UNIVERSITY OF CALIFORNIA SAN DIEGO

RNA-Seq Analysis of Pancreatic Ductal Adenocarcinoma:

Epigenetic Adaptation in Response of Drug Treatment

A Thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

in

Bioengineering

by

Yuxin Shi

Committee in Charge:

Professor Yingxiao Wang, Chair
Professor Stephanie Fraley
Professor Sheng Zhong

SIGNATURE

The Thesis of Yuxin Shi as it is listed on UCSD Academic Records is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California San Diego

2018

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I would first like to send my sincere thanks to Dr. Kathy Shaoying Lu for her support and patient guidance for my research. For the passing year, she had shared her expertise and knowledge with huge generosity and had given me tremendous encouragement in every step throughout the process which made this thesis become possible.

I would like to give special thanks to the PI Prof. Peter Yingxiao Wang for his support and being the chair of my defense committee. Prof. Wang provided me a lot of valuable points of view and suggestions when I work on my research. For Prof. Fraley and Prof. Zhong, I would also like to thank them for being my committee members and making special arrangement to attend my defense. It would be such an honor to receive their precious insights on my research and I do appreciate their feedback on my work.

In addition, I would like to thank them for all lab members who had given me so much help with my research, especially Dr. Yan Huang and Praopim Limsakul. Without their passionate participation and input, the project could not have been successfully conducted.

Lastly, I would like to thank my family and friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

ABSTRACT OF THE THESIS

RNA-Seq Analysis of Pancreatic Ductal Adenocarcinoma:

Epigenetic Modulation and Gemcitabine Adaptation

by

Yuxin Shi

Master of Science in Bioengineering

University of California San Diego, 2018

Professor Yingxiao Wang, Chair

Chemotherapy is the dominant treatment approach to many cancers. For decades, Gemcitabine (GEM) has been used as the first-line therapy for pancreatic ductal adenocarcinoma (PDAC), one of the most fatal solid tumor, but its competence is disappointingly constrained by

intrinsic or adaptive resistance. The mechanisms beneath such resistance have been intensively studied with great efforts and its correlations with many genes and pathways have been found. However, our knowledge has not been unified to provide abundant information for major advances, since the whole landscape in which drug adaptation and gene expressions are associated is not yet clear. Moreover, little had we known about the initiation of adaptation, which makes it more difficult to define better therapies and achieve better clinical outcomes.

Recently, epigenetic alterations have become potential prognostic biomarkers and offered vast options for PDAC detection. Previously, we had observed different histone 3 lysine 9 tri-methylation (H3K9me3) patterns among GEM sensitive and resistant PDAC cells during the GEM treatment, suggesting the role of histone modification in GEM resistance initiation and exhibition. In this paper, we would provide our analysis workflow to identify the candidate genes that might be modulated by H3K9me3 upon GEM treatment, and predict how they dynamically affect GEM sensitivity. Our results would be considered as a general framework to construct the regulatory network from epigenetic drug response to transcriptomic plasticity, and discover new opportunities for therapeutic development.

INTRODUCTION

Pancreatic Ductal Adenocarcinoma: Biological and Genetic Features

Pancreatic cancer is one of the most lethal and hard-to-treat cancers. As reported in Cancer Statistics 2018 by American Cancer Society, pancreatic cancer has already become the third leading cause of cancer-related death in the US (Siegel, Miller, & Jemal, 2018). Strikingly, its survival for all stages combined is only 8%, the lowest among all cancer types. In contrast to the steady increase in survival observed for most cancer types, there has been little improvement in pancreatic cancers. Less than 10% of the newly diagnosed pancreatic tumor can be localized, making the disease even harder to receive a timely and targeted treatment.

Among all pancreatic cancer types, Pancreatic ductal adenocarcinoma (PDAC) is by far the most common and aggressive one, representing more than 80% of all cases (Adamska, Domenichini, & Falasca, 2017; Hessmann, Johnsen, Siveke, & Ellenrieder, 2017). The biological features of PDAC have been well defined and confirmed in a broad amount of studies, and some of them are described as follows (Ryan, Hong, & Bardeesy, 2014; Wood & Hruban, 2018):

1. The propensity for both local invasion and distant metastasis. It has been recognized that metastatic seeding from PDAC can occur at the initial stage of development, and by the end, more than half of diagnosed patients are presented with distant metastasis. However, no visible and distinctive symptoms can be observed at the early stage and conventional biomarkers to rely on has not been found.
2. Heterogeneity of tumor composition. PDAC is unique among solid tumors for the development of dense fibrotic stroma surrounding the tumor. Hence the tumor does not behave as a homogeneous population: besides the tumor cells, its adjacent stromal

cells, immune cells, and extracellular matrix have all been reported to play their role in tumorigenesis, and contribute to metabolic aberration, immune dysfunction, and influence tumor progression and invasion.

