



# **At Cancer Cross-Roads -Investigating the Duality of ACE2 in the Pathophysiology of Human Cancers**

---

*Project thesis in partial fulfilment for the degree of*

**Masters in Bioinformatics and Computational Biology**

*presented by*

**Aldas Žarnauskas**

School of Biochemistry and Cell Biology

University College  
Cork

September 2023

Supervisors:

**Dr. Kellie Dean,  
Aideen McCabe,  
Prof. Justin V. McCarthy and  
James V. Harte**

## Table of Contents

1 Abbreviations	iv
2 Acknowledgements	vi
3 Abstract	vii
4 Introduction	9
4.1 Classical and Non-classical Axes of the Renin-Angiotensin System	9
4.2 Classical RAS Axes	10
4.3 Beyond the Classical Axis of the Renin-Angiotensin System (RAS) – New Roles for a Classical Pathway	12
4.4 Pathological Effects of Dysregulated Classical Renin-Angiotensin System (RAS)	15
4.5 The Non-Classical Renin-Angiotensin System (RAS)	15
4.6 ACE2 - a Key Regulator of the renin-angiotensin system (RAS) – Controls the Balance Between the Classical and Non-Classical axes	16
4.7 ACE2 Association with Diseases	17
4.8 ACE2 is a Prognostic and Predictive Marker and Therapeutic Target	18
4.9 The Role of ACE2 and the Renin-Angiotensin System (RAS) in Cancers	19
5 Aim and Hypothesis	21
6 Materials and Methods	22
6.1 Exploratory Data Analysis	22
6.2 Selection of Datasets and Data Acquisition	22
6.3 Processing of the Datasets	24
6.4 Normalised Counts Generation and Differential Gene Expression Analysis	26
6.5 Identification of Enriched Pathways with clusterProfiler	27
6.6 Methylation Analysis using SMART	28
6.7 Transcription Factor Enrichment Analysis with DoRothEA and VIPER	29
6.8 Estimation of Immune Cell Infiltration in RCC Subtypes using TIMER2.0	30
6.9 Single Sample Gene Set Enrichment Analysis using GSVA Package	31
6.10 Data Manipulation, Annotation and Visualisation	32
7 Results	33
7.1 ACE2 Transcript is Widely Distributed in the Human Body	33
7.2 ACE2 mRNA Expression to Protein Translation in the Kidney	34
7.3 ACE2 is Differentially Expressed in Renal Cell Carcinoma	35
7.4 ACE2 is Differentially Expressed in chRCC Compared to Matched ccRCC Normal Tissue	39
7.5 RAAS is Not Enriched in the Subtype Context	43
7.6 ACE2 is Hypomethylated in ccRCC	46
7.7 Transcription Factor Enrichment Analysis Does Not Reveal Major Epigenetic Regulators of RAAS Genes in RCC	48

7.8	ACE2 Correlates with Immune Infiltration in Renal Cell Carcinoma	51
7.9	RCC Samples do not Cluster by Subtype Based on Enrichment of Hallmark Cancer Pathways	54
7.10	The Role of ACE2 in Alternative Cancer Contexts: Stage, Metastasis and Treatment	57
7.10.1	ACE2 expression in cancer stage in clear cell renal cell carcinoma (ccRCC).	57
7.10.2	ACE2 expression and the progression of clear cell renal cell carcinoma to metastasis.	60
7.10.3	ACE2 expression in clear cell renal cell carcinoma (ccRCC) following prior treatment with radiotherapy.	63
8	Discussion and Future Perspectives	66
9	References	75
10	Supplementary Materials	86

## 1 Abbreviations

Abbreviation	Full Name
aa	Amino Acid
ACE1	Angiotensin Converting Enzyme 1
ACE1	Angiotensin Converting Enzyme 1 gene
ACE2	Angiotensin Converting Enzyme 2
ACE2	Angiotensin Converting Enzyme 2 gene
ACC	Adrenocortical Carcinoma
AGTR1	Angiotensin II Receptor Type 1
AGTR2	Angiotensin II Receptor Type 2
Ang I	Angiotensin I
Ang II	Angiotensin II
Ang (1-7)	Angiotensin (1-7)
Ang (1-9)	Angiotensin (1-9)
BLCA	Bladder Urothelial Carcinoma
BRD4	Bromodomain-containing protein
BRCA	Breast Invasive Carcinoma
CESC	Cervical and Endocervical Cancer
CHOL	Cholangiocarcinoma (Bile Duct Cancer)
chRCC	Kidney Chromophobe
ccRCC	Kidney Renal Clear Cell Carcinoma
COAD	Colon Adenocarcinoma
DLBC	Diffuse Large B-Cell Lymphoma
DGEA	Differential Gene Expression Analysis
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial-to-mesenchymal transition
ESCA	Esophageal Carcinoma
FDR	False discovery rate
GEO	Gene Expression Omnibus
GBM	Glioblastoma Multiforme
GSEA	Gene Set Enrichment Analysis
GTEX	The Genotype-Tissue Expression
HER2	Human Epidermal Growth Factor Receptor 2
HNSC	Head and Neck Squamous Cell Carcinoma
HPA	The Human Protein Atlas
IICA	Immune Infiltration Correlation Analysis
IHC	Immunohistochemical
KEGG	Kyoto Encyclopedia of Genes and Genomes
log2FC	Log2 Fold Change
LAML	Acute Myeloid Leukemia
LGG	Lower-Grade Glioma
LIHC	Liver Hepatocellular Carcinoma
LUAD	Lung Adenocarcinoma

Abbreviation	Full Name
LUSC	Lung Squamous Cell Carcinoma
MAS1	Mas1 Proto-Oncogene receptor
MME	Neprilysin
M-dendritic	Myeloid dendritic cell activated
MDSCs	myeloid-derived suppressor cells
NR3C2	Mineralocorticoid receptor
OS	Overall survival
OV	Ovarian Serous Cystadenocarcinoma
PAAD	Pancreatic Adenocarcinoma
PCA	Principal Component Analysis
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate Adenocarcinoma
pRCC	Kidney Renal Papillary Cell Carcinoma
READ	Rectum Adenocarcinoma
RAS	Renin-Angiotensin System
RAAS	Renin-Angiotensin-Aldosterone System
REN	Renin
SARC	Sarcoma
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SMART	The Shiny Methylation Analysis Resource Tool
ssGSEA	Single-sample Gene Set Enrichment Analysis
STAD	Stomach Adenocarcinoma
TCGA	The Cancer Genome Atlas
TFEA	Transcription Factor Enrichment Analysis
TF	Transcription factor
THCA	Thyroid Carcinoma
THYM	Thymoma
TGCT	Testicular Germ Cell Tumors
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma

## 2 Acknowledgements

Everybody that I have met so far contributed to this project, either directly or indirectly. I am thankful to all of them because all of us are interconnected. However, I am especially grateful for the support of two PhD students, James and Aideen, who guided me through this endeavour. They will be great science educators and science communicators.

### 3 Abstract

As a component of the non-classical renin-angiotensin system (RAS), angiotensin-converting enzyme 2 (ACE2) has a complex and dynamic role in human pathophysiology. Recent findings suggest that ACE2 may act as a potential biomarker for human cancers. In some cancers, such as breast cancer, higher levels of ACE2 expression are associated with better patient survival. However, context-dependent effects of ACE2 are emerging, with differing roles observed according to cancer subtype, stage and grade and responsiveness to therapy. Therefore, we aimed to further evaluate the role of ACE2 in human cancer.

Our study initially explored ACE2 expression across different cancer types using publicly available databases and identified renal cell carcinoma (RCC) as a potential cancer of interest: ACE2 was differentially expressed in RCC subtypes and higher levels of expression correlated with patient survival. Subsequently, using publicly available RNA-seq data with different contexts, we performed differential gene expression analysis (DGEA) and gene set enrichment analysis (GSEA) to investigate whether ACE2 - or the genes associated with the RAS - were differentially expressed or enriched in RCC. Furthermore, we performed methylation analysis (MA), transcription factor enrichment analysis (TFEA), immune infiltration correlation analysis (IICA), and single sample gene set enrichment analysis (ssGSEA) to explore changes in ACE2 expression.

DGEA showed that ACE2 is downregulated in one RCC subtype, chromophobe renal cell carcinoma (chRCC), compared to ccRCC-matched normal samples and other RCC subtypes. However, no significant changes in the expression of RAS-associated genes were observed; moreover, GSEA did not highlight RAS among the top 10 enriched pathways. Similarly, MA and TFEA did not pinpoint specific epigenetic markers or transcription factors influencing ACE2 expression. On the other hand, immune infiltration analysis indicated correlations between ACE2 expression and certain immune cells. Intriguingly, ssGSEA showed patient-to-patient variability in ACE2 expression within the same RCC subtype.

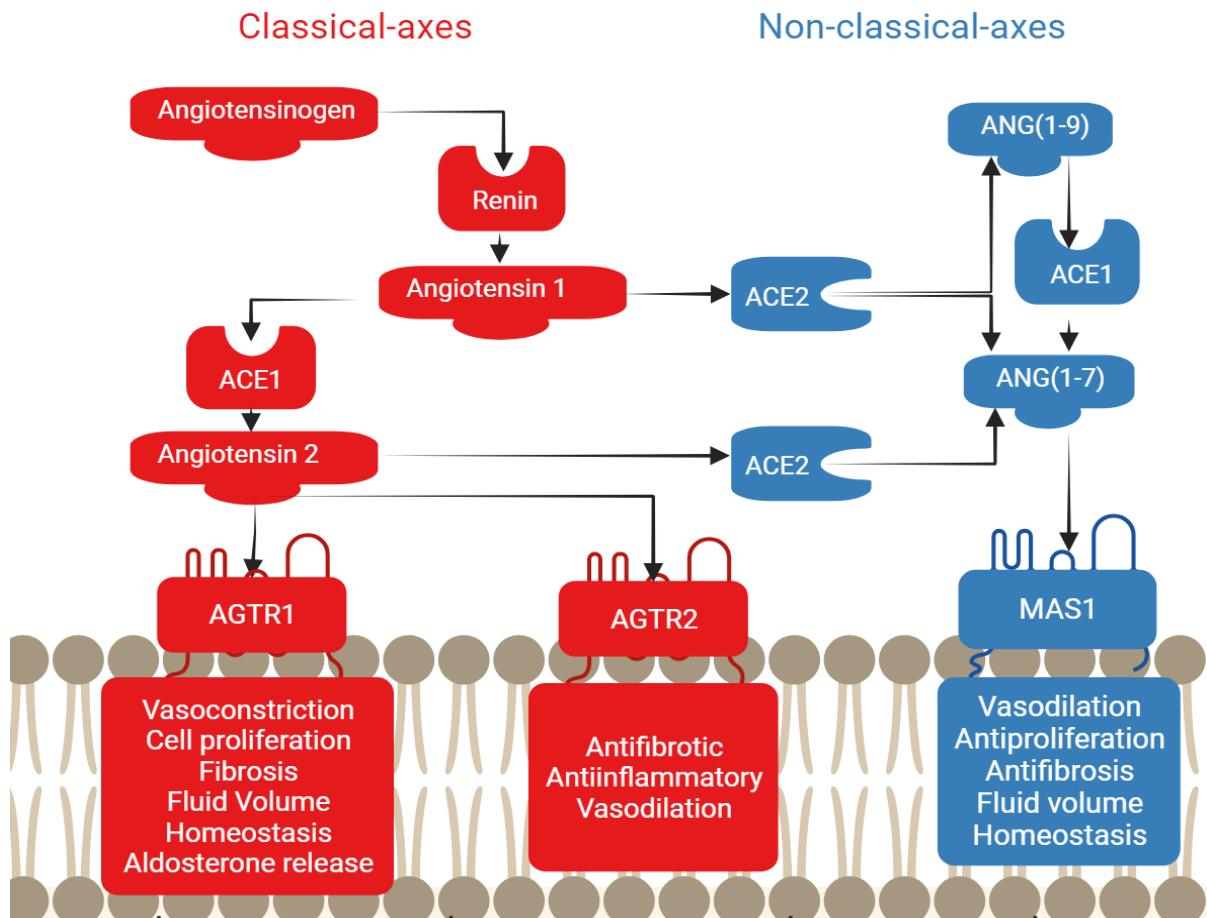
In summary, our exploratory study suggests a context-dependent role of ACE2 in RCC subtypes. Also, ssGSEA showed substantial patient variability in ACE2 expression within RCC subtypes. This underscores the potential utility of ACE2 as a promising biomarker in RCC patients, thereby justifying further investigation through functional experiments.

## 4 Introduction

The renin-angiotensin system (RAS) was first described in the early 20th century, with the discovery of renin and its role in regulating blood pressure (Basso *et al.*, 2001). Today, it is widely recognised that the RAS and its complex network of hormones and enzymes play a pivotal role in maintaining blood pressure, electrolyte, fluid homeostasis, and salt balance (Majid *et al.*, 2015). The RAS system elicits its regulating effects on a global scale through the cardiovascular system and on a local scale through endocrine and paracrine signalling in different body organs and tissues, including muscle, nervous system, bone, gonadal, gastrointestinal, immune, adipose, pancreatic, liver, and circulatory tissue (Lavoie *et al.*, 2003).

### 4.1 Classical and Non-classical Axes of the Renin-Angiotensin System

The RAS can be categorised into two main axes: the classical and non-classical axes (Figure 5.1). The two RAS axes have antagonistic effects, such as vasodilation and vasoconstriction which are in balance. The main components of the classical pathway are Renin (REN), Angiotensinogen (AGT), Angiotensin I (Ang I), Angiotensin II (Ang II), Angiotensin Converting Enzyme (ACE1), and Angiotensin II Receptors Type 1 and Type 2 (AGTR1 and AGTR2). In contrast, the main components of the non-classical axes are Angiotensin Converting Enzyme 2 (ACE2), Angiotensin (1-7) (Ang (1-7)), Angiotensin (1-9) (Ang (1-9)), and the Mas1 Proto-Oncogene receptor (MAS1) (Mirabito Colafella *et al.*, 2019).



**Figure 5.1.** The renin-angiotensin system. The two primary renin-angiotensin system axes are highlighted in colour: classical (red) and non-classical (blue). Renin produces angiotensin 1 (ang I) by the angiotensinogen cleavage. Subsequently, the cleavage of ang I by ACE1 activates the classical axes; whereas the cleavage of ang I by ACE2 activates the non-classical axes. Created with BioRender.com.

## 4.2 Classical RAS Axes

REN is an acid protease synthesised and stored within secretory vesicles found in the juxtaglomerular cells of the dense renal macula. Initially produced as a proenzyme, prorenin, undergoes cleavage to become active renin, which is subsequently released into the circulation. This highly AGT-specific enzyme plays a pivotal role in RAS by cleaving the N-terminal region of AGT to generate Ang I. The activity of plasma renin

is quantified through plasma renin activity, measured by assessing the concentration of Ang I present in the plasma (Riquier-Brison *et al.*, 2018).

AGT typically originates in the liver as a pre-angiotensinogen, a protein with an N-terminal signal peptide. This peptide is cotranslationally removed, resulting in AGT. AGT is the sole known substrate for REN and is subsequently secreted into circulation (Kaschina *et al.*, 2018). Once in circulation, AGT interacts with REN leading to the formation of Ang I, a 10-amino acid (aa) peptide. Further conversion of Ang I into Ang II, an active 8-aa peptide, occurs through the action of ACE1, initiating the classical RAS cascade. Interestingly, in the human heart, a majority of Ang I conversion to Ang II occurs through alternative pathways, particularly chymase, with variable contributions of ACE1 and chymase in Ang II formation across species and tissues (Keidar *et al.*, 2007).

Angiotensin II primarily targets angiotensin receptors AGTR1 and AGTR2, with its effects primarily mediated through AGTR1 receptors found in vascular smooth muscle and other tissues (Park *et al.*, 2014). These effects encompass promoting vasoconstriction, aldosterone release, water and sodium retention, and pro-inflammatory effects such as the release of inflammatory cytokines, tissue growth, and remodelling (Mirabito Colafella *et al.*, 2019). Additionally, Ang II stimulates vasopressin release, a hormone promoting water reabsorption in the kidneys (Keil *et al.*, 1975). Conversely, Ang I is cleaved by ACE2 to form Ang (1-9) or Ang (1-7), initiating the non-classical RAS axes.

ACE1, a transmembrane protein, plays a crucial role in the classical RAS pathway by cleaving the C-terminal dipeptide from Ang I, generating Ang II. Besides this role, ACE1 also participates in the non-classical axes, facilitating the conversion from Ang (1-9) to Ang (1-7). As a side note, Neprilysin (MME), also converts Ang (1-9) to Ang (1-7) (Pavo *et al.*, 2021). Thus, showing that RAS is not a closed system controlled by molecules in many other biological pathways. The RAS component, ACE1, also extends its effects beyond the RAS system, such as degrading bradykinin (vasodilator

peptide) (Mohater *et al.*, 2023) and increasing the significance of RAS to the human body.

The final components of the classical axes are the membrane-bound receptors, AGTR1 and AGTR2, which become active upon binding with Ang II. AGTR1 is widely expressed in numerous adult tissues, including blood vessels, the adrenal cortex, liver, kidney, and brain. AGTR2, on the other hand, is predominantly expressed during fetal development but also maintains a presence in select adult tissues, such as the adrenal medulla, uterus, and ovarian follicles. AGTR1 mainly induces vasoconstrictive, pro-inflammatory, and proliferative effects across various tissues and organs throughout the body. Conversely, AGTR2 seems to functionally oppose many of these actions. Consequently, alterations in the signalling of AGTR1 and AGTR2 can significantly impact whether cancer cells undergo apoptosis or survive in response to modifications in the RAS (Park *et al.*, 2014).

#### **4.3 Beyond the Classical Axis of the Renin-Angiotensin System (RAS) – New Roles for a Classical Pathway**

The RAS also interacts at multiple levels with other biological pathways: the Natriuretic Peptide system (Vasquez *et al.*, 2020), the Kallikrein-Kinin system (Bekassy *et al.*, 2022), and the Nitric Oxide system (Ranjbar *et al.*, 2022). Nevertheless, it is most commonly recognised in association with the aldosterone signalling pathway, frequently referred to as the renin-angiotensin-aldosterone system (RAAS), due to extensive cross-talk and shared components. Consequently, aldosterone serves as the ultimate product of the classical RAS cascade. Ang II triggers the release of aldosterone from the adrenal glands, the primary site of aldosterone synthesis. Upon secretion, aldosterone governs sodium and fluid homeostasis via the mineralocorticoid NR3C2 receptor. In conjunction with Ang II, aldosterone stimulates processes such as fibrosis, inflammation, cell proliferation, neovascularisation, and

oxidative stress (Holappa *et al.*, 2020). An extended list of genes involved in the RAAS are listed in Table 4.1. These RAAS-associated genes are often considered when exploring the effects of the RAS within the human body.

**Table 4.1.** The full list of genes comprising the renin-angiotensin-aldosterone system (RAAS).

Gene Symbol	Description
ACE	angiotensin I converting enzyme
ACE2	angiotensin converting enzyme 2
AGT	angiotensinogen
AGTR1	angiotensin II receptor type 1
AGTR2	angiotensin II receptor type 2
ANPEP	alanyl aminopeptidase, membrane
ATF1	activating transcription factor 1
ATF2	activating transcription factor 2
ATF4	activating transcription factor 4
ATF6B	activating transcription factor 6 beta
CALM1	calmodulin 1
CALM2	calmodulin 2
CALM3	calmodulin 3
CALML3	calmodulin like 3
CALML4	calmodulin like 4
CALML5	calmodulin like 5
CALML6	calmodulin like 6
CAMK1	calcium/calmodulin dependent protein kinase I
CAMK1D	calcium/calmodulin dependent protein kinase ID
CAMK1G	calcium/calmodulin dependent protein kinase IG
CAMK2A	calcium/calmodulin dependent protein kinase II alpha
CAMK2B	calcium/calmodulin dependent protein kinase II beta
CAMK2D	calcium/calmodulin dependent protein kinase II delta
CAMK2G	calcium/calmodulin dependent protein kinase II gamma
CAMK4	calcium/calmodulin dependent protein kinase IV
CMA1	chymase 1
CPA3	carboxypeptidase A3
CREB1	cAMP responsive element binding protein 1
CREB3	cAMP responsive element binding protein 3
CREB3L1	cAMP responsive element binding protein 3 like 1
CREB3L2	cAMP responsive element binding protein 3 like 2
CREB3L3	cAMP responsive element binding protein 3 like 3
CREB3L4	cAMP responsive element binding protein 3 like 4
CREB5	cAMP responsive element binding protein 5
CTSA	cathepsin A
CTSG	cathepsin G
CYP11A1	cytochrome P450 family 11 subfamily A member 1
CYP11B2	cytochrome P450 family 11 subfamily B member 2
CYP21A2	cytochrome P450 family 21 subfamily A member 2
ENPEP	glutamyl aminopeptidase
GNAQ	G protein subunit alpha q
HSD3B1	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1
HSD3B2	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2
ITPR1	inositol 1,4,5-trisphosphate receptor type 1
ITPR2	inositol 1,4,5-trisphosphate receptor type 2
ITPR3	inositol 1,4,5-trisphosphate receptor type 3
LNPEP	leucyl and cysteinyl aminopeptidase
MAS1	MAS1 proto-oncogene, G protein-coupled receptor
MME	membrane metalloendopeptidase
NLN	neurolysin
PLCB2	phospholipase C beta 2
REN	renin
STAR	steroidogenic acute regulatory protein
THOP1	thimet oligopeptidase 1

#### 4.4 Pathological Effects of Dysregulated Classical Renin-Angiotensin System (RAS)

Given its pivotal role in maintaining human body homeostasis, it comes as no surprise that the classical RAS axis has been implicated in a multitude of diseases and pathological conditions that differ from tissue to tissue. For example, in the brain, the classical axis promotes inflammation, cell death, demyelination, and alterations in cellular communication, which modifies plasticity and alters the development of cognitive processes (Jackson *et al.*, 2018), whereas, in the pancreas, it controls reduction in blood flow, diminishes insulin secretion and increases oxidative stress, leading to increased inflammation (Ramalingam *et al.*, 2017). However, on the contrary, it also regulates inflammation and apoptosis of the pancreatic islets. Moreover, this complex system has established links with a range of cardiometabolic risk factors, including high blood pressure, obesity, atherosclerosis, fatty liver, type 2 diabetes mellitus, and kidney fibrosis (Vargas Vargas *et al.*, 2022). For instance, the excessive production of Angiotensin II (Ang II) has been linked to conditions such as atherosclerosis, cardiac hypertrophy, renal fibrosis (Liu *et al.*, 2021) hypertension (Su *et al.*, 2021), kidney diseases like diabetic nephropathy (Fried *et al.*, 2013) and cardiovascular diseases including heart failure and atherosclerosis (Lin *et al.*, 2022).

#### 4.5 The Non-Classical Renin-Angiotensin System (RAS)

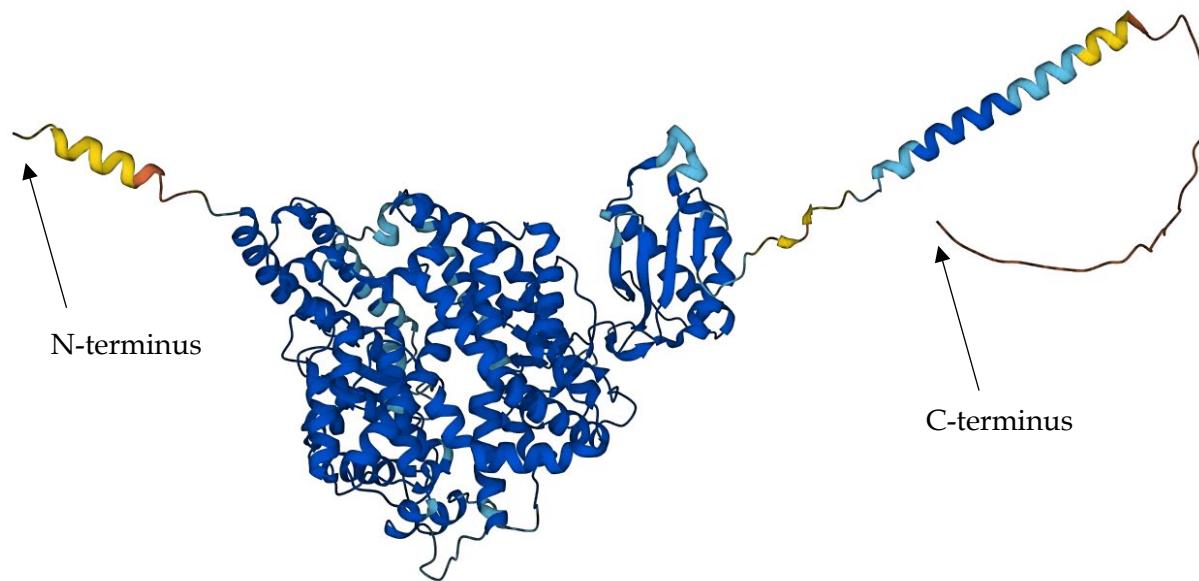
The non-classical axis serves to counterbalance the effects initiated by the classical axes, harmonising the body's equilibrium (Hiremath *et al.*, 2017). The central component of the non-classical axes, ACE2, plays a pivotal role in maintaining this equilibrium. ACE2 accomplishes this by cleaving and deactivating Ang II, a process involving the removal of a single amino acid, which transforms it into Ang (1-7). Ang (1-7), a bioactive peptide, binds to MAS1, thereby promoting anti-inflammatory

responses, vasodilation, and anti-fibrotic reactions that effectively counteract the effects of Ang II (Mirabito Colafella *et al.*, 2019). Furthermore, MAS1 receptor activation leads to a reduction in sympathetic tone, blood pressure, chronic hypertension, and fibrosis. Conversely, it enhances parasympathetic tone, baroreflex function, vasodilation, nitric oxide production, and natriuresis (Paz Ocaranza *et al.*, 2020). This underscores the critical role played by ACE2 regulation.

#### **4.6 ACE2 - a Key Regulator of the renin-angiotensin system (RAS) - Controls the Balance Between the Classical and Non-Classical axes**

ACE2 was first discovered in 2000 as a homologue of ACE. It was initially identified as a novel component of the renin-angiotensin system and was found to have a high affinity for Ang I and II, the primary effector molecules of the classical RAS pathway (Donoghue *et al.*, 2000; Tipnis *et al.*, 2000).

ACE2 is comprised of 805 amino acids, with a molecular weight of approximately 120 kDa. Its three-dimensional configuration includes a single transmembrane segment, an intracellular segment, and terminal N and C domains, with a single enzymatic active site that gives it characteristics different from those of ACE. It belongs to the peptidase M2 family and contains a single active site zinc metalloprotease domain. ACE2 functions as a carboxypeptidase, meaning it removes amino acids from the C-terminal end of peptides. This function allows ACE2 to specifically cleave and inactivate Ang II by removing a single amino acid, converting it into Ang (1-7) (Lubbe *et al.*, 2020). The structural representation of ACE2, generated using AlphaFold (Jumper *et al.*, 2021) is depicted in Figure 4.2.



