***C5orf15***	0.00036	2.08
737	***CHRNA5***	0.00027	2.16	794	***EPHB3***	0.00018	2.08

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
795	***ACTN1***	0.00002	2.08	852	TCP11L1	0.00001	2.00
796	***SEL1L***	0.00552	2.07	853	ZBTB47	0.00002	2.00
797	***DNAJC10***	0.01555	2.07	854	FKBP9	0.03660	2.00
798	MANSC1	0.00821	2.07	855	MMP17	0.00005	2.00
799	***WIPI1***	0.00002	2.07	856	DBN1	0.00000	1.99
800	***FAM200B***	0.00441	2.07	857	ADAM10	0.04147	1.99
801	***SPSB1***	0.00018	2.07	858	ZNRF2P1	0.01312	1.99
802	***AMACR***	0.00015	2.07	859	C9orf91	0.00001	1.99
803	***CTNNB1***	0.00052	2.06	860	RAB1A	0.01410	1.99
804	***PKIG***	0.00002	2.06	861	JAM3	0.00026	1.99
805	***ARL3***	0.03695	2.06	862	PHTF2	0.01695	1.99
806	***CPE***	0.04987	2.06	863	PLSCR4	0.01014	1.98
807	***STIM2***	0.00151	2.06	864	ARMCX3	0.00505	1.98
808	***GLCE***	0.03966	2.06	865	FAM171A2	0.00035	1.98
809	***LOC344887***	0.00006	2.05	866	ABR	0.00000	1.98
810	***PGBD1***	0.01579	2.05	867	IMPAD1	0.04600	1.98
811	***SPIN1***	0.00619	2.05	868	GDF11	0.00103	1.98
812	***RAP1B***	0.03106	2.05	869	ORAI2	0.00458	1.98
813	***CAMKK1***	0.00904	2.05	870	ME1	0.00095	1.98
814	***RBFOX2***	0.00332	2.05	871	LRCH1	0.01061	1.98
815	***LINC00294***	0.00243	2.04	872	LIMK2	0.00059	1.97
816	***PTPRE***	0.00140	2.04	873	DNAH17-AS1	0.03611	1.97
817	***MEF2A***	0.04740	2.04	874	CNN3	0.00210	1.97
818	***JAK1***	0.03132	2.04	875	ERAP1	0.04828	1.97
819	LOC541472	0.03679	2.03	876	TMCO1	0.02853	1.97
820	***SEC14L2***	0.00297	2.03	877	RAB23	0.04448	1.97
821	***OSBPL1A***	0.01181	2.03	878	VCL	0.00485	1.97
822	***ACTR3***	0.03040	2.03	879	PPIC	0.00314	1.97
823	***HSD17B12***	0.01203	2.03	880	ATP6V1G1	0.02128	1.97
824	***CTXN1***	0.00041	2.03	881	SIRPA	0.00253	1.97
825	***ATP6AP2***	0.00363	2.03	882	UNC5A	0.00687	1.97
826	***ARID2***	0.02821	2.03	883	PCSK7	0.00037	1.97
827	***VLDLR***	0.03854	2.02	884	TMSB4Y	0.03990	1.97
828	***LOC284454***	0.00406	2.02	885	TBKBP1	0.00115	1.96
829	***RSU1***	0.00001	2.02	886	GNAQ	0.02317	1.96
830	***GAS6***	0.00007	2.02	887	NXPE3	0.00219	1.96
831	***CCNE1***	0.00685	2.02	888	LAPTM4A	0.00241	1.96
832	***CERS6***	0.00632	2.02	889	PCDHA6	0.01841	1.96
833	***HSPB8***	0.00009	2.02	890	NOTCH2	0.00741	1.96
834	***SGPL1***	0.00111	2.02	891	UBXN4	0.00954	1.95
835	***ABI1***	0.01963	2.02	892	CARD6	0.00032	1.95
836	***TCF12***	0.03606	2.01	893	PALM	0.00053	1.95
837	***POMT2***	0.00021	2.01	894	B4GALT4	0.01572	1.95
838	***CHMP4C***	0.00709	2.01	895	LRRC49	0.02170	1.95
839	NOL4L	0.00015	2.01	896	BEND7	0.00058	1.94
840	***HACD1***	0.00004	2.01	897	TPST1	0.00001	1.94
841	***TMEM41B***	0.01276	2.01	898	LRRC8A	0.00018	1.94
842	DNAL1	0.02609	2.01	899	MR1	0.00037	1.94
843	***RASA1***	0.02085	2.01	900	RWDD2A	0.00111	1.93
844	***DCAF5***	0.00001	2.00	901	FAM161B	0.03804	1.93
845	***TAF9B***	0.00526	2.00	902	MAP1LC3A	0.00024	1.93
846	***ARPC5***	0.00462	2.00	903	PPM1A	0.04035	1.93
847	***CHN1***	0.01858	2.00	904	RNASE4	0.01653	1.93
848	MMP11	0.00512	2.00	905	FHL3	0.00001	1.93
849	RAP1GDS1	0.00006	2.00	906	TGOLN2	0.02938	1.93
850	NCOA1	0.01276	2.00	907	IGSF3	0.03779	1.93
851	NPTN	0.00182	2.00	908	FOXD1	0.04367	1.93

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
909	ARID1B	0.00386	1.92	966	DSC3	0.04310	1.85
910	NCKAP5L	0.00096	1.92	967	CD70	0.00000	1.85
911	VPS54	0.01981	1.92	968	ARCN1	0.01956	1.85
912	SGCE	0.03802	1.92	969	ITSN1	0.00274	1.85
913	TMEM50B	0.03731	1.92	970	PON2	0.00015	1.85
914	DIP2B	0.01656	1.91	971	OSTC	0.04346	1.84
915	BACE1	0.00057	1.91	972	SLC39A7	0.00205	1.84
916	AFF1	0.01438	1.91	973	LOC100506071	0.00230	1.84
917	CRTC1	0.00023	1.91	974	MORF4L1	0.00680	1.84
918	FARP1	0.00010	1.91	975	OLFM1	0.00009	1.84
919	KLF12	0.02802	1.91	976	MARK1	0.03339	1.84
920	SLC44A1	0.02902	1.91	977	PLAU	0.00064	1.83
921	LOC642852	0.02848	1.91	978	EFHC1	0.00198	1.83
922	HFE	0.00042	1.91	979	ITM2B	0.00004	1.83
923	ATG12	0.00757	1.90	980	NCOA4	0.00461	1.83
924	DNAJB6	0.00078	1.90	981	DDX26B	0.03110	1.83
925	SLC39A13	0.00000	1.90	982	LGMN	0.00048	1.83
926	SLC22A4	0.00207	1.90	983	SAMD11	0.00075	1.83
927	ABHD4	0.00011	1.89	984	PAPPA2	0.02941	1.83
928	JUN	0.02173	1.89	985	MBD5	0.03022	1.83
929	TM2D3	0.00064	1.89	986	AMOTL1	0.00491	1.83
930	CCDC93	0.00253	1.88	987	KCNK6	0.01686	1.83
931	PPFIA1	0.00095	1.88	988	TFDP2	0.01582	1.82
932	S1PR5	0.00147	1.88	989	ZFC3H1	0.02143	1.82
933	COL6A2	0.00043	1.88	990	TBC1D1	0.01505	1.82
934	FBXL3	0.04885	1.88	991	HOMER1	0.01642	1.82
935	MED13L	0.01526	1.88	992	LINC00648	0.00706	1.82
936	JADE1	0.01217	1.88	993	ZC3H12A	0.00001	1.81
937	CLN5	0.00062	1.88	994	LRRC16A	0.00516	1.81
938	KIF26B	0.03668	1.88	995	EHHADH	0.01442	1.81
939	PARD6G	0.00450	1.88	996	ATXN1	0.00683	1.81
940	CTIF	0.00012	1.88	997	PCGF3	0.01452	1.81
941	CROCC	0.01223	1.88	998	FAM43A	0.01935	1.81
942	LOC389831	0.02837	1.88	999	MBNL1-AS1	0.04020	1.81
943	NATD1	0.00014	1.87	1000	YWHAZ	0.04301	1.81
944	TMEM200B	0.04029	1.87	1001	CDC42SE1	0.00011	1.81
945	ATRNL1	0.04619	1.87	1002	POFUT2	0.00019	1.81
946	PCNXL4	0.04932	1.87	1003	TCAF1	0.00980	1.81
947	ZNF821	0.00063	1.87	1004	RNASEL	0.03264	1.80
948	STAU2	0.00731	1.87	1005	TAOK3	0.01575	1.80
949	MAN2A1	0.00795	1.87	1006	ERC1	0.00749	1.80
950	TMEM132A	0.00093	1.87	1007	PCED1B	0.00130	1.80
951	APP	0.01461	1.87	1008	RPS6KC1	0.01825	1.80
952	SPICE1	0.04980	1.87	1009	STX2	0.01322	1.80
953	HDX	0.00312	1.87	1010	ICAM1	0.01705	1.80
954	TXNDC15	0.00077	1.86	1011	HSPB1	0.00002	1.80
955	TRIO	0.00145	1.86	1012	TM7SF3	0.02699	1.79
956	CD46	0.00799	1.86	1013	SRC	0.00016	1.79
957	MSL3	0.00344	1.86	1014	PCNX	0.00081	1.79
958	B3GLCT	0.01755	1.86	1015	BBS9	0.00006	1.79
959	TMX4	0.01879	1.85	1016	ZFAND5	0.04114	1.79
960	GRAMD3	0.03010	1.85	1017	NLK	0.00001	1.79
961	SH3PXD2B	0.02060	1.85	1018	PDLIM3	0.00001	1.79
962	MT1X	0.00021	1.85	1019	PLSCR3	0.00527	1.79
963	WDR11	0.01067	1.85	1020	OSBPL10	0.00265	1.78
964	PHF21A	0.00553	1.85	1021	PSMD2	0.00025	1.78
965	SP100	0.01456	1.85	1022	SMURF1	0.00153	1.78