3. Tumor plasticity during progression. The majority of PDAC initiate from the same genetic lesion on KRAS oncogene. However, to complete the transformation into invasive carcinoma, other genetic mutations and microenvironment variation must cooperate, and the progressive accumulation of genetic and environmental remodeling allows PDAC to adapt to harsh conditions and increases proliferative ability.

Recent studies on the genomic and epigenetic markers of PDAC had given us more insight into the characteristics of PDAC. Like most solid malignancies, PDACs are driven by a group of genetic mutations("Genetics and Biology of Pancreatic Ductal Adenocarcinoma," 2016; Manji, Olive, Saenger, & Oberstein, 2017; Ryan et al., 2014). Large-scale sequencing on PDAC cells has reported that PDAC does not have a universal mutational landscape, and the mutation patterns can differ to a great extent among patients. For example, among the 184 cases of PDAC collected in The Cancer Genome Atlas (TCGA), the number of mutations observed in each patient ranges from 1 to more than 20000, and in total 13329 genes mutate and finally result in 32980 distinct mutations(Aguirre, Hruban, & Raphael, 2017).

However, despite such diversity, almost all PDACs share the same signature genetic lesions (Wood & Hruban, 2018). The most consistent mutations are occurring in four genes: KRAS, TP53, SMAD4, and CDKN2A. KRAS, as mentioned already, is an oncogene whose mutation affects more than 90% of PDAC (Ryan et al., 2014). Activation of KRAS gene stimulate a complicated cascade of tumorigenesis pathways and also play vital roles in aberrant

regulations of cell metabolism and autophagy. Loss of KRAS expression has been seen to result in massive cell death and arrested proliferation, leading to rapid tumor regression. The other three key genes, TP53, SMAD4, and CDKN2A, are all tumor suppressor genes, and loss of functions in those genes disrupt networks that normally restrain untoward growth, proliferation, survival, and invasion(Bailey et al., 2016; Borazanci et al., 2017; Deer et al., 2010).

Besides, some other gene mutations have been found in only a small fraction but shown a high correlation with certain phenotypes and clinical outcomes. For example, among the four subtypes of PDAC suggested by Bailey et al., Squamous PDAC tumors are significantly enriched for lysine demethylase 6A (KDM6A) and a poor prognosis, while the progenitor type has more TGFBR2 inactivating mutations and a relatively higher survival rate (Bailey et al., 2016).

In summary, diversity in PDAC genotypes and its ability to adapt, invade and migrate, are believed to give rise to its fatality. What's more, complexity in improving current therapies, especially overcome chemotherapy resistance, might also stem from the heterogeneity and plasticity of PDAC. Even a novel drug appears to deliver extraordinary outcomes during clinical trials, it might fail to bring the same benefits when applied to a larger population and for a longer time. However, with a profound understanding of its nature, we will be able to improve today's approaches and create effective personalized therapies designed around each tumor's unique qualities.

Gemcitabine: Strengths and Weaknesses

Adjuvant treatments, including chemotherapy and radiotherapy, are important alternative therapies besides resection surgery for PDAC (Alvarellos et al., 2014; De Sousa Cavalcante & Monteiro, 2014). Although surgery is theoretically curative, only patients without a locally

advanced or metastatic tumor are considered for resection. Therefore, adjuvant treatments have been the mainstay practice and improvements of current approaches have always been in great need.

Gemcitabine (GEM) is the first FDA-approved and so far the most widely used chemotherapeutic agent for PDAC (Binenbaum, Na'Ara, & Gil, 2015). It is an analog of deoxycytidine that interferes with DNA synthesis by interrupting DNA elongation and causes irreversible DNA damage(Alvarellos et al., 2014). Other modes of actions of GEM include self-potentiation by blocking the nucleoside salvage pathway and apoptosis induction through pathways such as the MAPK pathway.

However, GEM and GEM-based combination therapies prolong life expectancy only moderately when comparing to other agents, and resistance is the major impediment for GEM therapies deliver satisfactory results (De Sousa Cavalcante & Monteiro, 2014; Fryer, Barlett, Galustian, & Dalgleish, 2011; M. P. Kim & Gallick, 2008). Such resistance can be extrinsic, intrinsic or acquired and is tightly associated with the characteristics of PDAC (Binenbaum et al., 2015; Liang et al., 2017; Samulitis et al., 2015).