**Figure 4.2.** The AlphaFold-predicted ACE2 structure, with varying shades of blue, indicates the confidence levels associated with different regions. The N and C terminus are indicated with arrows. AlphaFold assigns a per-residue confidence score (pLDDT) that ranges from 0 to 100, and regions scoring below 50 pLDDT may indicate potential structural instability when considered in isolation. Dark blue areas denote very high confidence ( $p\text{LDDT} > 90$ ), while light blue areas reflect confident predictions ( $90 > p\text{LDDT} > 70$ ). Yellow regions represent lower confidence ( $70 > p\text{LDDT} > 50$ ), and orange areas signify very low confidence ( $p\text{LDDT} < 50$ ). Source organism: Homo sapiens; UniProt: Q9BYF1.

#### 4.7 ACE2 Association with Diseases

In recent years, ACE2 received the most attention as the receptor for the severe acute respiratory syndrome coronavirus 2 virus (SARS-CoV-2), responsible for the COVID-19 pandemic. SARS-CoV-2 utilises ACE2 to infiltrate lung cells and initiate infection, thereby linking higher ACE2 expression to more adverse outcomes in affected individuals. The ACE2 structure facilitates its binding with the SARS-CoV-2 virus, establishing it as a critical target for viral entry (Mészáros *et al.*, 2021; Fagyas *et al.*,

2022). ACE2 activity also exhibits an increase in heart failure patients and those experiencing acute myocardial infarction, with a notable correlation to the severity of heart disease. Within the realm of renal diseases, ACE2 plays a pivotal role in chronic kidney disease, impacting patients in both early stages and those undergoing dialysis. Additionally, elevated ACE2 activity is observed in individuals with type 1 diabetes and vascular complications. Furthermore, evidence suggests ACE2 activity upregulation during the acute phase of ST-elevation myocardial infarction. It has also been postulated that ACE2 may counteract the adverse effects of Angiotensin II in individuals with cardiovascular disease, diabetes, and hypertension (Anguiano *et al.*, 2015).

#### **4.8 ACE2 is a Prognostic and Predictive Marker and Therapeutic Target**

As previously discussed, ACE2 activity demonstrates potential associations with cardiovascular diseases, including heart failure and type 1 diabetes. ACE2 has demonstrated the capacity to mitigate myocardial injury and exert diverse effects, including anti-inflammatory, antioxidant, anti-apoptotic, and anti-cardiomyocyte fibrosis actions (Zhu *et al.*, 2021). Consequently, ACE2 emerges as a promising biomarker for predicting the risk and prognosis of cardiovascular disease. Utilising ACE2 activity as a predictive tool can aid in early intervention and the development of personalised treatment strategies for these conditions (Anguiano *et al.*, 2015).

Furthermore, ACE2 exhibits strong associations with chronic kidney disease and diabetic nephropathy, suggesting its potential as a renoprotective biomarker (Anguiano *et al.*, 2017). Additionally, the linkages between ACE2 and COVID-19 viral infections have prompted further investigations, highlighting ACE2 as a compelling therapeutic target for COVID-19 (Wang *et al.*, 2022).

#### 4.9 The Role of ACE2 and the Renin-Angiotensin System (RAS) in Cancers

Studies suggest that the dysregulation of the classical RAAS pathway may contribute to the pathogenesis of certain cancers. Ang II has been implicated in tumour metastasis by facilitating tumour adhesion to vascular endothelial cells, promoting migration, invasion, and angiogenesis (Kidoguchi *et al.*, 2023). Moreover, MAS1 has been associated with tumour induction in animal models, with studies revealing its upregulation in specific cancer types such as breast cancer, ovarian cancer, and gastric cancer, where MAS1 activation has been shown to induce tumorigenesis (Paz Ocaranza *et al.*, 2020). On the contrary, the depletion of the non-classical axes components results in vascular dysfunction (Qaradakhi *et al.*, 2020). However, the recent COVID-19 pandemic has brought the ACE2 protein to the forefront as a key player in the RAS and various pathological conditions (Ni *et al.*, 2020; Beyerstedt *et al.*, 2021).

Higher ACE2 expression in clear cell renal cell carcinoma is correlated with better patient prognosis, as evidenced by several studies (Khanna *et al.*, 2021; Niu *et al.*, 2021; T. Wang *et al.*, 2021). Conversely, ACE2 overexpression has shown promise in inhibiting the growth, metastasis, and angiogenesis of tumour cells in lung, breast, colon, and pancreatic cancers (Huang *et al.*, 2021). Across various cancer types, ACE2 overexpression has consistently demonstrated its potential as an anti-tumour agent, repressing tumour cell proliferation, invasion, epithelial-to-mesenchymal transition (EMT), and metastasis (Pathania *et al.*, 2021). Notably, ACE2 expression exhibits a positive correlation with low EMT scores in aerodigestive and respiratory cancer cell lines, as well as in normal and cancer patients. In specific cancer types such as uterine corpus endometrial carcinoma (UCEC) and papillary renal cell carcinoma (pRCC), ACE2 expression aligns with increased tumour infiltration and more favourable prognosis, linked to DNA hypomethylation within the ACE2 promoter (Pathania *et al.*, 2021). Moreover, restoration of ACE2 protein was associated with suppressed cell proliferation and motility and increased sensitivity to hypoxia-induced injury of pancreatic cancer cell lines (Zhou *et al.*, 2011).

However, context-dependent effects of ACE2 are apparent, with differing roles in various breast cancer subtypes. While some studies suggest that ACE2 upregulation inhibits angiogenesis, invasion, and metastasis in breast cancer (Zhang *et al.*, 2019), contradictory findings propose that ACE2 upregulation could be indicative of a poor prognosis, particularly in certain patient groups (Bhari *et al.*, 2020). Notably, the intricate relationship between ACE2 and breast cancer varies among molecular subtypes, with ACE2 upregulation in HER2-positive breast cancer associated with EGFR expression and a less favourable outcome (Nair *et al.*, 2021). The dual nature of ACE2's role in tumorigenesis underscores its contextual dependence, highlighting the need for further research to elucidate its function as either a promoter or suppressor of human cancers across different tissue types and contexts.

## 5 Aim and Hypothesis

We hypothesised that angiotensin-converting enzyme 2 (ACE2), a key component of the renin-angiotensin system (RAS), may influence the development and progression of cancers. Emerging evidence suggests that elevated ACE2 levels may have a preventative role in cancer development (Khanna *et al.*, 2021; Niu *et al.*, 2021; T. Wang *et al.*, 2021).

Nevertheless, the role of ACE2 in cancer remains unclear, with conflicting findings in breast cancer studies. It is important to acknowledge that the context-dependent effects of ACE2 are emerging and these manifest differently in various breast cancer subtypes. While some studies indicate that ACE2 upregulation inhibits angiogenesis, invasion, and metastasis in breast cancer (Zhang *et al.*, 2019), contradictory research suggests that ACE2 upregulation could signal a poorer prognosis - especially in HER2-positive breast cancer patient cohorts (Nair *et al.*, 2021). Notably, the intricate relationship between ACE2 and breast cancer varies among molecular subtypes, with ACE2 upregulation in HER2-positive breast cancer being associated with epidermal growth factor receptor (EGFR) expression and an overall unfavourable outcome (Nair *et al.*, 2021).

The dualistic nature of ACE2 in tumorigenesis underscores a context-dependent behaviour, highlighting the necessity for further research to elucidate whether it acts as a promoter or suppressor of human cancers across diverse tissue types and contextual settings.

Therefore, we aimed to evaluate the expression of ACE2 in different contexts in different human cancers. Our exploratory analyses were designed to identify whether ACE2 expression could be potentially employed as a diagnostic, prognostic, or therapeutic marker in the pathophysiology of human cancers.

## 6 Materials and Methods

### 6.1 Exploratory Data Analysis

The Human Protein Atlas (HPA) database (*The Human Protein Atlas*, 2023) was utilised to assess the summarised ACE2 expression at mRNA and protein levels in humans. This analysis aimed to gain insights into the predominant expression sites of ACE2 in humans and to determine if ACE2 is translated into proteins.

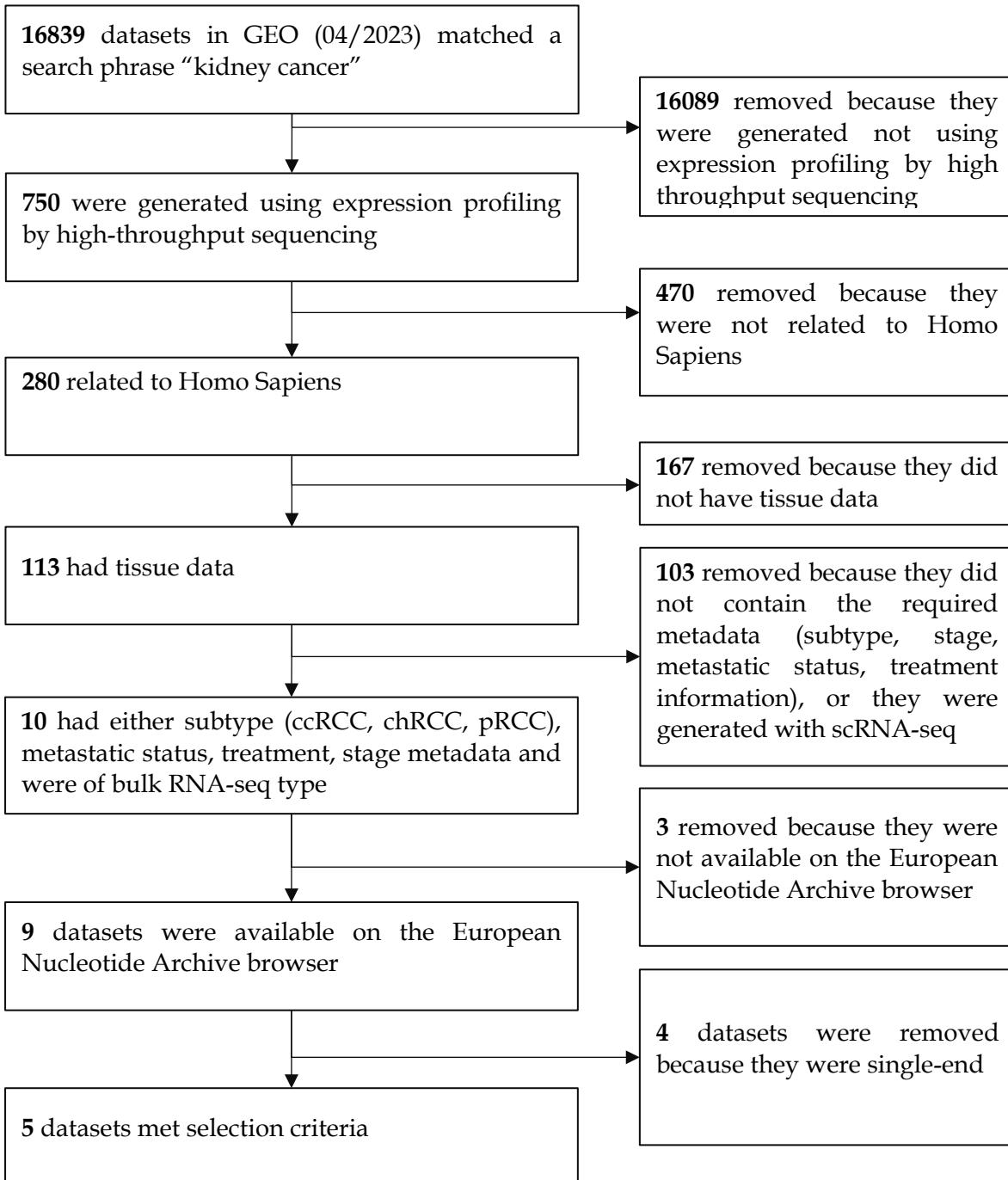
Furthermore, initial exploratory data analysis of ACE2 expression in cancers was conducted using cBioportal (Cerami *et al.*, 2012), UALCAN (Chandrashekhar *et al.*, 2017), Gene Expression Profiling Interactive Analysis 2.0 (GEPIA2) (Tang *et al.*, 2019) and Kaplan-Maier (KM) Plotter (Lánczky *et al.*, 2021) databases.

GEPIA2 were utilised to investigate ACE2 expression based on various contexts, including subtype, and stage. In addition, GEPIA2 contains data from The Cancer Genome Atlas (TCGA) project (Weinstein *et al.*, 2013) as well as The Genotype-Tissue Expression (GTEx) project (Lonsdale *et al.*, 2013). Additionally, survival analysis was performed using GEPIA2: patients were stratified according to ACE2 expression, with the median ACE2 expression selected as a cut-off.

### 6.2 Selection of Datasets and Data Acquisition

Renal cell carcinoma (RCC) was identified as the cancer of interest following the exploratory analysis because ACE2 was differentially expressed and associated with patient survival in multiple RCC subtypes. To advance the investigation, RNA-seq datasets for kidney cancer were selected from the Gene Expression Omnibus (GEO) (Edgar, Domrachev and Lash, 2002), encompassing data uploaded up to April 30<sup>th</sup>, 2023. The criteria for dataset selection are shown in Figure 6.1. The selected datasets were obtained from the European Nucleotide

Archive (ENA) (Leinonen *et al.*, 2011). The list of datasets used in this project is detailed in Table 6.1.



**Figure 6.1.** Selection of datasets criteria used in GEO to filter out not suitable datasets for this project.

**Table 6.1.** The list of datasets used in this project. The metadata column specifies the type of information which was used to categorise samples.

Accession number	GEO number	Samples Size	Metadata	Reference
PRJNA284822	GSE69197	6	Subtype	(Ho <i>et al.</i> , 2015)
PRJNA601152	GSE143630	44	Stage, Metastatic progression	(Shih <i>et al.</i> , 2020)
PRJNA641885	GSE153262	16	Treatment	(Chow <i>et al.</i> , 2020)
PRJNA779050	GSE188486	28	Subtype	(Nassar <i>et al.</i> , 2023)
PRJNA898782	GSE217386	20	Subtype, Stage	(Qin <i>et al.</i> , 2023; Wang <i>et al.</i> , 2023)

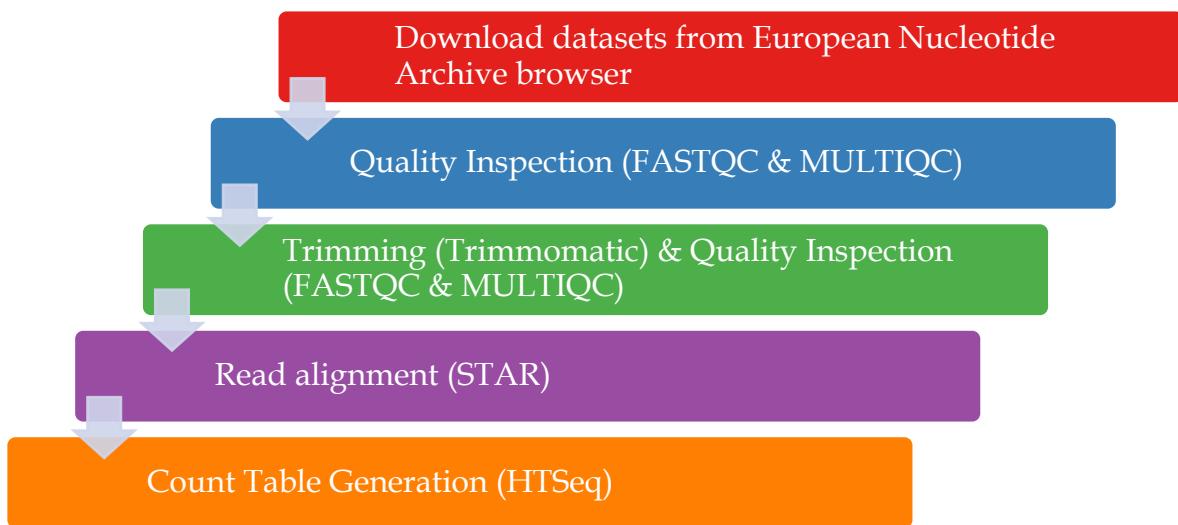
### 6.3 Processing of the Datasets

Further processing of the datasets, as illustrated in **Figure 6.2**, was conducted on a UNIX system-based server running Ubuntu version 22.04. Server access was provided by Professor P. Baranov (School of Biochemistry and Cell Biology, University College Cork, Ireland). The initial assessment of read quality was carried out using FASTQC, version 0.1.3 (Andrew, 2010). Subsequently, the FASTQC reports were summarised using MULTIQC, version 1.15 (Ewels *et al.*, 2016). The examination of the reads encompassed considerations of quality, adapter content, and other relevant quality parameters, all of which met satisfactory standards and were used in subsequent analyses. Access to all MULTIQC reports can be found on the GitHub page: [https://github.com/aldaszarnauskas/mastersThesis/tree/main/multiQC\\_reports](https://github.com/aldaszarnauskas/mastersThesis/tree/main/multiQC_reports). An illustrative MULTIQC report is presented in Supplementary Figure 10.4.

Following the quality inspection, the human genome and its annotation were retrieved from the Ensemble database (Martin *et al.*, 2023), specifically versions GRCh38 and GRCh38.109.

These were employed to prepare the human genome for read alignment using the STAR sequence aligner, version 2.7.0a (Dobin *et al.*, 2013). Genome index generation adhered to the default parameters described in the STAR manual 2.7.11a. However, the --sjdbOverhang parameter varied according to the read length of each dataset, with its value set as the length of the read minus one in digits.

Subsequently, we utilised the generated genome indices for read alignment of the five selected datasets listed in Table 6.1. Similar to the previous step, default parameters for read alignment, as detailed in the STAR manual, were used, except for the --alignEndsType parameter, which was consistently set to "EndToEnd" for all alignments that force STAR to do end-to-end read alignment and be more precise in matching a read to the Humans genome. Unsorted BAM files were generated to alleviate computational demands during alignment. These unsorted BAM files were then sorted using SAMtools (Li *et al.*, 2009) version 1.18, following the guidelines in the SAMtools manual (Pollard, accessed 15-09-2023). Count generation was performed with HTSeq (Anders *et al.*, 2015) version 2.0.3. Similarly, the default HTSeq parameters, detailed in the HTSeq documentation 2.0.3, were used. The -s parameter was specified as "yes" or "no" depending on whether the reads were paired-end or single-end. A bash script detailing the processing steps for the reads described above can be accessed on the GitHub page: <https://github.com/aldaszarnauskas/mastersThesis/blob/main/CountGenerationBASH.txt>



**Figure 6.2.** High-level workflow of the RNA-seq processing.

#### 6.4 Normalised Counts Generation and Differential Gene Expression Analysis

The estimation of immune cell infiltration level was carried out using R packages ESTIMATE (Yoshihara *et al.*, 2013) (version 2.0.0) and CePa (Gu *et al.*, 2013) (version 0.8.0). Tumour cell infiltration was estimated due to the utilisation of bulk RNA-seq datasets in the analysis. The bulk mRNA-seq data encompassed mRNA transcripts originating from a mixed population of cells sampled, including but not limited to cancer cells, stromal cells, and immune infiltrating cells. Thus, the incorporation of immune infiltrating cells into the analysis was deemed essential, as previous research has demonstrated that these infiltrating immune cells constitute the predominant fraction of normal cells within tumour tissues. These immune cells not only introduce variability into molecular studies but also play a significant role in cancer biology (Yoshihara *et al.*, 2013). Consequently, immune cell infiltration levels were estimated, and the resultant scaled values were employed in subsequent analyses for the generation of normalised counts and the execution of Differential Gene Expression Analysis (DGEA) with DESeq2 (Love *et al.*, 2014) (version 3.17). Notably, the calculation of tumour infiltration was performed independently for each dataset.

In scenarios where follow-up analyses, such as Transcription Factor Enrichment Analysis (TFEA), and single-sample Gene Set Enrichment Analysis (ssGSEA), involved multiple datasets and necessitated the consideration of batch effects, a batch effect removal step with limma (version 3.17) was implemented (Ritchie *et al.*, 2015). Supplementary Figure 10.6 A-B demonstrates a successful batch removal.

Furthermore, an additional filtering step was integrated into the DESeq2 pipeline, whereby genes with fewer than ten counts in three or more samples were excluded. Furthermore, during DGEA with DESeq2, the design formula was adjusted to account for data originating from multiple datasets, incorporating tumour infiltration values, and specifying the comparison group. An example of the design formula:  $\sim \text{batch} + \text{purityscaled} + \text{subtype}$ . The  $\log_{2}\text{FC} > 1$  and adjusted p-value below a false discovery rate (FDR) cutoff of 0.05 were used as significance cut-off values.

Additionally, the reference level is specified in the DESeq object before DGEA. The reference level is a group of samples that is compared against other groups or likewise contrasts. However, such a DESeq2 design yields comparisons between the reference level and contrasts, but not the comparisons within contrasts. To obtain the comparisons within contrasts, the nbinomWaldtest() function of the DESeq2 package was employed to alter the reference level of the DESeq object.

Furthermore, a shrinkage step was applied post-DGEA, as it enhances DGEA outcomes by moderating log-fold change estimates for genes with low read counts and high variance. This reduction in the number of falsely reported significant differentially expressed genes contributes to the robustness of the results (Love *et al.*, 2014).

## 6.5 Identification of Enriched Pathways with clusterProfiler

Gene Set Enrichment Analysis (GSEA) was used to obtain the enrichment score of the gene sets from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa *et al.*, 2000). The list of differentially expressed genes, obtained in Section 7.4 and sorted in

descending order based on their Log2 Fold Change (log2FC), was used. This GSEA analysis was conducted using clusterProfiler (Yu *et al.*, 2012; Wu *et al.*, 2021) (version 3.17). To identify biologically significant enrichments, a significance threshold of P-adjusted value  $\leq 0.05$  was utilised.

## 6.6 Methylation Analysis using SMART

The Shiny Methylation Analysis Resource Tool (SMART) serves as an interactive web server designed for the analysis of DNA methylation patterns (Li *et al.*, 2019), harnessing data sourced from the TCGA project. It also offers a versatile platform for the examination and visualisation of DNA methylation data across various cancers and their corresponding normal samples.

ACE2 DNA methylation patterns across RCC subtypes were exploited with SMART. It's worth noting that data for matched normal tissues of chRCC were not available. The p-value was set to  $< 0.05$  and calculated using the Wilcoxon rank sum test.

Additionally, it's important to mention that the methylation data presented here represents an average calculated from the combination of ten probes specifically designed to measure methylation levels, as specified in **Table 6.2**. The methylation values were represented as Beta-values, and the aggregation method employed to derive the ACE2 methylation score was based on the median of these values.

**Table 6.2.** The list of methylation probes used to measure ACE2 methylation level in SMART.

Probe	Chromosome	Start	End	Strand
cg23232263	chrX	15561359	15561360	+
cg05039749	chrX	15565389	15565390	-
cg05748796	chrX	15601214	15601215	+
cg16734967	chrX	15601980	15601981	-
cg08559914	chrX	15602117	15602118	-
cg18877734	chrX	15602961	15602962	-
cg21598868	chrX	15603044	15603045	-
cg18458833	chrX	15603354	15603355	-

## 6.7 Transcription Factor Enrichment Analysis with DoRothEA and VIPER

Transcription Factor Enrichment Analysis was undertaken to assess the enrichment score of transcription factors (TFs) within samples and to compare the activity of TFs across different groups. To execute this analysis, the DoRothEA package in R (Garcia-Alonso *et al.*, 2019; Badia-I-Mompel *et al.*, 2022; Müller-Dott *et al.*, 2023) (version 3.17), along with viper (Alvarez *et al.*, 2016) (version 3.17) and bcellviper (Alvarez, 2023) (version 1.36.0) was applied.

DoRothEA is a valuable resource that provides insights into the regulatory relationships between TFs and their target genes, often referred to as regulons. It also offers confidence levels associated with these regulons, ranging from A (highest confidence) to E (lowest confidence). In the analysis, regulons with high confidence levels were used, specifically A and B, to enhance the reliability of the results. Additionally, regulons associated with the regulation of ACE2 were selected, even though these were classified as E confidence. This decision was driven by the significance of ACE2 as the gene of interest in this project. Furthermore, the DoRothEA database was supplemented with additional regulons known to regulate ACE2, although not present in the database. These TFs include STAT3 (Liang *et al.*, 2022), HNF1 $\alpha$  and HNF1 $\beta$  (Pedersen *et al.*, 2013), and FOXA2 (Pedersen *et al.*, 2017, p. 2), as supported by the literature.

Statistical comparisons of enrichment scores between groups were carried out using the script written by Aideen McCabe (School of Biochemistry and Cell Biology, University College Cork, Ireland) that utilised rstatix package (version 0.7.2) and the car package (version 3.1-2). A suitable statistical method, based on the TF activity across sample distribution and variance, was selected to compare its enrichment levels among subtypes. A significance cutoff of p-value < 0.05 was applied to identify statistically significant differences.

## 6.8 Estimation of Immune Cell Infiltration in RCC Subtypes using TIMER2.0

The correlation analysis between the ACE2 gene and immune cell infiltration levels was conducted using the TIMER2.0 database (Li *et al.*, 2016, 2017, 2020). This database allows users to select a gene of interest and explore the correlation between its expression and immune infiltration across cancer types. Among various algorithms available in TIMER2.0 for assessing gene-immune cell correlation, we utilised the CIBERSORT algorithm. CIBERSORT, a gene expression-based deconvolution algorithm, offers remarkable sensitivity and specificity in discriminating human tissue and accurately estimating immune composition in cancer samples (Chen *et al.*, 2018).

To quantify the correlation between infiltrate estimation values and ACE2 gene expression, we employed the purity-adjusted Spearman's rho ( $r$ ). Additionally, we determined the correlation between ACE2 and tumour purity. The choice of "purity adjustment" aligns with TIMER2.0 recommendations since tumour purity significantly influences this analysis. In most cases, immune cell types exhibit a negative correlation with tumour purity (Rhee *et al.*, 2018). The resulting correlation values were interpreted as follows:  $p < 0.05$  and  $r > 0.2$  indicated a positive correlation,  $p < 0.05$  and  $r < -0.2$  signified a negative correlation, while  $p > 0.05$  indicated no significant correlation (Li *et al.*, 2020).

Furthermore, nine immune infiltrate cells were selected based on their relevance to RCC, as substantiated by the literature: Macrophage M1, B cell memory (B-memory), T cell regulatory (Tregs), T cell follicular helper (Tf-helper), T cell CD8+ (CD8), T cell CD4+ naïve (CD4),

Neutrophil, Myeloid dendritic cell activated (M-dendritic), and Monocyte. T cell regulatory (Tregs) and T cell follicular helper (Tf-helper) were linked to a poor prognosis in RCC, while Monocytes were associated with a more favourable prognosis, as supported by studies (Zhu *et al.*, 2019; Y. Wang *et al.*, 2021). Additionally, T cell CD8+ (CD8) was found to be associated with prolonged overall survival in chRCC, whereas in pRCC, Macrophage M1 exhibited a favourable outcome in terms of survival (S. Zhang *et al.*, 2019). In addition, the association of B cell memory (B-memory), T cell CD4+ naïve (CD4), Neutrophils, and activated Myeloid dendritic cells (M-dendritic) was already associated with ACE2 in pRCC (Yang, Li, Hu, *et al.*, 2020).