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1023	VMP1	0.01821	1.78	1080	DLG5	0.01201	1.72
1024	SV2A	0.00318	1.78	1081	DNMBP	0.01132	1.72
1025	IL11RA	0.01191	1.78	1082	MYO1E	0.02679	1.72
1026	GALNT2	0.01100	1.78	1083	AVL9	0.00549	1.72
1027	EVC2	0.00219	1.78	1084	TRAK2	0.01289	1.72
1028	PNMA1	0.01070	1.78	1085	BCAR3	0.02964	1.72
1029	RASA3	0.00033	1.78	1086	MAGED1	0.00956	1.72
1030	HOXB2	0.04536	1.78	1087	SPPL2A	0.03668	1.72
1031	FGD4	0.01036	1.77	1088	GSN	0.00075	1.72
1032	PLCG1	0.00037	1.77	1089	KCNMA1	0.00082	1.71
1033	ZNF561	0.04950	1.77	1090	BBS4	0.00000	1.71
1034	ABL1	0.00024	1.77	1091	SLC38A9	0.02791	1.71
1035	ADAMTS7	0.03069	1.77	1092	ABTB2	0.00039	1.71
1036	CDYL	0.00001	1.77	1093	LMBRD1	0.02935	1.71
1037	KIAA1841	0.00290	1.77	1094	ELOVL5	0.00457	1.71
1038	AP5M1	0.03769	1.77	1095	PLXNB1	0.00012	1.71
1039	MAP1LC3B	0.01791	1.77	1096	HCG18	0.00254	1.71
1040	DNAJB5	0.00000	1.77	1097	KLHL20	0.00320	1.71
1041	HDAC7	0.00001	1.77	1098	PRSS23	0.03126	1.70
1042	SPATS2L	0.02866	1.76	1099	SLC6A8	0.04355	1.70
1043	SEPT2	0.00914	1.76	1100	FEZ1	0.00024	1.70
1044	SDC3	0.00210	1.76	1101	SEC31A	0.00195	1.70
1045	AP3B1	0.02501	1.76	1102	GPR176	0.02187	1.70
1046	XXYLT1	0.00605	1.76	1103	P3H1	0.02245	1.70
1047	FAM126B	0.04070	1.76	1104	PRKD1	0.02187	1.70
1048	MEX3D	0.02735	1.76	1105	ZYX	0.00039	1.70
1049	ERCC6-PGBD3	0.00918	1.76	1106	SMTN	0.00296	1.70
1050	LOC101927204	0.03383	1.75	1107	ZNF319	0.00138	1.69
1051	C9orf3	0.00006	1.75	1108	TCTN2	0.00854	1.69
1052	CCDC92	0.00237	1.75	1109	MARCH5	0.00352	1.69
1053	CCDC109B	0.00249	1.75	1110	APH1B	0.00990	1.69
1054	PTK2	0.00092	1.75	1111	GLIS3	0.00002	1.69
1055	FBXL5	0.00175	1.75	1112	ORMDL3	0.00033	1.68
1056	GPC2	0.00897	1.75	1113	CD40	0.00204	1.68
1057	NDRG4	0.04772	1.74	1114	RALB	0.02418	1.68
1058	FCHSD2	0.00003	1.74	1115	RAB3GAP2	0.04655	1.68
1059	ENO3	0.00046	1.74	1116	PARP8	0.01375	1.68
1060	CAPN5	0.00381	1.74	1117	SUMF1	0.00935	1.68
1061	TM2D2	0.00042	1.74	1118	S1PR2	0.01706	1.68
1062	SYNPO	0.00248	1.74	1119	IRAK4	0.01465	1.68
1063	FOSL2	0.02506	1.74	1120	TRAPPC10	0.02350	1.68
1064	LAPTM4B	0.01477	1.74	1121	MEIS3	0.00977	1.67
1065	VOPP1	0.00047	1.74	1122	CAP1	0.00064	1.67
1066	CADM4	0.03641	1.74	1123	BMP1	0.00493	1.67
1067	ZSCAN30	0.04695	1.74	1124	TUSC3	0.01948	1.67
1068	CSGALNACT2	0.01112	1.74	1125	FRMD5	0.01385	1.67
1069	BTN2A2	0.00513	1.74	1126	AKTIP	0.04109	1.67
1070	SETD7	0.01327	1.73	1127	DEGS1	0.00054	1.67
1071	UBE2J1	0.02266	1.73	1128	SERP1	0.03864	1.67
1072	MESDC2	0.00652	1.73	1129	TINAGL1	0.01032	1.67
1073	SBDS	0.00946	1.73	1130	FANCF	0.03363	1.66
1074	CNOT4	0.00079	1.73	1131	TSPAN9	0.02150	1.66
1075	MTMR3	0.00007	1.73	1132	TCTEX1D2	0.02011	1.66
1076	PDCD6IP	0.01720	1.73	1133	COG3	0.00688	1.66
1077	HBP1	0.01562	1.73	1134	FIBCD1	0.00548	1.66
1078	ENY2	0.00637	1.73	1135	WSB1	0.01302	1.66
1079	USP54	0.00041	1.73	1136	SMAP1	0.00093	1.66

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1137	CSNK1G3	0.01876	1.66	1194	MT2A	0.00311	1.58
1138	DNAH5	0.00541	1.66	1195	HERPUD2	0.02346	1.58
1139	B3GNT9	0.01431	1.66	1196	PAPSS1	0.01466	1.58
1140	AZI2	0.02527	1.66	1197	UBE2K	0.02795	1.58
1141	DSTN	0.00286	1.66	1198	STXBP4	0.00615	1.58
1142	PEX2	0.04158	1.66	1199	VIM	0.00182	1.58
1143	SCPEP1	0.02767	1.65	1200	LEPROT	0.04245	1.57
1144	GOLPH3L	0.02816	1.65	1201	C11orf30	0.02790	1.57
1145	HLA-DMB	0.03160	1.65	1202	MAPK7	0.00361	1.57
1146	LTBP4	0.00912	1.65	1203	NGF	0.00065	1.57
1147	VTI1A	0.00586	1.65	1204	TMEM59	0.00593	1.57
1148	SSBP3	0.00065	1.65	1205	FAM219B	0.01608	1.56
1149	MAPKBP1	0.00018	1.65	1206	IRF2BPL	0.04650	1.56
1150	FMNL3	0.00022	1.65	1207	OCIAD1	0.00642	1.56
1151	CLDND1	0.02297	1.65	1208	NUTM2B-AS1	0.00541	1.56
1152	MPZL1	0.00512	1.65	1209	PSD3	0.04071	1.56
1153	TMED9	0.01710	1.64	1210	DDAH2	0.00143	1.56
1154	SIK2	0.04044	1.64	1211	BFAR	0.00038	1.55
1155	ZMYM4	0.01632	1.64	1212	TTC30A	0.00015	1.55
1156	C7orf73	0.01282	1.64	1213	C15orf57	0.00234	1.55
1157	DHX32	0.00977	1.64	1214	CALM2	0.01498	1.55
1158	USP35	0.00043	1.64	1215	SH3KBP1	0.03938	1.55
1159	PPP3CB	0.00369	1.63	1216	ZDHHC6	0.02076	1.55
1160	RUFY1	0.01465	1.63	1217	STEAP3	0.00015	1.55
1161	ATP1B3	0.02918	1.63	1218	ARPC2	0.03536	1.54
1162	DGCR8	0.00148	1.63	1219	KDELR3	0.00083	1.54
1163	TMEM50A	0.00316	1.63	1220	TAF7	0.03481	1.54
1164	BTBD10	0.04671	1.63	1221	PCYT1A	0.00219	1.54
1165	SLC29A4	0.02121	1.63	1222	LIPA	0.03606	1.54
1166	SVIL-AS1	0.01353	1.62	1223	UBE2F	0.01997	1.54
1167	LIMD2	0.00005	1.62	1224	BASP1	0.00507	1.53
1168	TTC8	0.01102	1.62	1225	LMBR1L	0.03551	1.53
1169	MLPH	0.00146	1.62	1226	NMB	0.02143	1.53
1170	FUBP1	0.03377	1.62	1227	CHCHD7	0.00448	1.53
1171	COCH	0.00807	1.62	1228	PLOD1	0.02671	1.53
1172	LEPROTL1	0.00658	1.61	1229	TFG	0.01189	1.52
1173	CTDSP2	0.00589	1.61	1230	CAPZB	0.00008	1.52
1174	PHACTR4	0.00050	1.61	1231	KCTD10	0.03441	1.52
1175	MSN	0.00023	1.61	1232	GALNT18	0.03514	1.52
1176	AFAP1	0.00385	1.61	1233	GNAS	0.00339	1.52
1177	DFNA5	0.00868	1.61	1234	ZNF706	0.04575	1.52
1178	FAAP100	0.00000	1.61	1235	ABCC5	0.00552	1.52
1179	ACO1	0.00742	1.61	1236	EVL	0.00898	1.52
1180	TBC1D2B	0.02526	1.61	1237	ADORA1	0.01199	1.52
1181	MAP7D1	0.00017	1.60	1238	MANBA	0.04460	1.51
1182	SLC25A37	0.00007	1.60	1239	GALNT11	0.00192	1.51
1183	SERPINH1	0.02243	1.60	1240	WBP1L	0.00238	1.51
1184	GLT8D1	0.02662	1.60	1241	APPL2	0.02222	1.51
1185	DGKD	0.00845	1.60	1242	CD151	0.00167	1.50
1186	MAN1A2	0.00123	1.60	1243	CMIP	0.03898	1.50
1187	STK24	0.00182	1.60	1244	ADCY6	0.00862	1.50
1188	SEPT15	0.02952	1.60	1245	ZSWIM6	0.03321	1.49
1189	CTNNA1	0.03597	1.59	1246	FHL1	0.03759	1.49
1190	NDEL1	0.00224	1.59	1247	KLC4	0.00612	1.49
1191	FLNA	0.04853	1.59	1248	MFSD11	0.01478	1.49
1192	PTPRA	0.02996	1.59	1249	ALS2	0.03601	1.49
1193	SPHK1	0.00184	1.58	1250	TNFAIP1	0.01899	1.49

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1251	NDFIP1	0.01127	1.49	1308	DOK4	0.02404	1.35
1252	TAF1B	0.02219	1.49	1309	RBM4B	0.02646	1.35
1253	IFNGR2	0.00072	1.49	1310	FAM32A	0.02772	1.35
1254	STX12	0.00234	1.49	1311	RBM4	0.03585	1.34
1255	CCAR1	0.02328	1.48	1312	HOMER3	0.00826	1.33
1256	RBL2	0.03152	1.48	1313	RARA	0.01170	1.33
1257	DGKA	0.01121	1.48	1314	SMYD3	0.01228	1.33
1258	CDC23	0.00507	1.48	1315	ICA1	0.01172	1.32
1259	CERCAM	0.00354	1.48	1316	LINC00152	0.00839	1.31
1260	CERK	0.01141	1.48	1317	PDE6D	0.00182	1.31
1261	PTPN9	0.01323	1.48	1318	TRIM16L	0.03597	1.30
1262	CTPS1	0.01987	1.48	1319	TSPAN15	0.02433	1.30
1263	CALCOCO1	0.00114	1.47	1320	SNX1	0.01356	1.27
1264	PLA2G15	0.00325	1.47	1321	CETN2	0.04094	1.27
1265	ATG9A	0.03711	1.47	1322	ERCC3	0.03520	1.25
1266	RNF14	0.00276	1.47	1323	FUNDC2	0.02693	1.22
1267	AP3M2	0.01588	1.47	1324	REEP5	0.04121	1.20
1268	IRF7	0.00008	1.47	1325	MRPS15	0.03257	-1.23
1269	TRAF3	0.03432	1.46	1326	MRPL43	0.01586	-1.23
1270	PACSIN2	0.00932	1.46	1327	PNKP	0.02443	-1.24
1271	HLA-E	0.03840	1.46	1328	PPAP2C	0.01119	-1.24
1272	ATG16L1	0.00631	1.45	1329	DHX30	0.01783	-1.24
1273	KIFC3	0.00090	1.45	1330	KRTCAP2	0.01602	-1.24
1274	SYDE1	0.00017	1.45	1331	COX5A	0.00503	-1.24
1275	LDB1	0.00002	1.45	1332	RPL19	0.02029	-1.25
1276	SFXN3	0.00272	1.45	1333	SNRPA	0.01751	-1.25
1277	ATP8B2	0.04536	1.45	1334	C19orf43	0.00755	-1.25
1278	ZSWIM4	0.01764	1.44	1335	THAP4	0.02035	-1.25
1279	SNURF	0.01367	1.44	1336	HCFC1R1	0.00566	-1.25
1280	PAFAH1B1	0.04647	1.44	1337	COMMD4	0.00882	-1.25
1281	CHPF2	0.03349	1.44	1338	RPL13A	0.00800	-1.26
1282	TWF2	0.00067	1.44	1339	COX8A	0.01377	-1.26
1283	TMTC4	0.03144	1.44	1340	SH3GLB2	0.03471	-1.26
1284	CACFD1	0.00303	1.44	1341	COQ9	0.02244	-1.26
1285	FSCN1	0.00250	1.44	1342	FADS3	0.02254	-1.26
1286	WWC3	0.04937	1.44	1343	TRMT2A	0.01970	-1.26
1287	TMEM44	0.00037	1.43	1344	TOMM22	0.03149	-1.26
1288	GRN	0.04885	1.43	1345	ELP5	0.03170	-1.26
1289	PIAS3	0.01023	1.43	1346	SMYD5	0.01925	-1.26
1290	RFWD2	0.01116	1.43	1347	RANGRF	0.00101	-1.26
1291	CEP68	0.00035	1.42	1348	RAC3	0.02790	-1.26
1292	RNF215	0.03179	1.41	1349	FIS1	0.02617	-1.26
1293	NPLOC4	0.03103	1.41	1350	POR	0.03129	-1.26
1294	UBTD1	0.00867	1.41	1351	RPS26	0.02686	-1.27
1295	STX1A	0.03749	1.41	1352	FBL	0.04886	-1.27
1296	PPP1R21	0.03812	1.40	1353	ATRAID	0.02200	-1.27
1297	DNASE1L1	0.00424	1.40	1354	POLR2H	0.04792	-1.27
1298	PPP1R18	0.00056	1.40	1355	EBP	0.03841	-1.27
1299	MAGED2	0.02023	1.40	1356	PSME2	0.00461	-1.27
1300	ANXA2R	0.00144	1.39	1357	E4F1	0.01253	-1.27
1301	ACOX3	0.00134	1.38	1358	ADRM1	0.01500	-1.27
1302	ABCD4	0.00927	1.37	1359	PSMB3	0.02432	-1.27
1303	RRAS	0.00744	1.37	1360	RPL32	0.02295	-1.27
1304	RUSC2	0.00026	1.37	1361	MRPS11	0.01387	-1.28
1305	RXRB	0.01045	1.36	1362	FIBP	0.01299	-1.28
1306	LOC220729	0.04024	1.36	1363	CPNE1	0.00827	-1.28
1307	PROCR	0.02743	1.35	1364	POLR2E	0.02142	-1.28

