2023

Cancer/Testis Gene Expression Changes in Metastatic Cancer

Clara M. Mosentine
Michigan Technological University, cmmosent@mtu.edu

CANCER/TESTIS GENE EXPRESSION CHANGES IN METASTATIC CANCER

By

Clara M. Mosentine

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Biological Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2023

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Biological Sciences.

Department of Biological Sciences

Thesis Advisor:	*Paul Goetsch*
Committee Member:	*Xiaohu Tang*
Committee Member:	*Rupali Datta*
Department Chair:	*Chandrashekhar P. Joshi*

Table of Contents

List of Figures iv

List of Tables v

Author Contribution Statement vii

Acknowledgements viii

Abstract ix

1 Introduction 1
- 1.1 Cancer/Testis Genes 1
- 1.2 Cancer/Testis Antigens and Medical Applications 2
- 1.3 Stages of Cancer 3
- 1.4 Goals and Hypothesis 4

2 Materials and Methods 5
- 2.1 Data used 5
- 2.2 Differential Expression Analysis 7
- 2.3 Data Cleaning 7
- 2.4 ANOVA Analysis 8

3 Results 9
- 3.1 deCT Gene Analysis 9
- 3.2 Confirmed CT Genes Found 11
- 3.3 deCT Genes Differentially Expressed via Stage 14
- 3.4 Overall Metastatic Behavior in CT genes 24
- 3.5 Genes of Interest 25
  - 3.5.1 TTK 35
  - 3.5.2 NCAPH 38
  - 3.5.3 KIF18A 40
  - 3.5.4 CDCA5 42
  - 3.5.5 CCNB2 44
  - 3.5.6 PBK 46
  - 3.5.7 KPNA2 48

4 Discussion 50

5 Reference List 51

List of Figures

Figure 3.1. Number of deCT Genes in Tested Tissue Types. 10

Figure 3.2. Number of Confirmed CT Genes in Tested Tissue Types. 1**Error! Bookmark not defined.**

Figure 3.3. Number of Shared Genes of Interest between Cancer Types. 26

Figure 3.4. Stagewise Expression of TTK. 37

Figure 3.5. Stagewise Expression of NCAPH. 39

Figure 3.6. Stagewise Expression of KIF18A. 41

Figure 3.7. Stagewise Expression of CDCA5. 43

Figure 3.8. Stagewise Expression of CCNB2. 45

Figure 3.9. Stagewise Expression of PBK. 47

Figure 3.10. Stagewise Expression of KPNA2. 49

List of Tables

Table 2.1. Full List of the Cancer Types Analyzed and the Number of Samples of Each Cancer Type. 6

Table 3.1. deCT Genes Differentially Expressed via Stage in Breast Infiltrating Duct Carcinoma. 15

Table 3.2. deCT Genes Differentially Expressed via Stage in Intestine Adenocarcinoma. 16

Table 3.3. deCT Genes Differentially Expressed via Stage in Kidney Clear Cell Adenocarcinoma. 17

Table 3.4. deCT Genes Differentially Expressed via Stage in Kidney Papillary Adenocarcinoma. 20

Table 3.5. deCT Genes Differentially Expressed via Stage in Kidney Renal Cell Carcinoma. 22

Table 3.6. Genes Exhibiting Stagewise Differential Expression in 4 of the 7 Cancer Types Tested. 27

Table 3.7. Genes Exhibiting Stagewise Differential Expression in 3 of the 7 Cancer Types Tested. 29

Table 3.8. Genes Exhibiting Stagewise Differential Expression in 2 of the 7 Cancer Types Tested. 31

Table 3.9. Genes Exhibiting Stagewise Differential Expression in 1 of the 7 Cancer Types Tested. 33

Author Contribution Statement

It should be known that Noah Mason aided significantly in the implementation of the DESeq2 pipeline, and in the creation of the annotated_Expression pipeline. Paul Goetsch was also invaluable as a source of guidance in both the analysis and writing of this study.

Acknowledgements

In no particular order, I would like to thank my friends, family, and colleagues for helping me through this past year. No one is an island, and I want to acknowledge everyone I have leaned on for support or help. I would first like to thank my family for showing me unconditional support over the past year. I would like to thank my grandmother, who paid my tuition. I would like to thank Destiny Wu and Evelyn Pae for listening to my struggles and helping me through them with empathy and compassion. I would also like to thank Noah Mason, who collaborated with me over the summer to create the analyses pipelines that eventually became precursors to those used in this study. Without his help, this study would not be nearly as complete. Finally, I would like to thank my advisor Dr. Paul Goetsch, who has given me this opportunity, and the guidance necessary to complete this step in my academic journey.

Thank you, all of you!

Abstract

Metastasis is the movement of cancerous cells to new parts of the body, often through the blood or lymph systems. Metastasis is classified as stage IV cancer, a prognosis that is significantly more difficult to effectively treat compared to earlier cancer stages. We are interested in assessing whether expression of Cancer/testis (CT) genes, a class of genes that are predominantly expressed in germ cells while also being abnormally expressed in a large percentage of cancers, is associated with cancer metastasis. Germ cells make up an organism's reproductive system, such as the testis and ovaries, and exhibit cellular immortality and, in the case of testis, self-proliferative behavior in the form of spermatogenesis. Cancer cells adopt similar germ and self-proliferative behaviors, making CT gene analysis useful to isolate the genes and pathways involved in tumorigenesis. By assessing for differential expression between testis germline, normal tissue, and cancerous tissue, we identified 50 CT genes whose expression increases in stage IV tumors. Importantly, we determined that the majority of known CT genes were not associated with metastasis, suggesting that the most CT genes are activated in early cancer stages. However, many of the 50 CT genes that were associated with metastasis have previously been linked to metastatic and aggressive cancer behavior. Our analysis will direct more attention towards these CT genes linked with metastasis. Ideally, understanding how these CT genes are overactivated in stage IV tumors may shed light on a new avenue to develop new diagnostics or treatments aimed to improve quality of life, longevity, and prognosis for patients with stage IV cancer prognoses.

1 Introduction

Our analysis investigates the relationship between cancer/testis genes and the metastatic behaviors characterized in stage IV cancers. Cancer/testis (CT) genes are a class of genes predominantly expressed in germ cells while also being abnormally expressed in a large percentage of cancers. CT genes are associated with tumorigenesis, as cancer cells adopt germ cell-like behavior. For example, testis contain germline reproductive cells that undergo meiosis to generate sperm gametes. Importantly, meiotic behavior is associated with cancer metastasis, the traveling of cancer cells from the primary tumor to new locations in the body. If CT genes are already associated with tumorigenesis, could an important subset also be associated with metastasis?

CT genes proven to be associated with metastasis would be ideal targets for future studies on cancer dysregulation. Currently, after a cancer metastasizes, prognosis becomes significantly poorer for the patient. We expect that CT genes enriched in metastatic cancers will be important targets for the development of future diagnostic or therapeutic procedures, with the aim to reduce morbidity and increasing life expectancy of patients diagnosed with stage IV cancers.

1.1 Cancer/Testis Genes

Cancer/testis (CT) genes are genes that are predominantly expressed in both cancer and testis cells, when compared to other cell types in the body (Craig *et al.*, 2021; da

Silva *et al.*, 2017). CT genes are, by definition, strongly associated with cancer. First described in 1991 by van der Bruggen, over a thousand CT genes have been identified to date (Chen *et al.*, 1994; Chen *et al.*, 1997; da Silva *et al.*, 2017; De Smet *et al.*, 1994; Dobrynin *et al.*, 2013; van der Bruggen *et al.*, 1991; Wang *et al.*, 2016).

A previous CT gene analysis identified genes responsible for behaviors shared by both cancer cells and testis cells (Wang *et al.*, 2016). The testis contain male germline reproductive cells, which undergo meiosis in the form of spermatogenesis, the act of creating haploid sperm gametes (Guo *et al.*, 2018). Meiotic gene activation and behaviors have been observed in germ cell tumors, and at least one meiotic gene supports cell migration and invasion, a hallmark of metastatic activity in cancer (Feichtinger & McFarlane, 2019; Tan *et al*., 2023). In addition, meiotic genes drive genome instability and cancer progression through the promotion of proliferative behavior in tumor cells (Lingg *et al*., 2022). This study aims to analyze the role of CT genes in metastasis, as the meiotic behavior in testis germline may promote cell proliferation and motility.

1.2 Cancer/Testis Antigens and Medical Applications

Cancer/testis (CT) antigens are a group of antigens found in tumors, which are normally expressed in testis germ cells, but not in other somatic tissues (Scanlan *et al*., 2002). In cancer treatments, CT antigens are used to create immune responses toward the tumors they are found in (Scanlan *et al*., 2002). These treatments are known as immunotherapy, where tumor-specific antigens are targeted by the patient's immune system (Zhang & Zhang, 2020).

CT antigens are particularly useful in immunotherapy, as they are by definition present in cancerous cells but not in most normal tissue (Soldatova *et al.*, 2019). This allows for treatment to target cancer specifically, without harming other systems in the body (Soldatova *et al.*, 2019; Xu *et al.*, 2019). Due to the effectiveness of targeting CT antigens in immunotherapy, understanding their regulation in cancer is a major step towards the development of new cancer drugs and therapies.

1.3 Stages of Cancer

Within cancer diagnoses, stages are used to indicate the severity of the tumor found. While different cancer diagnoses use various substages, most follow the same guiding definitions for the overall stages of cancer progression.

Stages I to III are all characterized by tumor size, and the localized spread of the primary tumor. Stage I cancer is a small tumor that has not spread to surrounding tissues or lymph nodes. Stage II cancer has grown in size, but still has not spread. Stage III cancer is a larger tumor that may have spread to surrounding tissues or lymph nodes (American Society of Clinical Oncology, 2021).

Finally, stage IV is where metastasis is observed; the cancer has moved from its point of origin to a new organ in the body (other than the lymph nodes). Metastasis is characterized by cancerous cells detaching from the primary tumor and traveling through the blood or lymph system to a new location, where they then begin to proliferate into a new tumor, sometimes called a “secondary” or “metastatic” cancer.