## 6.9 Single Sample Gene Set Enrichment Analysis using GSVA Package

Single sample Gene Set Enrichment Analysis (ssGSEA) assess the enrichment score of the gene sets across individual patient samples. Following best practices, the normalised mRNA expression data, as recommended by GSVA R package authors (Hänelmann *et al.*, 2013), served as the input for conducting ssGSEA. The ssGSEA analysis was executed using GSVA (version 3.17), Biobase (Huber *et al.*, 2015) (version 3.17), GSEABase (Morgan *et al.*, 2023) (version 3.17). To comprehensively assess the expression patterns of samples across various pathways, the HALLMARK pathways sourced from MSigDB were used (Liberzon *et al.*, 2011).

Four specific HALLMARK pathways known for their relevance to ccRCC, namely HALLMARK\_ANGIOGENESIS, HALLMARK\_G2M\_CHECKPOINT, HALLMARK\_INFLAMMATORY\_RESPONSE, and HALLMARK\_HYPOXIA, as substantiated by existing literature (Petitprez *et al.*, 2021). Additionally, considering the project's context, the Renin-angiotensin pathway gene set was included. These selected pathways were investigated individually to gain deeper insights into their impact on RCC subtypes.

## 6.10 Data Manipulation, Annotation and Visualisation

All data manipulation, annotation, and visualisation were carried out using R version 4.3.1 and RStudio version 2023.06.0+421. Data manipulation tasks were accomplished with the dplyr (Hadley Wickham and Romain François and Lionel Henry and Kirill Müller and Davis Vaughan, 2023) (version 1.28.0), tibble (version 3.2.1), tidyverse (version 2.0.0), andforcats (version 1.0.0) packages. Annotation procedures relied on the following packages: biomaRt (Smedley *et al.*, 2009) (version 2.56.1), msigdbr (Mootha *et al.*, 2003; Subramanian *et al.*, 2005) (version 7.5.2), AnnotationDbi (version 1.62.2), org.Hs.eg.db (version 3.17.0), GEOquery (version 2.68.0), HelpersMG (version 6.0). For data visualisation, pheatmap (Kolde, 2012) (version 1.0.12), ggplotify (0.1.2), patchwork (version 1.1.3), ggplot2 (Wickham, 2016) (version 3.4.3), RColorBrewer (version 1.1-3), ComplexHeatmap (Gu *et al.*, 2016; Gu, 2022) (version 2.16.0), ggstance (version 0.3.6), gridExtra (version 2.3), ggrepel (version 0.9.3), and ggsignif (version 0.6.4) packages were used.

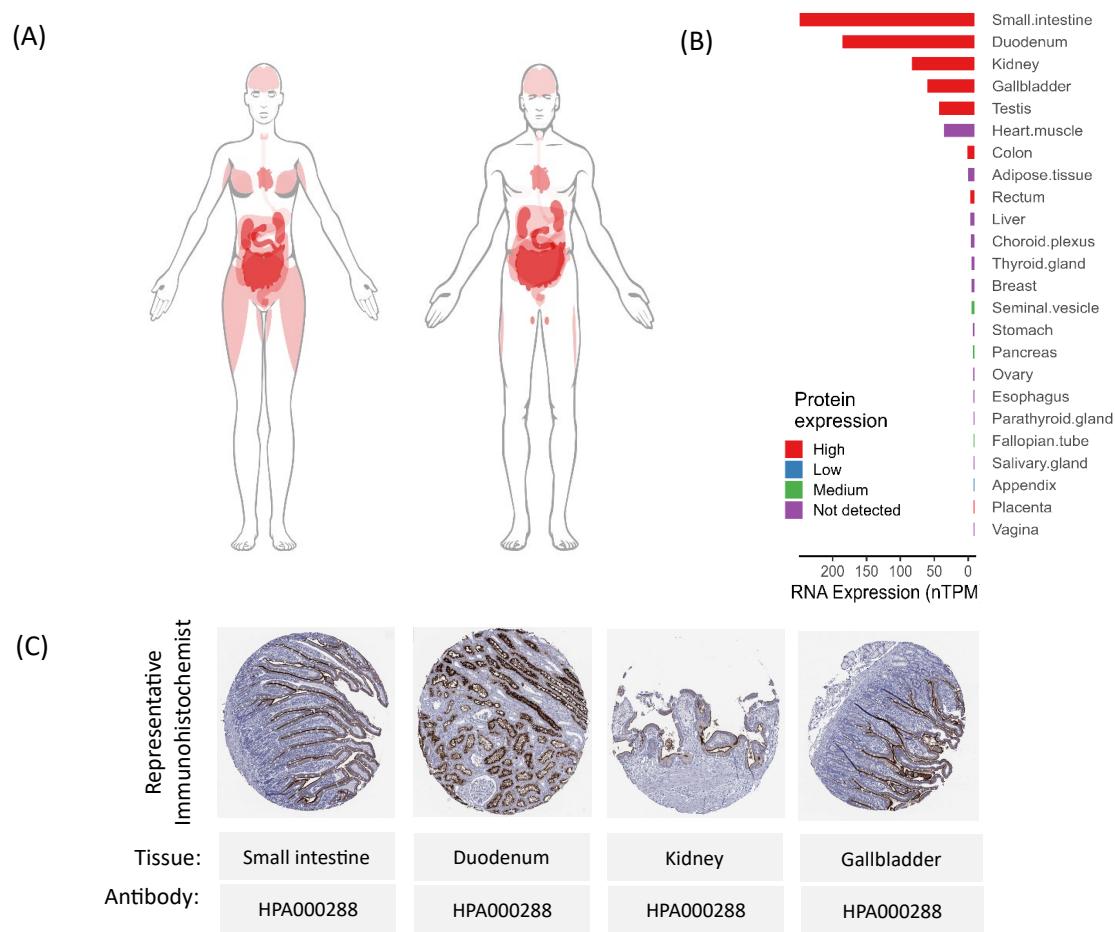
Furthermore, the scripts of all analyses completed in R and the scripts used to visualise results can be viewed on the following GitHub link: <https://github.com/aldaszarnauskas/mastersThesis>.

## 7 Results

### 7.1 ACE2 Transcript is Widely Distributed in the Human Body

To understand the pathomechanistic role of angiotensin-converting enzyme 2 (ACE2) in human disease, it is important to understand the tissue expression of *ACE2* under physiological conditions. The Human Protein Atlas (HPA) database serves as a resource containing information about human proteins through a combination of transcriptomics and antibody-based proteomics (*The Human Protein Atlas*, 2023). Here, we leveraged the HPA database to evaluate *ACE2* transcript and protein levels in human tissues (Figure 7.1).

According to the HPA database, *ACE2* is widely distributed in male and female tissue (Figure 7.1 A). Notably, the highest transcript levels of *ACE2* were found in the small, duodenum, kidney and gallbladder (Figure 7.1 B). The robust transcript expression in these tissues correlates with the high level of translated *ACE2* protein, which was detected by antibody-based proteomics; representative immunohistochemical (IHC) sections showing translation of *ACE2* transcripts are shown in Figure 7.1 C. Additionally, *ACE2* was detectable at high levels in the testis, colon, and rectum (Figure 7.1 B); although the *ACE2* transcript has been documented in other sites, the transcript levels were lower than 1 normalised transcript per million (nTPM).

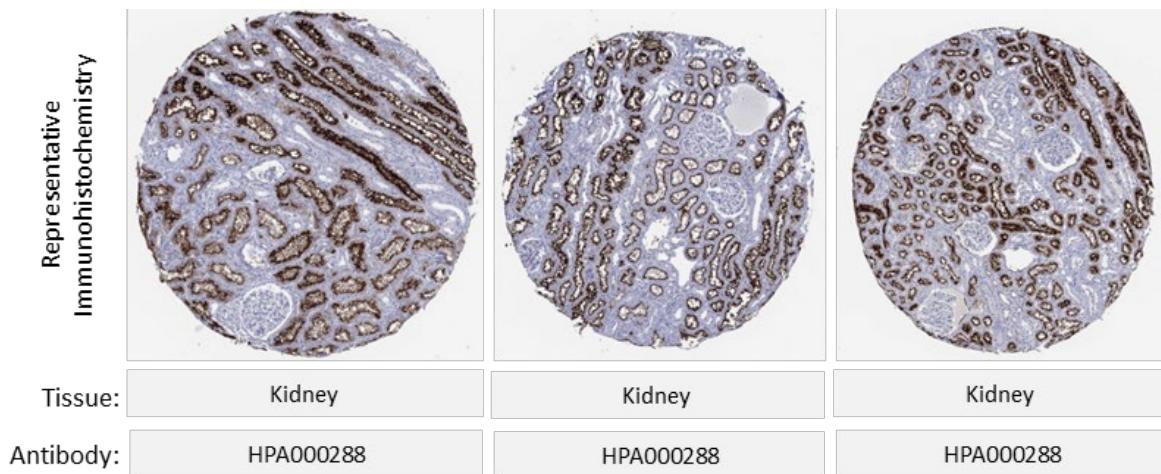


**Figure 7.1.** Summarised tissue expression of ACE2 in humans. (A) Anatomograms of ACE2 tissue expression in human tissues. (B) ACE2 expression summary, as normalised transcript per million expression (nTPM), in 55 different human tissues. Data derived from the HPA consensus dataset, combining the HPA and The Genotype-Tissue Expression (GTEx) transcriptomics datasets. (C) Representative IHC sections derived from the HPA showing ACE2 expression in the four tissues with the highest mRNA expression: small intestine, duodenum, kidney, and gallbladder. An anti-ACE2 antibody (HPA000288: Atlas Antibodies, Sigma-Aldrich) immunohistochemically detected ACE2 expression in tissue sections.

## 7.2 ACE2 mRNA Expression to Protein Translation in the Kidney

Additional IHC staining, in Figure 7.2, of kidney tissue sections using ACE2-specific antibodies provides a better insight into the localisation and abundance of ACE2 within kidney tissues. Notably, all sections acquired strong staining in Bowman's capsule and proximal tubules of microvilli with moderate staining in proximal tubules of cellular bodies. These sections demonstrate the specific regions where ACE2

expression is most abundant in the kidney. Thus showcasing which kidney parts lack ACE2 protein and its potential reno-protective properties (Mizuiri *et al.*, 2015).



**Figure 7.2.** Representative IHC sections derived from the HPA showing ACE2 expression in the kidney. An anti-ACE2 antibody (HPA000288: Atlas Antibodies, Sigma-Aldrich) immunohistochemically detected ACE2 expression in tissue sections.

### 7.3 ACE2 is Differentially Expressed in Renal Cell Carcinoma

Next, we leveraged publicly available databases to interrogate the tissue expression of ACE2 in various human cancers. We conducted an exploratory data analysis of ACE2 expression in various cancer types, utilising data from Gene Expression Profiling Interactive Analysis 2.0 (GEPIA2).

In GEPIA2 we explored the differential expression of ACE2 in human cancer and found nine cancers with differentially expressed ACE2. ACE2 was significantly upregulated in colorectal adenocarcinoma (COAD), papillary renal cell carcinoma (pRCC), pancreatic adenocarcinoma (PAAD), rectum adenocarcinoma (READ) and stomach adenocarcinoma (STAD). In contrast, ACE2 was significantly downregulated in chromophobe renal cell carcinoma (chRCC), sarcoma (SARC), thyroid carcinoma (TCGT), and thyroid adenocarcinoma (THCA) (Supplementary Figure 10.1)

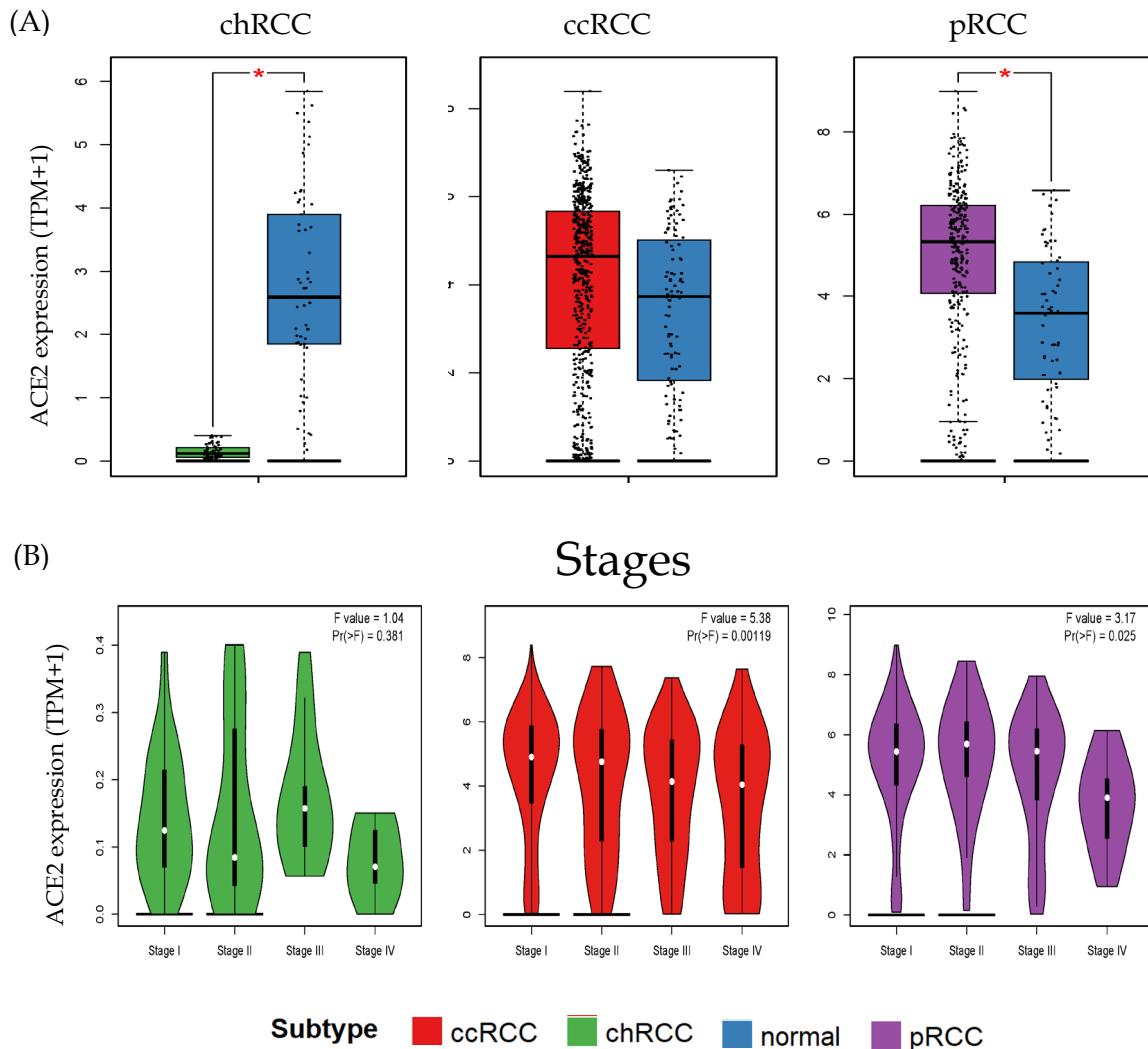
Significant differences in ACE2 expression were determined using the following cutoffs: log2FC = 1 and p-value < 0.05.

Moreover, higher ACE2 expression correlated with improved overall survival (OS) in clear cell renal cell carcinoma (ccRCC), ovarian cancer (OV), and liver hepatocellular carcinoma (LIHC) cancer types; whereas, lower ACE2 expression improved OS in lower-grade glioma (LGG) cancer (Figure 7.4 A and Supplementary Figure 10.2). Additionally, when assessing disease-free survival (DFS), higher ACE2 expression was beneficial in ccRCC, lung squamous cell carcinoma (LUSC), uterine carcinosarcoma (UCS), and OV cancer (Figure 7.4 B and Supplementary Figure 10.3).

Interestingly, the GEPIA2 database categorises renal cell carcinoma (RCC) according to the three most prevalent subtypes: chRCC, ccRCC, and pRCC (Muglia *et al.*, 2015). Differential expression analysis revealed significant differences in ACE2 expression levels between chRCC and pRCC when compared to normal tissues (Figure 7.3 A). Specifically, ACE2 was significantly downregulated in chRCC, while it was significantly upregulated in pRCC compared to matched normal (p-value < 0.05 & log2FC > 1 significance thresholds). Notably, ACE2 was not significantly differentially expressed in ccRCC (Figure 7.3 A).

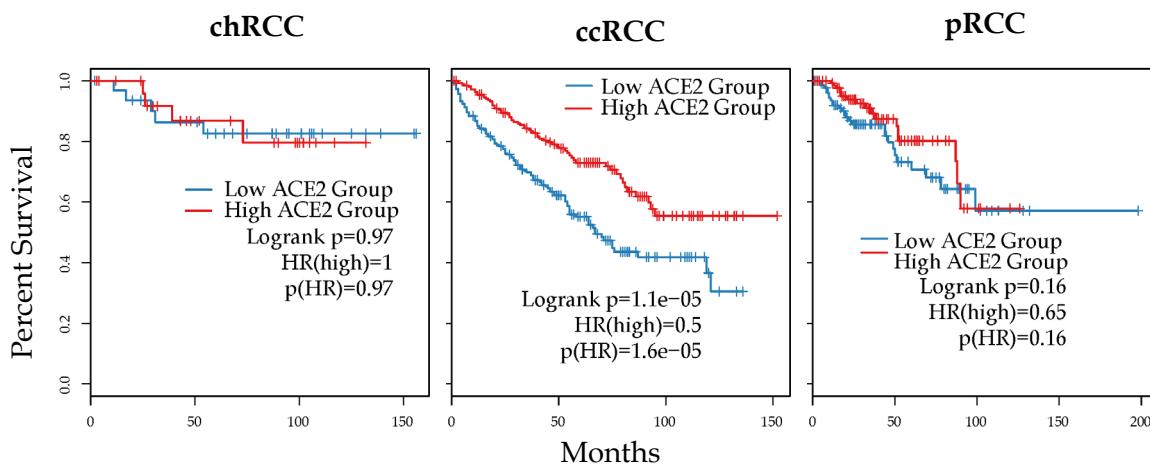
The examination of ACE2 expression across the four major cancer stages using GEPIA2 revealed that it varied for ccRCC and pRCC, although no such difference was found in the context of chRCC (Figure 7.3 B). Furthermore, a higher expression of ACE2 in ccRCC was associated with a more favourable prognosis (Figure 7.4 A-B), while no significant survival difference was observed based on ACE2 levels in chRCC and pRCC., (Figure 7.4 A-B).

## Cancer vs Normal

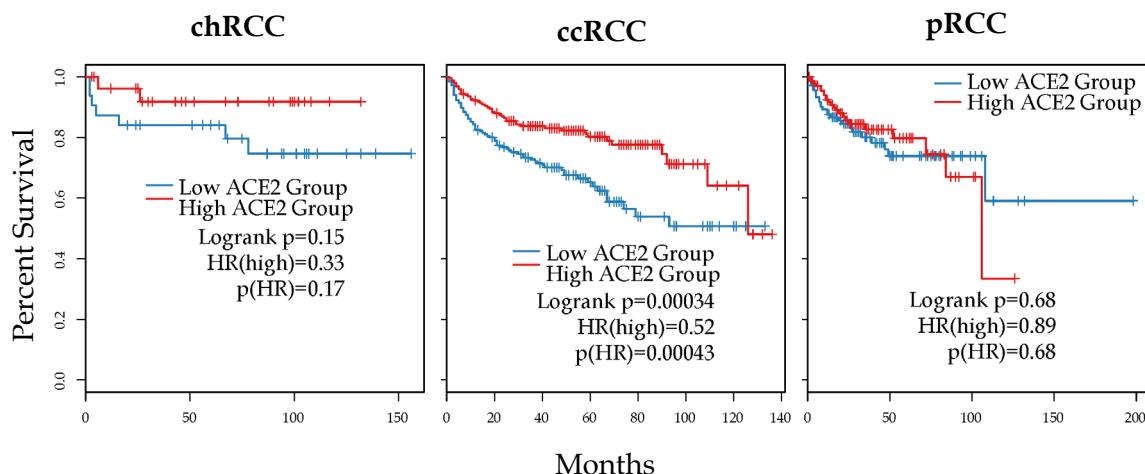


**Figure 7.3.** Exploratory analyses of *ACE2* expression within three RCC subtypes: chRCC, ccRCC, and pRCC. (A) Comparison of *ACE2* expression in RCC subtypes and their respective paired normal kidney tissues. There were 66 tumour and 53 normal samples in chRCC group, 523 tumour and 100 normal samples in ccRCC group and 286 tumour and 60 normal samples in pRCC group. (B) *ACE2* expression levels across RCC subtypes categorized by the four major stages. The data utilized for these analyses were sourced from GEPIA and with a criteria of  $\log_2FC = 1$  and a  $p$ -value of 0.05, with datasets originating from The Cancer Genome Atlas (TCGA) matched normal and GTEx data, and *ACE2* expression grouping into low and high expression was determined by the median. In both panels *ACE2* expression is shown as transcripts per million (TPM).

## Overall Survival



## Disease Free Survival



**Figure 7.4.** Survival analysis conducted for overall and disease-free survival in three RCC subtypes (chRCC, ccRCC, and pRCC) subtypes based on low and high ACE2 expression levels. (A) Overall Survival; (B) Disease free survival. For both survival analyses, there were  $n(\text{high})=141$  and  $n(\text{low})=141$  in pRCC group,  $n(\text{high})=258$  and  $n(\text{low})=258$  in ccRCC, and  $n(\text{high})=28$  and  $n(\text{low})=32$  in chRCC group. The data utilized for these analyses were sourced from GEPIA and with a criteria of p-value of 0.05, and ACE2 expression grouping into low ( $n(\text{low})$ ) and high ( $n(\text{high})$ ) expression was determined by the median.

## 7.4 ACE2 is Differentially Expressed in chRCC Compared to Matched ccRCC Normal Tissue

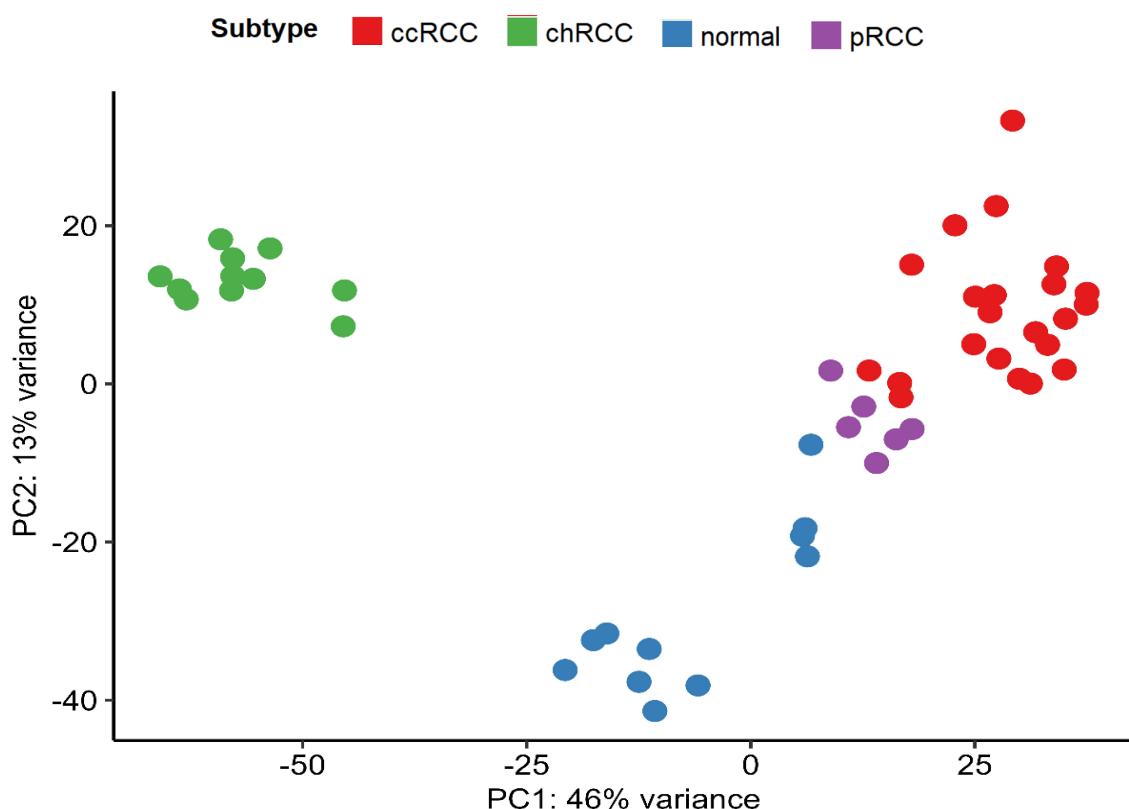
To further examine the role of ACE2 expression in RCC, we identified RNA-seq datasets for patients with RCC using the Gene Expression Omnibus (GEO) (Table 6.1) and subsequently obtained each dataset at the European Nucleotide Archive (ENA). As cancer-associated contexts appear to be important factors that dictate the role ACE2 plays in human cancer, datasets encompassing different cancer-associated contexts were selected for further investigation: selected contexts included RCC subtypes, cancer stages, metastasis progression status and prior treatment. Such metadata allowed for a more in-depth investigation of ACE2 expression in RCC, which has not been previously presented in the literature.

Stage, metastasis progression and treatment will be discussed in later sections; however, in this subsection, we will elaborate on the analysis results in the subtype context. Besides, the following analyses were done using the following RNA-seq datasets: GSE217386, GSE188486, and GSE217386.