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1365	UQCC1	0.01928	-1.28	1422	HIGD2A	0.00011	-1.33
1366	C11orf98	0.01838	-1.28	1423	NIT1	0.02995	-1.33
1367	NSMCE1	0.02638	-1.28	1424	RPS10	0.04319	-1.33
1368	NADSYN1	0.01797	-1.28	1425	SLC37A4	0.00166	-1.33
1369	RPS13	0.01611	-1.28	1426	TRIM11	0.00728	-1.33
1370	MTFP1	0.03826	-1.28	1427	NDUFB10	0.03450	-1.33
1371	TRAF7	0.00643	-1.28	1428	DAGLB	0.01784	-1.33
1372	NOP56	0.01432	-1.28	1429	FLII	0.01764	-1.33
1373	DEF8	0.01951	-1.29	1430	NR1H2	0.01100	-1.33
1374	CPSF4	0.03446	-1.29	1431	COQ4	0.00814	-1.33
1375	IMPDH2	0.00640	-1.29	1432	ASPSCR1	0.00809	-1.33
1376	C19orf48	0.00285	-1.29	1433	ZDHHC12	0.00161	-1.33
1377	ADSL	0.00285	-1.29	1434	PMS2P1	0.04644	-1.33
1378	MRPL27	0.03886	-1.29	1435	SLC25A39	0.00145	-1.33
1379	GAPDH	0.01147	-1.29	1436	ARPC1B	0.01220	-1.33
1380	HAUS5	0.02383	-1.29	1437	ZNRD1	0.00042	-1.33
1381	EIF2D	0.02435	-1.29	1438	SEPT9	0.00886	-1.33
1382	RANGAP1	0.04355	-1.29	1439	NCAPH2	0.01489	-1.33
1383	SURF1	0.04252	-1.29	1440	NTMT1	0.00447	-1.33
1384	PLEKHJ1	0.01375	-1.29	1441	MYO19	0.00061	-1.34
1385	ANAPC2	0.01269	-1.30	1442	ECHS1	0.01662	-1.34
1386	ABHD14B	0.03406	-1.30	1443	FUK	0.03960	-1.34
1387	PHB2	0.00084	-1.30	1444	BTBD2	0.00534	-1.34
1388	METTL22	0.02239	-1.30	1445	GLTSCR2	0.04560	-1.34
1389	UBE2M	0.02058	-1.30	1446	DALRD3	0.01976	-1.34
1390	GEMIN4	0.04158	-1.30	1447	TACO1	0.00248	-1.34
1391	OAZ1	0.00613	-1.30	1448	CAMK1	0.01223	-1.34
1392	PNKD	0.00845	-1.31	1449	CDK4	0.01214	-1.34
1393	IDH3G	0.02200	-1.31	1450	RPL29	0.03241	-1.34
1394	TEX264	0.01142	-1.31	1451	SLC25A1	0.00252	-1.34
1395	UBQLN4	0.02520	-1.31	1452	SHPK	0.00581	-1.34
1396	EIF3G	0.02782	-1.31	1453	DUS2	0.02829	-1.34
1397	RPL38	0.01892	-1.31	1454	SSNA1	0.01203	-1.34
1398	PSMG3	0.00450	-1.31	1455	DHPS	0.00451	-1.34
1399	ALAS1	0.01127	-1.31	1456	NSMF	0.01128	-1.34
1400	HINT2	0.01102	-1.31	1457	C9orf142	0.02079	-1.35
1401	ORAI3	0.00852	-1.31	1458	PCID2	0.00872	-1.35
1402	MRPL24	0.01009	-1.31	1459	CISD3	0.01301	-1.35
1403	FDX1L	0.00313	-1.32	1460	VPS52	0.00143	-1.35
1404	TRIM47	0.02275	-1.32	1461	MGMT	0.04381	-1.35
1405	FBXW5	0.01081	-1.32	1462	HS1BP3	0.00925	-1.35
1406	CSNK2B	0.03908	-1.32	1463	CDA	0.00040	-1.35
1407	PGM1	0.01176	-1.32	1464	KIF22	0.00067	-1.35
1408	DBNDD1	0.00039	-1.32	1465	RRNAD1	0.00118	-1.35
1409	TSPO	0.00259	-1.32	1466	RPS6KB2	0.00865	-1.35
1410	BOLA1	0.01743	-1.32	1467	TUBGCP2	0.00182	-1.35
1411	NOC2L	0.02508	-1.32	1468	C19orf70	0.04811	-1.35
1412	EIF3K	0.00256	-1.32	1469	CHMP2A	0.00126	-1.35
1413	FDPS	0.01237	-1.32	1470	NUBP2	0.02131	-1.35
1414	MRPL2	0.00007	-1.32	1471	RUSC1	0.00141	-1.35
1415	RPUSD2	0.01054	-1.32	1472	DESI1	0.03891	-1.35
1416	ADAM15	0.00065	-1.32	1473	SRM	0.04586	-1.35
1417	POLD1	0.00673	-1.32	1474	FAU	0.00820	-1.36
1418	ACADVL	0.02278	-1.32	1475	ALDH16A1	0.03987	-1.36
1419	EEF2	0.01449	-1.32	1476	PPP2R4	0.00014	-1.36
1420	LCMT1	0.04241	-1.32	1477	NOP2	0.00757	-1.36
1421	LTBR	0.04867	-1.32	1478	CDC20	0.02180	-1.36

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1479	HAGH	0.00029	-1.36	1536	MTG2	0.00015	-1.39
1480	TEAD4	0.00830	-1.36	1537	ITPA	0.00454	-1.39
1481	ATP5B	0.00019	-1.36	1538	FANCG	0.00218	-1.39
1482	OXLD1	0.00161	-1.36	1539	RPS21	0.00088	-1.39
1483	DANCR	0.04243	-1.36	1540	SARS2	0.00002	-1.39
1484	ATAD3A	0.00880	-1.36	1541	PAQR7	0.00822	-1.39
1485	TPRA1	0.01553	-1.36	1542	PGP	0.04258	-1.39
1486	MED11	0.00660	-1.36	1543	RPSA	0.02348	-1.39
1487	SLC25A6	0.00059	-1.36	1544	STOML2	0.00022	-1.39
1488	CEP72	0.00899	-1.36	1545	MRPL10	0.00438	-1.39
1489	WDR34	0.00241	-1.36	1546	GRWD1	0.00241	-1.39
1490	CLIC1	0.00079	-1.36	1547	TIMM17B	0.00120	-1.39
1491	DNAJA3	0.02993	-1.36	1548	AAAS	0.00046	-1.39
1492	NDOR1	0.00222	-1.36	1549	TMEM141	0.01340	-1.39
1493	VARS2	0.00390	-1.36	1550	CUL9	0.03677	-1.39
1494	ZMYND19	0.03685	-1.36	1551	COA4	0.01438	-1.39
1495	DVL2	0.00206	-1.36	1552	OGDH	0.00004	-1.39
1496	GPX4	0.00172	-1.36	1553	C20orf27	0.00157	-1.39
1497	LSM2	0.02333	-1.37	1554	PMM1	0.03821	-1.40
1498	RPLP0	0.01325	-1.37	1555	TAP1	0.00410	-1.40
1499	DCAF7	0.00071	-1.37	1556	UBE2S	0.03172	-1.40
1500	TKFC	0.00963	-1.37	1557	DPM2	0.00247	-1.40
1501	RRP1	0.04394	-1.37	1558	ECSIT	0.01667	-1.40
1502	ANXA5	0.00059	-1.37	1559	MVD	0.00141	-1.40
1503	SUV39H1	0.01142	-1.37	1560	RFC2	0.02307	-1.40
1504	NR2C2AP	0.01344	-1.37	1561	MRPL41	0.00056	-1.40
1505	EIF3D	0.03428	-1.37	1562	POLR2L	0.00047	-1.40
1506	RNH1	0.00126	-1.37	1563	PHGDH	0.02581	-1.40
1507	RPL7A	0.00402	-1.37	1564	TSPAN4	0.00040	-1.40
1508	SMPD2	0.02399	-1.37	1565	BCKDHA	0.00021	-1.40
1509	ACOT8	0.01335	-1.37	1566	POLDIP2	0.00004	-1.40
1510	AURKAIP1	0.00032	-1.37	1567	QTRT1	0.00701	-1.40
1511	RPL18	0.00371	-1.37	1568	GPI	0.01291	-1.40
1512	RAB40C	0.01415	-1.37	1569	OGFOD2	0.00006	-1.40
1513	SNX17	0.01970	-1.37	1570	POLRMT	0.01626	-1.40
1514	MRPL17	0.00138	-1.37	1571	IFRD2	0.00038	-1.40
1515	PPP2R3B	0.00756	-1.37	1572	NINJ1	0.00651	-1.40
1516	SIVA1	0.00387	-1.37	1573	TRAPPC2L	0.01682	-1.40
1517	ENO1	0.00112	-1.37	1574	MRPL37	0.00005	-1.40
1518	MPG	0.00239	-1.37	1575	RECQL4	0.00143	-1.40
1519	DUSP9	0.01406	-1.37	1576	PLEKHH3	0.00035	-1.40
1520	GPX1	0.00217	-1.38	1577	ANKRD13D	0.00298	-1.40
1521	CSTB	0.00019	-1.38	1578	C14orf80	0.00075	-1.40
1522	GINS2	0.01500	-1.38	1579	RPS9	0.01391	-1.40
1523	TTC9C	0.04758	-1.38	1580	PHF19	0.01582	-1.41
1524	GPANK1	0.00082	-1.38	1581	PELP1	0.01321	-1.41
1525	TBC1D2	0.00277	-1.38	1582	NELFE	0.00196	-1.41
1526	APOA1BP	0.00355	-1.38	1583	WDR74	0.00057	-1.41
1527	ARHGAP27	0.01328	-1.38	1584	FLOT2	0.00102	-1.41
1528	SLC9A3R2	0.02325	-1.38	1585	SNX21	0.00267	-1.41
1529	PYCR2	0.00003	-1.38	1586	RPL35	0.01852	-1.41
1530	NUDT22	0.00041	-1.38	1587	TMEM147	0.00064	-1.41
1531	MRPS26	0.00498	-1.38	1588	NOB1	0.00513	-1.41
1532	WDR54	0.00046	-1.38	1589	MRPL38	0.00087	-1.41
1533	PPAN	0.00367	-1.38	1590	IPO4	0.00090	-1.41
1534	MROH1	0.00344	-1.38	1591	MYBL2	0.00115	-1.41
1535	NOC4L	0.01824	-1.38	1592	IGFBP2	0.00188	-1.41