The extrinsic resistance is primarily referred to the impaired GEM delivery (Binenbaum et al., 2015; Liang et al., 2017). As mentioned above, PDAC is characterized to grow within extensive fibrotic tissue and a hypovascular and poorly perfused ECM. Thus, compounds like GEM need to overcome this barrier first before accumulating in PDAC cells, and because of its activity on normal tissues, increasing the dose of GEM is not a practical solution.

The internal resistance can also be stroma-mediated, where the stroma stimulates pathways to counteract its toxicity (Liang et al., 2017; Sherman et al., 2017). Another evidence of intrinsic resistance is the diversity of drug response in PDAC subtypes. As reported in the

article by Collisson et al. in 2011, the log10(IC50) of GEM could span from -4 to 10 among 18 PDAC subtypes (Collisson et al., 2011). Plus, the response to GEM can be highly specific for phenotypes and different transcription profiles (Bailey et al., 2016; Dreyer, Chang, Bailey, & Biankin, 2017). The third but not the least mechanism is the cancer stem cell model, in which a subset of cells naturally displays the chemo-resistant phenotype (Den, Gremeaux, Topal, & Vankelecom, 2012; Fitzgerald & McCubrey, 2014; Lee, Dosch, & Simeone, 2008).

To make the clinical situation more complex, resistance can develop over time, even in patients who usually have a good initial response to GEM along or combined with other agents (De Sousa Cavalcante & Monteiro, 2014; Jia & Xie, 2015). Multiple mechanisms of GEM resistance have been identified for such adaptations. PDAC cells have observed to show plasticity in altering its gene expression profiles, and eventually either become more similar to the intrinsically drug-resistant cells or activate other pathways to compensate for the GEM-induced damage(Alvarellos et al., 2014; Biancur & Kimmelman, 2018; De Sousa Cavalcante & Monteiro, 2014; López-Casas & López-Fernández, 2010; Santa Pau, Real, & Valencia, 2014). For instance, epithelial-to-mesenchymal transition (EMT) towards stem-cell-like phenotypes have been proved to drive GEM resistance(Arumugam et al., 2009). Meanwhile, signaling pathways dominating cell growth, proliferation, differentiation, apoptosis, and invasion, migration and angiogenesis, such as the notable PIk3/Akt, EGFR, Hedgehog, NOTCH, MAPK, and NFκB pathways, appear to either directly or indirectly influence chemosensitivity (Jia & Xie, 2015; Lee et al., 2008; Shields, Dangi-Garimella, Krantz, Bentrem, & Munshi, 2011).

In all, overcoming GEM resistance remains challenging. The mechanisms of GEM resistance, although intensively studied in recent years still largely unclear due to the complexity of its origin and the tumor features. Targeting on only a single pathway/mechanism is unlikely to

bring major improvement in GEM sensitivity since the multiple pathways usually intertwine and synergize in drug resistance (Jia & Xie, 2015). Therefore, further studies are urgently needed to draw the network involving all players in GEM resistance and develop new strategies to target the most vital ones.

Histone Modification in PDAC: Its Roles and Therapeutic Potential

As discussed above, GEM resistance can originate from ineffective drug delivery, impaired GEM mechanisms in tumor cells, and adaptative responses to GEM therapies. Several cell pathways vital for these events have been identified, such as GEM transportation by hENTs and hCNTs, GEM activation and degradation by DCK and CDA, as well as nucleoside salvage pathways (Ciccolini, Serdjebi, Peters, & Giovannetti, 2016; Kroep et al., 2002; Nordh, Ansari, & Andersson, 2014; Saiki et al., 2012). Also, determinants of the tumor and stromal microenvironment are crucial for GEM resistance initiation. Whole genome sequencing on different subtypes of PDAC has already identified some candidates, such as Heat shock protein 27 (Mori-Iwamoto et al., 2008), ZEB1(Arumugam et al., 2009), MUC1(Shukla et al., 2017), and more genes that involved in correlated cell signaling, differentiation, catabolism, and migration.

What worth notice is that not all genes involved in these events are mutated; in fact, the majority of them are found differentially expressed among phenotypes without any genetic alternatives (Wood & Hruban, 2018). Meanwhile, dysregulation of key epigenetic factors and chromatin-modifying proteins, particularly those responsible for post-translational histone editing, are frequently found in human pancreatic cancer (Bailey et al., 2016; Manji et al., 2017). Therefore, epigenetic regulation was suggested to involve or even govern innate resistance and plasticity of the tumor.