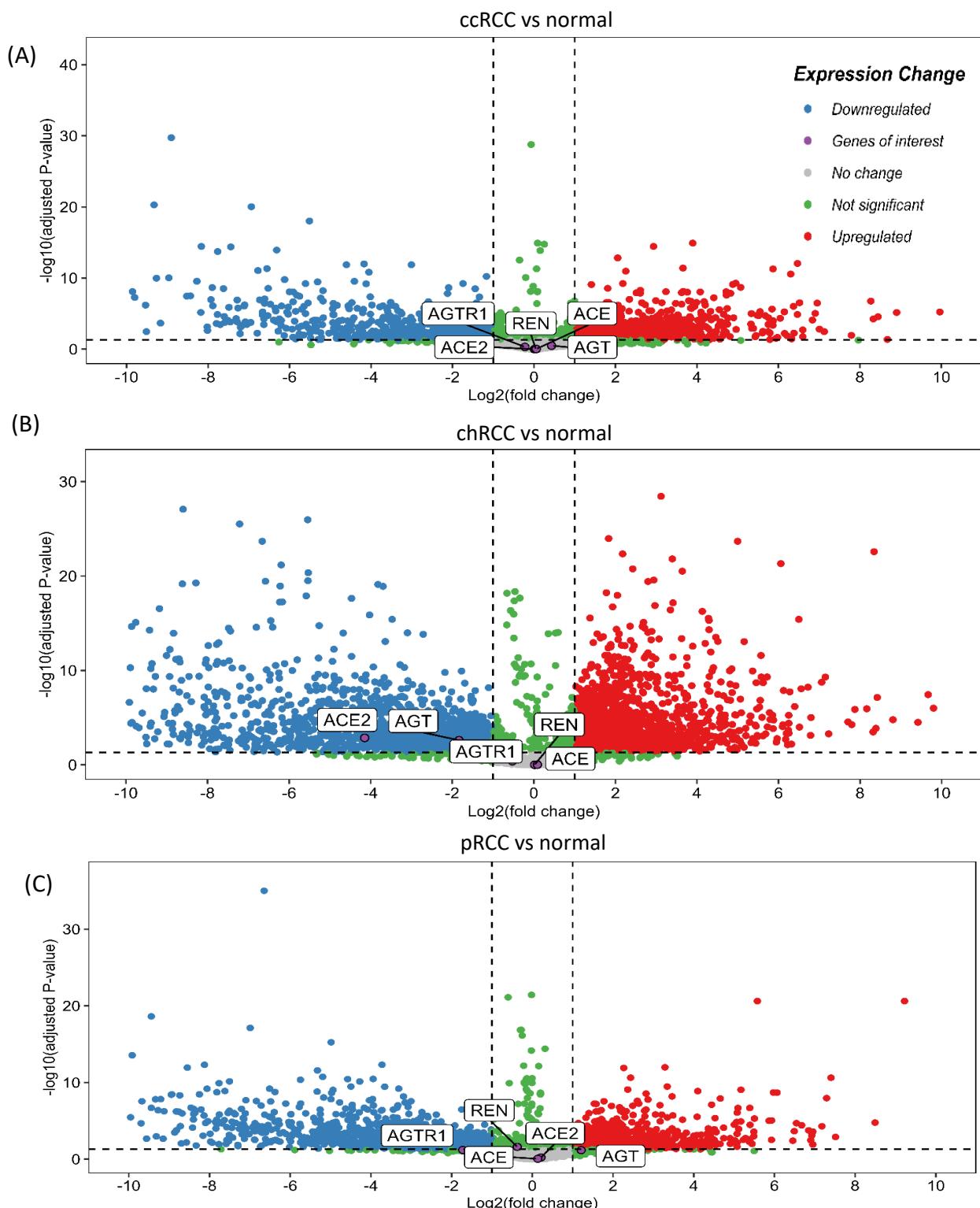
To begin with, firstly, a principal component analysis (PCA) (Figure 7.5) was carried out, encapsulating gene expression per sample via two principal components, revealing sample clustering based on subtype.

Subsequently, we conducted differential expression analysis using DESeq2, both between subtypes and matched ccRCC normal tissues, and within subtypes. The overall distribution of gene expression in these differential gene expression comparisons is showcased in Figure 7.6 and Supplementary Figure 10.5. The volcano plots show many differentially expressed genes in all comparisons which suggest that RCC subtypes and normal differ in global gene expression patterns. ACE2 exhibited consistent downregulation in all chRCC samples when compared to other subtypes. Conversely, the ACE2 expression difference was not significant between chRCC and pRCC compared to normal and within those subtype comparisons (Table 7.1).

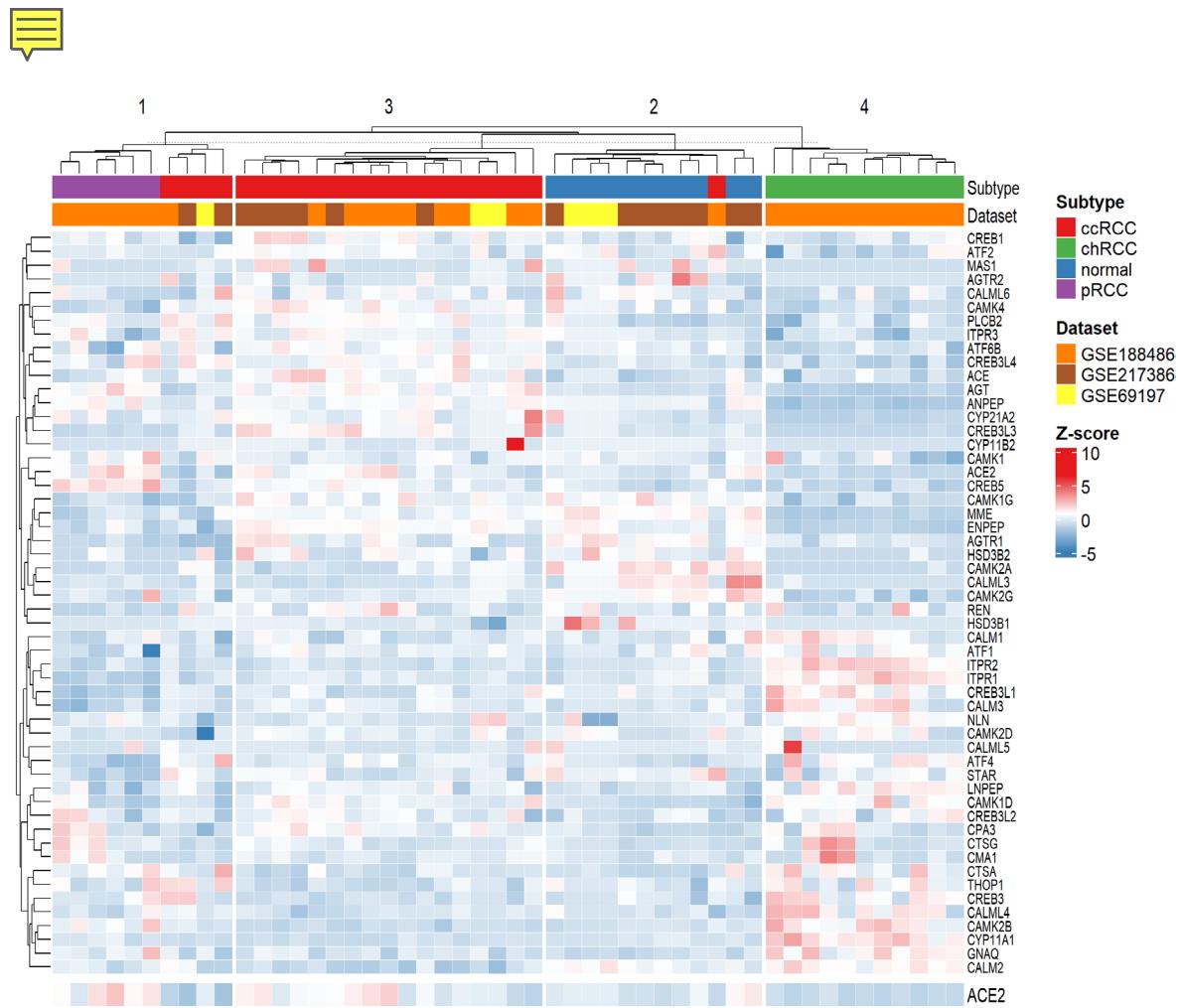
Next, we generated a heatmap (Figure 7.7) to visualise the expression patterns of the normalised mRNA expression z-scores of all genes within the renin-angiotensin-aldosterone system (RAAS) (Table 4.1). In the heatmap, we can see sample-to-sample variability within pRCC and ccRCC subtypes. In contrast, *ACE2* expression is visually similar if we compare only chRCC or normal samples with each other. In addition, employing k-means clustering with  $k=4$ , we observed distinct cluster formations: two clusters exclusively comprised samples from ccRCC and chRCC, while the third exhibited a mixture of predominantly pRCC samples with a few ccRCC samples and the fourth normal samples with one ccRCC sample.



**Figure 7.5.** PCA plot summarizing gene expression of ccRCC matched normal and RCC subtype samples into two principal components. Clustering pattern by subtype can be observed.



**Figure 7.6.** Volcano Plots depicting differential gene expression analyses: (A) ccRCC vs. normal, (B) chRCC vs. normal, and (C) pRCC vs. normal. Differential expression of five genes – *ACE2*, *ACE*, *AGTR1*, *REN* and *AGT* – from the two main RAS axes are highlighted in boxes. Other three genes – *AGTR2*, *CYPB11B2* and *MAS1* – that are also assigned to the main two RAS axes are not displayed because they were either not expressed or their expression data was filtered during the DGEA process due to low counts.



**Figure 7.7.** Heatmap displaying the normalized z-scores of all genes within the renin-angiotensin-aldosterone system (RAAS). For clarity, ACE2 expression is highlighted at the bottom of the heatmap. Visually, ACE2 is underexpressed in chRCC and has patient intervariability for ccRCC, pRCC and normal samples.

**Table 7.1.** P-adjusted and log2FC of ACE2 expression from a DGEA between RCC subtypes.

Comparison	P-adjusted	Log Fold 2 change
chRCC VS normal	0.0014 **	-4.148
ccRCC VS normal	0.9956	0.010
pRCC VS normal	0.5905	0.216
ccRCC VS chRCC	0.0022 **	3.053
pRCC VS chRCC	0.0001 ***	6.342
pRCC VS ccRCC	0.5916	0.114

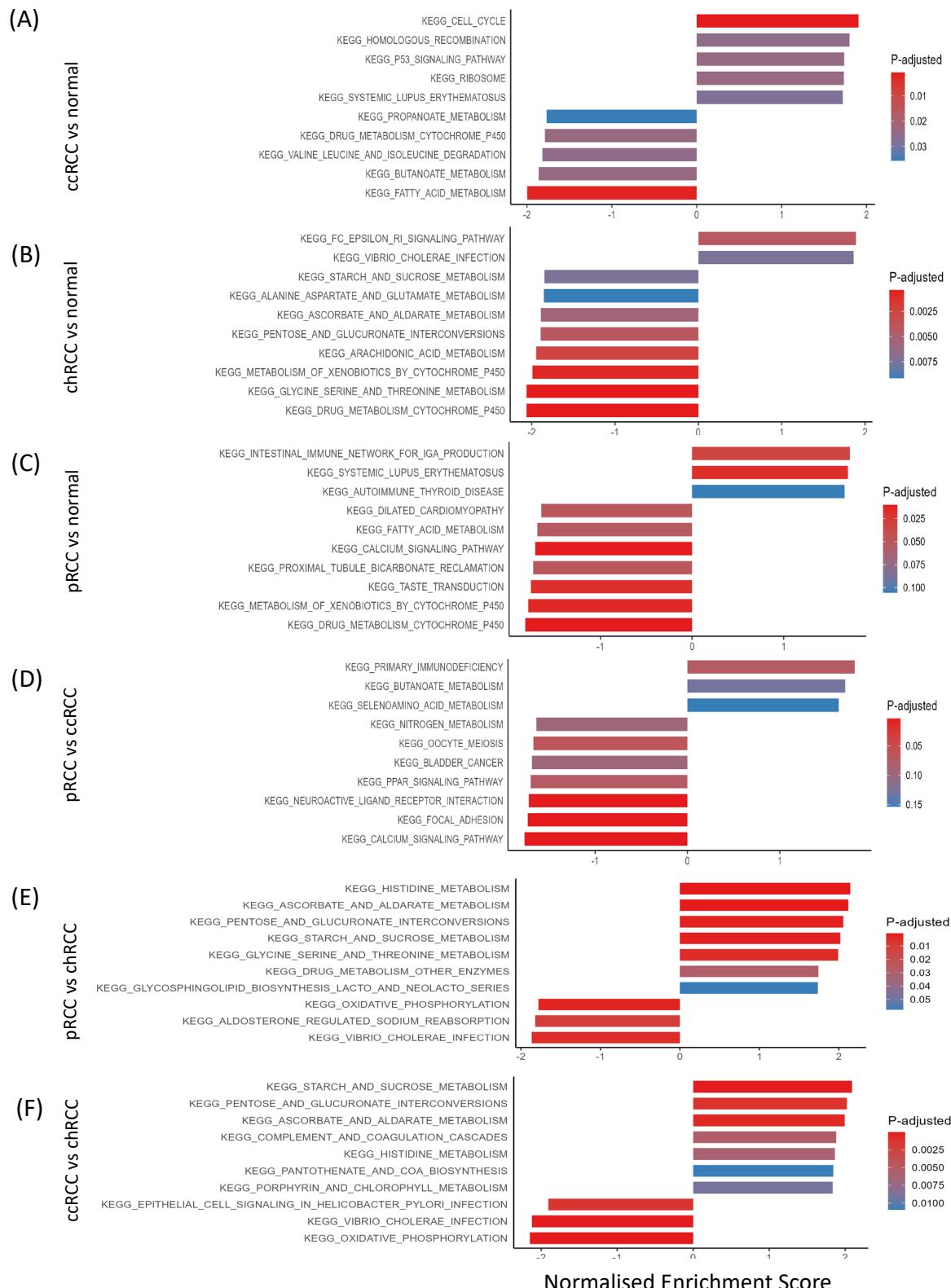
P-Adjusted Values (p-adj) and log2FC for specified type comparisons are included. Significance values: ns: p > 0.05; \*: p-adj <= 0.05; \*\*: p-adj <= 0.01; \*\*\*: p-adj <= 0.001.

## 7.5 RAAS is Not Enriched in the Subtype Context

As explained in the previous subsection, we decided to focus on the subtype context with our further investigation. Subsequently, we tested if statistically significant *ACE2* differential expression in the subtype context is due to dysregulation of the *ACE2* gene that leads to its differential expression in RCC subtypes or if this is due to dysregulation of the whole RAAS that as a result influences *ACE2* expression.

Gene set enrichment analysis (GSEA) is a tool for assessing whether a set of genes collectively exhibits differential expression patterns. We utilised the clusterProfiler R package, using Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets sourced from the MSigDB database, and applied it to the previously obtained DGEA results (Section 7.3). Notably, the KEGG database encompasses diverse gene sets corresponding to various biological pathways. Specifically, we focused on the renin-angiotensin system gene set, denoted as "KEGG\_RENIN\_ANGIOTENSIN\_SYSTEM," for our analysis.

GSEA revealed that RAAS was not highly enriched (Table 7.2) and did not rank among the top ten enriched pathways for comparisons between subtypes and normal tissues (Figure 7.8 A-C) and within subtype comparisons (Figure 7.8 D-F).



**Figure 7.8.** The gene set enrichment plots (A-F) show top ten most enriched pathways for comparisons between RCC subtypes against normal samples, (A) ccRCC vs normal, (B) chRCC vs normal, and (C) pRCC vs normal, and compared to each other (D) pRCC vs ccRCC, (E) pRCC vs chRCC, and (F) ccRCC vs chRCC. Notably, renin-angiotensin system is not among top 10 enriched pathways in neither of the comparisons.

**Table 7.2.** GSEA results. Normalised enrichment score (NES) and adjusted p-value for the renin-angiotensin system across comparisons between RCC subtypes and normal samples. None of the comparisons passed the p-adjusted significance cut-off < 0.05.

Type of comparison	NES	P-adjusted
chRCC VS normal	-1.37127	0.2810
ccRCC VS normal	0.490921	0.9951
ccRCC VS chRCC	-0.7452	0.8805
pRCC VS chRCC	1.217996	0.5951
pRCC VS ccRCC	1.575236	0.1755
pRCC VS normal	1.587347	0.1928

Interestingly, KEGG\_CELL\_CYCLE (Figure 7.8 A), KEGG\_PRIMARY\_IMMUNODEFICIENCY (Figure 7.8 D), and KEGG\_ALDOSTERONE\_REGULATED\_SODIUM\_REABSORPTION (Figure 7.8 E) showed up among the top 10 enriched pathways. Positive cell cycle enrichment in ccRCC compared to normal suggests that cell cycle dysregulation may play a role in ccRCC carcinogenesis. Moreover, the positive enrichment of the gene set typical to immunodeficiency between pRCC and ccRCC may suggest that immune infiltration is at low quantities in pRCC compared to ccRCC.

Lastly, the negative enrichment of aldosterone-controlled sodium reabsorption between pRCC and chRCC raises the question of why it is negatively enriched in pRCC. One of the possible reasons could be that the RAAS system became dysregulated because aldosterone is the final product of the classical RAAS axes. On the other hand, some pathways were common between pairs of comparisons along with many enriched metabolic pathways (Table 7.3).

**Table 7.3.** Commonly enriched pathways between a pair of comparisons, as determined by GSEA results.

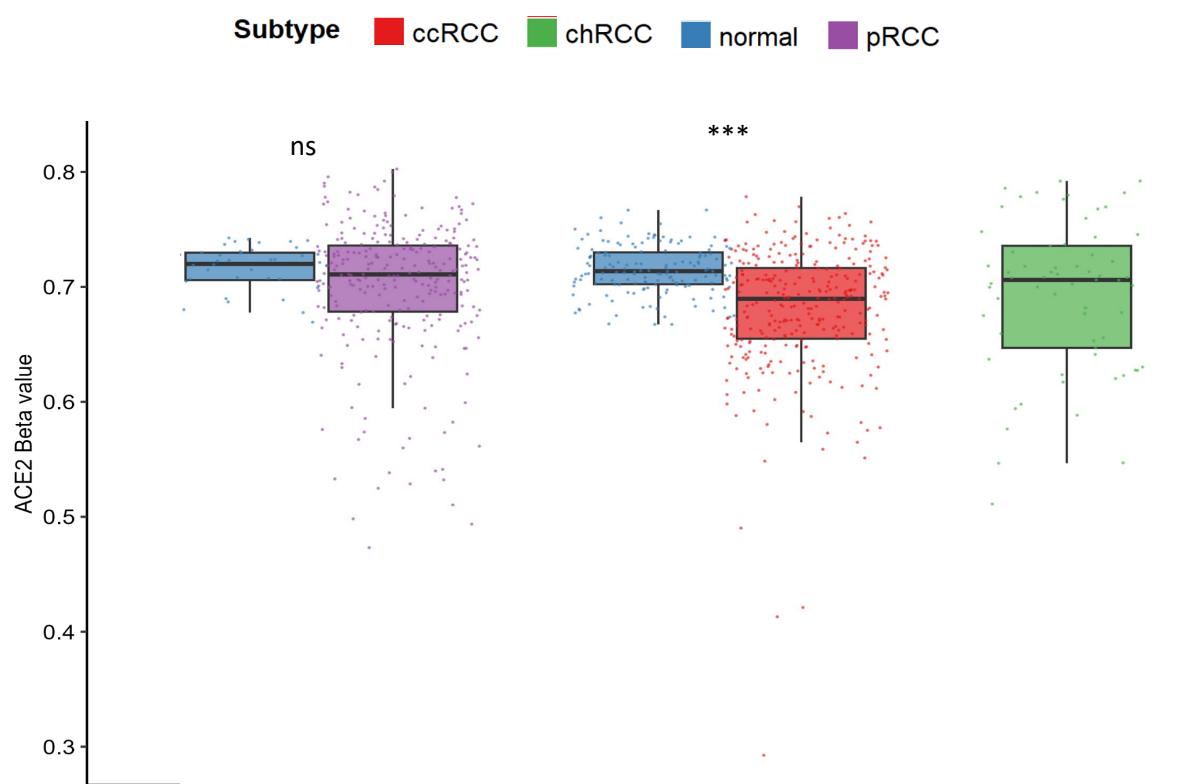
Nr	Comparison	Common Pathways
1	(ccRCC vs normal) VS (chRCC vs normal)	KEGG_DRUG_METABOLISM_CYTOCHROME_P450
2	(ccRCC vs normal) VS (pRCC vs normal)	KEGG_FATTY_ACID_METABOLISM, KEGG_DRUG_METABOLISM_CYTOCHROME_P450, KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS
3	(ccRCC vs normal) VS (pRCC vs normal)	KEGG_BUTANOATE_METABOLISM
4	(chRCC vs normal) VS (pRCC vs normal)	KEGG_DRUG_METABOLISM_CYTOCHROME_P450, KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450
5	(pRCC vs normal) VS (pRCC vs ccRCC)	KEGG_CALCIUM_SIGNALING_PATHWAY
6	(pRCC vs chRCC) VS (ccRCC vs chRCC)	KEGG_HISTIDINE_METABOLISM, KEGG_ASCORBATE_AND_ALDARATE_METABOLISM, KEGG_PENTOSE_AND_GLUCURONATE_INTERCONVERSIONS, KEGG_STARCH_AND_SUCROSE_METABOLISM, KEGG_VIBRIO_CHOLERAES_INFECTION, KEGG_OXIDATIVE_PHOSPHORYLATION

## 7.6 ACE2 is Hypomethylated in ccRCC

After observing significant differential expression between RCC subtypes in the DGEA, we performed methylation analysis (MA) to see if the methylation of the *ACE2* gene in RCC subtypes is significantly different when compared to normal kidney tissue. Methylation of a gene is one of the epigenetic gene control mechanisms that upregulates or downregulates a gene if it is hypomethylated or hypermethylated (Li *et al.*, 2019). Hence, if the *ACE2* methylation pattern matches the pattern of *ACE2* expression observed in the DGEA, then this may suggest that methylation could be a driver of differential expression in RCC subtypes. As a side note, a matching pattern means that if we observe a gene upregulation in cancer vs. normal, then we expect to see a gene hypomethylation in that cancer. Conversely, if a gene is downregulated in a cancer, then we expect to see its hypermethylation.

The Shiny Methylation Analysis Resource Tool (SMART) database provides data on the methylation levels of human genes in various cancer types including chRCC, ccRCC, and pRCC. Hence, we conducted MA of ACE2 methylation levels in these subtypes, comparing them to matched normal samples (Figure 7.9). It is important to note that matched normal samples for the chRCC subtype were not available.

Notably, a statistically significant difference in methylation levels was observed between ccRCC and their matched normal samples. Nonetheless, *ACE2* was not significantly expressed in ccRCC compared to normal. Therefore, the methylation pattern and differential expression of *ACE2* in ccRCC compared to normal do not match. In contrast, the methylation of *ACE2* in pRCC compared to normal was not significant which aligns with the DGEA results (Table 7.1).



**Figure 7.9.** Methylation levels of ACE2 in RCC subtypes compared to matched normal kidney samples. Methylation data was retrieved from the SMART database. Significance levels are denoted as follows: "ns" for  $p > 0.05$ , "\*" for  $p \leq 0.05$ , "\*\*" for  $p \leq 0.01$  and "\*\*\*" for  $p \leq 0.001$ . Beta-value indicates the level of methylation of a gene, 0 means that a gene has no methylation, whereas the beta-value of 1 specifies that a gene is completely methylated.

## 7.7 Transcription Factor Enrichment Analysis Does Not Reveal Major Epigenetic Regulators of RAAS Genes in RCC

To supplement the interrogation of epigenetic factors that may regulate *ACE2* or RAAS in RCC subtypes, we investigated the enrichment scores of transcription factors (TFs). Transcription factors are proteins that bind genes and activate or suppress the transcription of a gene.

DoRothEA R package contains information about interactions between TFs and genes, also known as regulons. In addition, such information contains details if a regulon activates or suppresses the transcription of a gene. Furthermore, it also contains the confidence score of a regulon which signifies the reliability of this regulation, ranging from A (highest confidence) to E. DoRothEA determines the confidence level of a regulon using multiple sources, including literature-curated resources, ChIP-seq peaks, motif analysis, as well as an inference from gene expression data (Müller-Dott *et al.*, 2023). It assigns the A confidence if all the resources mentioned above agree that a certain TF regulates a gene. In contrast, it assigns an E confidence if only one resource is available to proof an interaction between a TF and a gene.

Accordingly, we harnessed DoRothEA alongside VIPER, another R package, to perform transcription factor enrichment analysis (TFEA) and quantify the activity of TFs in RCC subtypes and chRCC-matched normal samples. However, in the TFEA, we selectively included TFs that regulate genes at the A and B confidence levels, to increase the reliability of TFEA results. Additionally, we incorporated TFs that regulate *ACE2*, despite their classification at the E confidence level. Furthermore, based on literature findings, we introduced several additional TFs, namely STAT3, HNF1 $\alpha$ , HNF1 $\beta$ , and FOXA2, as potential regulators of *ACE2*.

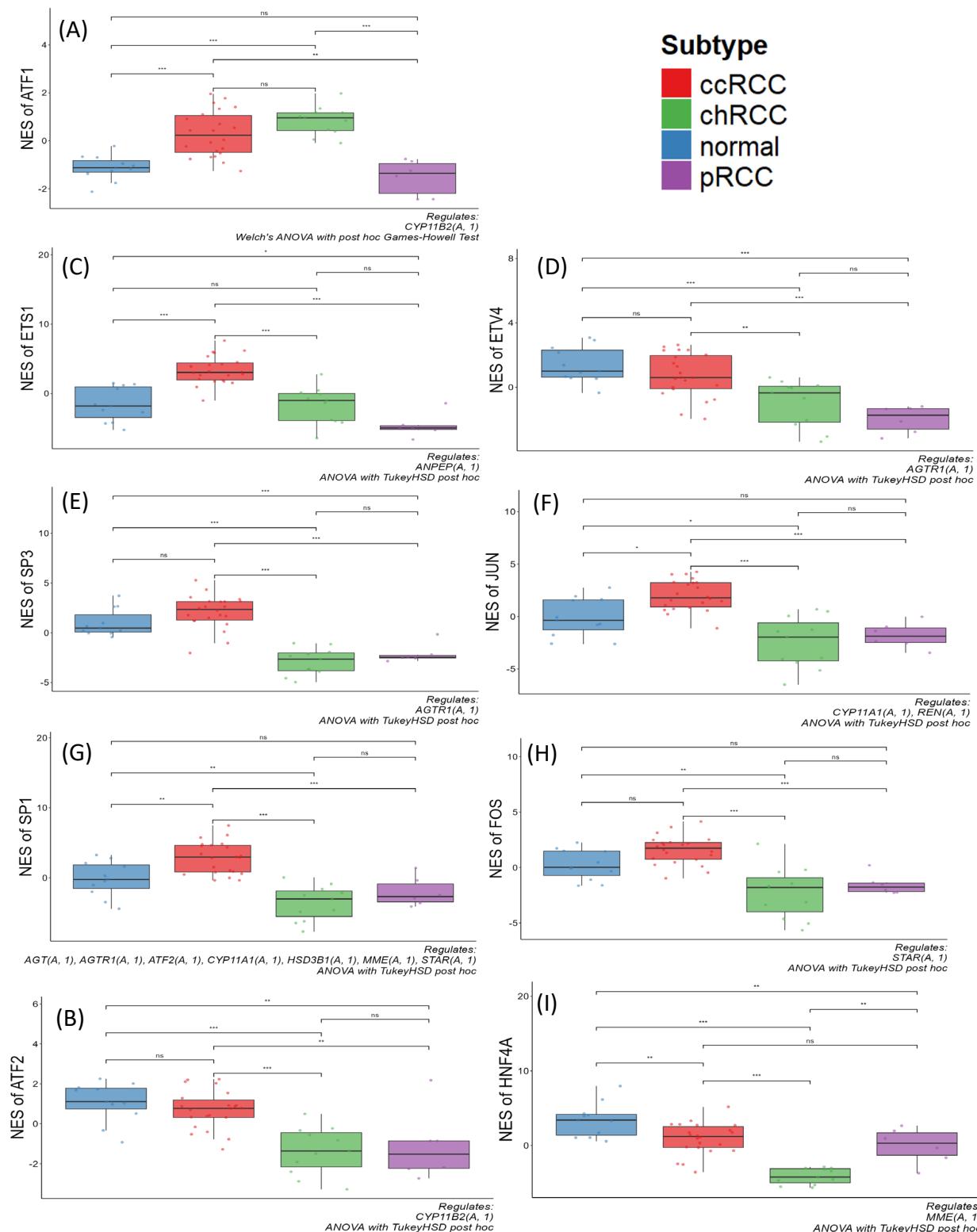
Subsequently, in the interpretation of TFEA results, we included *ACE2* and RAAS-related TFs that were significantly enriched in at least 4 out of 6 comparisons in the subtype context. Eight TFs were enriched in 4 comparisons (Figure 7.10 A-H) and one in 5 comparisons (Figure 7.10 I).