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1593	SNX8	0.00301	-1.42	1650	YDJC	0.00958	-1.46
1594	FOXRED1	0.00671	-1.42	1651	PSMB10	0.01360	-1.46
1595	RNASEK	0.00012	-1.42	1652	HAUS7	0.00465	-1.46
1596	CHCHD5	0.00591	-1.42	1653	COQ6	0.00097	-1.46
1597	GALM	0.01672	-1.42	1654	CARD10	0.00896	-1.46
1598	TMCO6	0.03786	-1.42	1655	SHROOM3	0.02183	-1.46
1599	RTEL1	0.00015	-1.42	1656	SLC50A1	0.00001	-1.46
1600	CCDC86	0.00780	-1.42	1657	APRT	0.00008	-1.46
1601	SDHAF1	0.00579	-1.42	1658	SRI	0.00002	-1.46
1602	CST3	0.00013	-1.42	1659	VRK3	0.00095	-1.46
1603	MRPL54	0.01171	-1.43	1660	FAH	0.01882	-1.46
1604	MEPCE	0.00001	-1.43	1661	CALHM2	0.00080	-1.46
1605	ARL2	0.00505	-1.43	1662	NAPRT	0.00432	-1.46
1606	ACD	0.00131	-1.43	1663	PMPCA	0.00013	-1.46
1607	RPS16	0.01155	-1.43	1664	AKR7A2	0.00064	-1.46
1608	GTPBP6	0.00398	-1.43	1665	TRUB2	0.00013	-1.46
1609	UQCC3	0.00203	-1.43	1666	IMP3	0.00047	-1.47
1610	DRAP1	0.00334	-1.43	1667	SELO	0.01264	-1.47
1611	CYC1	0.00072	-1.43	1668	ARRDC1	0.00002	-1.47
1612	ATP5G1	0.00056	-1.43	1669	RPL18A	0.00165	-1.47
1613	C16orf59	0.04394	-1.43	1670	SDCCAG3	0.00393	-1.47
1614	MMP24	0.00619	-1.43	1671	ADPRHL1	0.02405	-1.47
1615	C19orf60	0.02894	-1.44	1672	AIMP2	0.00003	-1.47
1616	MRPS24	0.00237	-1.44	1673	RPS18	0.00625	-1.47
1617	TIMM50	0.00056	-1.44	1674	DDX41	0.00001	-1.47
1618	TRMT112	0.00005	-1.44	1675	DDX56	0.00011	-1.47
1619	HGH1	0.00050	-1.44	1676	IRF3	0.00133	-1.47
1620	PIM2	0.00484	-1.44	1677	NMRAL1	0.00001	-1.47
1621	TUBB4B	0.00049	-1.44	1678	WBP2	0.00150	-1.47
1622	RPS28	0.00011	-1.44	1679	MYBBP1A	0.00336	-1.47
1623	RPS19	0.00360	-1.44	1680	ZBTB48	0.00083	-1.47
1624	DBNDD2	0.00486	-1.44	1681	THAP7	0.00027	-1.47
1625	LSM4	0.00003	-1.44	1682	RBKS	0.00397	-1.47
1626	SURF2	0.02123	-1.44	1683	ICT1	0.00271	-1.47
1627	FAM64A	0.00498	-1.44	1684	MAPK11	0.03404	-1.47
1628	RPS6KA1	0.00082	-1.44	1685	ADCK2	0.00305	-1.47
1629	MRPL4	0.00197	-1.44	1686	SNU13	0.00000	-1.48
1630	CIB1	0.00010	-1.44	1687	FANCA	0.00083	-1.48
1631	C19orf24	0.00328	-1.45	1688	RGS3	0.00008	-1.48
1632	MAD2L2	0.00009	-1.45	1689	ETFB	0.00225	-1.48
1633	NDUFV1	0.00006	-1.45	1690	INF2	0.00236	-1.48
1634	RNF31	0.00077	-1.45	1691	CRAT	0.00001	-1.48
1635	MRPS12	0.00047	-1.45	1692	EMG1	0.00084	-1.48
1636	C17orf53	0.00097	-1.45	1693	NSUN5	0.00885	-1.48
1637	RBP7	0.04121	-1.45	1694	PGAM5	0.00014	-1.48
1638	KLHDC4	0.00004	-1.45	1695	POLD2	0.00004	-1.48
1639	PCBD1	0.00110	-1.45	1696	EIF4EBP1	0.00129	-1.48
1640	SFXN4	0.00091	-1.45	1697	CECR5	0.00152	-1.48
1641	BCAS4	0.00377	-1.45	1698	RPL13	0.00042	-1.48
1642	SHMT2	0.00056	-1.45	1699	CDT1	0.00462	-1.48
1643	C8orf82	0.01210	-1.45	1700	SMARCD2	0.00010	-1.48
1644	PGD	0.00236	-1.45	1701	AHCY	0.00001	-1.49
1645	PRMT7	0.00577	-1.45	1702	TST	0.00075	-1.49
1646	RPL37A	0.00026	-1.45	1703	BOP1	0.00378	-1.49
1647	RRP9	0.00723	-1.45	1704	INPP5B	0.00007	-1.49
1648	TUFM	0.00003	-1.46	1705	VARS	0.00034	-1.49
1649	STUB1	0.00536	-1.46	1706	MCM7	0.00017	-1.49

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1707	RPL36	0.00100	-1.49	1764	RABGGTA	0.00004	-1.54
1708	RUVBL2	0.00003	-1.49	1765	MRPS18B	0.00045	-1.54
1709	ZP3	0.00016	-1.49	1766	CLPP	0.00010	-1.54
1710	HAX1	0.00011	-1.49	1767	MRPL12	0.00021	-1.54
1711	ACBD4	0.00311	-1.49	1768	ABHD12	0.00005	-1.54
1712	LIME1	0.00132	-1.49	1769	THOC6	0.00002	-1.54
1713	HSD3B7	0.01781	-1.50	1770	PTRH1	0.00253	-1.54
1714	METTL12	0.02788	-1.50	1771	B3GNTL1	0.03439	-1.54
1715	CLU	0.00223	-1.50	1772	HECW1	0.03426	-1.54
1716	PNPLA2	0.00267	-1.50	1773	COASY	0.00014	-1.55
1717	RPLP2	0.00003	-1.50	1774	TRAF4	0.00189	-1.55
1718	FANCE	0.00316	-1.50	1775	KCNJ11	0.03942	-1.55
1719	PHLDB3	0.01705	-1.50	1776	FARSA	0.00001	-1.55
1720	QPCTL	0.00103	-1.50	1777	KIAA0319	0.00010	-1.55
1721	ACSF3	0.00132	-1.50	1778	UBE2O	0.00110	-1.55
1722	NDUFAF3	0.00009	-1.50	1779	RPL23A	0.00000	-1.55
1723	RNASEH2A	0.00004	-1.50	1780	SPC24	0.00535	-1.55
1724	GTPBP3	0.00440	-1.50	1781	IMP4	0.00007	-1.56
1725	CABYR	0.00016	-1.50	1782	DPP7	0.00001	-1.56
1726	METTL1	0.00021	-1.50	1783	SLC25A22	0.00498	-1.56
1727	C19orf33	0.00004	-1.50	1784	TMC6	0.01095	-1.56
1728	MOCOS	0.00419	-1.51	1785	CEBPB	0.00220	-1.56
1729	FMNL1	0.03285	-1.51	1786	THAP7-AS1	0.00001	-1.56
1730	IRF5	0.00020	-1.51	1787	ZNHIT2	0.00063	-1.56
1731	NLE1	0.00863	-1.51	1788	KHK	0.00001	-1.56
1732	SLC27A3	0.00556	-1.51	1789	FAM83A	0.00016	-1.56
1733	PKMYT1	0.00619	-1.51	1790	SLC25A11	0.00010	-1.57
1734	HERC4	0.00321	-1.51	1791	TACC2	0.00145	-1.57
1735	MDH2	0.00001	-1.51	1792	MIF4GD	0.00025	-1.57
1736	NOL6	0.00150	-1.51	1793	WDR62	0.00003	-1.57
1737	SNHG15	0.04632	-1.51	1794	NANS	0.00073	-1.57
1738	LONP1	0.00346	-1.51	1795	LINC00116	0.00007	-1.57
1739	TARS2	0.00595	-1.51	1796	RPS5	0.00281	-1.57
1740	TLCD1	0.00069	-1.51	1797	FBXO2	0.03998	-1.58
1741	NDUFA7	0.00000	-1.52	1798	TPRN	0.00015	-1.58
1742	MFSD3	0.00168	-1.52	1799	OSGIN1	0.01070	-1.58
1743	STRA13	0.00403	-1.52	1800	EXOSC4	0.00283	-1.58
1744	OVCA2	0.00000	-1.52	1801	ALKBH2	0.00022	-1.58
1745	FKBP4	0.02259	-1.52	1802	PDXP	0.00000	-1.58
1746	ACO2	0.00000	-1.52	1803	NECAB3	0.00002	-1.58
1747	C9orf114	0.00000	-1.52	1804	UBE2C	0.00033	-1.58
1748	ANXA11	0.00000	-1.52	1805	PSMB8	0.00082	-1.58
1749	RPL8	0.00019	-1.52	1806	ETHE1	0.00817	-1.58
1750	NQO2	0.00050	-1.53	1807	CTU2	0.00036	-1.58
1751	EIF6	0.00012	-1.53	1808	DPM3	0.00007	-1.59
1752	PLCD3	0.00098	-1.53	1809	MRPS34	0.00001	-1.59
1753	SCARB1	0.00068	-1.53	1810	E2F1	0.00030	-1.59
1754	DUS1L	0.00004	-1.53	1811	TK1	0.00071	-1.59
1755	RPP25	0.03033	-1.53	1812	SAC3D1	0.00001	-1.59
1756	TNS3	0.00941	-1.53	1813	STK16	0.00000	-1.59
1757	CCDC51	0.00275	-1.53	1814	HDDC3	0.00073	-1.59
1758	FAM96B	0.00000	-1.53	1815	SFXN2	0.00093	-1.60
1759	DHDH	0.02378	-1.53	1816	PDCD2L	0.00053	-1.60
1760	TBRG4	0.00029	-1.53	1817	F12	0.00046	-1.60
1761	RIMS4	0.00018	-1.54	1818	SYTL1	0.00497	-1.60
1762	BRI3	0.00001	-1.54	1819	ZNF593	0.00019	-1.60
1763	RILP	0.00128	-1.54	1820	CCDC85C	0.01004	-1.60