So far, numerous studies have made effort to unveil the association between the epigenetic profile and phenotype of PDAC. One consensus is that changes in global levels of histone modifications are reliable predictors of clinical outcomes (Brusslan et al., 2015; Manuyakorn et al., 2010; Park et al., 2008; Wei et al., 2008). As an example, Wei et al. reported the correlation between low expression of H3K27me3 and shorter overall survival time, and the following studies have confirmed that reprogramming of H3K27me3 by mutant KDM6A can repress crucial tumor suppressor genes (Wei et al., 2008). A large amount of research then follows and emphasizes the role of chromatin structure dysregulation in tumorigenesis. McDonald and his colleagues reported global redistribution of H3K9 methylation is linked to both cell metabolic activities and migration. A recent review also highlights histone modification to directly deactivate tumor suppressor genes and give rise to abnormal cell division, stressed metabolism and metastasis (McDonald et al., 2017; McDonald, Wu, Timp, Doi, & Feinberg, 2011).

Since epigenetic, especially histone modifications have been validated to affect outcomes for PDAC, manipulating histone editors have offered exciting possibilities to novel therapies that overcome chemoresistance. In recent years histone acetylation inhibitors have shown its significant potential to use as a single agent or in combined usage to cover the shortage of existing chemotherapies (Lane & Chabner, 2009; Minucci & Pelicci, 2006). In addition, in early 2018 researchers applied selective histone methyltransferase inhibitors to PDAC cells to reduce the H3K9me3 level (Lu et al., 2018). They observed the silencing effect on the biomarkers of apoptosis was erased, and sequentially, GEM sensitivity of treated groups was significantly increased. Therefore, we expect epigenetic manipulation of PDAC transcriptome, both via histone methylation and acetylation, to be an effective approach to overcome GEM resistance.

In our previous studies, we utilized H3K9me3 FRET biosensor to visualize the global change of chromatin reconstruction in PDAC upon GEM treatment. As a result, we detected different patterns of H3K9me3 level in GEM sensitive and resistant clones along the treatment, and similar experiments had not been reported in existing publications. To our surprise, both cell types became brighter as observed by fluorescence microscopy during the first 24 hours of treatment, indicating a significant enrichment in H3K9 methylation and correspondingly a global trend of gene deactivation. Then gradually the methylation melted to a state even lower than initial ones in GEM resistance cells, while GEM sensitive ones showed similar decline but in a much smaller move. These results brought up some further questions:

- What are the dynamic epigenetic changes upon GEM treatment?
- What are the genes that have been modulated by such epigenetic change?
- How these modulations differ in GEM sensitive/resistant phenotypes?
- What is the correlation between gene modulations and GEM adaptation?

Since very little had been found in the literature that could elucidate these questions, we applied the same dose of GEM to sensitive and resistance PDAC cells for three days, and then used RNA-seq to evaluate the differential gene expression between two cell types and capture the dynamic response across the time course. In the following sections, we would analyze the transcriptome profiles with multiple bioinformatics strategies and interpreted the results together with previous findings. In the end, we would discuss the limitations and pitfalls of our experiment and make some suggestions on following experiments.

To sum up, although great advancements have been made to build up the relationship between histone modification, GEM resistance and clinical outcomes of PDAC, our knowledge is not yet unified. Therefore, the aim of this study is to test the hypothesis that abnormal methylation of H3K9 on a group of mediators underlies the different dynamics of GEM response. We were hoping to answer if the global H3K9me3 level could indicate specific pathway activation, and how those events would possibly cause the resistance to initiate. In this paper, we would discuss the preliminary results of this project based on the first RNA-seq experiment and give suggestions for validations and future works.

RESULTS

GEM sensitive and resistant PDAC cells already exhibit different transcriptome profile

Before the experiment was conducted, GEM sensitive and resistant cells were cultured in GEM-free and 100 nM GEM culture media for sufficient period of time and proliferated at a similar speed, therefore they were expected to have a stable gene expression profile. However, though they had similar proliferation rate as measured through flow cytometry, we were unable to justify that whether GEM resistant cells were adapted to drug exposure by differentially expressing a special set of genes, or they were recovered from the treatment and regain a similar profile as GEM sensitive ones.

Therefore, comparing the RNA-Seq components extracted from samples before treatment would help us to understand if any changes had already appeared on RNA transcript expression patterns after long time GEM exposure. Here the RNA-Seq results from the two pre-treatment samples were compared and genes with an absolute log2 fold change greater than 1.5 were collected as differentially expressed. In total 327 genes were detected to be significantly upregulated and 373 genes downregulated in GEM resistant cells with respect to GEM sensitive ones.

To obtain further insight into the function of these genes, Gene Ontology (GO) enrichment analysis was conducted on these up and down-regulated genes respectively. We uploaded these two gene sets to Metascape and analysis was done based on GO biological processes (Tripathi et al., 2015). According to the results, transcripts with higher expression in the GEM resistant cells were enriched for glutathione metabolism, lysosome and purine metabolism, whereas GEM sensitive cells were overrepresented for Hippo signaling pathway, ECM-receptor interaction and TGF-beta signaling pathway.