Gene expression has been observed across the stages of cancer in several studies. A study on esophageal squamous cell carcinoma found that a few select genes such as P160ROCK and JNK2 were activated in the early stages of cancer. (Zhou *et al*., 2003). Another study looking for stage-specific genes in cancer found several genes that were enriched in relevant cancerous pathways (Aouiche *et al*., 2019). Finally, some CT genes, such as CT45A1, MAGE-C2, GAGE, XAGE1, and CAGE, have already been proven to act as oncogenic triggers for cancer metastasis (Shang *et al*., 2014; Gjerstorff *et al.*, 2015). However, a systematic analysis of CT genes and their association with cancer has yet to be performed.

1.4 Goals and Hypothesis

The goal of this study is to analyze confirmed CT genes for links to metastatic behavior. I hypothesize that the majority of CT genes show little correlation to metastasis. However, the few that do may be of interest to future metastatic cancer studies. These genes could be targeted to research and create new treatments to fight late-stage cancer.

2 Materials and Methods

2.1 Data used

This analysis utilized TCGA and GTEX sample data gathered from the Genomic Data Commons (Grossman *et al*., 2016). While every cancer and tissue types are available to be analyzed, for the purposes of this study, only those listed in Table 2.1 were used. The RNA-Seq data, which shows the expression of each gene, was gathered for each sample. The metadata of each sample was also collected, including the sample name, primary diagnosis, gender, condition, tissue type, tumor stage, case ID, and mutation.

Since several cancer types were reported under different names, we created a new variable called “lumped diagnosis,” which contained homogenized diagnosis names between all relevant samples. Only cancer diagnoses with the largest number of samples were used. These diagnoses and their sample sizes are shown below (Table 2.1).

Table 2.1. Full List of Cancer Types Analyzed and the Number of Samples of Each Cancer Type.

Tissue Type	**Cancer Diagnosis**	**Number of Samples**
Breast	Infiltrating duct carcinoma, NOS	958
Intestine	Adenocarcinoma, NOS	691
Kidney	Clear cell adenocarcinoma, NOS	594
Kidney	Papillary adenocarcinoma, NOS	300
Kidney	Renal cell carcinoma, NOS	282
Lung	Adenocarcinoma, NOS	690
Lung	Squamous cell carcinoma, NOS	528

2.2 Differential Expression Analysis

The genes in each tissue were first compared using the R package DESeq2 (Love *et al*., 2014). This compared GTEX, noncancerous tissue data (either breast, intestine, kidney, or lung) to GTEX testis data. We selected any gene that showed a log2 fold-change of more than 1 (meaning the genes were upregulated in testis), and a p-adjusted value of less than 0.01. This analysis created a list of upregulated testis genes that were more highly expressed in testis, as compared to the normal tissue.

Next, the TCGA data for the specific cancer type was compared to the GTEX tissue data. For example, we compared breast cancer RNA-seq samples from the TCGA to normal breast tissue RNA-seq samples from the GTEX. The same fold and p-adjusted value cut-offs were used. This analysis created a list of upregulated cancer genes that were more highly expressed in cancerous tissue, as compared to normal tissue. The upregulated testis genes and upregulated cancer genes were then compared. Only genes which were upregulated in both cancer and testis, which we named “differentially expressed CT genes” or “deCT genes”, were selected and used in our downstream analysis.

2.3 Data Cleaning

We used our Annotated_Expression program to affix metadata to each sample, including primary diagnosis, gender, condition of sample (cancerous or normal), tissue

type, and tumor stage. We filtered out samples that lacked desired metadata to ensure null values and zeroes did not affect any calculations or final results.

2.4 ANOVA Analysis

At the beginning of the ANOVA analysis, we intersected the resultant deCT genes with a curated list of confirmed CT genes. We did this to ensure that all genes used in the final analysis were CT genes. This both aided in ensuring the analysis was on pre-existing CT genes (allowing comparison between my work and previous studies), and limited the number of results to a more reasonable amount to sort through manually.

The identified deCT gene set that only includes genes that were upregulated in both cancer and testis, as compared to normal tissue, were analyzed for differential expression across cancer stages. We performed a one-way ANOVA test, which identified if any of the four stages of cancer were differentially expressed, as compared to each other stage. The resultant P values from the one-way ANOVA test were then converted to P adjusted values using the Bonferroni method. Genes determined to be differentially expressed in the stages of cancer, with a P adjusted value of < 0.01, were then sorted through manually.

3 Results

3.1 deCT Gene Analysis

We first identified upregulated genes shared in both testis and cancer after each were compared to normal tissue. We named these “deCT genes,” which stands for “differentially expressed CT” genes. Out of a total of 42,685 genes, we identified 11,649 unique deCT genes upregulated in at least one of the four tissues analyzed (breast, intestine, kidney, and lung (figure 3-1). In breast cancers, including infiltrating duct carcinomas, we identified 5,472 deCT genes. In lung cancers, including squamous cell carcinoma and adenocarcinoma, we identified 5,072 deCT genes. In intestinal cancers, including adenocarcinoma, we identified 6,800 deCT genes. In kidney cancers such as clear cell adenocarcinoma, papillary adenocarcinoma, and renal cell carcinoma, we identified 3,283 deCT genes. Importantly, 684 deCT genes were observed in all four tissues (figure 3-1), making up 5.9% of the identified deCT gene set. We also found 5,809 deCT genes that were only in one tissue type. This made up 49.9% of all the deCT genes found. This indicated that a large percentage of deCT genes were tissue specific.

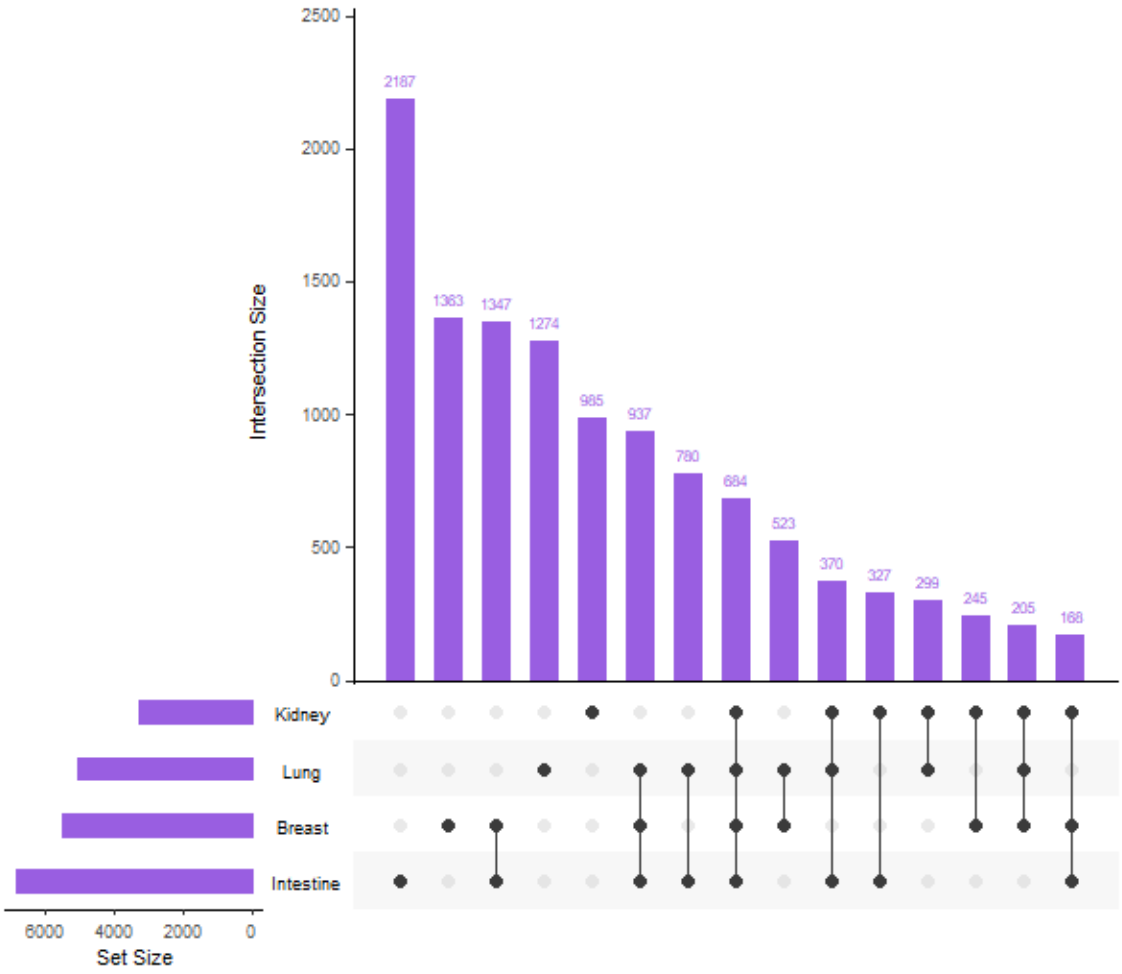

Figure 3-1 Number of deCT Genes in Tested Tissue Types

This figure shows the number of deCT genes shared between the analyzed tissues. The black circles connected indicate the number of genes listed above are shared between those tissues.

3.2 Confirmed CT Genes Found

We were interested in evaluating whether our deCT gene analysis identified known CT genes identified in other studies (van der Bruggen *et al*., 1991; Chen *et al*., 1994; De Smet *et* al., 1994; Chen *et al*., 1997; Dobrynin *et al*., 2013; Wang *et al*., 2016; da Silva *et al*., 2017). We intersected our deCT gene set with known CT genes to ensure our downstream analysis stayed within the bounds of previously confirmed CT genes. Out of the 1,449 confirmed CT genes described in the literature, 481 were identified as differentially expressed between testis, cancer, and normal tissue in our analysis of breast, intestine, kidney, and lung cancer. Therefore, our analysis accounted for about 33% of known CT genes.

Within these 481 confirmed CT genes, we observed 115 genes were tissue specific. Of these 115 unrepeated genes, infiltrating duct adenocarcinoma of the breast contributed 97, or 84.3%. Outside of this, the number of CT genes found in each diagnosis was roughly the same, as seen in figure 3-2. Altogether, we expect that extending our deCT gene analysis to other tissues in the future will account for more known CT genes observed previously in the literature.