To understand if those nine TFs may regulate genes in the RAAS, firstly, we must inspect if their activity matches the gene expression pattern across subtypes of the genes that they regulate. For example, if a TF is enriched in cancer compared to normal, then the expected gene expression pattern of the genes that it activates is gene upregulation in cancer. Conversely, a gene is expected to be downregulated in cancer if that TF suppresses a gene and vice versa if a TF is negatively enriched in cancer.

However, none of those TFs, except HNF4A, had a matching enrichment pattern as the gene they regulate. HNF4A was significantly enriched in 5 out of 6 comparisons. It regulates Neprilysin (MME). It is known that the downregulation of MME increases the bioavailability of natriuretic peptides and bradykinin which results in the antihypertensive response. Subsequently, it was suggested that MME is downregulated to counterbalance the reduction of Ang (1-7) (Mehranfard *et al.*, 2021).

The only TF regulating *ACE2* that was significantly enriched, but only in 2 out of 6 comparisons was FOXA2 (data not shown). Nonetheless, its enrichment pattern did not align with the *ACE2* expression pattern across subtypes and it also does not give a deeper understanding of what caused differential expression between RCC subtypes.

In summary, the TFEA analysis did not yield substantial conclusions regarding the relationships between certain TFs and *ACE2* expression, nor did it shed light on the regulation of these TFs in the context of RAAS pathway genes.



**Figure 7.10.** Normalised enrichment scores of 9 TFs across four sample types: chRCC, ccRCC, pRCC and matched cRCC normal. These 9 TFs displayed significantly different normalised enrichment score in at least 4 of 6 comparisons. Each boxplot includes details about the genes regulated by the respective TF and the statistical method used to assess the significance of differences in TF enrichment scores between subtypes. Significance levels are denoted as follows: "ns" for  $p > 0.05$ , "\*\*" for  $p < 0.05$ , "\*\*\*" for  $p < 0.01$ , and "\*\*\*\*" for  $p < 0.001$ .

## 7.8 ACE2 Correlates with Immune Infiltration in Renal Cell Carcinoma

In the following analysis, we investigated *ACE2* correlation with immune infiltrates and aimed to understand why high *ACE2* expression gives a survival advantage in ccRCC patients and no advantage in chRCC and pRCC (Figure 7.4 A-B). It is known that immune infiltration plays an important role in RCC development (Wang *et al.*, 2019), affects the tumour microenvironment (Tang *et al.*, 2023) and impacts RCC patients survival (Chen *et al.*, 2021). Moreover, it was also shown that RCC components play a major role in immunomodulation (Laghla *et al.*, 2021), whereas *ACE2* sensitises breast cancer to immunotherapy and chemotherapy by remodelling the tumour microenvironment (Mei *et al.*, 2022). In contrast, abnormal *ACE2* expression was associated with immune infiltration that worsens the prognosis of COVID-19 patients in UCEC and pRCC (Yang, Li and Zhou, 2020).

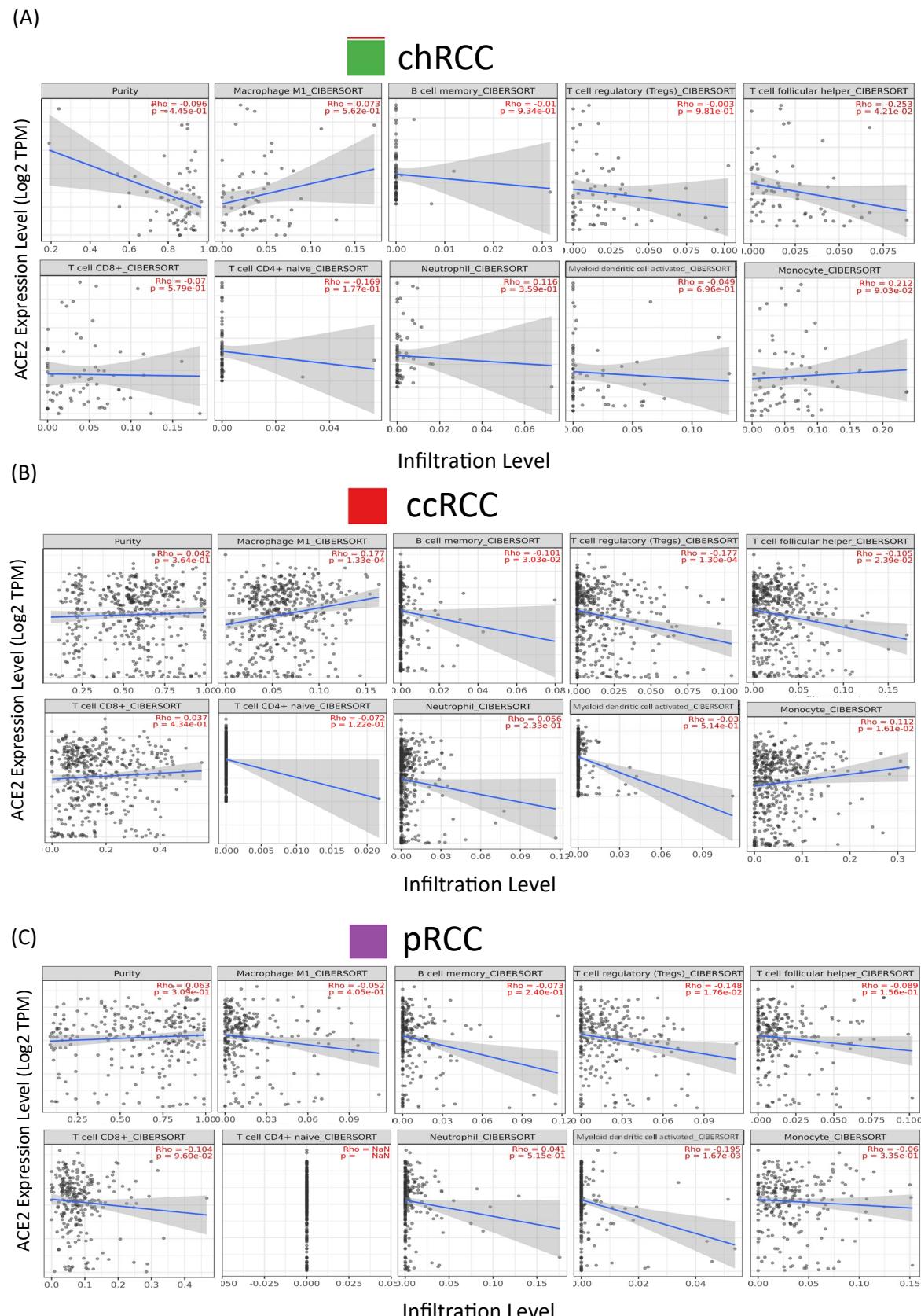
Nonetheless, the survival advantage of RCC subtype patients is affected differently by different immune infiltrates. T cell regulatory (Tregs) and T cell follicular helper (Tf-helper) were linked to a poor prognosis in RCC, while Monocytes were associated with a more favourable prognosis, as supported by studies (Zhu *et al.*, 2019; Y. Wang *et al.*, 2021). Additionally, T cell CD8+ (CD8) was found to be associated with prolonged overall survival in chRCC, whereas in pRCC, Macrophage M1 exhibited a favourable outcome in terms of survival (S. Zhang *et al.*, 2019). In addition, the association of B cell memory (B-memory), T cell CD4+ naïve (CD4), Neutrophils, and activated Myeloid dendritic cells (M-dendritic) was already associated with *ACE2* in pRCC (Yang, Li, Hu, *et al.*, 2020).

Consequently, to investigate *ACE2* role in immunomodulation we selected the nine immune infiltrates mentioned above: Macrophage M1, B cell memory (B-memory), T cell regulatory (Tregs), T cell follicular helper (Tf-helper), T cell CD8+ (CD8), T cell CD4+ naïve (CD4), Neutrophil, Myeloid dendritic cell activated (M-dendritic), and Monocyte.

We interrogated TIMER2.0 for assessing *ACE2* correlation with immune infiltration in chRCC, ccRCC, and pRCC. Using the CIBERSORT algorithm, we obtained correlation scores for *ACE2* with tumour purity and those nine immune cell types (Figure 7.11).

In chRCC, *ACE2* exhibited a positive correlation with Tf-helper cells ( $RHO = 0.253, p = 0.0421$ ) (Figure 7.11 A). In ccRCC, *ACE2* displayed a positive correlation with Macrophages M1 ( $RHO = 0.177, p = 0.000133$ ) and Monocytes ( $RHO = 0.112, p = 0.0161$ ), while showing a negative correlation with B-memory ( $RHO = -0.101, p = 0.0303$ ), Tregs ( $RHO = -0.177, p = 0.00013$ ), and Tf-helper ( $RHO = -0.105, p = 0.0239$ ) (Figure 7.11 B). In addition, *ACE2* exhibited negative correlations with Tregs ( $RHO = -0.148, p = 0.0176$ ) and M-dendritic cells ( $RHO = -0.195, p = 0.00167$ ) in pRCC (Figure 7.11 C).

These findings provide initial insights into the associations between *ACE2* expression and immune cell infiltration across different RCC subtypes, shedding light on the potential immunological dynamics within these cancers.



**Figure 7.11.** Correlation between ACE2 expression levels (Log2 TPM) with tumour purity and nine distinct immune cell types in RCC subtypes: (A) chRCC, (B) ccRCC and (C) pRCC. Spearman's correlation ( $r$ ) was used to calculate correlation scores with  $p < 0.05$  and  $r > 0.2$  indicating a positive correlation and  $p < 0.05$  and  $r < -0.2$  indicating a negative correlation.

## 7.9 RCC Samples do not Cluster by Subtype Based on the Enrichment of Hallmark Cancer Pathways

In the exploratory analysis (Figure 7.3 A) and in the heatmap of the normalised count z-scores of RAAS components (Figure 7.7), we observed variability in the ACE2 expression of patients within the same RCC subtypes. Accordingly, we investigated if patients differ in the enrichment score of cancer hallmark pathways within the same subtype and compared to patients of other subtypes.

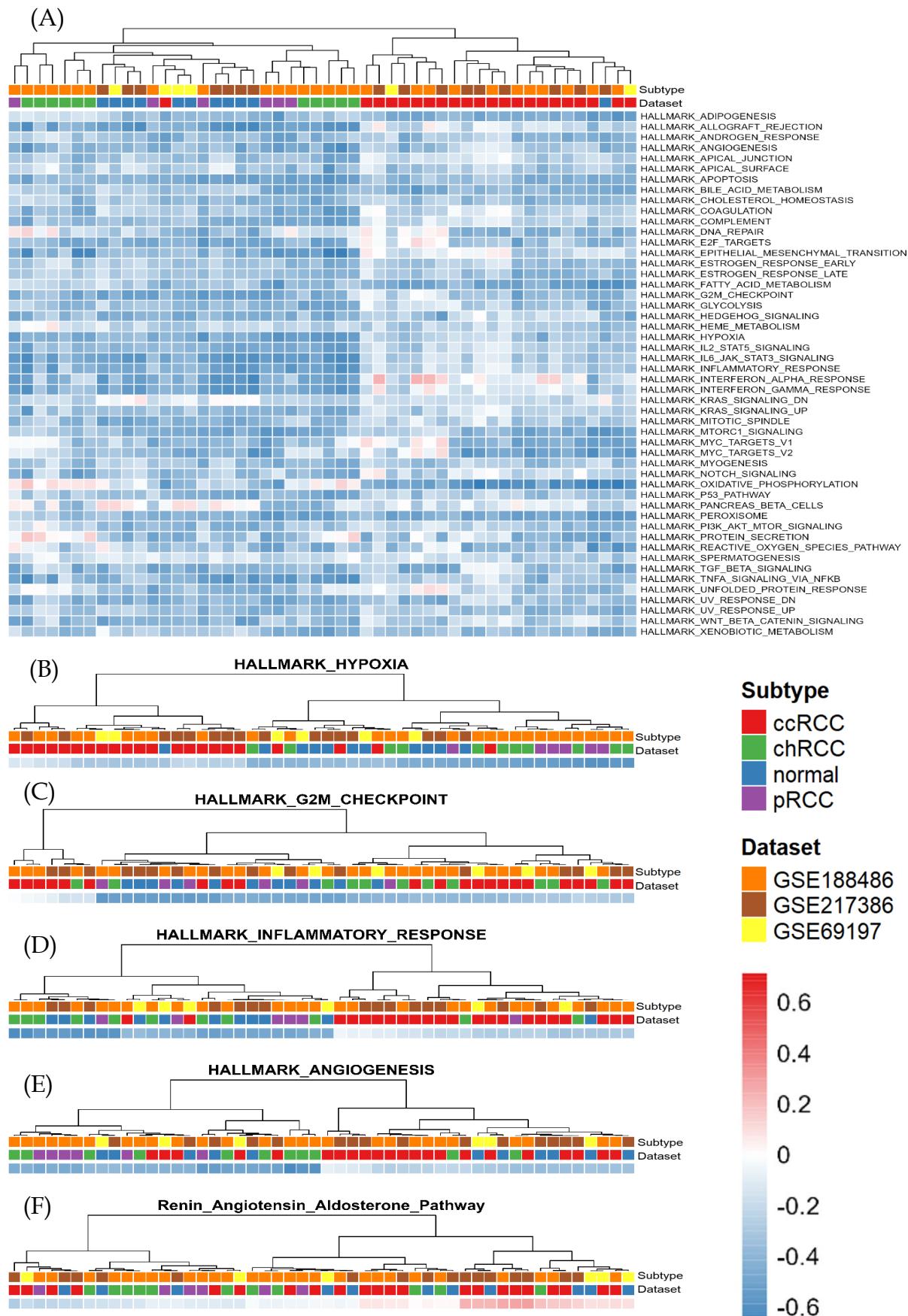
Single sample gene set enrichment analysis (ssGSEA) closely resembles that of GSEA. Both methods assess whether a group of genes collectively shows enrichment between different gene sets. However, unlike GSEA, which measures the enrichment of gene sets between two groups of samples, for example, subtypes, ssGSEA evaluates gene set enrichment individually for each sample.

We utilised the GSVA R package to conduct ssGSEA on samples of three RCC subtypes: chRCC, ccRCC, and pRCC. We employed the HALLMARK gene sets sourced from the MSigDB database, representing hallmark pathways typically associated with cancers. We explored how the samples clustered based on the collective enrichment of all hallmark pathways (Figure 7.12 A). Based on the latter results, we see one cluster with 22 samples, primarily consisting of 21 ccRCC samples and one ccRCC-matched normal sample. However, no further clustering is observed.

We additionally investigated how the samples clustered when considering individual hallmark pathways due to their relevance to ccRCC as indicated by (Petitprez *et al.*, 2021). Specifically, we selected HALLMARK\_HYPOXIA, HALLMARK\_G2M\_CHECKPOINT, HALLMARK\_INFLAMMATORY\_RESPONSE, and HALLMARK\_ANGIOGENESIS. In addition, we also included Renin\_Angiotensin\_Aldosterone\_Pathway in such an analysis (Table 4.1).

Similarly, ccRCC samples also predominantly comprised one of the two major clusters for HALLMARK\_HYPOXIA and HALLMARK\_INFLAMMATORY\_RESPONSE (Figure 7.12 B, D). Thus showing that these two pathways are enriched more in ccRCC

compared to chRCC, pRCC, and normal samples. Conversely, samples did not group according to subtypes based on enrichment scores of HALLMARK\_ANGIOGENESIS, HALLMARK\_G2M\_CHECKPOINT or Renin\_Angiotensin\_Aldosterone\_Pathway (Figure 7.12 C, E-F). This coincides with the patient-to-patient variability in ACE2 expression between the same RCC subtypes mentioned at the beginning of the subsection.



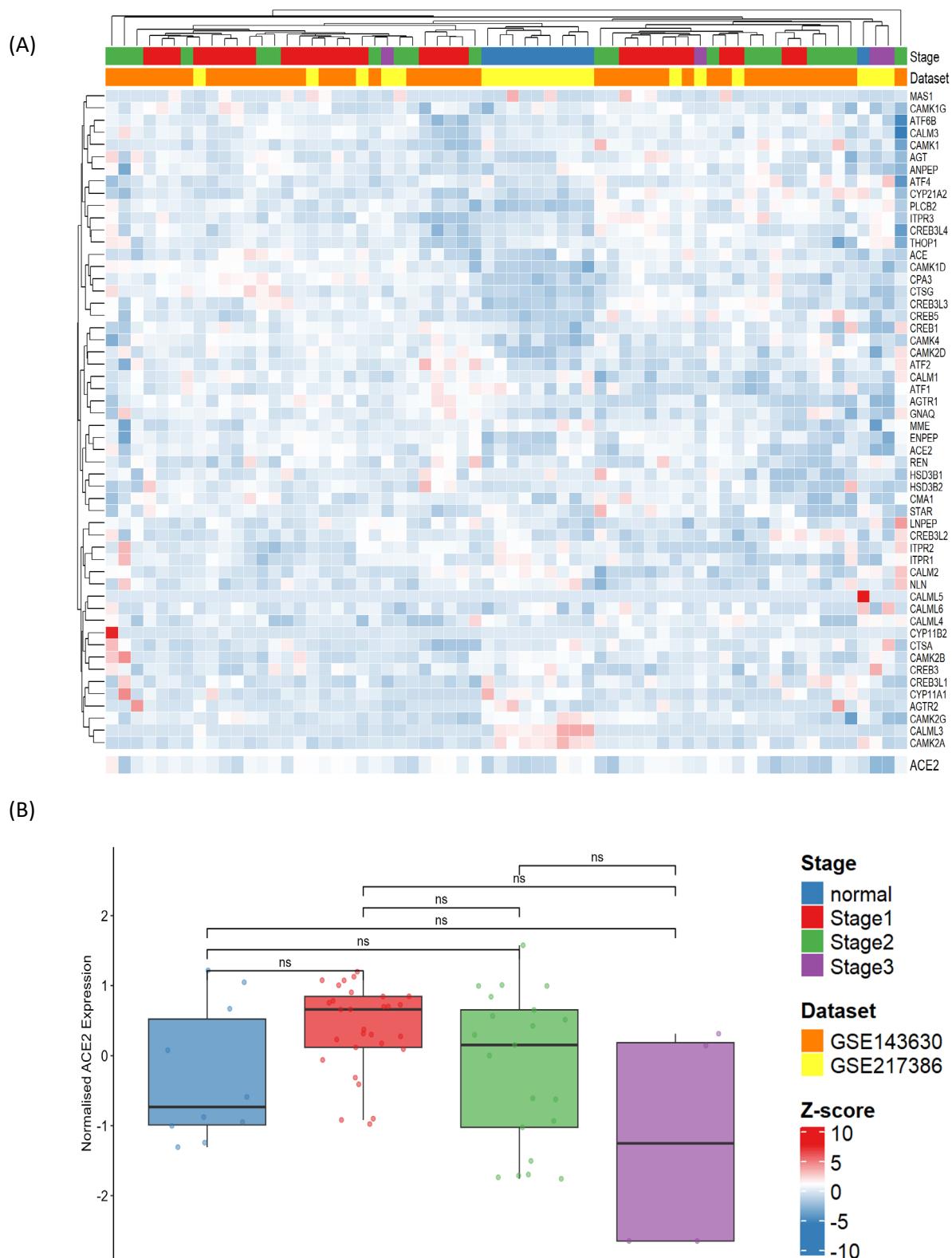
**Figure 7.12.** ssGSEA showing the enrichment scores of the cancer hallmark pathways and RAAS in patient samples across RCC patient samples. (A) Enrichment scores of all cancer hallmark pathways. (B-F) Enrichment scores for four hallmark pathways and RAAS for each patient.

## 7.10 The Role of ACE2 in Alternative Cancer Contexts: Stage, Metastasis and Treatment

In addition to our examination of *ACE2* expression in the context of RCC subtypes, we extended our analysis to three additional contexts: stage, progression to metastasis, and treatment. However, as mentioned in the section previously, due to limited data availability for these cancer-associated contexts in patients with chRCC and pRCC, our analyses were applied to only ccRCC datasets as follows: GSE143630, GSE153262, GSE217386.

### 7.10.1 ACE2 expression in cancer stage in clear cell renal cell carcinoma (ccRCC).

We assessed if *ACE2* expression differs across the three major cancer stages and normal samples. The initial heatmap of normalised *ACE2* and RAS-associated genes z-scores shows no clustering by stage (Figure 7.13 A). Similarly, PCA also did not display any clustering by stage (Supplementary Figure 10.6 B). Furthermore, looking only at *ACE2* expression we observed no significant difference in *ACE2* expression between stages of ccRCC (Figure 7.13 B and Table 7.4)



**Figure 7.13.** Investigation of RAAS and *ACE2* expression in RCC stage. (A) Heatmap displaying the normalized z-scores of all genes within the renin-angiotensin-aldosterone system (RAAS). For clarity, *ACE2* expression is highlighted at the bottom of the heatmap. Visually, *ACE2* is underexpressed in all samples. (B) Boxplot of *ACE2* expression across cancer stages. No statistical significance was obtained. In addition, statistical testing was performed with DESeq2 while using the log2FC threshold of 1 and p-adjusted < 0.05.

**Table 7.4.** P-adjusted and log2FC of ACE2 expression from a DGEA between stages.

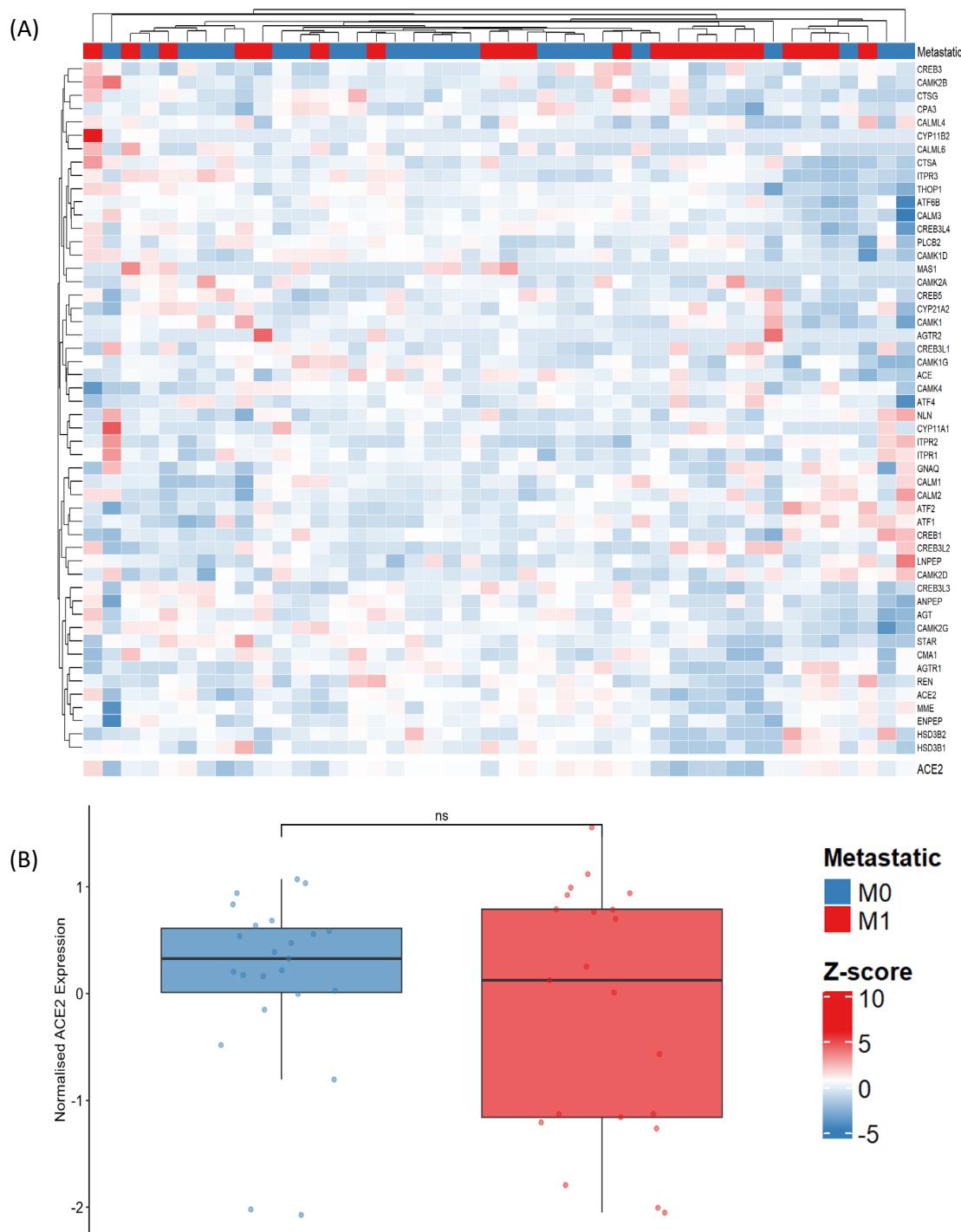
Comparison	P-adjusted	Log2FoldChange
Stage1 vs Normal	0.30105	0.258249
Stage2 vs Normal	0.574384	0.111521
Stage3 vs Normal	0.818471	0.10026
Stage2 vs Stage1	0.73123	-0.05271
Stage3 vs Stage1	0.655148	-0.08786
Stage3 vs Stage2	0.866416	-0.0305

P-Adjusted Values (p-adj) and log2FC for specified type comparisons are included. Significance values: ns: p > 0.05; \*: p-adj <= 0.05; \*\*: p-adj <= 0.01; \*\*\*: p-adj <= 0.001.

### 7.10.2 ACE2 expression and the progression of clear cell renal cell carcinoma to metastasis.

Next, we investigated the expression of *ACE2* and RAS-associated genes in samples collected from patients classified as either stage 1 or stage 2 ccRCC at the time pf sample collection, who later on remained free from metastasis (M0) or developed metastasis (M1). In line with our previous observations, no distinct clustering by M0 or M1 status was evident (Figure 7.14 A and Supplementary Figure 10.6 C).

Similarly, when examining *ACE2* expression between M0 and M1, we also observed no significant differences (Figure 7.14 B and Table 7.5), suggesting that *ACE2* and RAS-associated gene expression may not be predictive of progression to metachronous metastasis



**Figure 7.14.** Investigation of RAAS and ACE2 expression M0 and M1. (A) Heatmap displaying the normalized z-scores of all genes within the renin-angiotensin-aldosterone system (RAAS). For clarity, ACE2 expression is highlighted at the bottom of the heatmap. Visually, ACE2 expression do not cluster by sample type. (B) Boxplot of ACE2 expression showing differences between M0 and M1. No statistical significance was obtained. In addition, statistical testing was performed with DESeq2 while using the log2FC threshold of 1 and p-adjusted < 0.05.