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1821	PSME1	0.00014	-1.60	1878	SQRDL	0.00013	-1.68
1822	CDC45	0.00112	-1.60	1879	HIST2H2AC	0.04016	-1.68
1823	TTLL12	0.00000	-1.60	1880	RAD23A	0.00000	-1.68
1824	IFI35	0.01523	-1.61	1881	CRACR2B	0.02851	-1.68
1825	BIN1	0.00006	-1.61	1882	GSTO2	0.00000	-1.69
1826	MVP	0.00000	-1.61	1883	ISG15	0.00355	-1.69
1827	ITPK1	0.00014	-1.61	1884	NME3	0.00006	-1.69
1828	BAIAP2	0.00005	-1.61	1885	PCYT2	0.00001	-1.69
1829	NUDT14	0.00003	-1.61	1886	ELFN2	0.00826	-1.70
1830	HIST2H2AB	0.04245	-1.62	1887	MST1R	0.00012	-1.70
1831	RPL12	0.00010	-1.62	1888	ARHGEF16	0.00192	-1.70
1832	NDRG1	0.01397	-1.62	1889	REPIN1	0.00095	-1.70
1833	DDX28	0.00063	-1.62	1890	MRM1	0.00055	-1.70
1834	MLX	0.00010	-1.62	1891	DAGLA	0.02320	-1.70
1835	RPS2	0.00001	-1.62	1892	GHDC	0.02538	-1.71
1836	ESRRA	0.00028	-1.62	1893	TTC39C	0.00000	-1.71
1837	MUTYH	0.00069	-1.63	1894	MSH5	0.00000	-1.71
1838	LMTK3	0.04097	-1.63	1895	PITX1	0.00007	-1.71
1839	FAM207A	0.00011	-1.63	1896	LAMA5	0.03966	-1.71
1840	ZNF524	0.00993	-1.63	1897	PAM16	0.00043	-1.72
1841	GREB1L	0.03545	-1.63	1898	DTX2	0.00002	-1.72
1842	DCXR	0.00263	-1.63	1899	MCTP2	0.00675	-1.72
1843	AGFG2	0.01782	-1.63	1900	CHEK2	0.00000	-1.72
1844	MCM2	0.00003	-1.63	1901	FLOT1	0.00000	-1.73
1845	JADE2	0.00614	-1.63	1902	A4GALT	0.00083	-1.73
1846	CCDC85B	0.00136	-1.63	1903	TERC	0.02303	-1.73
1847	CENPM	0.00011	-1.64	1904	PTP4A3	0.00016	-1.73
1848	PUSL1	0.00052	-1.64	1905	RPARP-AS1	0.02728	-1.73
1849	TYSND1	0.00004	-1.64	1906	HIST1H1C	0.00674	-1.74
1850	XRCC3	0.00002	-1.64	1907	BCL3	0.00089	-1.74
1851	CLIP2	0.00003	-1.64	1908	ADRBK1	0.00001	-1.74
1852	APEH	0.00000	-1.64	1909	C9orf89	0.00001	-1.74
1853	E2F2	0.02004	-1.64	1910	HPSE	0.00111	-1.74
1854	TSEN34	0.00000	-1.64	1911	ARRB2	0.00000	-1.74
1855	CYB5A	0.00071	-1.64	1912	IGFLR1	0.01817	-1.74
1856	BLVRB	0.00051	-1.65	1913	COTL1	0.00001	-1.74
1857	PDE9A	0.00060	-1.65	1914	STON2	0.00451	-1.74
1858	ECI1	0.00001	-1.65	1915	ACY1	0.00010	-1.74
1859	HIST1H4C	0.02845	-1.66	1916	NDUFC2-KCTD14	0.02746	-1.75
1860	MYO18A	0.00010	-1.66	1917	TNFSF13	0.01660	-1.75
1861	UCP2	0.00402	-1.66	1918	MRPS2	0.00001	-1.76
1862	SSH3	0.00001	-1.66	1919	C16orf13	0.00000	-1.76
1863	MPP1	0.00184	-1.66	1920	PKN1	0.00000	-1.77
1864	ALDH4A1	0.00035	-1.66	1921	PARD6A	0.00004	-1.77
1865	POC1A	0.00024	-1.66	1922	CARS2	0.00327	-1.77
1866	TSEN54	0.00010	-1.67	1923	BIK	0.00000	-1.78
1867	ST6GALNAC4	0.00000	-1.67	1924	LMNA	0.00000	-1.78
1868	SLC25A10	0.00010	-1.67	1925	TRAP1	0.00000	-1.78
1869	PPARGC1B	0.00083	-1.67	1926	PYGB	0.00000	-1.78
1870	DUS3L	0.00014	-1.67	1927	SBNO2	0.00000	-1.78
1871	RPP25L	0.00029	-1.67	1928	SH2D5	0.02872	-1.78
1872	SGSM3	0.00001	-1.67	1929	LDOC1	0.00060	-1.79
1873	TYMSOS	0.03198	-1.67	1930	ARMC7	0.00055	-1.79
1874	HMBS	0.00000	-1.67	1931	VPS9D1-AS1	0.02401	-1.79
1875	DHRS13	0.00579	-1.68	1932	TIGD3	0.01431	-1.79
1876	MCM5	0.00001	-1.68	1933	LYPD3	0.00003	-1.80
1877	ATP8B3	0.00004	-1.68	1934	TTC39A	0.00116	-1.80

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
1935	NOP14-AS1	0.00000	-1.80	1992	TRERF1	0.00044	-1.91
1936	HIST1H1E	0.00376	-1.80	1993	CCDC103	0.00173	-1.91
1937	MFI2	0.00014	-1.80	1994	LPAR1	0.00000	-1.91
1938	TNFAIP2	0.00005	-1.81	1995	NOXA1	0.01636	-1.92
1939	TMEM205	0.00004	-1.81	1996	PRRT3-AS1	0.02224	-1.92
1940	HMGA1	0.00012	-1.81	1997	RIN1	0.01202	-1.92
1941	ZNF511	0.00000	-1.81	1998	HIST1H1B	0.00096	-1.92
1942	HIST1H3A	0.00254	-1.81	1999	HIST1H2AC	0.00187	-1.93
1943	CYSRT1	0.00165	-1.81	2000	TPCN1	0.00003	-1.93
1944	FBXO6	0.02392	-1.82	2001	HIST1H2BC	0.00194	-1.93
1945	PPL	0.00327	-1.82	2002	ACADS	0.00003	-1.93
1946	SPATA13	0.00004	-1.82	2003	HIST1H2AK	0.00250	-1.94
1947	APOL6	0.00057	-1.82	2004	ARHGAP8	0.00045	-1.94
1948	TMTC2	0.00214	-1.82	2005	ALDH3A1	0.01362	-1.94
1949	SH3RF2	0.00002	-1.83	2006	GMPR	0.00016	-1.94
1950	HIST1H4B	0.01004	-1.83	2007	STMN3	0.00344	-1.94
1951	FAHD2B	0.00002	-1.83	2008	TRIM62	0.00024	-1.94
1952	GCHFR	0.00003	-1.83	2009	HIST1H2AI	0.00111	-1.94
1953	SCO2	0.00001	-1.84	2010	LRRC61	0.00000	-1.95
1954	HIST1H2AM	0.00126	-1.84	2011	BCL2L1	0.00000	-1.95
1955	AIF1L	0.00057	-1.84	2012	RPH3AL	0.00000	-1.96
1956	MRPS6	0.00001	-1.84	2013	HSBP1L1	0.00023	-1.97
1957	PARD6B	0.00426	-1.84	2014	ITGB2	0.00720	-1.97
1958	LOC81691	0.00003	-1.84	2015	WSCD1	0.00005	-1.97
1959	GAL3ST1	0.01345	-1.84	2016	HIST1H2AB	0.00041	-1.97
1960	WNT10A	0.00795	-1.85	2017	CASP9	0.00000	-1.97
1961	RHOF	0.00000	-1.85	2018	MSLN	0.00154	-1.97
1962	ADAM11	0.00837	-1.85	2019	HIST1H4A	0.00065	-1.97
1963	GALE	0.00002	-1.85	2020	ABALON	0.00000	-1.97
1964	DNTTIP1	0.00000	-1.86	2021	PBX4	0.02333	-1.98
1965	RGL3	0.00020	-1.86	2022	HIST1H2BL	0.00071	-1.98
1966	HIST1H2AH	0.00071	-1.86	2023	HIST1H2AL	0.00929	-1.98
1967	DDN	0.02301	-1.86	2024	CLDN7	0.00036	-1.98
1968	GPRC5A	0.00000	-1.86	2025	PPP1R35	0.00000	-1.98
1969	RAC2	0.00092	-1.87	2026	PIP5KL1	0.04207	-1.99
1970	SERPINA1	0.00001	-1.87	2027	PKP3	0.00000	-2.00
1971	SYT12	0.00171	-1.87	2028	***FOXJ1***	0.00000	-2.00
1972	CRB3	0.00001	-1.87	2029	***COMTD1***	0.00011	-2.00
1973	FOXN3-AS1	0.00209	-1.87	2030	***TBL1X***	0.00653	-2.00
1974	TMEM61	0.04116	-1.87	2031	***RITA1***	0.00000	-2.00
1975	HIST1H4D	0.00512	-1.87	2032	***HIST1H2AD***	0.00720	-2.01
1976	FAM86DP	0.00556	-1.87	2033	***BDH1***	0.00007	-2.01
1977	NXN	0.00000	-1.88	2034	***MALL***	0.00000	-2.01
1978	PON3	0.01277	-1.88	2035	***BTBD11***	0.00007	-2.01
1979	DENND3	0.00000	-1.88	2036	FLJ37035	0.00349	-2.02
1980	FAM86EP	0.00449	-1.88	2037	***GLS2***	0.01743	-2.03
1981	KRT19	0.00000	-1.89	2038	***TNNT1***	0.00000	-2.03
1982	HIST1H3B	0.01141	-1.89	2039	***ANKRD39***	0.00001	-2.03
1983	ITGB2-AS1	0.01961	-1.89	2040	***PPIF***	0.00000	-2.03
1984	HIST1H4H	0.00050	-1.89	2041	***ACSS2***	0.00000	-2.04
1985	IFITM2	0.00050	-1.89	2042	PPFIBP2	0.02615	-2.04
1986	SEMA4B	0.00002	-1.90	2043	RBM47	0.00193	-2.04
1987	ECH1	0.00000	-1.90	2044	CRYM	0.01124	-2.04
1988	C16orf45	0.02399	-1.90	2045	***TGM2***	0.00000	-2.05
1989	HIST2H3D	0.02298	-1.90	2046	***TMC5***	0.00058	-2.05
1990	TFPI	0.00169	-1.90	2047	***BAG1***	0.00000	-2.06
1991	ITPKA	0.00044	-1.91	2048	***PSMB9***	0.00018	-2.06