In breast cancers, including infiltrating duct carcinomas, we identified 321 confirmed CT genes. In lung cancers, including squamous cell carcinoma and adenocarcinoma, we identified 253 confirmed CT genes. In intestinal cancers, including adenocarcinoma, we identified 199 confirmed CT genes. In kidney cancers, including

clear cell adenocarcinoma, papillary adenocarcinoma and renal cell carcinoma, we identified 226 confirmed CT genes.

While we found the most confirmed CT genes in breast cancers, we still found at least 199 confirmed CT genes within each other tissues. This indicates that our deCT gene analysis was effective in finding confirmed CT genes.

We also found 71 of the 481 confirmed CT genes that were in each of the 4 tissues analyzed (breast, intestine, kidney, and lung). This meant around 14.8% of the confirmed CT genes found were tissue nonspecific and expressed in cancer regardless of diagnosis or tissue type.

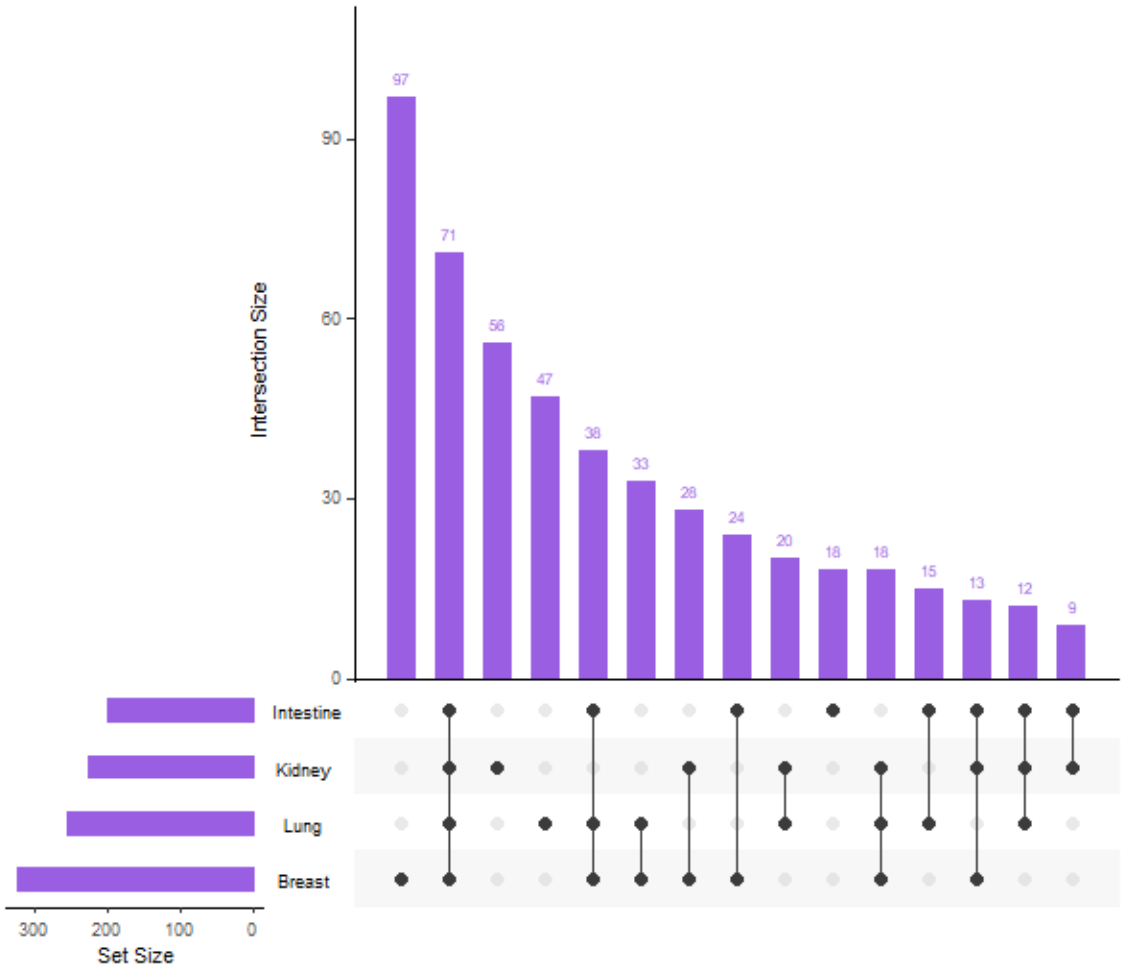

Figure 3-2 Number of Confirmed CT Genes in Tested Tissue Types

This figure shows the number of deCT genes that were confirmed CT genes. The number of genes is shown as they were found in the analyzed tissues. The black circles connected indicate the number of genes listed above are shared between those tissues.

3.3 deCT Genes Differentially Expressed via Stage

To identify the subset of deCT genes associated with stage IV cancer, we performed an ANOVA analysis on 481 confirmed CT genes observed in our deCT analysis and previously in the literature. We used ANOVA to compare the expression of each gene as it changed throughout the stages of cancer. This resulted in a list of genes with differential expression in the stages of cancer significant enough to result in a p-adjusted value of less than 0.01. The full list of these genes are provided in the tables below (Table 3.1-3.5). Genes with significant changes in expression or low p-adjusted values also had their expression graphed via stage.

Both diagnoses in lung (adenocarcinoma and squamous cell carcinoma) resulted in no results, indicating that none of the genes tested were differentially expressed across the stages of cancer.

Table 3.2. deCT Genes Differentially Expressed via Stage in Breast Infiltrating Duct Carcinoma.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**
ENSG00000093009	CDC45
ENSG00000104147	OIP5
ENSG00000112742	TTK
ENSG00000121152	NCAPH
ENSG00000121211	MND1
ENSG00000121621	KIF18A
ENSG00000146670	CDCA5
ENSG00000157456	CCNB2
ENSG00000168078	PBK
ENSG00000182459	TEX19
ENSG00000182481	KPNA2
ENSG00000204019	CT83

Table 3.2. deCT Genes Differentially Expressed via Stage in Intestine Adenocarcinoma.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**
ENSG00000168078	PBK
ENSG00000175294	CATSPER1

Table 3.3. deCT Genes Differentially Expressed via Stage in Kidney Clear Cell Adenocarcinoma.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**
ENSG00000004848	ARX
ENSG00000091651	ORC6
ENSG00000093009	CDC45
ENSG00000104147	OIP5
ENSG00000105011	ASF1B
ENSG00000105467	SYNGR4
ENSG00000112742	TTK
ENSG00000121152	NCAPH
ENSG00000121211	MND1
ENSG00000121621	KIF18A
ENSG00000123975	CKS2
ENSG00000124196	GTSF1L

ENSG00000133101	CCNA1
ENSG00000136327	NKX2-8
ENSG00000139800	ZIC5
ENSG00000146670	CDCA5
ENSG00000152086	TUBA3E
ENSG00000152503	TRIM36
ENSG00000153044	CENPH
ENSG00000157456	CCNB2
ENSG00000159224	GIP
ENSG00000168078	PBK
ENSG00000170965	PLAC1
ENSG00000171540	OTP
ENSG00000171872	KLF17
ENSG00000171956	FOXB1
ENSG00000175294	CATSPER1

ENSG00000176177	ENTHD1
ENSG00000179046	TRIML2
ENSG00000182481	KPNA2
ENSG00000185662	SMIM23
ENSG00000205078	SYCE1L
ENSG00000220891	LL22NC03-63E9.3
ENSG00000248099	INSL3
ENSG00000258713	C20orf141

Table 3.4. deCT Genes Differentially Expressed via Stage in Kidney Papillary Adenocarcinoma.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**
ENSG00000068985	PAGE1
ENSG00000091651	ORC6
ENSG00000093009	CDC45
ENSG00000101003	GINS1
ENSG00000104147	OIP5
ENSG00000105011	ASF1B
ENSG00000112742	TTK
ENSG00000121152	NCAPH
ENSG00000121211	MND1
ENSG00000121621	KIF18A
ENSG00000123975	CKS2
ENSG00000133863	TEX15

ENSG00000136492	BRIP1
ENSG00000139800	ZIC5
ENSG00000146670	CDCA5
ENSG00000153044	CENPH
ENSG00000157456	CCNB2
ENSG00000161609	KASH5
ENSG00000168078	PBK
ENSG00000170965	PLAC1
ENSG00000171864	PRND
ENSG00000179046	TRIML2
ENSG00000182459	TEX19
ENSG00000182481	KPNA2
ENSG00000198754	OXCT2
ENSG00000258713	C20orf141

Table 3.5. deCT Genes Differentially Expressed via Stage in Kidney Renal Cell Carcinoma.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**
ENSG00000112742	TTK
ENSG00000117148	ACTL8
ENSG00000121152	NCAPH
ENSG00000121621	KIF18A
ENSG00000133101	CCNA1
ENSG00000136327	NKX2-8
ENSG00000139800	ZIC5
ENSG00000146670	CDCA5
ENSG00000152503	TRIM36
ENSG00000157456	CCNB2
ENSG00000160994	CCDC105
ENSG00000164399	IL3

ENSG00000171872	KLF17
ENSG00000173908	KRT28
ENSG00000176177	ENTHD1
ENSG00000182481	KPNA2
ENSG00000185972	CCIN
ENSG00000220891	LL22NCO3-62E9.3
ENSG00000241369	LINC01192
ENSG00000258713	C20orf141

3.4 Overall Metastatic Behavior in CT genes

Our goal in this analysis was to determine if the majority of CT genes experienced differential expression throughout the stages of cancer. If a CT gene was upregulated in stage IV cancers, it would be associated with metastasis. We selected any deCT genes that showed differential expression in any stage of cancer.

In breast infiltrating duct carcinoma, 12 of the 321 confirmed CT genes, or 3.7%, displayed significant differential expression between tumor stages. In lung adenocarcinoma, 0 of the 253 confirmed CT genes, or 0%, displayed significant differential expression between tumor stages. In lung squamous cell carcinoma, 0 of the 253 confirmed CT genes, or 0%, displayed significant differential expression between tumor stages. In intestinal adenocarcinoma, 2 of the 199 confirmed CT genes, or 1%, displayed significant differential expression between tumor stages. In kidney clear cell adenocarcinoma, 35 of the 226 confirmed CT genes, or 15.4%, displayed significant differential expression between tumor stages. In kidney papillary adenocarcinoma, 26 of the 226 confirmed CT genes, or 11.5%, displayed significant differential expression between tumor stages. In kidney renal cell carcinoma, 20 of the 226 confirmed CT genes, or 8.8%, displayed significant differential expression between tumor stages.