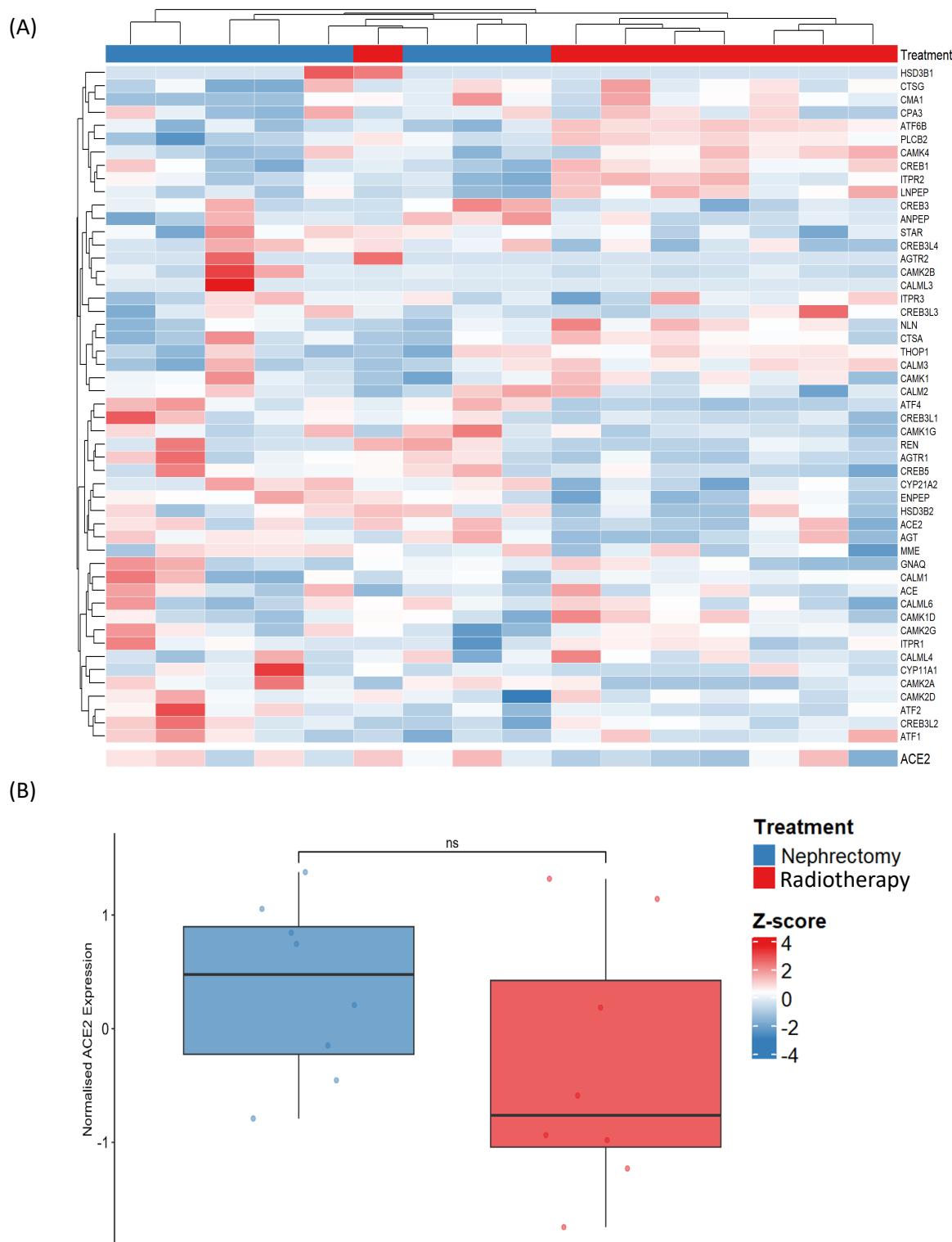
**Table 7.5.** P-adjusted and log2FC of ACE2 expression from a DGEA between patients that did not progress to metastasis compared to patients that did progress to metastasis.

Comparison	P-adjusted	Log2FoldChange
M1_vs_M0	0.983348	-0.00057
P-Adjusted Values (p-adj) and log2FC for specified type comparisons are included. Significance values: ns: p > 0.05; *: p-adj <= 0.05; **: p-adj <= 0.01; ***: p-adj <= 0.001.		

### 7.10.3 ACE2 expression in clear cell renal cell carcinoma (ccRCC) following prior treatment with radiotherapy.

Our final analysis focused on investigating potential differences in *ACE2* and RAS-associated gene expression and between patient groups that underwent different treatments: patients were either treated by nephrectomy alone or nephrectomy following radiotherapy (referred to hereafter as radiotherapy). Interestingly, when considering all RAS-associated gene sets, two distinct clusters were evident; these clusters were predominantly segregated according to the presence or absence of radiotherapy in the treatment of ccRCC (Figure 7.15 A). When assessing the PCA plot (Supplementary Figure 10.6 D), we also observed the formation of patient samples into two distinct clusters that correspond to the type of treatment. Overall, this may suggest a change in the RAS following radiotherapy. However, DGEA did not show *ACE2* among the differentially expressed genes between these two treatments (Figure 7.15 B and Table 7.6).

In summary, our comprehensive analysis across different contexts – stage, metastasis progression, and treatment – revealed no statistically significant differences in *ACE2* expression.



**Figure 7.15.** Investigation of RAAS and ACE2 expression between patients treated with nephrectomy and nephrectomy followed by radiotherapy. (A) Heatmap displaying the normalized z-scores of all genes within the Renin-Angiotensin-Aldosterone System (RAAS). For clarity, ACE2 expression is highlighted at the bottom of the heatmap. Visually, ACE2 is underexpressed patients treated with radiotherapy compared to patients treated only with nephrectomy. (B) Boxplot of ACE2 expression in patient groups treated only with nephrectomy and patients treated with nephrectomy followed by radiotherapy. No statistical significance was obtained. In addition, statistical testing was performed with DESeq2 while using the log2FC threshold of 1 and p-adjusted < 0.05.

**Table 7.6.** P-adjusted and log2FC of ACE2 expression from a DGEA between patients treated with nephrectomy and patients treated with nephrectomy followed by radiotherapy.

Comparison	P-adjusted	Log2FoldChange
Nephrectomy vs Radiotherapy	0.871662	-0.03873
P-Adjusted Values (p-adj) and log2FC for specified type comparisons are included. Significance values: ns: p > 0.05; *: p-adj <= 0.05; **: p-adj <= 0.01; ***: p-adj <= 0.001.		

## 8 Discussion and Future Perspectives

Angiotensin-converting enzyme 2 (ACE2) was discovered only two decades ago, independently identified by Donoghue *et al* (2000) and Tipnis *et al.* (2000). Yet, within such a short timeframe, ACE2 has emerged as an important molecule with dualistic effects in cancer biology.

Some research indicates that ACE2 overexpression inhibits the growth, metastasis, and angiogenesis of tumour cells in a variety of cancers, including lung, breast, colon, and pancreatic cancers (Huang *et al.*, 2021). Additionally, complementary studies hypothesised various mechanisms through which ACE2 may suppress cancers. For instance, the upregulation of ACE2 has been demonstrated to suppress cell proliferation and motility while increasing the sensitivity to hypoxia-induced injury in pancreatic cancer cell lines (Zhou *et al.*, 2011). Narayan *et al.* have proposed that ACE2 exhibits its anti-tumour effects in thyroid cancer by regulating angiotensin II (Ang II) (2020). Furthermore, Gottschalk *et al.* have suggested that ACE2 levels are elevated in lung cancer to catalyse the synthesis of tumour growth suppressive angiotensin (1-7) (Ang (1-7)) (2021). Therefore, the duality of ACE2's role in cancer is becoming increasingly evident.

Krishnan *et al.* suggests that Ang (1-7) may prevent metastasis in prostate cancer cell lines (2013), whereas Zheng *et al.* correlate Ang (1-7) synthesis with pro-metastatic properties in a clear cell renal cell carcinoma (ccRCC) cell line (2015). In addition, Khanna *et al.* have identified that ACE2 abrogates the resistance of VEGFR inhibitors and proposed ACE2 as a potential therapeutic target for clear cell renal cell carcinoma (ccRCC) (2021). Studies in patients with colorectal cancer suggest that higher ACE2 expression is associated with improved patient overall survival (OS) (Li *et al.*, 2021) and anti-tumorigenic effects in mouse models (Huang *et al.*, 2023). However, conflicting evidence also suggests that elevated levels of ACE2 and bromodomain-containing protein 4 (BRD4), a transcriptional and epigenetic regulator (Zhang *et al.*,

2023), impact DNA damage/repair and apoptosis, and is predictive of poor patient survival in colorectal cancer (Zhang *et al.*, 2023).

This evidence is complemented by the aforementioned evidence of a dualistic role in breast cancer (Zhang *et al.*, 2019; Bhari *et al.*, 2020; Nair *et al.*, 2021), which suggests that ACE2 may play a cancer-context-dependent role in human cancers. However, these studies do not pinpoint the precise mechanisms underlying the dualistic effects of ACE2 in cancers.

Consequently, we conducted a preliminary bioinformatics-based analysis of ACE2 and renin-angiotensin system (RAS) associated gene expression in cancers. Our objective was to understand if ACE2 plays a cancer-context-dependent role. The potential implication of this project extends the idea that ACE2 may serve as a promising biomarker or therapeutic target in the realm of cancer research.

Initially, after the exploratory analyses (Section 7.3), we chose renal cell carcinoma (RCC) as the focal point of our investigation. Three RCC subtypes - chromophobe renal cell carcinoma (chRCC), ccRCC, and papillary renal cell carcinoma (pRCC) - had either differentially expressed ACE2 or had associated better patient outcomes with elevated ACE2 levels. Such intriguing results suggested that ACE2 may also play a cancer-context-dependent role in RCC. Interestingly, other studies had already pinpointed ACE2 expression as a promising biomarker in ccRCC (Niu *et al.*, 2021; T. Wang *et al.*, 2021) and pRCC (Pathania *et al.*, 2021). Furthermore, the downregulation of ACE2 by SARS-CoV-2 worsens the prognosis of ccRCC and pRCC patients via metabolism and immunoregulation (Tang *et al.*, 2021).

Constricted by RNA-seq data availability on the European Nucleotide Archive (ENA), we explored ACE2 and RAS-associated gene expression in four different RCC contexts: subtype, stage, metastasis progression, and treatment. Notably, patient samples with stage, metastasis progression and treatment were comprised only of ccRCC patient samples. Additionally, we could obtain only ccRCC-matched normal samples but not the matched normal samples for the other two subtypes.

Consequently, we compared the gene expression of ccRCC-matched normal samples against all three RCC subtypes.

Our initial differential gene expression analysis (DGEA) did not reveal significant differences in ACE2 expression between ccRCC and pRCC compared to normal samples. However, we observed significant ACE2 downregulation in chRCC compared to normal, as well as pRCC and ccRCC samples (Table 7.1). These findings align with those of Larrinaga *et al.*, who observed ACE2 downregulation in chRCC but not ccRCC when compared to normal renal tissue. Interestingly, ACE2 protein levels were downregulated in both subtypes, chRCC and ccRCC in the latter study (2010). Thereby, suggesting that protein modifications could occur at a post-transcriptional level and illustrating the importance of not relying solely on mRNA levels as an index of change in protein level (Kasinath *et al.*, 2006).

Furthermore, in our investigation of ACE2 expression and RAS-associated genes in additional cancer-associated contexts beyond subtype - including stage, metastasis progression and treatment - we did not detect significant alterations in ACE2 expression. Nevertheless, Zuo *et al.* have demonstrated that ACE2 increases during chemotherapy in breast cancer rendering patients less sensitive to the chemotherapy (2022). However, we did not find RNA-seq data of patients with chemotherapy and without chemotherapy. Hence, we could not test this hypothesis that the ACE2 expression changes during chemotherapy. On the other hand, we tested if ACE2 expression changes between patients with only nephrectomy alone and nephrectomy followed by radiotherapy. However, we did not find significant changes in ACE2 expression in this context.

The outcomes of our initial DGEA support the idea hypothesis that ACE2 may play a role within the context of RCC subtypes, but not necessarily in the other three cancer-associated contexts. Consequently, we conducted additional analyses, including gene set enrichment analysis (GSEA), transcription factor enrichment analysis (TFEA), and methylation analysis (MA), to gain deeper insights into the factors contributing to

*ACE2*'s differential expression between chRCC compared to normal, ccRCC, and pRCC samples.

In our GSEA analysis, our primary focus was on assessing the enrichment score of the renin-angiotensin system (RAS). A few studies demonstrated that positive enrichment of RAS corresponds to higher *ACE2* expression. They suggest that *ACE2* might be upregulated to counteract the effects of the classical RAS, associated with poor survival in RCC patients, as noted in studies by Miyajima et al. (2015) and Nuzzo et al. (2022). The dysregulation of RAS in the RCC microenvironment promotes inflammation and angiogenesis, which in turn facilitates tumour cell growth (Mourão *et al.*, 2023). Notably, previous studies have already demonstrated RAS enrichment in other cancer types such as lung adenocarcinoma (LUAD), breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), prostate adenocarcinoma (PRAD), stomach adenocarcinoma (STAD) and liver hepatocellular carcinoma (LIHC) (Hoang *et al.*, 2020).

Correspondingly, *ACE2* was also upregulated in COAD and STAD but not BRCA, PRAD, and LIHC (Supplementary Figure 10.1). These results suggest that *ACE2* may be elevated in COAD and STAD due to positively enriched RAS. However, our GSEA did not indicate significant RAS enrichment in three RCC subtypes and implies that RAS may not be dysregulated in RCC (Table 7.2). Instead, GSEA results align with the existing evidence of metabolic pathway dysregulation (Figure 7.8), in alignment with existing datasets for patients with chRCC and ccRCC (Larrinaga *et al.*, 2010).

Epigenetics plays an important role in gene expression regulation, which is a culmination of various mechanisms including histone modification, DNA methylation, and transcription factor control (Cavalli *et al.*, 2019). To further understand the mechanisms responsible for changes in *ACE2* expression in different RCC subtypes, we conducted two epigenetic analyses: MA and TFEA.

The MA results showed an intriguing observation of significant hypomethylation of *ACE2* in ccRCC, while no significant change was noted in pRCC. Unfortunately, we were unable to examine the global methylation status of the *ACE2* promoter in patients

with chRCC due to the unavailability of methylation data for this subtype. It is worth noting that another study exploring ACE2 methylation in cancers reported contrasting findings: ACE2 was hypomethylated in pRCC and displayed no significant change in ccRCC (Chai *et al.*, 2020). The differences in ACE2 methylation status between our results and those of Chai *et al.* (2020) could potentially be attributed to the number of methylation probes used in our analysis (ten probes) versus their study (four probes). In addition, the same study also observed hypomethylated ACE2 in four other human cancers - COAD, pRCC, pancreatic adenocarcinoma (PAAD) and rectum adenocarcinoma (READ), which also had elevated ACE2 expression in our exploratory analyses (Supplementary Figure 10.1). Similarly, Chai *et al.* (2020) reported ACE2 hypermethylation in testicular germ cell tumours (TGCT), which also corresponds to ACE2 downregulation in TGCT in our exploratory analysis.

We also conducted TFEA because multiple studies have shown that ACE2 expression is regulated by transcription factors (TF) (Pedersen *et al.*, 2013, 2017; Beacon *et al.*, 2021) which we included in our analysis. However, in our analysis, we did not identify enriched transcription factors that exhibited a consistent pattern across subtypes corresponding to ACE2 expression. This absence of a clear TF signature may be attributed to a lack of TFs that regulate ACE2 at high confidence.

To assess the impact of ACE2 on patient survival prognoses, we examined ACE2 correlation patterns with immune infiltrates across the three different RCC subtypes. ACE2 is known to be associated with critical processes such as angiogenesis, which have a role in modulating immune infiltration (Kim *et al.*, 2022). Notably, previous bioinformatics analyses have shown that ACE2 exhibits a strong positive correlation with immune-promoting signatures across various cancer types (Zhang *et al.*, 2020).

Conversely, an alternative bioinformatics study presented a contrasting perspective, suggesting that upregulated ACE2 activity was linked to reduced overall activity of the anti-cancer immune response in ccRCC (Li *et al.*, 2022). It is worth noting that RCC is recognised as one of the most immune-infiltrated tumours in a pan-cancer context (Vuong *et al.*, 2019). RCC is considered immunogenic but is also known to induce

immune dysfunction, largely by recruiting immune-inhibitory cells like regulatory T cells (Tregs) and myeloid-derived suppressor cells (MDSCs) into the tumour microenvironment (Díaz-Montero *et al.*, 2020).

With this in mind, we conducted an immune infiltration correlation analysis (IICA) to uncover potential associations between immune infiltrates and ACE2 in these RCC subtypes. Intriguingly, our analysis revealed a negative correlation between ACE2 and Tregs, follicular helper T (Tf-helper) cells, and memory B cells (B-memory) (Figure 7.11 B), cell types that have been associated with adverse outcomes in ccRCC patients (Pan *et al.*, 2020). These findings offer a partial explanation for the improved survival outcomes observed in ccRCC patients with high ACE2 expression (Figure 7.4). Moreover, in the other two subtypes, chRCC displayed positive correlations with Tf-helper cells, while pRCC exhibited negative correlations with Tregs and dendritic cells. Thus, our findings suggest that ACE2 may not play a predominant role in the modulation of immune infiltration within these subtypes. Collectively these results elucidate a potential mechanism through which ACE2 may positively influence the prognosis and survival of ccRCC patients but not chRCC and pRCC. Nonetheless, this IICA reveals only ACE2 correlation patterns with the nine immune infiltrates that we have selected.

A striking observation that emerged is the significant patient-to-patient variability in ACE2 expression, irrespective of cancer-associated contexts such as subtype, stage, or progression to metastasis. Consequently, we performed single sample gene set enrichment analysis (ssGSEA) to measure the enrichment score of RAS-associated genes for each patient.

Firstly, we evaluated patient samples in the context of cancer hallmarks defined by MSigDB (Figure 7.12 A). While the heatmap representing all hallmark pathways did exhibit some sample clustering based on subtypes, this clustering was relatively limited. Similarly, when we performed ssGSEA with individual hallmark pathways, we observed only partial clustering based on subtype (Figure 7.12 B, D). However, it

is noteworthy that we did not observe any significant patient sample clustering by subtype when plotting the enrichment scores of RAS-associated genes (Figure 7.12 F).

This interpatient variability in *ACE2* expression could potentially be attributed to distinct immune infiltrate signatures across patients. Şenbabaoğlu *et al.* (2016) have shown that ccRCC patients form three distinct groups based on immune infiltration: T-cell enriched, heterogeneously infiltrated and non-infiltrated groups. A similar investigation clustered ccRCC patients into one immune-infiltrating group and three non-immune-infiltrating groups (Vuong *et al.*, 2019). Consequently, such results suggest the potential for patient stratification into groups characterised by high or low *ACE2* expression, with treatment approaches tailored accordingly. This precision medicine approach, which considers individual patient characteristics, has already been applied successfully in the treatment of various cancers. An illustrative example is the treatment of breast cancer patients; HER2-positive patients are treated at a high success rate with trastuzumab, a monoclonal antibody specific to the Human Epidermal Growth Factor Receptor (HER2) (Maximiano *et al.*, 2016). Intriguingly, *ACE2* expression in HER2-positive cancers has been associated with poorer patient outcomes (Nair *et al.*, 2021). However, further research is needed to validate and expand upon these findings.

The bioinformatic evaluation of *ACE2* expression in RCC outlined herein has several limitations that should be considered. Firstly, our examination of *ACE2* and RAS-associated gene expression spans several different cancer-associated contexts: subtype, stage, metastasis progression, and treatment. However, it is important to acknowledge that the datasets for the latter three contexts were derived exclusively from ccRCC patients. Consequently, our assessment of *ACE2* expression - and the potential role of *ACE2* in those cancer-associated contexts - provides insights primarily into ccRCC patients. To provide a more comprehensive understanding, it would be beneficial to include *ACE2* expression analysis using samples from pRCC, chRCC, and many other, rarer RCC subtypes such as Xp11 translocation RCC (Athanaszio *et al.*, 2021) to capture a broader view of *ACE2* and RAS-associated gene

expression in these subtypes across various contexts. To our knowledge, these datasets are not currently available.

Additionally, we did not stratify ccRCC patients into high and low---ACE2-expressing patients in our exploratory analyses. Our ssGSEA analysis revealed substantial inter-patient variability in ACE2 expression among RCC subtypes. This high variability in ACE2 expression, particularly in ccRCC samples, introduces challenges in identifying biologically meaningful changes in gene expression, as genes with high log2FC changes may not attain statistical significance, given the inherent volatility of ACE2 expression within the group (Xiao *et al.*, 2014). This challenge was exacerbated by our relatively small sample size (Table 6.1), which limited the statistical power needed to detect differentially expressed genes with confidence (Ching *et al.*, 2014).

Another major limitation comes from the availability of matched normal samples. While we possessed ccRCC-matched normal samples, we lacked equivalent matched normal samples for the chRCC and pRCC subtypes. Consequently, we utilised ccRCC-matched normal samples to assess gene expression against the latter two subtypes against normal kidney samples. This approach, while informative, introduces potential bias and may not accurately represent the gene expression profiles of chRCC and pRCC in their respective contexts.

In summary, our project offers an initial examination of ACE2 and RAS-associated gene expression across three distinct RCC subtypes, while exploring the hypothesis of ACE2 exhibiting context-dependent roles in cancer. However, our findings primarily highlight differential ACE2 expression among subtypes, with limited variations observed in other contextual factors. Consequently, our results do not strongly support the idea of ACE2 acting as a context-dependent molecule in RCC.

Nonetheless, a noteworthy observation emerges concerning the substantial patient variability within RCC subtypes. Particularly, this observation holds significance for ccRCC patients, where elevated ACE2 expression correlates with improved patient prognosis. This underscores the potential utility of ACE2 as a promising biomarker in

ccRCC patients, thereby justifying further investigation through functional experiments.

## 9 References

- Alvarez, M.J. *et al.* (2016) ‘Functional characterization of somatic mutations in cancer using network-based inference of protein activity’, *Nature Genetics*, 48(8), pp. 838–847. Available at: <https://doi.org/10.1038/ng.3593>.
- Alvarez, M.J. (2023) *bcellViper*, *Bioconductor*. Available at: <http://bioconductor.org/packages/bcellViper/> (Accessed: 2 September 2023).
- Anders, S. *et al.* (2015) ‘HTSeq—a Python framework to work with high-throughput sequencing data’, *Bioinformatics*, 31(2), pp. 166–169. Available at: <https://doi.org/10.1093/bioinformatics/btu638>.
- Andrew, S. (2010) *Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data*. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed: 2 September 2023).
- Anguiano, L. *et al.* (2015) ‘Circulating angiotensin-converting enzyme 2 activity in patients with chronic kidney disease without previous history of cardiovascular disease’, *Nephrology Dialysis Transplantation*, 30(7), pp. 1176–1185. Available at: <https://doi.org/10.1093/ndt/gfv025>.
- Anguiano, L. *et al.* (2017) ‘Circulating ACE2 in Cardiovascular and Kidney Diseases’, *Current Medicinal Chemistry*, 24(30), pp. 3231–3241. Available at: <https://doi.org/10.2174/0929867324666170414162841>.
- Athanazio, D.A. *et al.* (2021) ‘Classification of renal cell tumors – current concepts and use of ancillary tests: recommendations of the Brazilian Society of Pathology’, *Surgical and Experimental Pathology*, 4(1), p. 4. Available at: <https://doi.org/10.1186/s42047-020-00084-x>.
- Badia-I-Mompel, P. *et al.* (2022) ‘decoupleR: ensemble of computational methods to infer biological activities from omics data’, *Bioinformatics Advances*, 2(1), p. vbac016. Available at: <https://doi.org/10.1093/bioadv/vbac016>.
- Basso, N. *et al.* (2001) ‘History about the discovery of the renin-angiotensin system’, *Hypertension (Dallas, Tex.: 1979)*, 38(6), pp. 1246–1249. Available at: <https://doi.org/10.1161/hy1201.101214>.
- Beacon, T.H. *et al.* (2021) ‘Epigenetic regulation of ACE2, the receptor of the SARS-CoV-2 virus1’, *Genome*, 64(4), pp. 386–399. Available at: <https://doi.org/10.1139/gen-2020-0124>.
- Bekassy, Z. *et al.* (2022) ‘Crosstalk between the renin-angiotensin, complement and kallikrein-kinin systems in inflammation’, *Nature Reviews. Immunology*, 22(7), pp. 411–428. Available at: <https://doi.org/10.1038/s41577-021-00634-8>.
- Beyerstedt, S. *et al.* (2021) ‘COVID-19: angiotensin-converting enzyme 2 (ACE2) expression and tissue susceptibility to SARS-CoV-2 infection’, *European Journal of Clinical Microbiology & Infectious Diseases*, 40(5), pp. 905–919. Available at: <https://doi.org/10.1007/s10096-020-04138-6>.
- Bhari, V.K. *et al.* (2020) ‘SARS-CoV-2 cell receptor gene ACE2 -mediated immunomodulation in breast cancer subtypes’, *Biochemistry and Biophysics Reports*, 24, p. 100844. Available at: <https://doi.org/10.1016/j.bbrep.2020.100844>.

Cavalli, G. et al. (2019) 'Advances in epigenetics link genetics to the environment and disease', *Nature*, 571(7766), pp. 489–499. Available at: <https://doi.org/10.1038/s41586-019-1411-0>.

Cerami1, E. et al. (2012) 'The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data', *Cancer discovery*, 2(5), pp. 401–404. Available at: <https://doi.org/10.1158/2159-8290.CD-12-0095>.

Chai, P. et al. (2020) 'Genetic alteration, RNA expression, and DNA methylation profiling of coronavirus disease 2019 (COVID-19) receptor ACE2 in malignancies: a pan-cancer analysis', *Journal of Hematology & Oncology*, 13, p. 43. Available at: <https://doi.org/10.1186/s13045-020-00883-5>.

Chandrashekhar, D.S. et al. (2017) 'UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses', *Neoplasia (New York, N.Y.)*, 19(8), pp. 649–658. Available at: <https://doi.org/10.1016/j.neo.2017.05.002>.

Chen, B. et al. (2018) 'Profiling tumor infiltrating immune cells with CIBERSORT', *Methods in molecular biology (Clifton, N.J.)*, 1711, pp. 243–259. Available at: [https://doi.org/10.1007/978-1-4939-7493-1\\_12](https://doi.org/10.1007/978-1-4939-7493-1_12).

Chen, L. et al. (2021) 'Gene expression-based immune infiltration analyses of renal cancer and their associations with survival outcome', *BMC Cancer*, 21(1), p. 595. Available at: <https://doi.org/10.1186/s12885-021-08244-2>.

Ching, T. et al. (2014) 'Power analysis and sample size estimation for RNA-Seq differential expression', *RNA*, 20(11), pp. 1684–1696. Available at: <https://doi.org/10.1261/rna.046011.114>.

Chow, J. et al. (2020) 'Radiation induces dynamic changes to the T cell repertoire in renal cell carcinoma patients', *Proceedings of the National Academy of Sciences of the United States of America*, 117(38), pp. 23721–23729. Available at: <https://doi.org/10.1073/pnas.2001933117>.