GEM sensitive and resistant PDAC cells responded differently to GEM treatment

As discussed in previous sections, intrinsic and acquired resistance to GEM has been attributed to not only aberrant transport of GEM into PDAC and activation/inactivation of GEM but also other biological pathways that compensate the GEM-induced DNA damage and/or suppress consequential apoptosis. Therefore, special attention was paid to these pathways when we were exploring the dynamic changes of the transcriptome in GEM sensitive and resistant cells.

Before diving into the patterns of certain genes, the similarity between the six samples was demonstrated via hierarchical clustering and principle component analysis(PCA), and the graphs are shown in Figure 1. Since there were no replicates for samples, the statistical test was not applicable and the results therefore would only be described but not interpreted. From the graph, it would be obvious that two cell lines were initially similar. However, at 48 hours, the two cell lines differed dramatically, indicating a sharp and significant shift of gene expression profile in both cell types. What was also remarkable is that at 72 hours, the two cell lines went closer to their initial states, and resistant cells were more similar to their starting status than sensitive cells. The graph of PCA pointed out that at 48 hours two cell lines altered their transcriptome in opposite directions, when GEM sensitive cells responded more vigorously. Similarly, the largest change also occurred at S_48 for H3K9me3 methylation (Figure 2), where both H3K9me3 methyltransferases and demethylases become less than the initial condition in GEM sensitive samples. Therefore, genes that most differentially expressed at 48 hours might be regulated by H3K9me3, however we could not know which genes were regulated respectively in the two cell lines. Thus, we would expect gene expression analysis to provide more detail on the differential response, especially for samples at 48 hours.

Differences in GEM transport

The main transporters that mediate most of GEM uptake are Solute Carrier Family(SLC) 29A1, 28A1 and 28A3, which were aliases of hENT1 and hCNTs mentioned in our Introduction (Alvarellos et al., 2014). To depict the dynamics of GEM uptake during the treatment, we took a closer look at the transcription results of these transporters. The expression levels of SLC28A1 and SLC28A3 were negligible compared to the predominant SLC29A1. Therefore, the mRNA reads of SLC29A1 was used as a representative to evaluate the transport level of GEM. As shown in Figure 3, the transporter SLC29A before treatment had almost the same level in GEM sensitive and resistant cells. What stood out from the plot is that following the application of GEM, the number of transporters both decreased in two cell lines at 48 hours, then rose again at 72hr but not fully returned to the initial level. This could be somehow counterintuitive since previous research gave opposite result that lower SLC29A1 level had been shown to associate with a poorer prognosis. But since the change was transient, the expression of SLC29 should be discussed in a dynamic but not stable context, and we would give our interpretation in the discussion section.

Differences in GEM activation/inactivation

Deoxycytidine kinase (DCK) and Cytidine deaminase (CDA) are two important enzymes are the gatekeepers of GEM metabolism. As illustrated in the graph adapted from Binenbaum et al., Phosphorylation of GEM by DCK is the first and rate-limiting step of GEM metabolism pathway (Binenbaum et al., 2015), and DCK is known to be the predominant enzyme in the salvage of dNTPs for DNA synthesis (Alvarellos et al., 2014). Due to its central role in gemcitabine metabolism, its deficiency is a major contributor to GEM resistance (Kroep et al., 2002; Saiki et al., 2012). Meanwhile, Cytidine deaminase (CDA) deaminating GEM is regarded

as the main mechanism for gemcitabine inactivation (Ciccolini et al., 2016; Kroep et al., 2002). Most of GEM in cytosol would eventually be degraded into less toxic form and excreted out of the cell.

Here the expression patterns of these two enzymes were illustrated in Figure 3. Within each cell type, DCK and CDA expression level went inversely, but the dynamics were distinct across cell types. In the sensitive cells, DCK was found to accumulate gradually, whereas CDA kept decreasing through the three-day treatment. At the same time, GEM resistant cells responded to the treatment with an even lower DCK level at 48 hours and a corresponding higher CDA level. Then at 72 hours, these two enzymes both had a slightly higher level than their pre-treatment state, indicating more active GEM processing was ongoing. The patterns of DCK and CDA were in good agreement with their complementary functions in GEM metabolism, and differences in these profiles suggested that sensitive and resistant cells might have utilized distinct strategies when flooded with GEM.