Our results showed that intestinal, breast, and lung cancers did not exhibit stage-wise differential expression in the majority of tested CT genes. However, all 3 kidney cancer types showed higher numbers of differential expression via stage. This indicates

that the kidney cancers analyzed (clear cell adenocarcinoma, papillary adenocarcinoma, and renal cell carcinoma) have more genes that are differentially expressed depending on cancer stage.

We looked at all 7 cancer types and their resultant confirmed CT genes that were differentially expressed in the stages of cancer. Between all of the analyzed cancer types, we identified 498 unique confirmed CT genes. Only 50 of those genes showed any connection in regulation to the stage of cancer, in any of the tested cancer types. 10% of the unique CT genes found were differentially expressed across the stages of cancer. This indicates that 90% of analyzed CT genes are not differentially expressed via stage.

3.5 Genes of Interest

In total, there were 95 genes between all cancer types that were marked as substantially differentiated via the stages of cancer. Many of these genes overlapped, being differentially expressed in multiple diagnoses at once. Of the five diagnoses that produced any results, 7 genes were observed in 4 total cancer types, 5 genes were observed in 3 cancer types, 14 genes were observed in 2 cancer types, and 24 genes were only observed in one cancer type (tables 3.6-3.9). As seen in figure 3-3, kidney diagnoses had the largest numbers of genes of interest which were not observed in other types of cancer from different tissues. In total, 23 (95.8%) of the 24 genes of interest were observed only in kidney cancers.

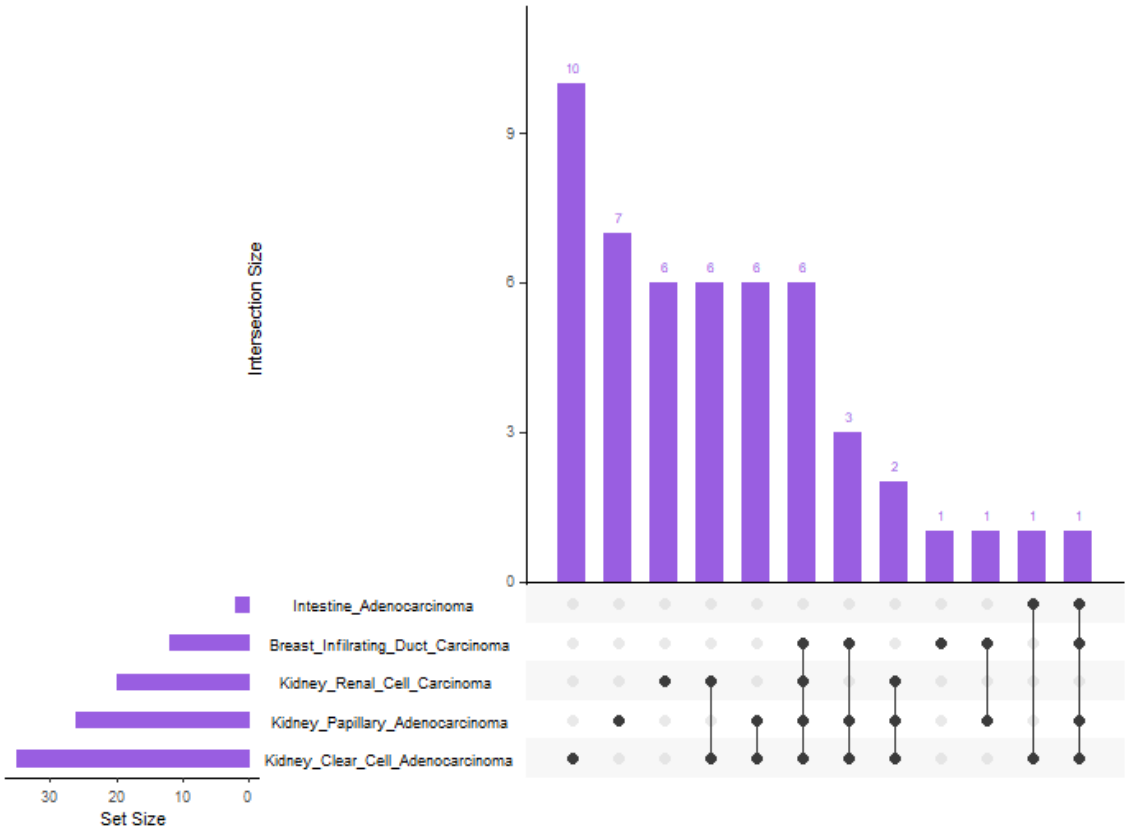

Figure 3-3 Number of Shared Genes of Interest between Cancer Types

This figure shows the number of genes shared between the analyzed cancer types. The black circles connected indicate the number of genes listed above are shared between those cancer types.

Table 3.6. Genes Exhibiting Stagewise Differential Expression in 4 of the 7 Cancer Types Tested.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**	**Cancer types gene is observed in**
ENSG00000112742	TTK	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000121152	NCAPH	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000121621	KIF18A	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000146670	CDCA5	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma,

		Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000157456	CCNB2	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000168078	PBK	Breast Infiltrating Duct Carcinoma, Intestine Adenocarcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000182481	KPNA2	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma

Table 3.7. Genes Exhibiting Stagewise Differential Expression in 3 of the 7 Cancer Types Tested.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**	**Cancer types gene is observed in**
ENSG00000093009	CDC45	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000104147	OIP5	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000121211	MND1	Breast Infiltrating Duct Carcinoma, Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000139800	ZIC5	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma

ENSG00000258713	C20orf141	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma

Table 3.8. Genes Exhibiting Stagewise Differential Expression in 2 of the 7 Cancer Types Tested.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**	**Cancer types gene is observed in**
ENSG00000091651	ORC6	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000105011	ASF1B	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000123975	CKS2	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000133101	CCNA1	Kidney Clear Cell Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000136327	NKX2-8	Kidney Clear Cell Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000152503	TRIM36	Kidney Papillary Adenocarcinoma, Kidney Renal Cell Carcinoma

ENSG00000153044	CENPH	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000170965	PLAC1	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000171872	KLF17	Kidney Clear Cell Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000175294	CATSPER1	Intestine Adenocarcinoma, Kidney Clear Cell Adenocarcinoma
ENSG00000176177	ENTHD1	Kidney Clear Cell Adenocarcinoma, Kidney Renal Cell Carcinoma
ENSG00000179046	TRIML2	Kidney Clear Cell Adenocarcinoma, Kidney Papillary Adenocarcinoma
ENSG00000182459	TEX19	Breast Infiltrating Duct Carcinoma, Kidney Papillary Adenocarcinoma
ENSG00000220891	LL22NC03-63E9.3	Kidney Clear Cell Adenocarcinoma, Kidney Renal Cell Carcinoma

Table 3.9. Genes Exhibiting Stagewise Differential Expression in 1 of the 7 Cancer Types Tested.

Ensembl gene ID (ENSG)	**Human Gene Nomenclature Committee (HGNC) name**	**Cancer diagnoses gene is observed in**
ENSG00000004848	ARX	Kidney Clear Cell Adenocarcinoma
ENSG00000068985	PAGE1	Kidney Papillary Adenocarcinoma
ENSG00000101003	GINS1	Kidney Papillary Adenocarcinoma
ENSG00000105467	SYNGR4	Kidney Clear Cell Adenocarcinoma
ENSG00000117148	ACTL8	Kidney Renal Cell Carcinoma
ENSG00000124196	GTSF1L	Kidney Clear Cell Adenocarcinoma
ENSG00000133863	TEX15	Kidney Papillary Adenocarcinoma
ENSG00000136492	BRIP1	Kidney Papillary Adenocarcinoma
ENSG00000152086	TUBA3E	Kidney Clear Cell Adenocarcinoma
ENSG00000159224	GIP	Kidney Clear Cell Adenocarcinoma
ENSG00000160994	CCDC105	Kidney Renal Cell Carcinoma

ENSG00000161609	KASH5	Kidney Papillary Adenocarcinoma
ENSG00000164399	IL5	Kidney Renal Cell Carcinoma
ENSG00000171540	OTP	Kidney Clear Cell Adenocarcinoma
ENSG00000171864	PRND	Kidney Papillary Adenocarcinoma
ENSG00000171956	FOXB1	Kidney Clear Cell Adenocarcinoma
ENSG00000173908	KRT28	Kidney Renal Cell Carcinoma
ENSG00000185662	SMIM23	Kidney Clear Cell Adenocarcinoma
ENSG00000185972	CCIN	Kidney Renal Cell Carcinoma
ENSG00000198754	OXCT2	Kidney Papillary Adenocarcinoma
ENSG00000204019	CT83	Breast Infiltrating Duct Carcinoma
ENSG00000205078	SYCE1L	Kidney Clear Cell Adenocarcinoma
ENSG00000241369	LINC01192	Kidney Renal Cell Carcinoma
ENSG00000248099	INSL3	Kidney Clear Cell Adenocarcinoma

Of particular interest are the 7 genes that were observed to be differentially expressed throughout the stages of cancer in 4 different cancer types. We hypothesize that these genes may be involved in a germ cellular mechanism that is adopted in most cancers. 6 of these genes were differentially expressed in the breast infiltrating duct carcinoma diagnosis group, kidney clear cell adenocarcinoma group, kidney papillary adenocarcinoma group, and kidney renal cell carcinoma group. The only gene not expressed in all 3 kidney diagnoses was PBK, which was differentially expressed in intestinal adenocarcinoma, breast infiltrating duct carcinoma, clear cell adenocarcinoma of the kidney, and papillary adenocarcinoma of the kidney. Importantly, the intestinal adenocarcinoma diagnosis group only returned 2 deCT. In contrast, the analyses of kidney cancers returned larger numbers of differentially expressed genes, and majority of repeatedly observed deCT genes occurred in our analysis of kidney cancers. The 7 common genes are also well documented in association with poor prognosis, metastasis, and aggressive tumor growth. This indicates that the analysis performed in our experiment was functional, and correctly selected genes associated with stagewise expression changes and metastasis. Below, some of the research on each of the 7 most repeated genes is examined more closely.