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2049	***LOC100288798***	0.02995	-2.07	2106	***ANKRD33B***	0.01142	-2.35
2050	***HIST1H2BD***	0.00023	-2.07	2107	***PGM5P2***	0.00038	-2.35
2051	C1orf226	0.02919	-2.07	2108	***EDN2***	0.00624	-2.35
2052	ARHGEF4	0.00580	-2.07	2109	***FOSL1***	0.00032	-2.36
2053	***THEM6***	0.00000	-2.07	2110	APLN	0.00203	-2.36
2054	***OSBP2***	0.00021	-2.07	2111	***ARHGEF2***	0.00000	-2.36
2055	***LY6E***	0.00000	-2.08	2112	***HIST3H2A***	0.00042	-2.37
2056	***ENTPD6***	0.00000	-2.08	2113	***TMEM173***	0.00000	-2.37
2057	***IL6R***	0.03289	-2.08	2114	DEGS2	0.04337	-2.37
2058	***HIST1H2BJ***	0.00021	-2.09	2115	TARID	0.00015	-2.38
2059	***CTSH***	0.00004	-2.09	2116	PRKCG	0.04621	-2.39
2060	***RASGEF1A***	0.00006	-2.09	2117	***OTUD3***	0.00000	-2.41
2061	***TRABD2A***	0.00000	-2.10	2118	***TRNP1***	0.00001	-2.42
2062	***HIST1H2BO***	0.00011	-2.11	2119	SYNE3	0.00737	-2.43
2063	***ISG20***	0.00000	-2.11	2120	***KRTCAP3***	0.02349	-2.43
2064	***PHLPP1***	0.00120	-2.11	2121	STAT4	0.00011	-2.43
2065	***IFI30***	0.00070	-2.11	2122	***HIST1H2AJ***	0.02619	-2.43
2066	***HIST1H4L***	0.00023	-2.12	2123	***NR0B1***	0.00321	-2.44
2067	***TMEM238***	0.00015	-2.12	2124	***GATA2***	0.00031	-2.45
2068	***MX1***	0.00001	-2.13	2125	***HIST1H2BF***	0.00012	-2.47
2069	TMEM215	0.02183	-2.13	2126	MAP3K5	0.00023	-2.47
2070	***TNFRSF18***	0.04657	-2.14	2127	***S100A6***	0.00000	-2.47
2071	***HIST1H3I***	0.00194	-2.14	2128	ACSS1	0.04839	-2.49
2072	TGFBR3	0.00174	-2.14	2129	***CDC42EP4***	0.00000	-2.49
2073	***ADAP1***	0.00301	-2.15	2130	***EML2***	0.00000	-2.50
2074	***HIST3H2BB***	0.00019	-2.15	2131	***MAPKAPK3***	0.00000	-2.51
2075	***RAB3IL1***	0.00016	-2.16	2132	***VWA7***	0.00000	-2.51
2076	***BSPRY***	0.00009	-2.17	2133	***S100A3***	0.00000	-2.52
2077	***VSTM2L***	0.01780	-2.17	2134	***PPARG***	0.00000	-2.52
2078	***NTHL1***	0.00000	-2.18	2135	***HAS3***	0.00359	-2.52
2079	***LLGL2***	0.00000	-2.18	2136	***CHCHD10***	0.00000	-2.54
2080	***FNBP1***	0.00006	-2.18	2137	***ANKRD29***	0.00008	-2.54
2081	***UNC93B1***	0.00000	-2.19	2138	***SULT1A1***	0.00001	-2.54
2082	C10orf95	0.01075	-2.19	2139	***IRS2***	0.00010	-2.55
2083	***GCAT***	0.00000	-2.19	2140	***HIST1H2BN***	0.00006	-2.57
2084	PRPH	0.00521	-2.20	2141	***CPT1A***	0.00000	-2.60
2085	***FJX1***	0.00302	-2.20	2142	***NKD2***	0.00001	-2.60
2086	***ERBB3***	0.00010	-2.20	2143	***HIST1H2AE***	0.00002	-2.60
2087	PAG1	0.00097	-2.20	2144	***CPNE7***	0.00000	-2.63
2088	***KCNK5***	0.00000	-2.21	2145	ANK1	0.00270	-2.63
2089	***MAPK12***	0.00000	-2.21	2146	***ADAM8***	0.00033	-2.64
2090	***FBXO27***	0.00000	-2.23	2147	***KRT86***	0.00001	-2.65
2091	***APOBEC3B***	0.00292	-2.23	2148	***HOXA3***	0.00000	-2.67
2092	***SLC22A18***	0.00002	-2.23	2149	***PIM3***	0.00000	-2.67
2093	***CDC25B***	0.00000	-2.23	2150	***HIST1H3G***	0.00001	-2.67
2094	***P2RX5***	0.00000	-2.24	2151	***ABHD11***	0.00000	-2.68
2095	***HIST2H2BF***	0.04241	-2.24	2152	***EML2-AS1***	0.00014	-2.68
2096	***RAB26***	0.02736	-2.24	2153	***IGFBP6***	0.00000	-2.68
2097	***GRB7***	0.00000	-2.25	2154	***EVA1C***	0.00012	-2.69
2098	***PLA2G16***	0.00000	-2.26	2155	***PDLIM2***	0.00003	-2.69
2099	LONRF2	0.00046	-2.26	2156	***TSPAN1***	0.00002	-2.69
2100	ZCCHC2	0.00001	-2.27	2157	***ID1***	0.00001	-2.70
2101	***HIST1H3D***	0.00023	-2.29	2158	LINC01348	0.00688	-2.70
2102	SDR16C5	0.00938	-2.31	2159	***RAB17***	0.00000	-2.71
2103	***HIST1H3H***	0.00006	-2.32	2160	***ADORA2B***	0.00000	-2.72
2104	***TESC***	0.00000	-2.33	2161	CPLX1	0.02503	-2.72
2105	NPAS1	0.00144	-2.33	2162	***MBP***	0.00052	-2.76

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2163	***HIST1H2BG***	0.00009	-2.77	2220	***ALDH3B1***	0.00000	-3.48
2164	PDE4B	0.00005	-2.77	2221	***CYP4F11***	0.00000	-3.50
2165	***HPCAL1***	0.00000	-2.79	2222	***HIST1H2BI***	0.00000	-3.51
2166	***UPP1***	0.00013	-2.80	2223	KCNK3	0.00015	-3.52
2167	***AKR1B10***	0.00005	-2.81	2224	***S100A4***	0.00001	-3.59
2168	***PSCA***	0.00140	-2.81	2225	***TACSTD2***	0.00000	-3.60
2169	***HIST1H2BM***	0.00002	-2.81	2226	***ASS1***	0.00000	-3.65
2170	***FAM195A***	0.00000	-2.82	2227	***MUC1***	0.00001	-3.67
2171	***ADGRG2***	0.00002	-2.82	2228	***MTSS1L***	0.00000	-3.68
2172	***IMPA2***	0.00000	-2.83	2229	***ANKRD2***	0.00003	-3.69
2173	***SGPP2***	0.00071	-2.83	2230	ZC3H12D	0.01523	-3.70
2174	***PC***	0.00000	-2.84	2231	***MPV17L***	0.00000	-3.77
2175	S1PR4	0.00091	-2.84	2232	***OASL***	0.00732	-3.78
2176	***HPDL***	0.00000	-2.84	2233	***WBSCR27***	0.00000	-3.84
2177	***CKB***	0.00000	-2.85	2234	MMP28	0.04081	-3.86
2178	NALCN	0.00000	-2.86	2235	***SECTM1***	0.00000	-3.92
2179	***NLRP12***	0.00002	-2.87	2236	***PITPNM3***	0.00000	-3.93
2180	GFAP	0.00006	-2.87	2237	CEMIP	0.04870	-3.98
2181	***SLC16A5***	0.00000	-2.88	2238	***DUSP5***	0.00098	-4.07
2182	***TBC1D8***	0.00000	-2.88	2239	***CYP2S1***	0.00000	-4.09
2183	***GOLT1A***	0.00000	-2.89	2240	***WFDC21P***	0.00116	-4.19
2184	***WFDC2***	0.01320	-2.89	2241	***LOC100288181***	0.00000	-4.20
2185	***MROH6***	0.00000	-2.92	2242	***CXCL2***	0.00000	-4.22
2186	***AVPI1***	0.00000	-2.93	2243	***SOCS2***	0.00019	-4.24
2187	LINC00346	0.00038	-2.93	2244	***CLDN4***	0.00000	-4.25
2188	***HMHA1***	0.00000	-2.98	2245	***TJP3***	0.00001	-4.40
2189	PCSK9	0.00044	-2.99	2246	***CXCL8***	0.00421	-4.40
2190	***TFAP2C***	0.00017	-3.01	2247	***RHOV***	0.00000	-4.42
2191	***AREG***	0.04056	-3.01	2248	DHX58	0.00139	-4.48
2192	***PLIN2***	0.00002	-3.02	2249	***FA2H***	0.00000	-4.58
2193	***HIST1H1D***	0.00215	-3.02	2250	***SFTA1P***	0.00004	-4.60
2194	***HIST1H2BE***	0.02192	-3.02	2251	***CITED2***	0.00001	-4.61
2195	***CABLES1***	0.00001	-3.02	2252	LOC100133669	0.00972	-4.68
2196	***RAB38***	0.00000	-3.02	2253	***GPAT3***	0.00000	-4.72
2197	***CCDC64***	0.00002	-3.04	2254	***LAMA3***	0.00000	-4.73
2198	***DBP***	0.00091	-3.06	2255	FCMR	0.00000	-4.74
2199	***CCNA1***	0.01418	-3.07	2256	TNFRSF1B	0.00001	-4.76
2200	PLXNA2	0.00001	-3.07	2257	WDR86	0.00058	-4.82
2201	***ELMO3***	0.00000	-3.07	2258	POM121L9P	0.00111	-4.87
2202	***RASD1***	0.01265	-3.08	2259	***WISP2***	0.00000	-4.88
2203	SRPX2	0.04008	-3.10	2260	***SLC22A31***	0.00000	-4.94
2204	***C1QL1***	0.00001	-3.11	2261	GRAMD2	0.00001	-5.01
2205	UPK3B	0.00017	-3.11	2262	ATOH8	0.00006	-5.10
2206	TERT	0.00313	-3.12	2263	***KCNK15***	0.00000	-5.18
2207	***CSF1***	0.00000	-3.16	2264	***ELF3***	0.00000	-5.32
2208	***C10orf54***	0.00000	-3.17	2265	***SLCO4A1***	0.00000	-5.37
2209	C3	0.00888	-3.17	2266	***LY6K***	0.00000	-5.37
2210	TNFRSF11A	0.00479	-3.18	2267	***SUSD2***	0.00027	-5.42
2211	CD14	0.00347	-3.20	2268	***PTPRH***	0.00000	-5.46
2212	***EFHD2***	0.00000	-3.20	2269	***ADIRF***	0.00301	-5.55
2213	***LOC100506860***	0.00078	-3.27	2270	CD22	0.00041	-5.59
2214	***TMEM45B***	0.00000	-3.29	2271	***HHIPL2***	0.00000	-5.65
2215	***CLDN3***	0.00000	-3.33	2272	***SLPI***	0.00024	-5.69
2216	***AKR1C1***	0.00041	-3.37	2273	***PRODH***	0.00000	-5.76
2217	***CRABP2***	0.00000	-3.43	2274	***FBP1***	0.00000	-5.90
2218	***PLEKHG4***	0.00000	-3.44	2275	***RASL11A***	0.00000	-6.12
2219	KRT15	0.00450	-3.44	2276	DNAH11	0.00008	-6.14