Díaz-Montero, C.M. et al. (2020) 'The immunology of renal cell carcinoma', *Nature Reviews Nephrology*, 16(12), pp. 721–735. Available at: <https://doi.org/10.1038/s41581-020-0316-3>.

Dobin, A. et al. (2013) 'STAR: ultrafast universal RNA-seq aligner', *Bioinformatics (Oxford, England)*, 29(1), pp. 15–21. Available at: <https://doi.org/10.1093/bioinformatics/bts635>.

Donoghue, M. et al. (2000) 'A Novel Angiotensin-Converting Enzyme–Related Carboxypeptidase (ACE2) Converts Angiotensin I to Angiotensin 1-9', *Circulation Research*, 87(5), pp. e1–e9. Available at: <https://doi.org/10.1161/01.RES.87.5.e1>.

Ewels, P. et al. (2016) 'MultiQC: summarize analysis results for multiple tools and samples in a single report', *Bioinformatics (Oxford, England)*, 32(19), pp. 3047–3048. Available at: <https://doi.org/10.1093/bioinformatics/btw354>.

Fagyas, M. et al. (2022) 'Circulating ACE2 activity predicts mortality and disease severity in hospitalized COVID-19 patients', *International Journal of Infectious Diseases*, 115, pp. 8–16. Available at: <https://doi.org/10.1016/j.ijid.2021.11.028>.

Fried, L.F. et al. (2013) 'Combined angiotensin inhibition for the treatment of diabetic nephropathy', *The New England Journal of Medicine*, 369(20), pp. 1892–1903. Available at: <https://doi.org/10.1056/NEJMoa1303154>.

Garcia-Alonso, L. et al. (2019) 'Benchmark and integration of resources for the estimation of human transcription factor activities', *Genome Research*, 29(8), pp. 1363–1375. Available at: <https://doi.org/10.1101/gr.240663.118>.

Gottschalk, G. et al. (2021) 'ACE2: At the crossroad of COVID-19 and lung cancer', *Gene Reports*, 23, p. 101077. Available at: <https://doi.org/10.1016/j.genrep.2021.101077>.

Gu, Z. et al. (2013) 'CePa: an R package for finding significant pathways weighted by multiple network centralities', *Bioinformatics*, 29(5), pp. 658–660. Available at: <https://doi.org/10.1093/bioinformatics/btt008>.

Gu, Z. et al. (2016) 'Complex heatmaps reveal patterns and correlations in multidimensional genomic data', *Bioinformatics*, 32(18), pp. 2847–2849. Available at: <https://doi.org/10.1093/bioinformatics/btw313>.

Gu, Z. (2022) 'Complex heatmap visualization', *iMeta*, 1(3), p. e43. Available at: <https://doi.org/10.1002/imt2.43>.

Hadley Wickham and Romain François and Lionel Henry and Kirill Müller and Davis Vaughan (2023) 'dplyr: A Grammar of Data Manipulation'. Available at: <https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>.

Hänzelmann, S. et al. (2013) 'GSVA: gene set variation analysis for microarray and RNA-Seq data', *BMC Bioinformatics*, 14(1), p. 7. Available at: <https://doi.org/10.1186/1471-2105-14-7>.

Ho, T.H. et al. (2015) 'A Multidisciplinary Biospecimen Bank of Renal Cell Carcinomas Compatible with Discovery Platforms at Mayo Clinic, Scottsdale, Arizona', *PloS One*, 10(7), p. e0132831. Available at: <https://doi.org/10.1371/journal.pone.0132831>.

Hoang, T. et al. (2020) 'Genetic Susceptibility of ACE2 and TMPRSS2 in Six Common Cancers and Possible Impacts on COVID-19', *Cancer Research and Treatment : Official Journal of Korean Cancer Association*, 53(3), pp. 650–656. Available at: <https://doi.org/10.4143/crt.2020.950>.

Holappa, M. et al. (2020) 'Local ocular renin-angiotensin-aldosterone system: any connection with intraocular pressure? A comprehensive review', *Annals of Medicine*, 52(5), pp. 191–206. Available at: <https://doi.org/10.1080/07853890.2020.1758341>.

Huang, X. et al. (2023) 'IDDF2023-ABS-0255 Ace2 deficiency promotes colorectal cancer by remodeling gut microbiota', *Gut*, 72(Suppl 1), pp. A119–A119. Available at: <https://doi.org/10.1136/gutjnl-2023-IDDF.104>.

Huber, W. et al. (2015) 'Orchestrating high-throughput genomic analysis with Bioconductor', *Nature Methods*, 12(2), pp. 115–121. Available at: <https://doi.org/10.1038/nmeth.3252>.

Jackson, L. et al. (2018) 'Within the Brain: The Renin Angiotensin System', *International Journal of Molecular Sciences*, 19(3), p. 876. Available at: <https://doi.org/10.3390/ijms19030876>.

Jumper, J. et al. (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature*, 596(7873), pp. 583–589. Available at: <https://doi.org/10.1038/s41586-021-03819-2>.

Kanehisa, M. et al. (2000) 'KEGG: Kyoto Encyclopedia of Genes and Genomes', *Nucleic Acids Research*, 28(1), pp. 27–30.

Kaschina, E. et al. (2018) 'Hypertension and the Renin–Angiotensin–Aldosterone System☆', in I. Huhtaniemi et al. (eds) *Encyclopedia of Endocrine Diseases (Second Edition)*. Oxford: Academic Press, pp. 505–510. Available at: <https://doi.org/10.1016/B978-0-12-801238-3.03969-6>.

Kasinath, B.S. et al. (2006) 'mRNA translation: unexplored territory in renal science', *Journal of the American Society of Nephrology: JASN*, 17(12), pp. 3281–3292. Available at: <https://doi.org/10.1681/ASN.2006050488>.

Keidar, S. et al. (2007) 'ACE2 of the heart: From angiotensin I to angiotensin (1–7)', *Cardiovascular Research*, 73(3), pp. 463–469. Available at: <https://doi.org/10.1016/j.cardiores.2006.09.006>.

Keil, L.C. et al. (1975) 'Release of Vasopressin by Angiotensin II1', *Endocrinology*, 96(4), pp. 1063–1065. Available at: <https://doi.org/10.1210/endo-96-4-1063>.

Khanna, P. et al. (2021) 'ACE2 abrogates tumor resistance to VEGFR inhibitors suggesting angiotensin-(1-7) as a therapy for clear cell renal cell carcinoma', *Science Translational Medicine*, 13(577), p. eabc0170. Available at: <https://doi.org/10.1126/scitranslmed.abc0170>.

Kidoguchi, S. et al. (2023) 'Blood pressure control with renin–angiotensin system inhibitors in hypertension patients with cancer - good or bad?', *Hypertension Research*, 46(2), pp. 529–531. Available at: <https://doi.org/10.1038/s41440-022-01110-z>.

Kim, H.J. et al. (2022) 'Crosstalk between angiogenesis and immune regulation in the tumor microenvironment', *Archives of Pharmacal Research*, 45(6), pp. 401–416. Available at: <https://doi.org/10.1007/s12272-022-01389-z>.

Kolde, R. (2012) 'Pheatmap: pretty heatmaps', *R package version*, 1(2), p. 726.

Krishnan, B. et al. (2013) 'Angiotensin-(1-7) Attenuates Metastatic Prostate Cancer and Reduces Osteoclastogenesis', *The Prostate*, 73(1), p. 10.1002/pros.22542. Available at: <https://doi.org/10.1002/pros.22542>.

Laghlam, D. et al. (2021) 'Renin–Angiotensin–Aldosterone System and Immunomodulation: A State-of-the-Art Review', *Cells*, 10(7), p. 1767. Available at: <https://doi.org/10.3390/cells10071767>.

Lánczky, A. et al. (2021) 'Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation', *Journal of Medical Internet Research*, 23(7), p. e27633. Available at: <https://doi.org/10.2196/27633>.

Larrinaga, G. et al. (2010) 'Angiotensin-converting enzymes (ACE and ACE2) are downregulated in renal tumors', *Regulatory Peptides*, 165(2–3), pp. 218–223. Available at: <https://doi.org/10.1016/j.regpep.2010.07.170>.

Lavoie, J.L. et al. (2003) 'Minireview: Overview of the Renin-Angiotensin System—An Endocrine and Paracrine System', *Endocrinology*, 144(6), pp. 2179–2183. Available at: <https://doi.org/10.1210/en.2003-0150>.

Li, B. et al. (2016) 'Comprehensive analyses of tumor immunity: implications for cancer immunotherapy', *Genome Biology*, 17(1), p. 174. Available at: <https://doi.org/10.1186/s13059-016-1028-7>.

- Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25(16), pp. 2078–2079. Available at: <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, S. *et al.* (2021) 'Effects of ACE2 and Co-factors on the Mechanism of SARS-CoV-2 Infection Among Colorectal Cancer Patient'. Available at: <https://doi.org/10.21203/rs.3.rs-868323/v1>.
- Li, S. *et al.* (2022) 'The impact of ACE2 and co-factors on SARS-CoV-2 infection in colorectal cancer', *Clinical and Translational Medicine*, 12(7), p. e967. Available at: <https://doi.org/10.1002/ctm2.967>.
- Li, T. *et al.* (2017) 'TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells', *Cancer Research*, 77(21), pp. e108–e110. Available at: <https://doi.org/10.1158/0008-5472.CAN-17-0307>.
- Li, T. *et al.* (2020) 'TIMER2.0 for analysis of tumor-infiltrating immune cells', *Nucleic Acids Research*, 48(W1), pp. W509–W514. Available at: <https://doi.org/10.1093/nar/gkaa407>.
- Li, Y. *et al.* (2019) 'The SMART App: an interactive web application for comprehensive DNA methylation analysis and visualization', *Epigenetics & Chromatin*, 12(1), p. 71. Available at: <https://doi.org/10.1186/s13072-019-0316-3>.
- Liang, L. *et al.* (2022) 'Transcriptional regulation and small compound targeting of ACE2 in lung epithelial cells', *Acta Pharmacologica Sinica*, 43(11), pp. 2895–2904. Available at: <https://doi.org/10.1038/s41401-022-00906-6>.
- Liberzon, A. *et al.* (2011) 'Molecular signatures database (MSigDB) 3.0', *Bioinformatics*, 27(12), pp. 1739–1740. Available at: <https://doi.org/10.1093/bioinformatics/btr260>.
- Lin, H. *et al.* (2022) 'Kidney Angiotensin in Cardiovascular Disease: Formation and Drug Targeting', *Pharmacological Reviews*, 74(3), pp. 462–505. Available at: <https://doi.org/10.1124/pharmrev.120.000236>.
- Liu, L.-P. *et al.* (2021) 'New perspectives on angiotensin-converting enzyme 2 and its related diseases', *World Journal of Diabetes*, 12(6), pp. 839–854. Available at: <https://doi.org/10.4239/wjd.v12.i6.839>.
- Lonsdale, J. *et al.* (2013) 'The Genotype-Tissue Expression (GTEx) project', *Nature Genetics*, 45(6), pp. 580–585. Available at: <https://doi.org/10.1038/ng.2653>.
- Love, M.I. *et al.* (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550. Available at: <https://doi.org/10.1186/s13059-014-0550-8>.
- Lubbe, L. *et al.* (2020) 'ACE2 and ACE: structure-based insights into mechanism, regulation and receptor recognition by SARS-CoV', *Clinical Science (London, England : 1979)*, 134(21), pp. 2851–2871. Available at: <https://doi.org/10.1042/CS20200899>.
- Majid, D.S.A. *et al.* (2015) 'Salt-Sensitive Hypertension: Perspectives on Intrarenal Mechanisms', *Current hypertension reviews*, 11(1), pp. 38–48.
- Martin, F.J. *et al.* (2023) 'Ensembl 2023', *Nucleic Acids Research*, 51(D1), pp. D933–D941. Available at: <https://doi.org/10.1093/nar/gkac958>.
- Maximiano, S. *et al.* (2016) 'Trastuzumab in the Treatment of Breast Cancer', *BioDrugs*, 30(2), pp. 75–86. Available at: <https://doi.org/10.1007/s40259-016-0162-9>.

- Mehranfard, D. et al. (2021) 'Alterations in Gene Expression of Renin-Angiotensin System Components and Related Proteins in Colorectal Cancer', *Journal of the Renin-Angiotensin-Aldosterone System*, 2021, p. e9987115. Available at: <https://doi.org/10.1155/2021/9987115>.
- Mei, J. et al. (2022) 'Angiotensin-converting enzyme 2 identifies immuno-hot tumors suggesting angiotensin-(1–7) as a sensitizer for chemotherapy and immunotherapy in breast cancer', *Biological Procedures Online*, 24(1), p. 15. Available at: <https://doi.org/10.1186/s12575-022-00177-9>.
- Mészáros, B. et al. (2021) 'Short linear motif candidates in the cell entry system used by SARS-CoV-2 and their potential therapeutic implications', *Science Signaling*, 14(665), p. eabd0334. Available at: <https://doi.org/10.1126/scisignal.abd0334>.
- Mirabito Colafella, K.M. et al. (2019) 'The renin-angiotensin-aldosterone system and its therapeutic targets', *Experimental Eye Research*, 186, p. 107680. Available at: <https://doi.org/10.1016/j.exer.2019.05.020>.
- Miyajima, A. et al. (2015) 'Prognostic Impact of Renin–Angiotensin System Blockade on Renal Cell Carcinoma After Surgery', *Annals of Surgical Oncology*, 22(11), pp. 3751–3759. Available at: <https://doi.org/10.1245/s10434-015-4436-0>.
- Mizuiri, S. et al. (2015) 'ACE and ACE2 in kidney disease', *World Journal of Nephrology*, 4(1), pp. 74–82. Available at: <https://doi.org/10.5527/wjn.v4.i1.74>.
- Mohater, S. et al. (2023) 'Vitamin D improves hepatic alterations in ACE1 and ACE2 expression in experimentally induced metabolic syndrome', *Saudi Pharmaceutical Journal*, 31(9), p. 101709. Available at: <https://doi.org/10.1016/j.jsps.2023.101709>.
- Mootha, V.K. et al. (2003) 'PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes', *Nature Genetics*, 34(3), pp. 267–273. Available at: <https://doi.org/10.1038/ng1180>.
- Morgan, M. et al. (2023) 'GSEABase: Gene set enrichment data structures and methods'. Bioconductor version: Release (3.17). Available at: <https://doi.org/10.18129/B9.bioc.GSEABase>.
- Mourão, T.C. et al. (2023) 'Role of the Renin–Angiotensin System Components in Renal Cell Carcinoma: A Literature Review', *Current Urology Reports*, 24(7), pp. 345–353. Available at: <https://doi.org/10.1007/s11934-023-01160-x>.
- Muglia, V.F. et al. (2015) 'Renal cell carcinoma: histological classification and correlation with imaging findings', *Radiologia Brasileira*, 48(3), pp. 166–174. Available at: <https://doi.org/10.1590/0100-3984.2013.1927>.
- Müller-Dott, S. et al. (2023) 'Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities'. bioRxiv, p. 2023.03.30.534849. Available at: <https://doi.org/10.1101/2023.03.30.534849>.
- Nair, M.G. et al. (2021) 'High expression of ACE2 in HER2 subtype of breast cancer is a marker of poor prognosis', *Cancer Treatment and Research Communications*, 27, p. 100321. Available at: <https://doi.org/10.1016/j.ctarc.2021.100321>.

- Narayan, S.S. et al. (2020) 'Angiotensin converting enzymes ACE and ACE2 in thyroid cancer progression', *Neoplasma*, 67(2), pp. 402–409. Available at: [https://doi.org/10.4149/neo\\_2019\\_190506N405](https://doi.org/10.4149/neo_2019_190506N405).
- Nassar, A.H. et al. (2023) 'Epigenomic charting and functional annotation of risk loci in renal cell carcinoma', *Nature Communications*, 14(1), p. 346. Available at: <https://doi.org/10.1038/s41467-023-35833-5>.
- Ni, W. et al. (2020) 'Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19', *Critical Care*, 24(1), p. 422. Available at: <https://doi.org/10.1186/s13054-020-03120-0>.
- Niu, X. et al. (2021) 'ACE2 Is a Prognostic Biomarker and Associated with Immune Infiltration in Kidney Renal Clear Cell Carcinoma: Implication for COVID-19', *Journal of Oncology*, 2021, p. 8847307. Available at: <https://doi.org/10.1155/2021/8847307>.
- Nuzzo, P.V. et al. (2022) 'Impact of renin-angiotensin system inhibitors on outcomes in patients with metastatic renal cell carcinoma treated with immune-checkpoint inhibitors', *Clinical Genitourinary Cancer*, 20(4), pp. 301–306. Available at: <https://doi.org/10.1016/j.clgc.2022.04.012>.
- Pan, Q. et al. (2020) 'The immune infiltration in clear cell Renal Cell Carcinoma and their clinical implications: A study based on TCGA and GEO databases', *Journal of Cancer*, 11(11), pp. 3207–3215. Available at: <https://doi.org/10.7150/jca.37285>.
- Park, Y.-A. et al. (2014) 'Dual targeting of angiotensin receptors (AGTR1 and AGTR2) in epithelial ovarian carcinoma', *Gynecologic Oncology*, 135(1), pp. 108–117. Available at: <https://doi.org/10.1016/j.ygyno.2014.06.031>.
- Pathania, A.S. et al. (2021) 'COVID-19 and Cancer Comorbidity: Therapeutic Opportunities and Challenges', *Theranostics*, 11(2), pp. 731–753. Available at: <https://doi.org/10.7150/thno.51471>.
- Pavo, N. et al. (2021) 'Myocardial Angiotensin Metabolism in End-Stage Heart Failure', *Journal of the American College of Cardiology*, 77(14), pp. 1731–1743. Available at: <https://doi.org/10.1016/j.jacc.2021.01.052>.
- Paz Ocáranda, M. et al. (2020) 'Counter-regulatory renin–angiotensin system in cardiovascular disease', *Nature Reviews. Cardiology*, 17(2), pp. 116–129. Available at: <https://doi.org/10.1038/s41569-019-0244-8>.
- Pedersen, K.B. et al. (2013) 'The transcription factor HNF1 $\alpha$  induces expression of angiotensin-converting enzyme 2 (ACE2) in pancreatic islets from evolutionarily conserved promoter motifs', *Biochimica Et Biophysica Acta*, 1829(11), pp. 1225–1235. Available at: <https://doi.org/10.1016/j.bbagen.2013.09.007>.
- Pedersen, K.B. et al. (2017) 'Forkhead Box Transcription Factors of the FOXA Class Are Required for Basal Transcription of Angiotensin-Converting Enzyme 2', *Journal of the Endocrine Society*, 1(4), pp. 370–384. Available at: <https://doi.org/10.1210/js.2016-1071>.
- Petitprez, F. et al. (2021) 'Review of Prognostic Expression Markers for Clear Cell Renal Cell Carcinoma', *Frontiers in Oncology*, 11. Available at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.643065> (Accessed: 24 August 2023).

Pollard, M. (2023) *samtools-split manual page*. Available at: <https://www.htslib.org/doc/samtools-split.html> (Accessed: 15 September 2023).

Qaradakhi, T. et al. (2020) 'The potential actions of angiotensin-converting enzyme II (ACE2) activator diminazene aceturate (DIZE) in various diseases', *Clinical and Experimental Pharmacology & Physiology*, 47(5), pp. 751–758. Available at: <https://doi.org/10.1111/1440-1681.13251>.

Qin, X. et al. (2023) 'Widespread genomic/molecular alterations of DNA helicases and their clinical/therapeutic implications across human cancer', *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, 158, p. 114193. Available at: <https://doi.org/10.1016/j.biopha.2022.114193>.

Ramalingam, L. et al. (2017) 'The renin angiotensin system, oxidative stress and mitochondrial function in obesity and insulin resistance', *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1863(5), pp. 1106–1114. Available at: <https://doi.org/10.1016/j.bbadi.2016.07.019>.

Ranjbar, T. et al. (2022) 'The Renin-Angiotensin-Aldosterone System, Nitric Oxide, and Hydrogen Sulfide at the Crossroads of Hypertension and COVID-19: Racial Disparities and Outcomes', *International Journal of Molecular Sciences*, 23(22), p. 13895. Available at: <https://doi.org/10.3390/ijms232213895>.

Rhee, J.-K. et al. (2018) 'Impact of Tumor Purity on Immune Gene Expression and Clustering Analyses across Multiple Cancer Types', *Cancer Immunology Research*, 6(1), pp. 87–97. Available at: <https://doi.org/10.1158/2326-6066.CIR-17-0201>.

Riquier-Brison, A.D.M. et al. (2018) 'The macula densa prorenin receptor is essential in renin release and blood pressure control', *American Journal of Physiology. Renal Physiology*, 315(3), pp. F521–F534. Available at: <https://doi.org/10.1152/ajprenal.00029.2018>.

Ritchie, M.E. et al. (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43(7), p. e47. Available at: <https://doi.org/10.1093/nar/gkv007>.

Şenbabaoğlu, Y. et al. (2016) 'Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures', *Genome Biology*, 17(1), p. 231. Available at: <https://doi.org/10.1186/s13059-016-1092-z>.

Shih, A.J. et al. (2020) 'Prognostic Molecular Signatures for Metastatic Potential in Clinically Low-Risk Stage I and II Clear Cell Renal Cell Carcinomas', *Frontiers in Oncology*, 10, p. 1383. Available at: <https://doi.org/10.3389/fonc.2020.01383>.

Smedley, D. et al. (2009) 'BioMart – biological queries made easy', *BMC Genomics*, 10(1), p. 22. Available at: <https://doi.org/10.1186/1471-2164-10-22>.

Su, C. et al. (2021) 'Role of the central renin-angiotensin system in hypertension (Review)', *International Journal of Molecular Medicine*, 47(6), p. 95. Available at: <https://doi.org/10.3892/ijmm.2021.4928>.

Subramanian, A. et al. (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles', *Proceedings of the National Academy of Sciences*, 102(43), pp. 15545–15550. Available at: <https://doi.org/10.1073/pnas.0506580102>.

Tang, Q. et al. (2021) 'Downregulation of ACE2 expression by SARS-CoV-2 worsens the prognosis of KIRC and KIRP patients via metabolism and immunoregulation', *International Journal of Biological Sciences*, 17(8), pp. 1925–1939. Available at: <https://doi.org/10.7150/ijbs.57802>.

Tang, Q. et al. (2023) 'Identification of molecular subtypes based on chromatin regulator and tumor microenvironment infiltration characterization in papillary renal cell carcinoma', *Journal of Cancer Research and Clinical Oncology*, 149(1), pp. 231–245. Available at: <https://doi.org/10.1007/s00432-022-04482-4>.

Tang, Z. et al. (2019) 'GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis', *Nucleic Acids Research*, 47(W1), pp. W556–W560. Available at: <https://doi.org/10.1093/nar/gkz430>.

*The Human Protein Atlas* (2023). Available at: <https://www.proteinatlas.org/> (Accessed: 3 September 2023).

Tipnis, S.R. et al. (2000) 'A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase', *The Journal of Biological Chemistry*, 275(43), pp. 33238–33243. Available at: <https://doi.org/10.1074/jbc.M002615200>.

Vargas Vargas, R.A. et al. (2022) 'Renin–angiotensin system: Basic and clinical aspects—A general perspective', *Endocrinologia, Diabetes Y Nutricion*, 69(1), pp. 52–62. Available at: <https://doi.org/10.1016/j.endien.2022.01.005>.

Vasquez, N. et al. (2020) 'Angiotensin Receptor-Neprilysin Inhibitors and the Natriuretic Peptide Axis', *Current Heart Failure Reports*, 17(3), pp. 67–76. Available at: <https://doi.org/10.1007/s11897-020-00458-y>.

Vuong, L. et al. (2019) 'Tumor Microenvironment Dynamics in Clear-Cell Renal Cell Carcinoma', *Cancer Discovery*, 9(10), pp. 1349–1357. Available at: <https://doi.org/10.1158/2159-8290.CD-19-0499>.

Wang, C. et al. (2023) 'The chromosomal instability 25 gene signature is identified in clear cell renal cell carcinoma and serves as a predictor for survival and Sunitinib response', *Frontiers in Oncology*, 13, p. 1133902. Available at: <https://doi.org/10.3389/fonc.2023.1133902>.

Wang, J. et al. (2022) 'ACE2 Shedding and the Role in COVID-19', *Frontiers in Cellular and Infection Microbiology*, 11. Available at: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.789180> (Accessed: 17 September 2023).

Wang, T. et al. (2021) 'Downregulation of ACE2 is associated with advanced pathological features and poor prognosis in clear cell renal cell carcinoma', *Future Oncology [Preprint]*. Available at: <https://doi.org/10.2217/fon-2020-1164>.

Wang, Y. et al. (2019) 'Extent and characteristics of immune infiltration in clear cell renal cell carcinoma and the prognostic value', *Translational Andrology and Urology*, 8(6), pp. 609–618. Available at: <https://doi.org/10.21037/tau.2019.10.19>.

Wang, Y. et al. (2021) 'Immune Infiltration Landscape in Clear Cell Renal Cell Carcinoma Implications', *Frontiers in Oncology*, 10, p. 491621. Available at: <https://doi.org/10.3389/fonc.2020.491621>.

- Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at: <https://ggplot2.tidyverse.org> (Accessed: 16 September 2023).
- Wu, T. et al. (2021) 'clusterProfiler 4.0: A universal enrichment tool for interpreting omics data', *The Innovation*, 2(3). Available at: <https://doi.org/10.1016/j.xinn.2021.100141>.
- Xiao, Y. et al. (2014) 'A novel significance score for gene selection and ranking', *Bioinformatics*, 30(6), pp. 801–807. Available at: <https://doi.org/10.1093/bioinformatics/btr671>.
- Yang, J., Li, H., Hu, S., et al. (2020) 'ACE2 correlated with immune infiltration serves as a prognostic biomarker in endometrial carcinoma and renal papillary cell carcinoma: implication for COVID-19', *Aging (Albany NY)*, 12(8), pp. 6518–6535. Available at: <https://doi.org/10.18632/aging.103100>.
- Yang, J., Li, H. and Zhou, Y. (2020) 'The Abnormal ACE2 Expression Correlated with Immune Infiltration May Worsen Progression After SARS-CoV-2 Infection in UCEC and KIRP'. Rochester, NY. Available at: <https://doi.org/10.2139/ssrn.3555195>.
- Yoshihara, K. et al. (2013) 'Inferring tumour purity and stromal and immune cell admixture from expression data', *Nature Communications*, 4(1), p. 2612. Available at: <https://doi.org/10.1038/ncomms3612>.
- Yu, G. et al. (2012) 'clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters', *OMICS: A Journal of Integrative Biology*, 16(5), pp. 284–287. Available at: <https://doi.org/10.1089/omi.2011.0118>.
- Zhang, Q. et al. (2019) 'ACE2 inhibits breast cancer angiogenesis via suppressing the VEGFa/VEGFR2/ERK pathway', *Journal of experimental & clinical cancer research: CR*, 38(1), p. 173. Available at: <https://doi.org/10.1186/s13046-019-1156-5>.
- Zhang, S. et al. (2019) 'Immune infiltration in renal cell carcinoma', *Cancer Science*, 110(5), pp. 1564–1572. Available at: <https://doi.org/10.1111/cas.13996>.
- Zhang, S. et al. (2023) 'Targeting ACE2-BRD4 crosstalk in colorectal cancer and the deregulation of DNA repair and apoptosis', *npj Precision Oncology*, 7(1), pp. 1–7. Available at: <https://doi.org/10.1038/s41698-023-00361-4>.
- Zhang, Z. et al. (2020) 'The SARS-CoV-2 host cell receptor ACE2 correlates positively with immunotherapy response and is a potential protective factor for cancer progression', *Computational and Structural Biotechnology Journal*, 18, pp. 2438–2444. Available at: <https://doi.org/10.1016/j.csbj.2020.08.024>.
- Zheng, S. et al. (2015) 'Ang-(1-7) promotes the migration and invasion of human renal cell carcinoma cells via Mas-mediated AKT signaling pathway', *Biochemical and Biophysical Research Communications*, 460(2), pp. 333–340. Available at: <https://doi.org/10.1016/j.bbrc.2015.03.035>.
- Zhou, L. et al. (2011) 'Angiotensin-converting enzyme 2 acts as a potential molecular target for pancreatic cancer therapy', *Cancer Letters*, 307(1), pp. 18–25. Available at: <https://doi.org/10.1016/j.canlet.2011.03.011>.
- Zhu, G. et al. (2019) 'Profiles of tumor-infiltrating immune cells in renal cell carcinoma and their clinical implications', *Oncology Letters*, 18(5), pp. 5235–5242. Available at: <https://doi.org/10.3892/ol.2019.10896>.