3.5.1 TTK

The TTK gene (ENSG00000112742) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene product is TTK protein kinase, which is associated with significantly increased cell proliferation and migration in

human bladder cancer cells (Chen *et al.*, 2018). When this TTK protein kinase was knocked down in pancreatic cancer cells, cell proliferation was reduced, and apoptosis and necrosis rates were increased (Kaistha *et al*., 2014). A study in 2022 also showed that TTK expression predicted tumor proliferation and invasion in triple positive breast cancer (Gao *et al.*, 2022).

TTK's association with metastatic behaviors such as tumor invasion and migration is shown in figure 3-4, where we observed visible upregulation in stages III and IV of all kidney diagnoses. These findings support previous research linking TTK to metastatic behavior in several tissue varieties. Interestingly, breast infiltrating duct carcinoma displayed downregulation in stage I, and no significant differential expression in stage IV. We postulate that while TTK is a predictor of triple positive breast cancer, it may not contribute to metastatic behavior in other breast cancer diagnoses.

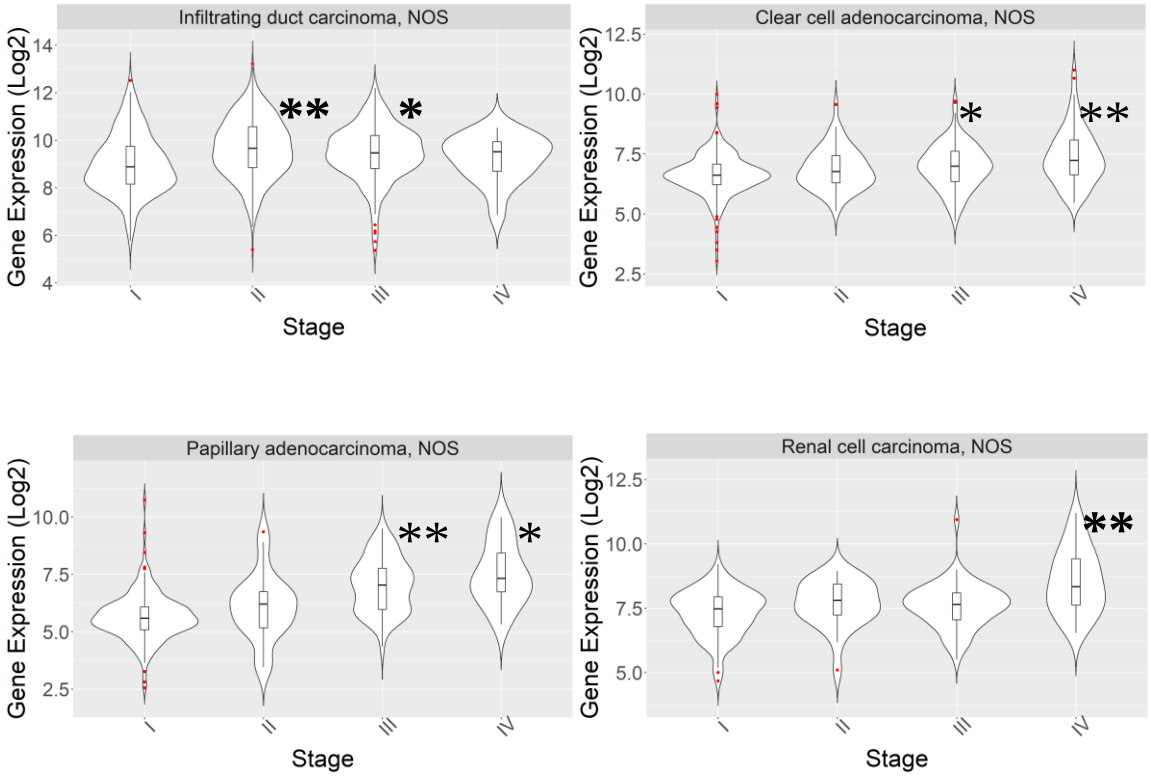

Figure 3-4 Stagewise Expression of TTK

This figure shows the expression changes of the gene TTK through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

z

3.5.2 NCAPH

The NCAPH gene (ENSG00000121152) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene product is NCAPH, also known as non-SMC condensing I complex subunit H. One study of NCAPH determined it has high odds ratios in stagewise comparisons, as well as prognosis comparisons and metastasis comparisons in endometrial cancer (Qui *et al.,* 2020). This means it is highly associated with metastasis and stagewise progression, as well as poor prognosis. The same study also labeled NCAPH as an oncogene. In addition, NCAPH has been proven to play a role in cell proliferation in colon cancer (Yin *et al*., 2017).

We observed visible upregulation in stages III and IV of all kidney diagnoses in figure 3-5. This supports previous research associating NCAPH with metastasis and stagewise progression. We also observed differential expression of NCAPH in stage I breast infiltrating duct carcinoma. This may link NCAPH with tumorigenesis in breast cancer.

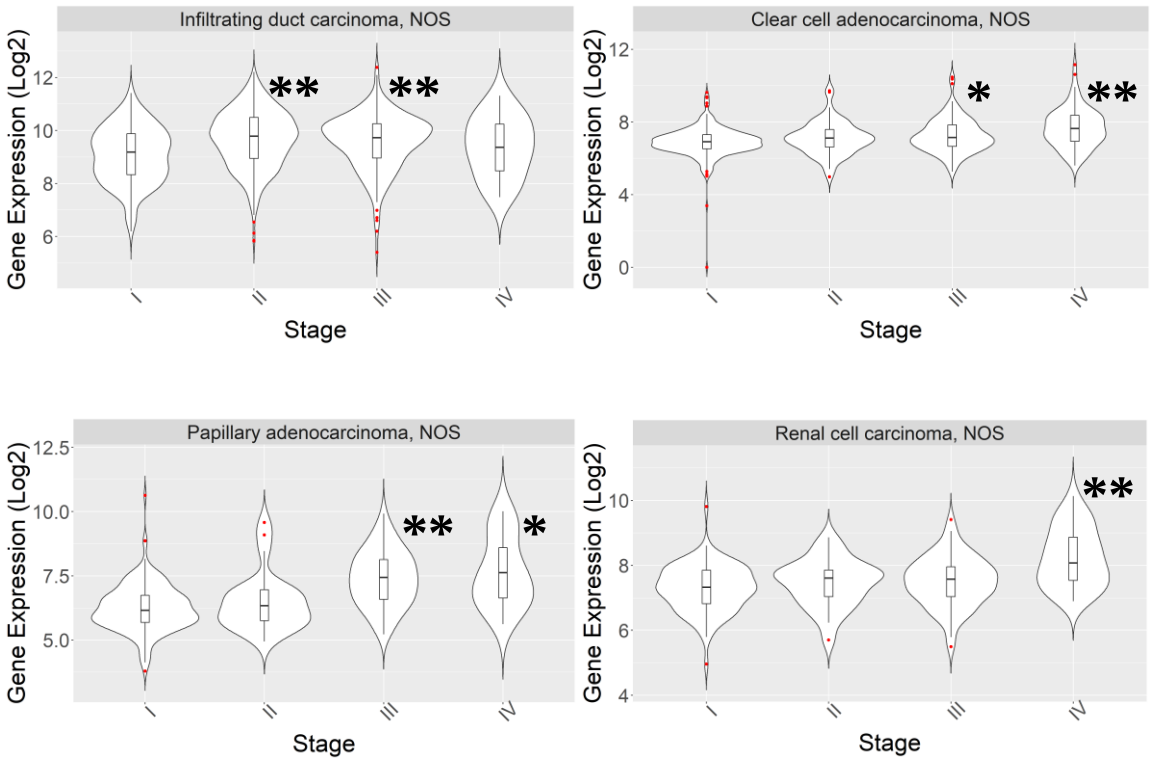

Figure 3-5 Stagewise Expression of NCAPH

This figure shows the expression changes of the gene NCAPH through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

3.5.3 KIF18A

The KIF18A gene (ENSG00000121621) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene product is KIF18A, a kinase that is strongly associated with a large variety of tumors. KIF18A has been associated with increased cell growth, migration, and invasion in liver cancer tissues such as hepatocellular carcinoma (Luo *et al*. 2018). It may promote metastatic behavior in liver cancer, and has also been proven to be involved in breast cancer carcinogenesis (Zhang *et al*., 2010).

KIF18A's association metastatic behaviors such as increased cell growth, migration, and invasion is shown in figure 3-6, where we observed visible upregulation in the stage IV of all kidney diagnoses. In particular, kidney papillary adenocarcinoma and renal cell carcinoma exhibited significant upregulation in stages III and IV. Breast infiltrating duct carcinoma displayed no significant regulation changes in the later stages of cancer, however.