Rank	Gene	Corrected p-value	FC	Rank	*Gene*	Corrected p-value	FC
2277	***CBLC***	0.00079	-6.18	2291	WBSCR28	0.00015	-9.88
2278	***ACOX2***	0.00002	-6.21	2292	***S100A9***	0.00005	-10.84
2279	***RARRES3***	0.00020	-6.57	2293	***C11orf86***	0.00000	-12.54
2280	***PLLP***	0.00000	-6.77	2294	***AQP3***	0.00000	-13.35
2281	***NAPSA***	0.00018	-6.82	2295	***SDPR***	0.00000	-13.61
2282	***SLCO4A1-AS1***	0.00000	-6.88	2296	***PTGES***	0.00000	-14.21
2283	ABCG2	0.00000	-7.25	2297	***GPX2***	0.03012	-14.46
2284	***FGFBP1***	0.00000	-7.28	2298	SCARA5	0.00000	-15.11
2285	***CYP4F3***	0.00001	-7.33	2299	***S100P***	0.00015	-20.06
2286	IL24	0.00466	-7.44	2300	ALPP	0.00000	-22.89
2287	***LCN2***	0.00000	-7.98	2301	***KRT4***	0.00010	-24.70
2288	***MB***	0.00007	-8.18	2302	KRT32	0.01098	-27.18
2289	***SCNN1A***	0.00000	-9.04	2303	ALPPL2	0.00000	-50.24
2290	MUC16	0.00029	-9.06				

A3.4 Significantly deregulated proteins in P5B3 upon stimulation with TGF-β

Rank	Protein	Corrected p-value	FC	Rank	Protein	Corrected p-value	FC
1	***VIME***	0.00000	181.84	56	***GELS***	0.00000	2.14
2	***FINC***	0.00000	19.38	57	***ANXA5***	0.00322	2.11
3	***BGH3***	0.00009	15.05	58	***HMCS1***	0.00008	2.11
4	***DPYL3***	0.00000	11.79	59	***TYB10***	0.00000	2.10
5	***CALD1***	0.00000	7.61	60	***ZYX***	0.02778	2.09
6	***TSP1***	0.00000	5.98	61	***GSTO1***	0.01402	2.09
7	***TENA***	0.00000	5.89	62	***ML12A***	0.00000	2.08
8	***PALLD***	0.00000	5.38	63	***TBB5***	0.00000	2.05
9	***FLNC***	0.00000	5.10	64	***MYL6***	0.00000	2.05
10	***LAMC2***	0.00040	5.00	65	***ATOX1***	0.00962	2.03
11	***LAMB3***	0.02921	4.51	66	***ZO1***	0.00000	2.02
12	CACO1	0.00726	4.29	67	STAT1	0.00018	2.02
13	***PDLI7***	0.00000	4.09	68	PML	0.00000	1.99
14	***TBB3***	0.00000	3.98	69	ACTZ	0.00000	1.98
15	***MYL9***	0.00000	3.80	70	AXA82	0.00001	1.98
16	***MT2***	0.00000	3.68	71	MRP	0.03424	1.98
17	***TGM2***	0.00001	3.68	72	ADHX	0.00011	1.96
18	***LMCD1***	0.00000	3.66	73	PLOD1	0.00935	1.96
19	***SH3K1***	0.00000	3.64	74	ARFG1	0.00004	1.96
20	***PACS1***	0.00000	3.24	75	ARPC5	0.00014	1.96
21	***TPM1***	0.00000	3.24	76	MT1E	0.00012	1.96
22	***INP4B***	0.00000	3.17	77	PI42C	0.00876	1.94
23	PRDBP	0.00000	3.11	78	TAGL2	0.00000	1.93
24	ITA2	0.00383	2.87	79	ARPC3	0.01371	1.93
25	***SH3L3***	0.00000	2.83	80	DCTN1	0.00731	1.90
26	***DDAH2***	0.00000	2.83	81	AKA12	0.00000	1.89
27	***P4HA2***	0.00000	2.78	82	K1C17	0.00084	1.86
28	***FSCN1***	0.00000	2.76	83	CATB	0.00000	1.84
29	TBB2A	0.00000	2.74	84	VAT1	0.00000	1.83
30	GOPC	0.02105	2.69	85	THIC	0.00000	1.81
31	***ANXA3***	0.00004	2.67	86	TPM4	0.00213	1.81
32	***CSRP2***	0.00000	2.66	87	PP14B	0.00001	1.77
33	***DREB***	0.00000	2.65	88	PLOD2	0.00000	1.74
34	***CNN2***	0.00000	2.56	89	TLN1	0.00000	1.73
35	***MYH9***	0.00000	2.54	90	UGPA	0.04892	1.71
36	***TPM2***	0.00000	2.54	91	CSRP1	0.00002	1.70
37	MAGD2	0.00302	2.54	92	PSMD2	0.00081	1.70
38	NAGK	0.00073	2.54	93	ARP2	0.00386	1.70
39	***NPC2***	0.00016	2.47	94	TB182	0.00000	1.70
40	***FLNA***	0.00062	2.44	95	AKAP2	0.00052	1.70
41	PLCG1	0.00068	2.44	96	ARC1B	0.00472	1.69
42	***TBA4A***	0.00000	2.37	97	GDIA	0.00000	1.68
43	***MOES***	0.00000	2.33	98	ILK	0.00092	1.67
44	PPR18	0.00000	2.33	99	MANF	0.00215	1.65
45	***TX1B3***	0.00000	2.29	100	CALU	0.00000	1.65
46	RAI14	0.00078	2.28	101	KCRB	0.03349	1.65
47	***HMGA1***	0.00003	2.27	102	TBA1C	0.00004	1.63
48	***BASP1***	0.00268	2.25	103	PYGB	0.00002	1.62
49	***FBLI1***	0.00000	2.24	104	CLIC4	0.00551	1.61
50	***ACTN1***	0.00000	2.20	105	DCTN2	0.00000	1.60
51	***HSPB1***	0.00000	2.19	106	UPP1	0.00053	1.60
52	ANXA6	0.00034	2.19	107	ACTB	0.00000	1.59
53	***PAWR***	0.00000	2.18	108	VINC	0.00217	1.58
54	***RRBP1***	0.00000	2.17	109	AP2A1	0.02731	1.58
55	***CNN3***	0.00011	2.17	110	MAP4	0.01799	1.56

Rank	Protein	Corrected p-value	FC	Rank	Protein	Corrected p-value	FC
111	ZO2	0.00000	1.54	168	RL18A	0.00445	-1.15
112	1433S	0.00000	1.54	169	RL5	0.02421	-1.16
113	CAP1	0.00008	1.53	170	AHSA1	0.00550	-1.16
114	ARP3	0.00014	1.53	171	LC7L2	0.01725	-1.17
115	STMN1	0.00013	1.53	172	MDHC	0.02684	-1.17
116	TPM3	0.00000	1.52	173	RL29	0.00554	-1.17
117	PA1B3	0.00001	1.52	174	RS8	0.00782	-1.19
118	SC31A	0.00083	1.50	175	TXND5	0.00480	-1.19
119	SNX12	0.00017	1.49	176	PSB1	0.04562	-1.19
120	CATZ	0.00000	1.48	177	PSA5	0.00164	-1.21
121	DAP1	0.00367	1.48	178	SODC	0.00044	-1.22
122	VIGLN	0.00005	1.47	179	RL15	0.01627	-1.22
123	BACH	0.00002	1.47	180	SRC8	0.00000	-1.23
124	DTD1	0.00005	1.47	181	TCPE	0.04756	-1.23
125	FKB10	0.00075	1.47	182	SAHH	0.00035	-1.24
126	COR1C	0.00000	1.46	183	PSB6	0.04428	-1.24
127	USO1	0.00144	1.44	184	ROA1	0.01048	-1.24
128	SERPH	0.00000	1.43	185	IRAK4	0.00034	-1.25
129	4EBP1	0.00066	1.43	186	RL21	0.02824	-1.25
130	COPG1	0.00003	1.43	187	PRSR2	0.00285	-1.25
131	CD2A1	0.00332	1.43	188	RUVB2	0.00066	-1.26
132	GNS	0.00044	1.43	189	CHM2B	0.00094	-1.26
133	HN1L	0.00475	1.42	190	SET	0.00135	-1.27
134	COPZ1	0.00881	1.42	191	PRDX1	0.00076	-1.28
135	PAK2	0.00063	1.41	192	DNJB1	0.00921	-1.28
136	NSF1C	0.00004	1.41	193	TCOF	0.00466	-1.28
137	PDLI5	0.00000	1.40	194	EPIPL	0.04719	-1.28
138	GRP78	0.00000	1.39	195	GUAD	0.03579	-1.29
139	COPB	0.00606	1.39	196	HS90A	0.00001	-1.29
140	S10AD	0.00001	1.38	197	RBM39	0.00735	-1.29
141	FA98B	0.01883	1.38	198	SNRPA	0.03255	-1.29
142	ERO1A	0.00001	1.38	199	AATC	0.01653	-1.30
143	TF65	0.00013	1.37	200	EIF3J	0.03683	-1.30
144	SCRN1	0.04790	1.36	201	LDHB	0.00004	-1.31
145	PP1A	0.00191	1.36	202	PQBP1	0.04068	-1.31
146	ANXA2	0.00000	1.35	203	RL13	0.01677	-1.32
147	TES	0.00003	1.34	204	SPEE	0.00005	-1.33
148	TYB4	0.04076	1.32	205	EFHD2	0.00401	-1.35
149	COF1	0.00171	1.31	206	RL36	0.02569	-1.36
150	1433B	0.00071	1.31	207	FA49B	0.00300	-1.36
151	MARE1	0.00026	1.31	208	AMRP	0.01499	-1.37
152	CPNS1	0.00940	1.29	209	PSA1	0.00001	-1.37
153	FKBP9	0.02160	1.29	210	ELOB	0.02459	-1.38
154	COPB2	0.00001	1.29	211	SNX2	0.00032	-1.38
155	G3BP1	0.01430	1.29	212	RSRC2	0.00244	-1.39
156	YTHD3	0.02052	1.28	213	SARG	0.02654	-1.43
157	VASP	0.00916	1.24	214	PA2G4	0.00000	-1.43
158	CLH1	0.02739	1.24	215	BIEA	0.00000	-1.45
159	SNX6	0.00204	1.23	216	RL12	0.02759	-1.45
160	PAXI	0.00054	1.22	217	PCY2	0.03423	-1.46
161	CAPZB	0.02199	1.21	218	CRIP1	0.00295	-1.47
162	CRIP2	0.01024	1.20	219	NASP	0.00004	-1.47
163	ANXA1	0.04906	1.20	220	LIMA1	0.00001	-1.49
164	EHD4	0.02634	1.20	221	YBOX3	0.01423	-1.50
165	ROA0	0.03391	1.19	222	KAD2	0.00573	-1.51
166	CAZA2	0.00012	1.18	223	MTND	0.00002	-1.52
167	SEPT2	0.00225	1.17	224	G6PD	0.00088	-1.52