Differences in competitive GEM inhibition

Ribonucleotide reductase RRM1 and RRM2 are responsible for the maintenance of deoxyribonucleotide concentration in the cytosol. Increasing RRM expression can expand the cellular pool of original dNTPs that competing with GEM during DNA synthesis, and has been proved to be inversely correlated with GEM efficiency (Alvarellos et al., 2014; De Sousa Cavalcante & Monteiro, 2014; Jia & Xie, 2015). In our experiment, GEM resistant cells already exhibited a higher expression level of RRMs before treated; and stimulation of GEM resulted in a reduced RRM expression in both cell lines. After 72 hours, the total RRMs in GEM sensitive and resistant cells all reached their greatest expression level. Although the final observation was consistent with our knowledge that GEM could directly inhibit RRM1 covalently and self-

enhanced (De Sousa Cavalcante & Monteiro, 2014), the temporary decline at 48 hours needs more careful interpretation.

Gene coexpression module

To illustrate the functional landscape of the transcriptome in GEM sensitive and resistant PDAC, WGCNA algorithms were applied to reconstruct the transcriptional network (Langfelder & Horvath, 2008). The reads of genes were normalized based on the library size of each sample and log-transformed to stabilize the variance. We selected the 5000 genes with the largest variance across samples as input to cluster genes by their expression pattern and build up the correlation matrix. During the analysis, both the correlations and their shared relationships across the whole gene set were considered and the score of Topological overlap measure (TOM) was calculated for each gene pair as a comprehensive measurement of similarity. Based on this score, hierarchical clustering identified 14 different modules from the selected set of genes. Genes in each module were regarded to express with high correlation and potentially participated in shared biological processes. The size and unique color label of each module are listed in Table.1.

To summarize the expression pattern of genes in each module, the Module Eigengenes (MEs), was calculated through principal component analysis. Next, correlations between MEs and each sample were derived and visualized as the matrix shown in Figure 5. In this case, if one module exhibits a great positive/negative correlation with one of the samples, genes in this module would be considered to be significantly up/down-regulated in this sample with respect to others. From the graph, we found some strong correlated module-sample pairs. For example, the turquoise and blue modules, although in opposite direction, were both highly associated with the sample of sensitive cells at 48 hours. Similarly, genes in the green module might be expected to

constantly activate in GEM sensitive cells, since this module is always positively correlated with sensitive samples but negatively with resistant ones.

Expression of the turquoise module associated with histone modification

To investigate the pattern of histone modification and its correlation with other pathways, the GO and KEGG enrichment analysis of individual module had been conducted as mentioned. Combined the results with the correlation matrix, we were able to describe the specific functional modifications occurred in each sample.

Surprisingly, we found that in the turquoise module, one of the most significant biological GO terms was "GO:0016569 covalent chromatin modification" (log transformed p-value = -8.808). Related to this annotation, we found that 11 enzymes that directly acetylate or methylate histones, especially H3K9 specific histone demethylase KDM4A, 4B and 4C (Greer & Shi, 2012). As shown in Figure 6, KEGG enrichment analysis on biological pathways revealed that genes in the turquoise module were also enriched in important cell regulatory pathways, such as mTOR, MAPK, PIk3, NOTCH and Hedgehog signaling pathways, which had been proved to be crucial for GEM resistance (Binenbaum et al., 2015). Also, there were other pathways, for example endocytosis, that were significantly enriched but less mentioned in existing literatures.

As indicated previously, gene expression profiles of GEM sensitive and resistant samples were most distinct at 48 hours. Besides, at 48 hours the global H3K9me3 intensity was greater in GEM sensitive cells from fluorescence imaging, and H3K9 methylation is believed as a marker of gene silencing. In addition, H3K9 demethylases and other tumor-related pathways were significantly down-regulated at the same time based on module-sample correlation and function annotation analysis. Putting all these facts together, we might propose that difference of

transcriptomic profiles in sensitive and resistant PDAC cells were closely related to H3K9 methylation, specifically via the epigenetic down-regulation of genes in those key pathways.

Identification of hub genes

To identify which genes might be directly modulated by H3K9 methylation and involved in GEM resistance, the gene set in the turquoise module was further analyzed by constructing the protein-protein interaction network (Figure 5). The complete set of 1445 genes in the turquoise module were uploaded to the STRING database, which is an extensive database of previously discovered protein interactions(Szklarczyk et al., 2017). The whole network was partitioned into clusters based on the strength of protein-protein connectivity, and the GO and KEGG enrichment analysis revealed the most significant function of each cluster. As a result, we identified 34 functional groups, and we found that genes related to histone modification were clustered together. Therefore, we selected the clusters whose functions were the mostly enriched ones in the whole network, and the cluster of endocytosis was chosen for the following analysis.