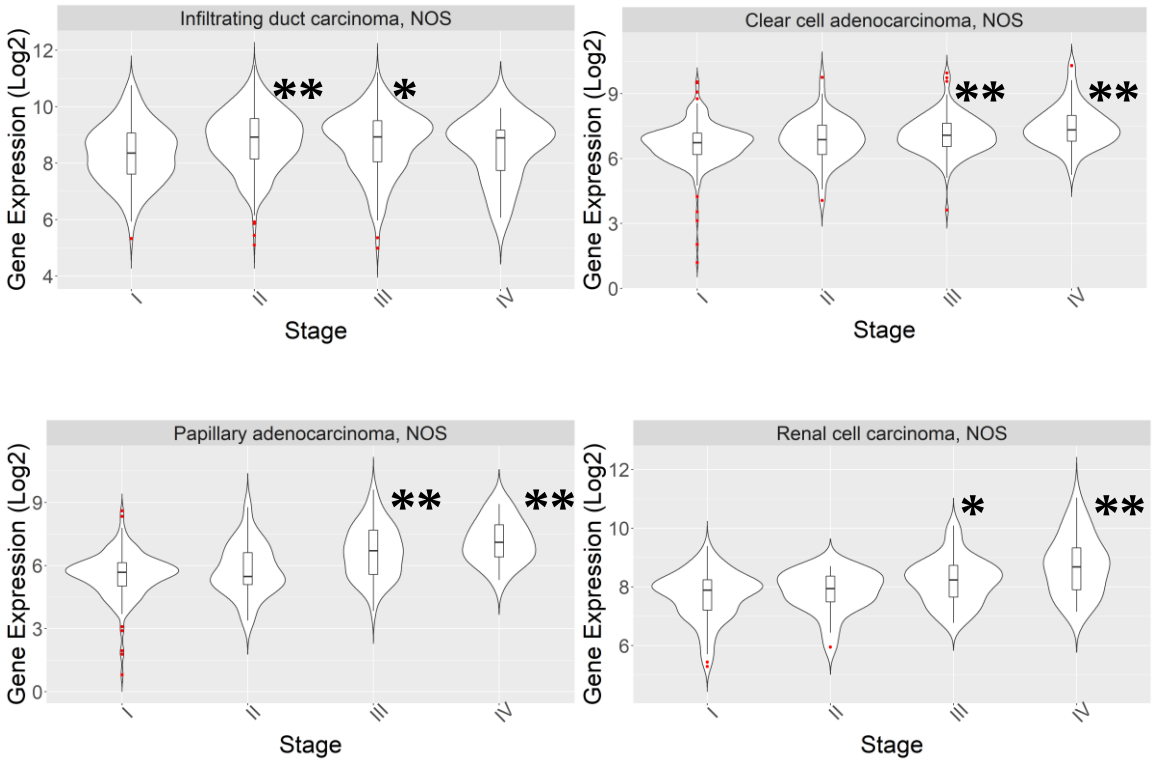

Figure 3-6 Stagewise Expression of KIF18A

This figure shows the expression changes of the gene KIF18A through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

3.5.4 CDCA5

The CDCA5 gene (ENSG00000146670) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene produces CDCA5, cell division cycle-associated 5. This protein has been associated with cell proliferation and the phosphorylation of the ERK signaling pathway, along with AKT. Inhibiting CDCA5 was found to similarly inhibit ERK, which then reduced cancer cell proliferation in both prostate cancer and hepatocellular carcinoma (Ji *et al*., 2021; Wang *et al*., 2018).

CDCA5's association with metastatic behaviors such as cell proliferation is shown in figure 3-7, where we observed visible upregulation in the stages III and IV of all kidney diagnoses. Breast infiltrating duct carcinoma also displayed graduation upregulation throughout stages of cancer.

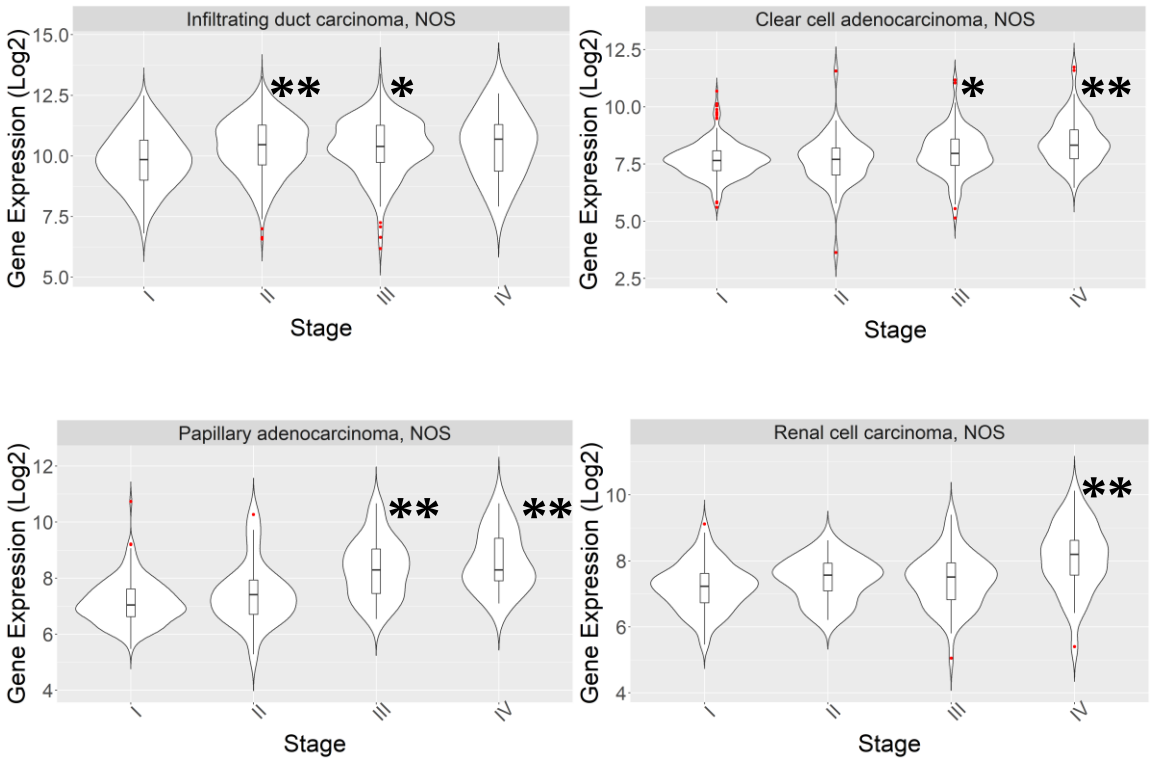

Figure 3-7 Stagewise Expression of CDCA5

This figure shows the expression changes of the gene CDCA5 through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

3.5.5 CCNB2

The CCNB2 gene (ENSG00000157456) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene product is cyclin B2 (CCNB2). Upregulation of CCNB2 was found to aid the process of lymphovascular invasion, a step of metastasis found in breast cancer. High expression of CCNB2 resulted metastasis and tumor aggressiveness (Aljohani *et al.*, 2022). CCNB2 was also shown to stimulate highly aggressive breast cancers known as triple-negative breast cancer, causing increased cell proliferation in an already highly aggressive and metastatic cancer diagnosis (Wu *et al.*, 2021).

CCNB2's association with metastasis and highly aggressive cancers is shown in figure 3-8, where we observed visible upregulation in stages III and IV of all kidney diagnoses. These findings support previous research linking CCNB2 to metastatic behavior in several tissue varieties. However, CCNB2 displayed no significant differential expression in stage IV breast infiltrating duct carcinoma, despite research proving its association with highly metastatic cancer diagnoses such as triple-negative breast cancer.

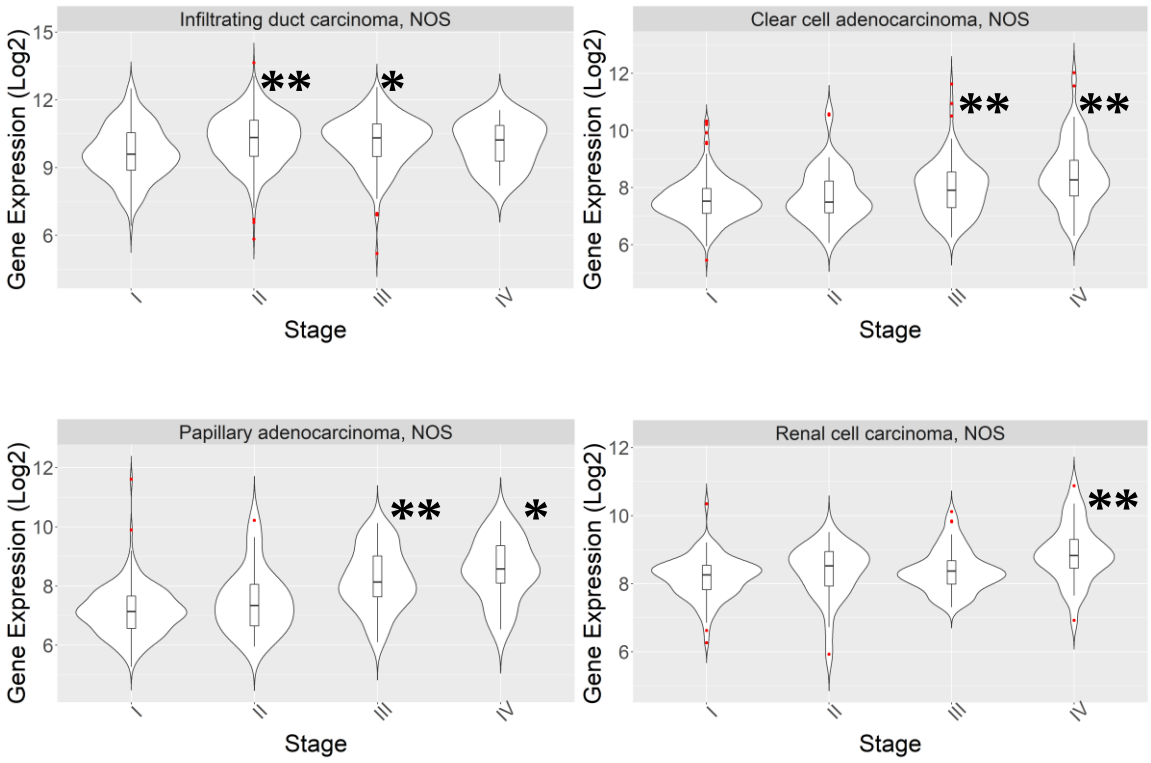

Figure 3-8 Stagewise Expression of CCNB2

This figure shows the expression changes of the gene CCNB2 through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

3.5.6 PBK

The PBK gene (ENSG00000168078) was differentially expressed via stage in intestinal adenocarcinoma, breast infiltrating duct carcinoma, clear cell adenocarcinoma of the kidney, and papillary adenocarcinoma of the kidney. The gene product is PDZ-binding kinase (PBK), which has been proven to promote metastasis in hepatocellular carcinoma (Yang *et al*., 2019). It has also been associated with poor prognosis and immune infiltration, where the tumor grows into surrounding, healthy tissue (Wen, *et al*., 2021).