Rank	Protein	Corrected p-value	FC	Rank	Protein	Corrected p-value	FC
225	RED	0.00106	-1.53	262	***K2C5***	0.00001	-2.08
226	PUR9	0.00012	-1.53	263	***IMDH2***	0.00000	-2.11
227	NUDT5	0.00007	-1.53	264	***FAS***	0.00000	-2.18
228	TKT	0.00000	-1.54	265	PTGR2	0.00001	-2.22
229	RNPS1	0.02061	-1.55	266	***AL1A3***	0.00005	-2.23
230	EF1B	0.00000	-1.56	267	DCXR	0.00591	-2.32
231	PGM2	0.01137	-1.57	268	***K2C8***	0.00000	-2.36
232	CAPG	0.00002	-1.57	269	***IL18***	0.01063	-2.37
233	SHLB2	0.00428	-1.58	270	K1C19	0.00000	-2.43
234	NDRG1	0.00073	-1.61	271	TOPK	0.00004	-2.43
235	NPM	0.00000	-1.62	272	EFHD1	0.00000	-2.55
236	TRXR1	0.00002	-1.62	273	***PDLI1***	0.00000	-2.68
237	FKBP4	0.00008	-1.64	274	***ADIRF***	0.00145	-2.71
238	CAH2	0.00017	-1.67	275	***FKBP5***	0.00000	-2.72
239	ACPH	0.00000	-1.67	276	KCRU	0.00000	-2.74
240	TACD2	0.00827	-1.68	277	***CADH1***	0.00000	-2.75
241	LG3BP	0.00070	-1.69	278	EVPL	0.00216	-2.85
242	CAYP1	0.00111	-1.69	279	***TRAD1***	0.00009	-2.88
243	EBP2	0.00000	-1.72	280	***PEPL***	0.00000	-2.91
244	C1TC	0.00102	-1.72	281	GSHR	0.00000	-3.04
245	BAIP2	0.00000	-1.74	282	***ASSY***	0.00000	-3.08
246	PRDX6	0.00000	-1.77	283	***SCEL***	0.00000	-3.52
247	PUR6	0.00000	-1.78	284	***TFR1***	0.00019	-3.65
248	PPT1	0.00537	-1.84	285	***SDPR***	0.00000	-3.80
249	GDIR2	0.00000	-1.85	286	***POMP***	0.00010	-3.87
250	IPYR	0.00000	-1.87	287	***NHRF1***	0.00000	-3.99
251	COMT	0.00024	-1.88	288	***GLNA***	0.00000	-4.10
252	LMO7	0.00027	-1.90	289	***SYUG***	0.00000	-4.19
253	HPCL1	0.02160	-1.91	290	***K1C15***	0.00096	-4.20
254	CI142	0.02238	-1.95	291	***S10AE***	0.00000	-4.44
255	PNCB	0.00000	-1.96	292	***LIMC1***	0.00005	-4.57
256	AHNK	0.00003	-1.96	293	***E41L1***	0.00000	-4.75
257	ANXA4	0.00007	-1.97	294	UPK3L	0.00000	-5.13
258	***K2C7***	0.02750	-2.00	295	***AGR2***	0.00007	-5.91
259	BRX1	0.00128	-2.02	296	***LY6D***	0.00083	-6.02
260	GOLP3	0.04358	-2.05	297	***K1C13***	0.00000	-39.42
261	***G6PI***	0.00000	-2.06				

A3.5 Significantly deregulated proteins in DU145 upon stimulation with TGF-β

Rank	Protein	Corrected p-value	FC	Rank	Protein	Corrected p-value	FC
1	***TAGL***	0.00000	8.51	56	TES	0.00000	1.63
2	***ITAV***	0.01456	8.06	57	LIMA1	0.00178	1.63
3	***QORX***	0.00000	6.71	58	CAP1	0.00002	1.62
4	***TPM1***	0.00000	6.65	59	TBB3	0.00012	1.60
5	***LMCD1***	0.00000	5.22	60	STK24	0.00015	1.59
6	***BGH3***	0.00000	4.90	61	NXP20	0.01418	1.58
7	***GAPR1***	0.02220	3.75	62	DREB	0.00000	1.58
8	PADI2	0.00021	3.74	63	TWF2	0.03928	1.57
9	***TSP1***	0.00004	3.56	64	COF1	0.01623	1.56
10	***TBA4A***	0.00000	3.52	65	MYH9	0.00001	1.56
11	INP4B	0.00702	3.51	66	TAGL2	0.00069	1.56
12	***NPC2***	0.00000	3.22	67	ML12A	0.00000	1.55
13	***P4HA1***	0.00000	3.19	68	VINC	0.00726	1.54
14	***MYL9***	0.01412	3.17	69	TIA1	0.04558	1.53
15	***FBLI1***	0.00000	2.89	70	GSTP1	0.00002	1.51
16	***P4HA2***	0.00000	2.81	71	CRIP2	0.02680	1.50
17	***MT2***	0.00000	2.56	72	FSCN1	0.00000	1.49
18	***HSPB1***	0.00000	2.54	73	HMGN1	0.02181	1.49
19	***DPYL3***	0.00000	2.52	74	SCRN1	0.01707	1.45
20	***PDLI7***	0.00009	2.45	75	PAWR	0.00543	1.44
21	***PLST***	0.00000	2.38	76	DPP9	0.00856	1.44
22	PACS1	0.00679	2.36	77	HMGB3	0.00018	1.44
23	***TYB4***	0.00000	2.34	78	FLNB	0.00168	1.43
24	***FKBP9***	0.00030	2.33	79	TX1B3	0.00008	1.41
25	***FLNA***	0.00000	2.32	80	VASP	0.00000	1.39
26	***ARFG1***	0.00664	2.25	81	ENPL	0.02118	1.38
27	***PLOD2***	0.00000	2.22	82	PICAL	0.00011	1.37
28	***FERM2***	0.00005	2.19	83	DCTN2	0.00887	1.37
29	***ACTN1***	0.00000	2.07	84	MYH10	0.00000	1.36
30	***PALLD***	0.00015	2.06	85	PAXI	0.00211	1.34
31	***SYNPO***	0.00114	2.04	86	SEPT7	0.00289	1.34
32	***CNN2***	0.00000	2.01	87	PDLI5	0.00438	1.34
33	PLOD1	0.00572	2.01	88	1433Z	0.04058	1.26
34	DNJB4	0.00269	1.98	89	LEG1	0.01802	1.26
35	CNN3	0.00044	1.89	90	PP1B	0.00012	1.25
36	AHNK2	0.00001	1.89	91	TBB2A	0.04118	1.24
37	ACTB	0.00121	1.88	92	KAP0	0.00430	1.23
38	MT1E	0.00002	1.87	93	MARE1	0.00915	1.18
39	SERPH	0.00000	1.86	94	DDX17	0.01824	-1.14
40	FHL2	0.00059	1.86	95	TXNL1	0.00875	-1.18
41	PTMS	0.00000	1.84	96	PSA3	0.03944	-1.19
42	ZYX	0.00000	1.80	97	CBR1	0.04640	-1.20
43	PPR18	0.00447	1.79	98	RL15	0.03649	-1.21
44	GRN	0.00066	1.78	99	SPTB2	0.00747	-1.22
45	F177A	0.00022	1.77	100	RL38	0.00080	-1.22
46	IMA3	0.01524	1.77	101	RL3	0.03381	-1.23
47	CALD1	0.03002	1.75	102	RUVB2	0.00016	-1.23
48	UBA6	0.00331	1.74	103	RL17	0.03033	-1.23
49	GELS	0.00657	1.72	104	RSSA	0.04133	-1.24
50	MYPT1	0.00021	1.68	105	EF1G	0.00070	-1.24
51	PPGB	0.00265	1.68	106	YBOX1	0.00223	-1.24
52	CATB	0.00005	1.67	107	4EBP1	0.01351	-1.24
53	MOES	0.00002	1.67	108	RS8	0.01484	-1.24
54	CALU	0.00000	1.66	109	SPEE	0.00006	-1.25
55	PDIA3	0.00104	1.64	110	PRDX6	0.01171	-1.26

Rank	Protein	Corrected p-value	FC	Rank	Protein	Corrected p-value	FC
111	RS4X	0.00095	-1.27	150	RL27	0.01246	-1.59
112	RL39	0.02376	-1.27	151	PYGB	0.00006	-1.59
113	PSMD7	0.04349	-1.27	152	BAG2	0.00244	-1.60
114	PSME2	0.03342	-1.27	153	PLSI	0.02119	-1.62
115	PCBP1	0.01309	-1.27	154	SYAC	0.00013	-1.63
116	SORCN	0.02600	-1.27	155	SIAS	0.00000	-1.63
117	FPPS	0.04770	-1.29	156	FLNC	0.00154	-1.64
118	CDV3	0.01212	-1.29	157	RANG	0.00185	-1.64
119	RL18	0.00180	-1.29	158	PDLI2	0.02387	-1.64
120	IF4B	0.00585	-1.29	159	PHB	0.00200	-1.65
121	PUR6	0.01484	-1.30	160	NDRG1	0.00014	-1.67
122	SAHH	0.00000	-1.31	161	SERC	0.00174	-1.69
123	NQO1	0.03717	-1.31	162	GSH1	0.00299	-1.70
124	ACLY	0.00000	-1.31	163	IDHC	0.00726	-1.73
125	IMUP	0.00037	-1.32	164	ASNS	0.00708	-1.75
126	AHSA1	0.00406	-1.33	165	HPRT	0.01196	-1.76
127	RL7	0.00911	-1.33	166	HSP71	0.00090	-1.76
128	SERA	0.02034	-1.34	167	SYWC	0.00000	-1.81
129	RL4	0.00073	-1.35	168	ENOB	0.00147	-1.88
130	EIF3G	0.00741	-1.35	169	MTND	0.00000	-1.88
131	NUDC	0.00005	-1.35	170	ANXA5	0.04123	-1.89
132	UBQL4	0.00025	-1.36	171	NAMPT	0.00015	-1.89
133	AATC	0.00015	-1.38	172	AHNK	0.03975	-1.90
134	IMDH2	0.00007	-1.39	173	K1C19	0.00001	-1.91
135	PA2G4	0.03605	-1.40	174	LMNA	0.00000	-1.97
136	GSHR	0.00478	-1.40	175	SYK	0.04154	-1.98
137	TRXR1	0.00008	-1.40	176	***S10A6***	0.00007	-2.09
138	UGDH	0.00526	-1.41	177	***AK1C2***	0.00098	-2.14
139	NUCL	0.00007	-1.41	178	AIM1	0.00015	-2.29
140	RL13A	0.00369	-1.42	179	***LMO7***	0.00055	-2.40
141	IF6	0.00250	-1.43	180	KCC2D	0.00006	-2.60
142	NPM	0.00007	-1.44	181	***UPP1***	0.00000	-2.69
143	DCPS	0.03955	-1.46	182	***ANXA1***	0.00000	-2.78
144	FKBP4	0.00023	-1.47	183	***KCRB***	0.00006	-2.91
145	HMGA1	0.00102	-1.48	184	***EFHD2***	0.00000	-2.92
146	HS90A	0.00004	-1.49	185	CL043	0.00634	-3.10
147	NH2L1	0.00469	-1.54	186	***TACD2***	0.00006	-3.66
148	RS11	0.03234	-1.55	187	***SDPR***	0.00000	-7.51
149	TGM2	0.00147	-1.57				