To find the hub genes, we paid special attention to the core genes that who forms the densest interaction with its neighbors, because genes with high local network density are leading the most highly interacted gene groups and have been observed to overpower these complexes. Thus, we picked genes with the highest local network density, and those are PIK3C2A and RAB5B in this cluster. Since PIK3C2A is not involved in endocytosis, RAB5B was chosen for analysis for further validation (Figure 7).

Evidence of RAB5B expression with PDAC

To find if the RAB5B expression has any effect on PDAC, we retrieved the PDAC mRNA sequencing data from TCGA and analyzed RAB5B expression using GEPIA(Tang et al., 2017). RAB5B has shown to be differentially expressed in PDAC and normal pancreatic tissue,

indicating that it might either function as an oncogene or might serve as a potential biomarker of PDAC. Moreover, survival analysis based on the same set of patient data has revealed a considerable correlation between high expression level of RAB5B, low overall survival rate (P-value = 0.11).

RAB5B appeared in the turquoise module and therefore had been shown to share expression pattern across samples with H3K9 demethylase KDM 4A,4B and 4C. Similarly, such correlations between RAB5B and common H3K9 demethylases were further validated using patient data from TCGA. All four H3K9-specific KDMs was shown to have to hold a tightly correlated expression with RAB5B, suggesting a regulatory influence of H3K9 methylation on RAB5B.

Thanks to the ChIP-Seq results contributed by Duaferia et al., the H3K9 methylation level on RAB5B on various PDAC subtypes and normal pancreatic tissue could be extracted and visualized in Figure 8 (Diaferia et al., 2016). The comparison showed that the coding region of RAB5B was far less H3K9 methylated in tumor samples and suggested that RAB5B might be directly activated via the epigenetic regulation mediated by H3K9me3 in PDAC.

To prove the RAB5B play a role in GEM resistance, additional downstream experiments would be necessary. For example, RAB5B could be knocked down or inhibited in GEM resistant cells or stimulated RAB5B expression in GEM sensitive cells, and its effects on GEM response could then be assessed to verify the correlation. Due to the limitation of time, the validations would be reserved for future work; however, some previous studies had provided some support to this hypothesis and would be discussed in the following section.

DISCUSSION

Dynamic changes of GEM transport and metabolism

Multiple independent studies, as well as several meta-analyses, had shown strong evidence that improved survival rates in PDAC treated with GEM are positively associated with SLC29A1. For instance, in a systematic review published in 2014, data from 855 patients in 10 individual studies were analyzed and 9 of 10 studies had shown a statistically significant longer overall survival with high SLC29A1-expression (Nordh et al., 2014). Notably, the genetic variants of SLC29A1 have not been observed to have a consistent association with GEM resistance, thus regulation on the translation level would be the possible mechanism for SLC29A1 to cause different clinical outcomes.

Similarly, High level of DCK has also been proved to be a reliable predictor of longer survival times in PDAC patients treated with GEM (Alvarellos et al., 2014; Binenbaum et al., 2015; De Sousa Cavalcante & Monteiro, 2014). Due to the central role of dCK in gemcitabine metabolism, deficiency of DCK had been validated to be a major contributor to both intrinsic and acquired GEM resistance. Concurrently, other studies had have shown that lower CDA activity also tends to be synchronized with faster response and better survival (Ciccolini et al., 2016; De Sousa Cavalcante & Monteiro, 2014; Jia & Xie, 2015).

While GEM directly inhibits the RRM1 and prevent deoxyribonucleoside to be recycled, unregulated RRM1 can boost normal DNA synthesis and promote DNA repair (Binenbaum et al., 2015; De Sousa Cavalcante & Monteiro, 2014; Jia & Xie, 2015; Santa Pau et al., 2014). This mechanism was shown to yield GEM resistance and RRM1 levels were inversely correlated with patient survival. On the top of that, in 2015 Tokunaga et al. used siRNA to suppress RRM1 gene

and successfully sensitized cancer cells to GEM (Tokunaga et al., 2015), which further proved the essential role of RRM1 in GEM compensation.

In our experiment, we extracted mRNA and built up the library for three time points, which allowed us to depict the trajectories of gene expression. Our results showed that SLC28A1, SLC28A2 and SLC28A3 were rarely expressed and SLC29A1 dominated in all six samples. The initial expression level of SLC29A1 was nearly identical in GEM sensitive and resistant cells, suggesting that GEM was transported equally into GEM sensitive and resistant cell lines before a higher dose were applied. This result might not lead to a solid conclusion due to the lack of replicates. In addition, as concluded in reviews by Alvarellos et al in 2014 and Adamska et al. in 2017, even though SLC29A1 has been proven to be a predictive marker for gemcitabine resistance and patient outcome, the correlation between SLC29A1 expression and intrinsic and acquired GEM resistance remains under debates and shows dependency on experimental context(Adamska et al., 2017; Alvarellos et al., 2014).