PBK's association with metastasis is shown in figure 3-9, where we observed visible differential expression in stages III and IV of kidney papillary adenocarcinoma and clear cell adenocarcinoma. Kidney papillary adenocarcinoma shows slight upregulation as stage progresses, and kidney renal cell carcinoma shows significant upregulation in stages III and IV of cancer. We also observed slight upregulation across cancer stage in breast infiltrating duct carcinoma. Intestinal adenocarcinoma exhibited downregulation as cancer progressed. These findings support previous research linking PBK to metastatic behavior.

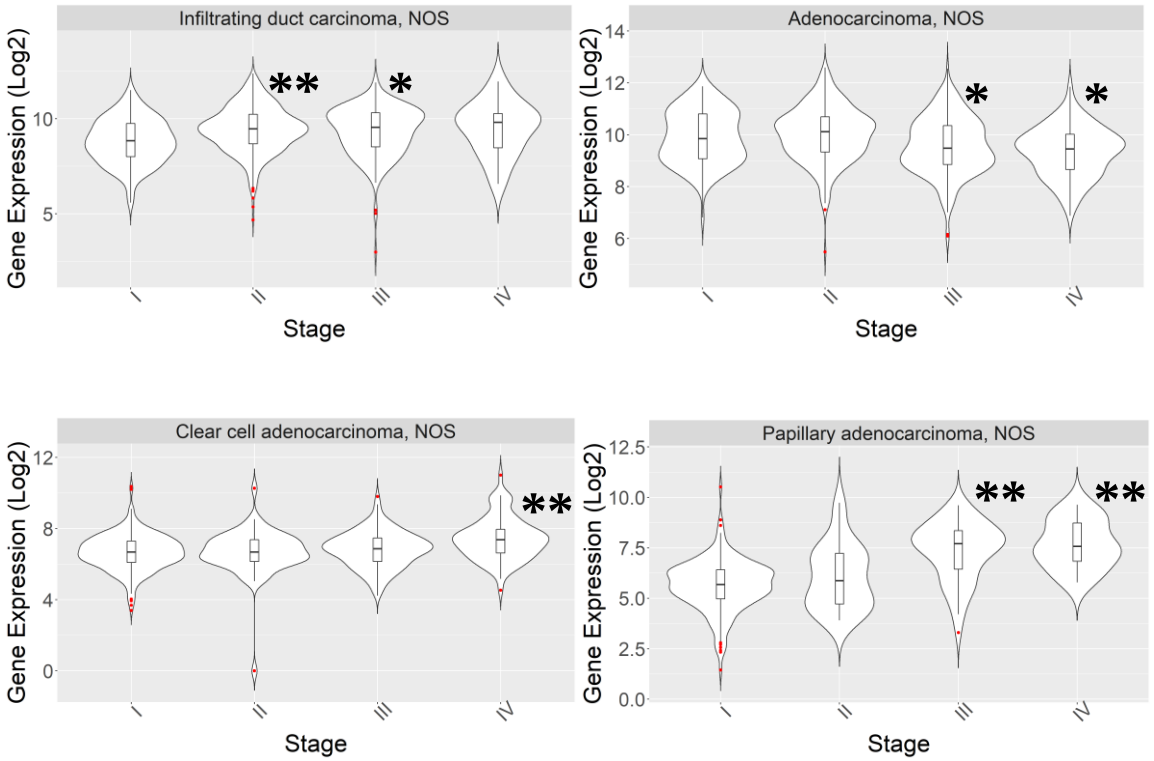

Figure 3-9 Stagewise Expression of PBK

This figure shows the expression changes of the gene PBK through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

3.5.7 KPNA2

The KPNA2 gene (ENSG00000182481) was differentially expressed via stage in all breast and kidney diagnoses tested in the analysis. The gene produces karyopherin a2 (KPNA2), which promotes cell proliferation and metastasis in kidney renal cell carcinoma, and may be a new target for colorectal cancer metastasis (Zheng *et al.*, 2021; Han *et al*., 2022). Similarly, KPNA2 has been shown to promote metastasis and growth in osteosarcomas as well (Zhou *et al*., 2023).

KPNA2's stage-wise expression is shown in figure 3-10, where we observed significant differential expression in stage IV of kidney renal cell carcinoma. We observed slight upregulation across the stages of cancer in kidney papillary adenocarcinoma. These findings support research connecting KPNA2 with metastasis and cell growth. However, little regulation change was observed in breast infiltrating duct carcinoma or clear cell adenocarcinoma of the kidney.

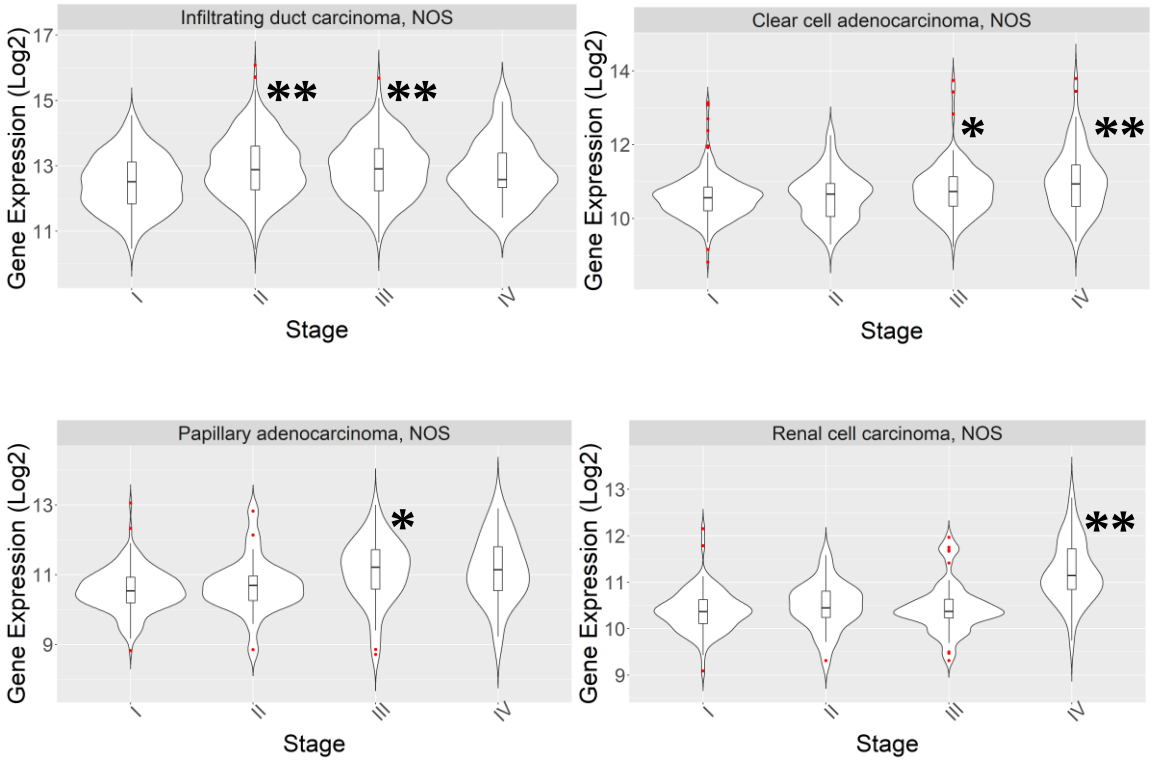

Figure 3-10 Stagewise Expression of KPNA2

This figure shows the expression changes of the gene KPNA2 through the stages of cancer. The width of each violin plot indicates the number of samples exhibiting that level of expression. Red dots indicate outliers.

* indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 0.01.

** indicates a T-Test comparing stage 1 to the labeled stage returned a P value of less than 10^{-4}.

4 Discussion

The goal of this study was to test if CT genes promoted or were associated with metastatic behavior. Only 50 unique CT genes with any differential expression via stage were found, out of the 481 found in total. This indicates 10% of the CT genes we isolated exhibited any stagewise differential expression, and the other 90% were not differentially expressed via stage.

We then conducted further analysis on the 50 unique CT genes with stage-wise differential expression. Within these 50 genes, we found 7 that were tissue nonspecific, and were differentially expressed in 4 of the 7 tested diagnoses. These 7 genes were TTK, NCAPH, KIF18A, CDCA5, CCNB2, PBK, and KPNA2. We found that these 7 genes were all linked to metastasis, invasion, or aggressive cell proliferation in previous studies. This indicates that our analysis succeeded in isolating genes associated with metastasis in stage IV cancer.

Both the tissue nonspecific and tissue specific genes isolated in this analysis could be of interest to future research focused on late-stage cancer. Treatments targeting genes associated with metastasis could result in lower mortality and morbidity rates in late-stage cancer patients. We hope that our results will help reveal ideal targets for future cancer research and therapies.

5 Reference List

1. Aljohani AI, Toss MS, El-Sharawy KA, Mirza S, Ball GR, Green AR, Rakha EA. (2022). Upregulation of Cyclin B2 (*CCNB2*) in breast cancer contributes to the development of lymphovascular invasion. Am J Cancer Res, 12(2), 469-489.
2. American Society of Clinical Oncology. (2021). Stages of Cancer. (https://www.cancer.net/navigating-cancer-care/diagnosing-cancer/stages-cancer).
3. Aouiche, C., Chen, B., & Shang, X. (2019). Predicting stage-specific cancer related genes and their dynamic modules by integrating multiple datasets. *BMC bioinformatics*, *20*(Suppl 7), 194.
4. Ayyoub, M., Scarlata, C., Hamaï, A., Pignon, P., Valmori, D. (2014) Expression of MAGE-A3/6 in primary breast cancer is associated with hormone receptor negative status, high histologic grade, and poor survival. J Immunother., 37(2), 73-76.
5. Chen, F., Wu, P., Hu, H., Tian, D., Jiang, N., & Wu, C. (2018). Protein kinase TTK promotes proliferation and migration and mediates epithelial-mesenchymal transition in human bladder cancer cells. *International journal of clinical and experimental pathology*, *11*(10), 4854–4861.
6. Chen, Y., Stockert, E., Chen, Y., Garin-Chesa, P., Rettig, W., van der Bruggen, P., Boon, T., and Old, L. (1994) Identification of the MAGE-1 gene product by monoclonal and polyclonal antibodies. Proc. Natl. Acad. Sci. U. S. A., 91, 1004-1008.