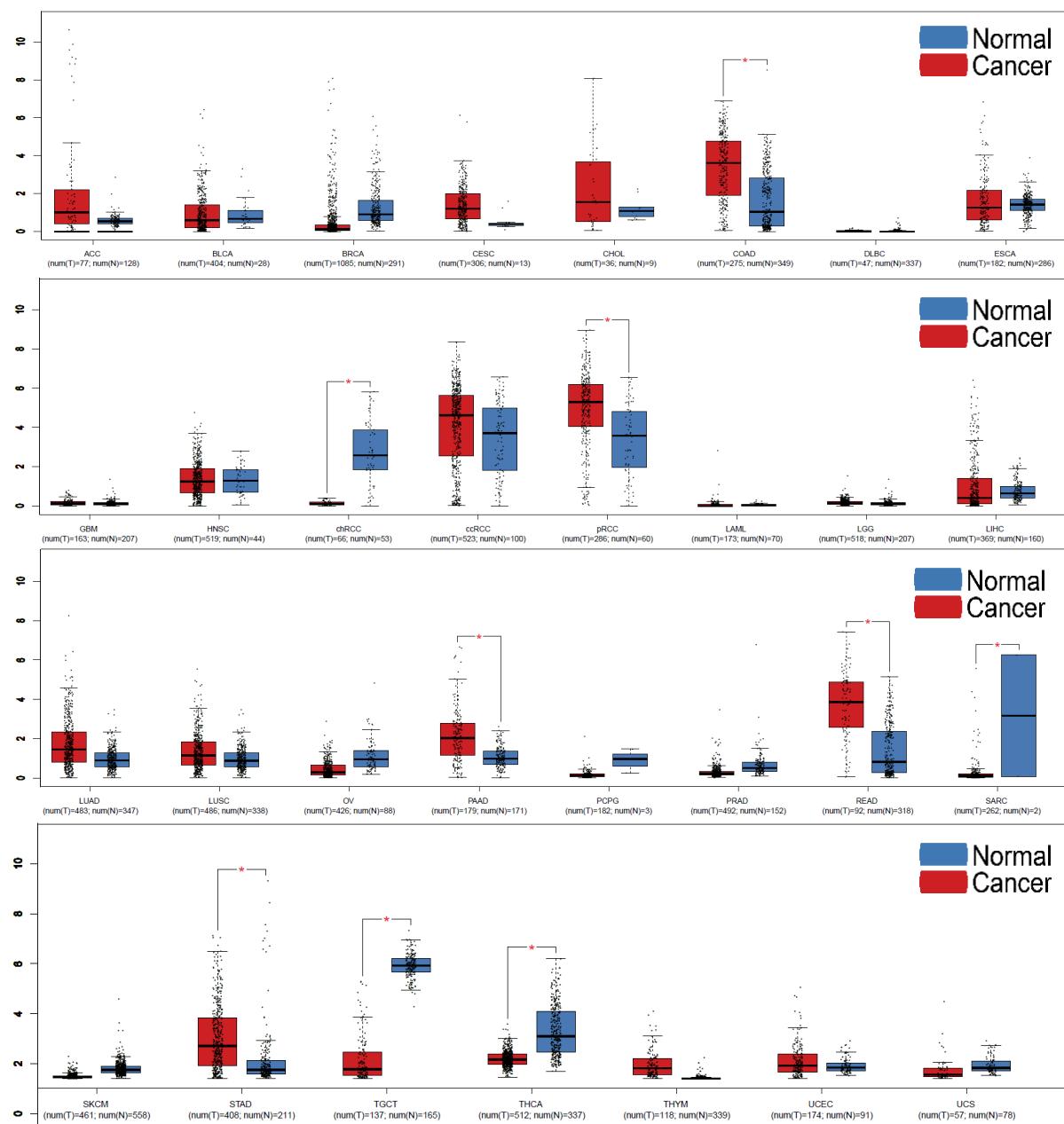
Zhu, H. et al. (2021) 'The role of SARS-CoV-2 target ACE2 in cardiovascular diseases', *Journal of Cellular and Molecular Medicine*, 25(3), pp. 1342–1349. Available at: <https://doi.org/10.1111/jcmm.16239>.

Zuo, X. et al. (2022) 'Chemotherapy induces ACE2 expression in breast cancer via the ROS-AKT-HIF-1 $\alpha$  signaling pathway: a potential prognostic marker for breast cancer patients receiving chemotherapy', *Journal of Translational Medicine*, 20(1), p. 509. Available at: <https://doi.org/10.1186/s12967-022-03716-w>.

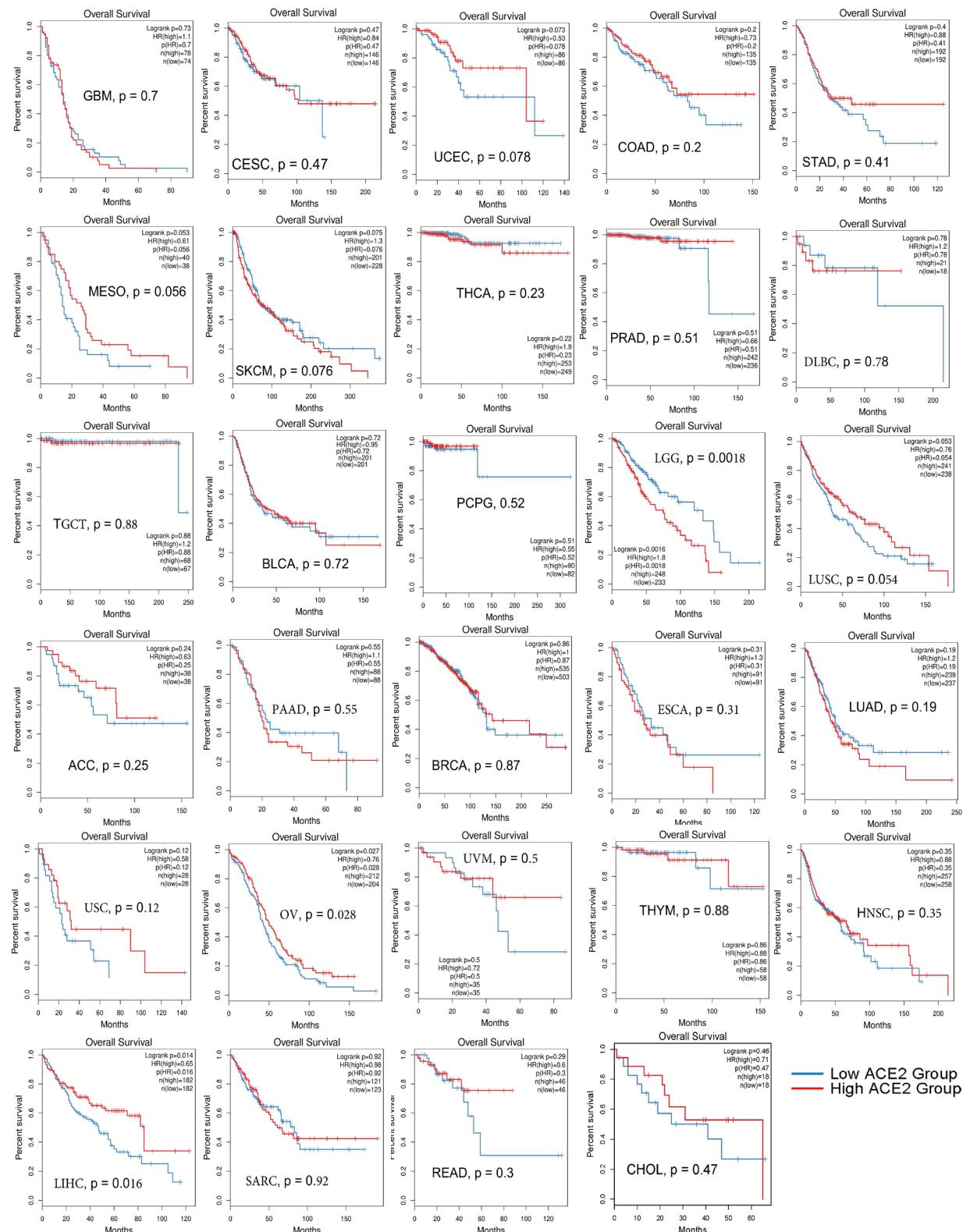
## 10 Supplementary Materials

**Supplementary Table 10.1** Abbreviations of cancers mentioned in this master's thesis

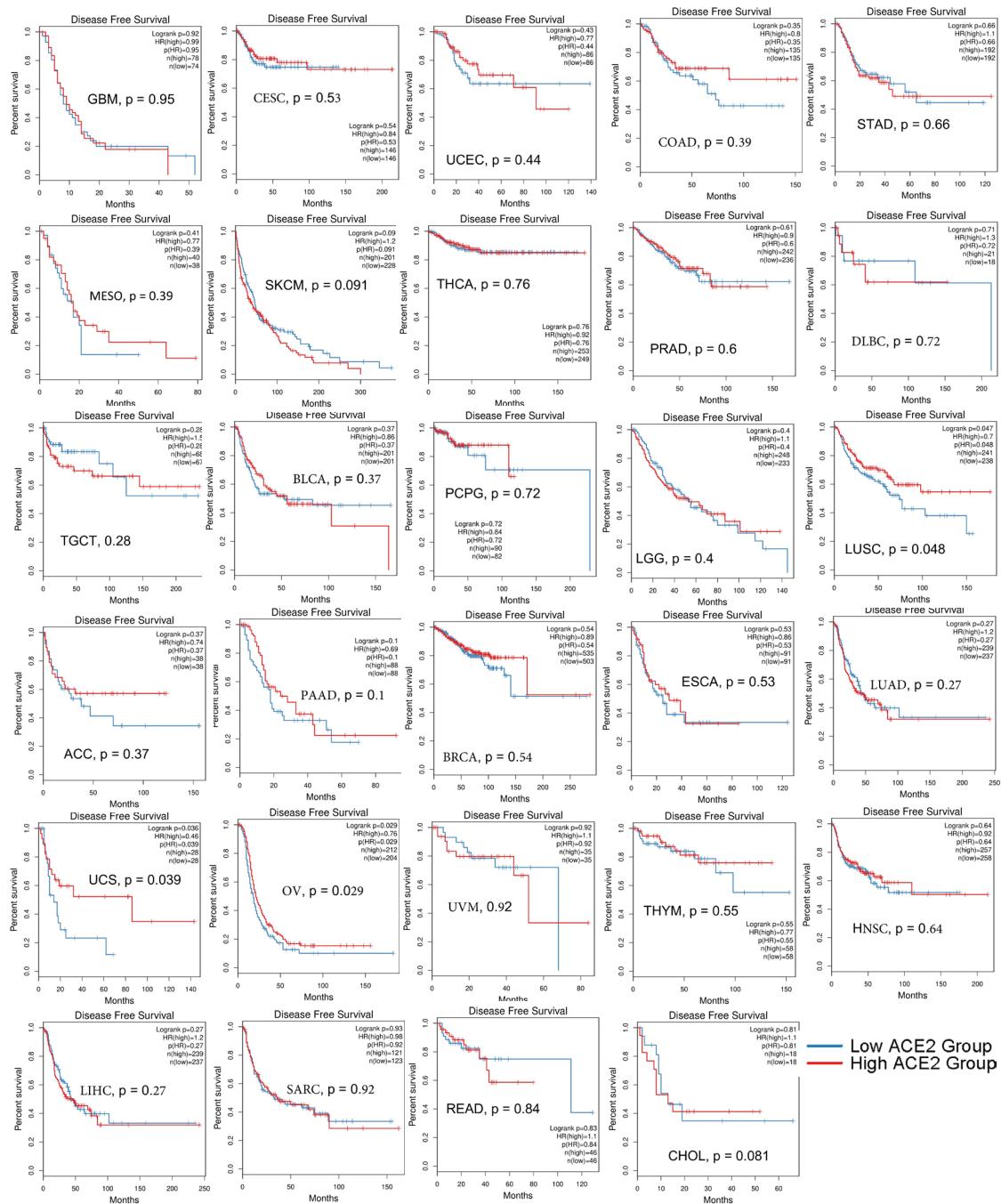
ACC: Adrenocortical Carcinoma
BLCA: Bladder Urothelial Carcinoma
BRCA: Breast Invasive Carcinoma
CESC: Cervical and Endocervical Cancer
CHOL: Cholangiocarcinoma (Bile Duct Cancer)
COAD: Colon Adenocarcinoma
DLBC: Diffuse Large B-Cell Lymphoma
ESCA: Esophageal Carcinoma
GBM: Glioblastoma Multiforme
HNSC: Head and Neck Squamous Cell Carcinoma
chRCC: Kidney Chromophobe
ccRCC: Kidney Renal Clear Cell Carcinoma
pRCC: Kidney Renal Papillary Cell Carcinoma
LAML: Acute Myeloid Leukemia
LGG: Lower-Grade Glioma
LIHC: Liver Hepatocellular Carcinoma
LUAD: Lung Adenocarcinoma
LUSC: Lung Squamous Cell Carcinoma
OV: Ovarian Serous Cystadenocarcinoma
PAAD: Pancreatic Adenocarcinoma
PCPG: Pheochromocytoma and Paraganglioma
PRAD: Prostate Adenocarcinoma
READ: Rectum Adenocarcinoma
SARC: Sarcoma
SKCM: Skin Cutaneous Melanoma
STAD: Stomach Adenocarcinoma
TGCT: Testicular Germ Cell Tumors
THCA: Thyroid Carcinoma
THYM: Thymoma
UCEC: Uterine Corpus Endometrial Carcinoma
UCS: Uterine Carcinosarcoma



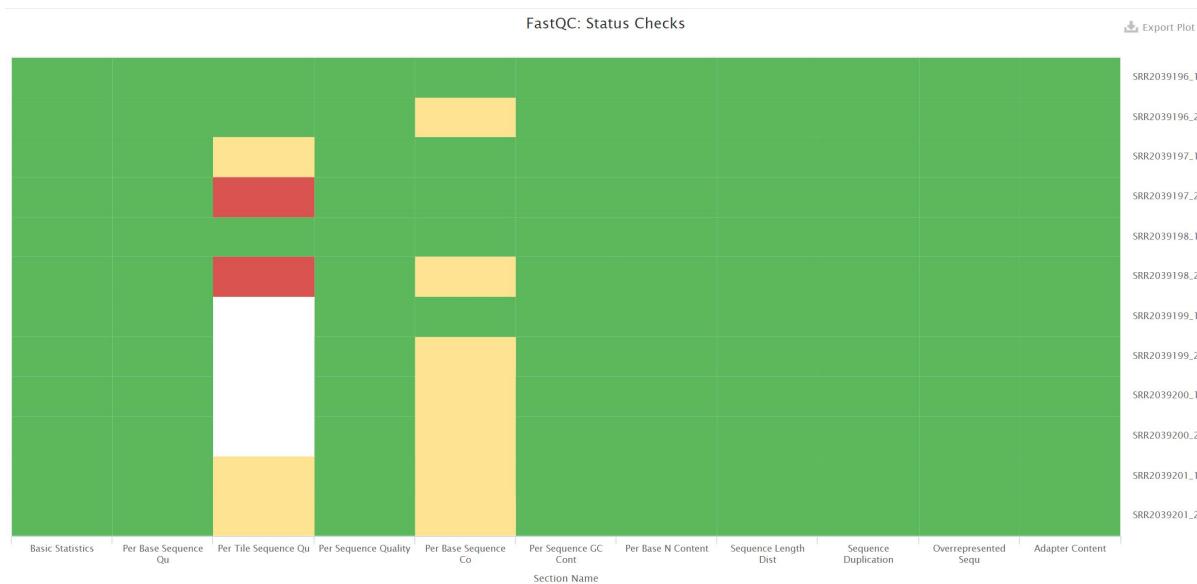
**Supplementary Figure 10.1.** The data for ACE2 expression in various cancers and their matched normal tissues was obtained from GEPIA2. Plots were generated using matched TCGA normal data and GTEx data. Significance was determined using a cutoff of a log-fold change of 1 and a p-value < 0.05. Significant differences are denoted in the graph with asterisks (\*) above the boxes. "num(T)" represents the number of tumor samples, and "num(N)" represents the number of normal samples included in the analysis.



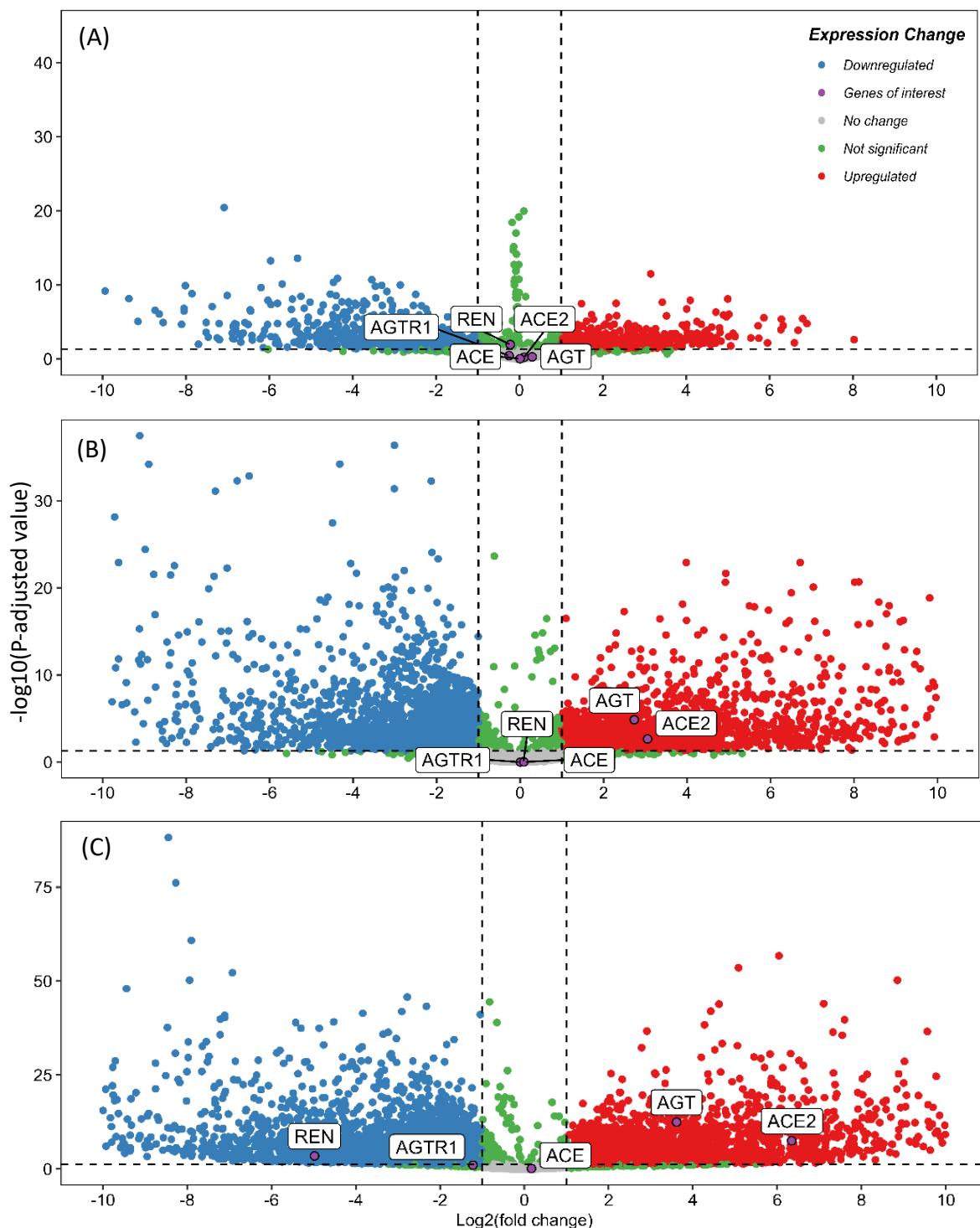
**Supplementary Figure 10.2.** Overall survival plots show the impact of high and low ACE2 expression on various cancers. The data utilised for these analyses were sourced from GEPIA and with a criteria of p-value of 0.05, and ACE2 expression grouping into low and high expression was determined by the median.



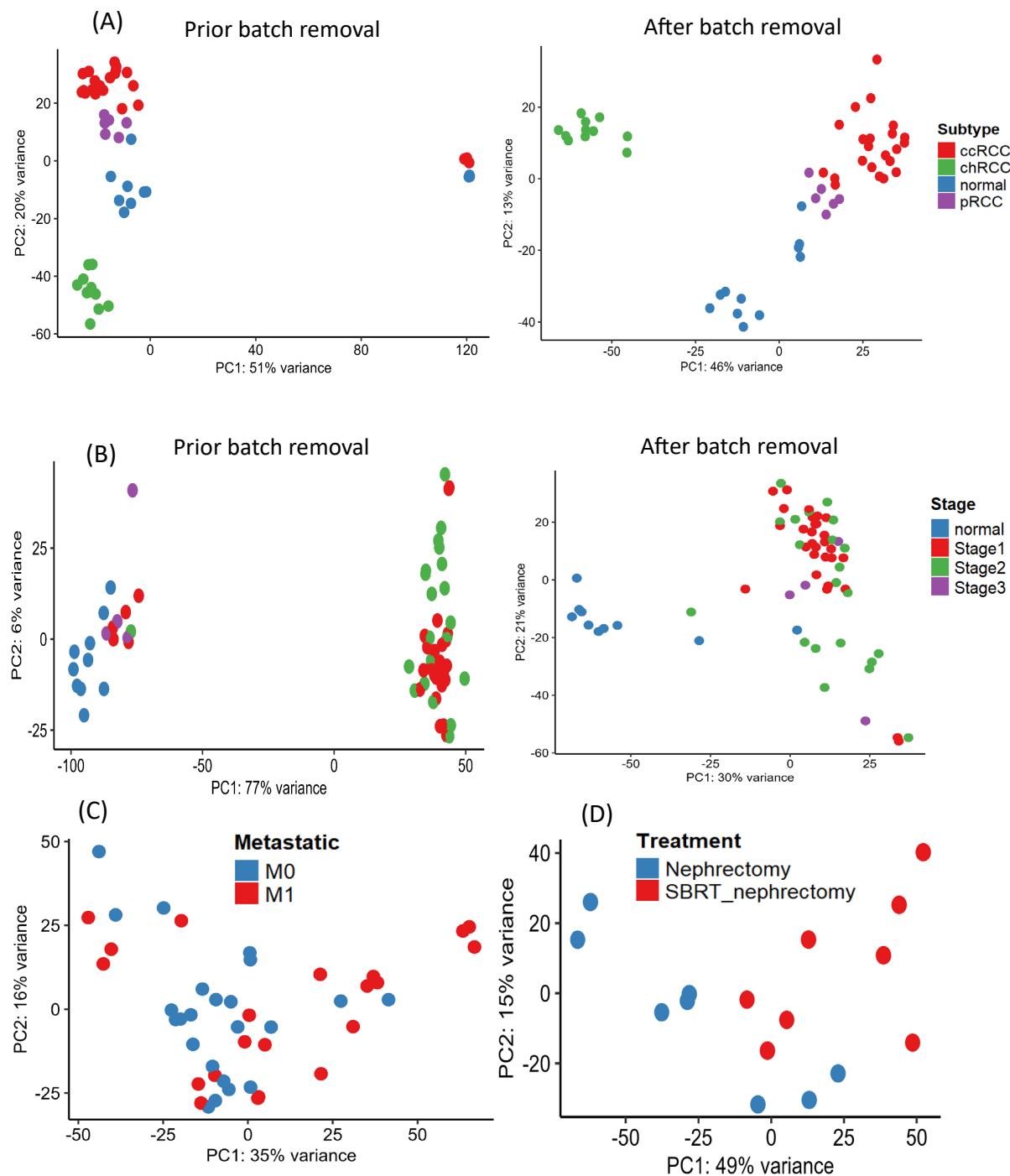
**Supplementary Figure 10.3.** Disease-free survival plots show the impact of high and low ACE2 expression on various cancers. The data utilised for these analyses were sourced from GEPIA and with a criteria of p-value of 0.05, and ACE2 expression grouping into low and high expression was determined by the median.



**Supplementary Figure 10.4.** Representative MULTIQC status check of the report for the GSE217386 dataset, providing a comprehensive summary of all quality parameters assessed within the dataset. The green colour indicates a satisfactory quality, the yellow colour alarms the user that the quality of the reads is average quality, the red colour signifies that the quality of the reads should be inspected further, whereas the white colour signifies that the quality of raw data was not assessed for that parameter.



**Supplementary Figure 10.5.** Volcano plots have been generated to visualise the overall gene expression patterns in three RCC subtypes. These comparisons include (A) pRCC vs. ccRCC, (B) ccRCC vs. chRCC, and (C) pRCC vs. chRCC. In these plots, specific thresholds have been applied, namely a log-fold change of 1 and a P-adjusted value less than 0.05, which are represented by dashed horizontal and vertical lines, respectively. Differential expression of five genes – ACE2, ACE, AGTR1, REN and AGT – from the two main RAS axes are highlighted in boxes. The other three genes – AGTR2, CYPB11B2 and MAS1 – that are also assigned to the main two RAS axes are not displayed because they were either not expressed or their expression data was filtered during the DGEA process due to low counts.



**Supplementary Figure 10.6.** Principal component analysis that summarises gene expression data into the two principal components. Gene expression was summarised in four different contexts: (A) subtype, (B) stage, (C) metastasis progression, and (D) treatment. Additional batch effect removal is shown for subtype and stage context which demonstrates that batch effects are not visually apparent in data after batch effect removal.