7. Chen, Y. T., Scanlan, M. J., Sahin, U., Tureci, O., Gure, A. O., Tsang, S., Williamson, B., Stockert, E., Pfreundschuh, M., and Old, L. J. (1997) A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proc. Natl. Acad. Sci. U. S. A., 94, 1914-1918.
8. Craig, A., Garcia-Lezana, T., Ruiz de Galarreta, M., Villacorta-Martin, C., Kozlova, G., et al. (2021) Transcriptomic characterization of cancer-testis antigens identifies MAGEA3 as a driver of tumor progression in hepatocellular carcinoma. PLOS Genetics, 17(6), e1009589.
9. da Silva, V., Fonseca, A., Fonseca, M., da Silva, T., Coelho, A., Kroll, J., de Souza, J., Stransky, B., de Souza, G., & de Souza, S. (2017) Genome-wide identification of cancer/testis genes and their association with prognosis in a pan-cancer analysis. Oncotarget, 8(54), 92966–92977.
10. De Smet, C., Lurquin, C., van der Bruggen, P., De Plaen, E., Brasseur, F., and Boon, T. (1994) Sequence and expression pattern of the human MAGE2 gene. Int. J. Immunogenet., 39, 121-129.
11. Dobrynin, P., Matyunina, E., Malov, S., & Kozlov, A. (2013) The novelty of human cancer/testis antigen encoding genes in evolution. Int. J. Genom., 2013, 105108.
12. Feichtinger, J., & McFarlane, R. J. (2019). Meiotic gene activation in somatic and germ cell tumours. *Andrology*, *7*(4), 415–427.
13. Fernald, K., and Kurokawa, M. (2013). Evading apoptosis in cancer. Trends. Cell Biol., 23(12), 620–633.

14. Gao YH, Qu SS, Cao LQ, Yao M. (2022) TTK predicts triple positive breast cancer prognosis and regulates tumor proliferation and invasion. Neoplasma., 69(2), 274-282.
15. Garnis, C., Buys, T.P. & Lam, W.L. (2004) Genetic alteration and gene expression modulation during cancer progression. *Mol. Cancer.*, 3, 9.
16. Gjerstorff, M. F., Andersen, M. H., & Ditzel, H. J. (2015). Oncogenic cancer/testis antigens: prime candidates for immunotherapy. *Oncotarget*, *6*(18), 15772–15787.
17. Grizzi, F., Franceschini, B., Hamrick, C. et al. (2007) Usefulness of cancer-testis antigens as biomarkers for the diagnosis and treatment of hepatocellular carcinoma. J. Transl. Med., 5, 3.
18. Grossman, Robert L., Heath, Allison P., Ferretti, Vincent, Varmus, Harold E., Lowy, Douglas R., Kibbe, Warren A., Staudt, Louis M. (2016) Toward a Shared Vision for Cancer Genomic Data. New England Journal of Medicine375:12, 1109-1112.
19. Guo, J., Grow, E., Mlcochova, H. et al. (2018) The adult human testis transcriptional cell atlas. Cell Res., 28, 1141–1157.
20. Han, F., Zhang, L., Liao, S. *et al.* (2022) The interaction between S100A2 and KPNA2 mediates NFYA nuclear import and is a novel therapeutic target for colorectal cancer metastasis. *Oncogene* 41, 657–670.
21. Ji, J., Shen, T., Li, Y., Liu, Y., Shang, Z., & Niu, Y. (2021). CDCA5 promotes the progression of prostate cancer by affecting the ERK signalling pathway. *Oncology reports*, *45*(3), 921–932.

22. Kaistha, B., Honstein, T., Müller, V. *et al.* (2014) Key role of dual specificity kinase *TTK* in proliferation and survival of pancreatic cancer cells. *Br J Cancer* 111, 1780–1787.

23. Lingg, L., Rottenberg, S., & Francica, P. (2022). Meiotic Genes and DNA Double Strand Break Repair in Cancer. *Frontiers in genetics*, *13*, 831620.

24. Love, M.I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12), 550.

25. Luo, W., Liao, M., Liao, Y., Chen, X., Huang, C., Fan, J., & Liao, W. (2018). The role of kinesin KIF18A in the invasion and metastasis of hepatocellular carcinoma. *World journal of surgical oncology*, *16*(1), 36.

26. Parmigiani, R. et al. (2006) Characterization of a cancer/testis (CT) antigen gene family capable of eliciting humoral response in cancer patients. Proc. Natl. Acad. Sci. U. S. A., 103(48), 18066-18071.

27. Pudova, E.A., Krasnov, G.S., Kobelyatskaya, A.A., Savvateeva, M.V., Fedorova, M.S., Pavlov, V.S., Nyushko, K.M., Kaprin, A.D., Alekseev, B.Y., Trofimov, D.Y., Sukhikh, G.T., Snezhkina, A.V., & Kudryavtseva, A.V. (2021) Gene Expression Changes and Associated Pathways Involved in the Profression of Prostate Cancer Advanced Stages. Frontiers in Genetics, 11, 1664-8021.

28. Qiu X, Gao Z, Shao J, Li H. (2020) NCAPH is upregulated in endometrial cancer and associated with poor clinicopathologic characteristics. Ann Hum Genet. 84(6), 437-446.

29. Roeder, C., Schuler-Thurner, B., Berchtold, S., Vieth, G., Driesch, Pv., Schuler, G., Lüftl, M. (2005) MAGE-A3 is a frequent tumor antigen of metastasized melanoma. Arch Dermatol Res., 296(7), 314-319.
30. Scanlan, M.J., Gure, A.O., Jungbluth, A.A., Old, L.J., Chen, Y.T. (2002) Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. Immunol Rev., 188, 22-32.
31. Shang, B., Gao, A., Pan, Y., Zhang, G., Tu, J., Zhou, Y., Yang, P., Cao, Z., Wei, Q., Ding, Y., Zhang, J., Zhao, Y., & Zhou, Q. (2014). CT45A1 acts as a new proto-oncogene to trigger tumorigenesis and cancer metastasis. *Cell death & disease*, *5*(6), e1285.
32. Soldatova, K., Kit, O., Tolmakh, R., Vladimirova, L., and Kutilin, D. (2019) Cancer-testis gene-expression features in various tumors. J. Clin. Oncol., 37(15), e14260.
33. Tan, K., Wang, K., Zhao, A. *et al.* (2023) Meiotic nuclear divisions 1 promotes proliferation and metastasis in hepatocellular carcinoma and is a potential diagnostic and therapeutic target gene. *Med Oncol* 40, 14.
34. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., Sanli, K., von Feilitzen, K., Oksvold, P., Lundberg, E., Hober, S., Nilsson, P., Mattsson, J., Schwenk, J.M., Brunnström, H., Glimelius, B., Sjöblom, T., Edqvist, P.H., Djureinovic, D., Micke, P., Lindskog, C., Mardinoglu, A., Ponten, F. (2017) A pathology atlas of the human cancer transcriptome. Science., 357(6352). Human Protein Atlas proteinatlas.org.

35. van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A., and Boon, T. (1991) A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science, 254, 1643-1647.
36. Wang, C., Gu, Y., Zhang, K. et al. (2016) Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. Nat. Commun., 7, 10499.
37. Wang, J., Xia, C., Pu, M., Dai, B., Yang, X., Shang, R. ... Dou, K. (2018). Silencing of CDCA5 inhibits cancer progression and serves as a prognostic biomarker for hepatocellular carcinoma. Oncology Reports, 40, 1875-1884.
38. Wen, H., Chen, Z., Li, M., Huang, Q., Deng, Y., Zheng, J., Xiong, M., Wang, P., & Zhang, W. (2021) An Integrative Pan-Cancer Analysis of PNK in Human Tumors. Front. Mol. Biosci, 8, 755911.
39. Wu, S., Su, R., & Jia, H. (2021) Cyclin B2 (CCNB2) Stimulates the Proliferation of Triple-Negative Breast Cancer (TNBC) cells *In Vitro* and *In Vivo*. Hindawi. 2021, 0278-0240.
40. Xu, J. et al. (2019) Systematic cancer-testis gene expression analysis identified CDCA5 as a potential therapeutic target in esophageal squamous cell carcinoma. EBioMedicine, 46, 54-65.
41. Yang, Q.X., Zhong, S., He, L., Jia, X.J., Tang, H., Cheng, S.T., Ren, J.H., Yu, H.B., Zhou, L., Zhou, H.Z., Ren, F., Hu, Z.W., Gong, R., Huang, A.L., Chen, J. (2019) PBK overexpression promotes metastasis of hepatocellular carcinoma via activating ETV4-uPAR signaling pathway. Cancer Lett. 452, 90-102.

42. Yin, L., Jiang, LP., Shen, QS. *et al.* (2017) NCAPH plays important roles in human colon cancer. *Cell Death Dis* 8, e2680.

43. Zhang, C., Zhu, C., Chen, H., Li, L., Guo, L., Jiang, W., Lu, S.H. (2010) Kif18A is involved in human breast carcinogenesis, Carcinogenesis, 31(9), 1676-1684.

44. Zhang, Y., & Zhang, Z. (2020). The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cellular & molecular immunology*, *17*(8), 807–821.

45. Zheng, S., Li, X., Deng, T., Liu, R., Bai, J., Zuo, T., Guo, Y., & Chen, J. (2021). KPNA2 promotes renal cell carcinoma proliferation and metastasis via NPM. *Journal of cellular and molecular medicine*, *25*(19), 9255–9267.

46. Zhou, J., Zhao, L. Q., Xiong, M. M., Wang, X. Q., Yang, G. R., Qiu, Z. L., Wu, M., & Liu, Z. H. (2003). Gene expression profiles at different stages of human esophageal squamous cell carcinoma. *World journal of gastroenterology*, *9*(1), 9–15.

47. Zhou, S., Zhuang, C., & Liu, R. (2023) KPNA2 promotes osteosarcoma growth and metastasis in a c-Myc-dependent manner via the hedgehog /GLT1 signaling pathway. Arabian Journal of Chemistry, 104805.

48. Zhu, P., and Fan, Z. (2018) Cancer stem cells and tumorigenesis. Biophys. Rep., 4(4), 178–188.