After 48 hours of intense GEM exposure, the sensitive cells shut down more influx of GEM compared to resistant ones. In this case, the effective concentration of GEM would be expected to reduce in GEM sensitive samples and so as the induced damages. At 72 hours, the level bounced back to a lower level compared to pre-treatment in both cell lines. One possible mechanism for such dynamics of SLC29A1 could be: 1) facing a huge amount of GEM, the tumor cells need to maintain the nucleoside uptake via reducing transporter expression 2) once adapted, the cells would recover the loss of SLC29A1 to support the transportation of other necessary nucleosides, and other pathways would be turned on for GEM processing and 3) resistant cells have become less vulnerable to greater chemical stimulus upon prolonged exposure, as pathways to compensate GEM-induced damages had already been activated.

The dynamic change of DCK and CDA status could also fit into this model. GEM sensitive cells kept enhancing DCK to deal with the vigorous flow of GEM, but they were not able to fully activate CDA for GEM degradation. On the contrary, GEM resistant cells shunt the drug to its deactivation path with a higher expression of CDA, and meanwhile turn down DCK to suppress the drug flow towards DNA synthesis. These results were in accordance with previous observations as well as clinical trials on prolonged fixed-dose treatment of GEM. In the study done by Grimison and his coworkers in 2007, they hypothesized an auto-induction of GEM metabolism: if the cells had experienced a long-time exposure to GEM, the threshold of GEM to saturate GEM phosphorylation by DCK would increase when reapplying the drug. Their results were supported by other studies which had reported adaptive DCK activation after GEM application (Grimison et al., 2007). Furthermore, CDA activity has been proved to directly correlate with the rate of GEM clearance (Gusella et al., 2011).

The patterns of RRMs agreed with this mechanism as well. Initially, GEM resistant cells already exhibited more RRMs to compete over the low drug concentration, while sensitive ones were naive from such stress. When a higher dose of GEM was added, RRMs were blocked directly by the drugs at 48 hours, but later GEM resistant cells quickly replenished and even raised RRMs level to overcome GEM in the competition of DNA elongation.

However, our results could be partially inconsistent with earlier findings on the expression levels of SLCs, DCK, CDA and RRMs after GEM application. For example, Farrell and her colleagues treated HPAFII cells for 72 hours with 10 μM GEM and reported that both SLC29A family and RRM family remained unchanged, while CDA doubles its amount through GEM treatment. This result differs from our findings, that SLCs declined slightly comparing to initial state, and CDA reached its lowest level at 72 hours. What leads to contradictory results

might be the limitation of samples in both experiments. In their experiment, only two time points were measured and the dynamics of cell response was not taken into consideration. Meanwhile, they failed to include the drug-tolerant phenotype in comparison, so no information was provided on the response of tumor with resistance. On the other hand, since we could not justify that the cell status had become stable and would maintain after 72 hours, it is also hard to deny the probability for CDA to boost in survived cells after a long time of GEM treatment. Also, the other research that we had reviewed paid little attention to the impact of GEM at the very beginning of treatment. Therefore, further tests conducted with larger sample size, longer time course and sufficient replicates would be favorable for us to validate our explanation on the dynamics of GEM processing.

With respect to the regulatory effect of H3K9me3 on GEM metabolisms, the ChIP-Seq results from Duaferia and colleagues had revealed distinct H3K9me3 patterns in PDAC compared to normal pancreatic tissues, which is shown in Figure 4. The differences were constant across all subtypes of PDAC, that SLC29A1, DCK, CDA and RRM1 might all be stimulated as a result of decreased H3K9 methylation. However, since the data could not address the dynamic aspect of cell response, it only gave us a hint that H3K9 methylation could be the direct cause for drug response, and be responsible for the GEM adaptation and tumor plasticity. Now we are in the process of investigating the results from our ChIP-Seq experiment using the same settings, and we hope the results could show evidence to solidify the causality of H3K9 methylation and differences between PDAC in their dynamic response to GEM therapy.

Histone modification potentially regulated pathways relevant to GEM resistance

WGCNA was performed to identify gene co-expression modules related with each sample and illustrated the differences of cell response to GEM in sensitive and resistant PDAC

cells. The gene clustering dendrogram and the module sample correlation matrix are shown in Figure 5. Since the relationship between epigenetic modification and drug resistance was of interest, we focused on the turquoise module, in which histone modification enzymes were enriched. GO and KEGG analysis on the turquoise module also showed that multiple pathways, that known to have strong correlations with GEM resistance, were found to share similar expression patterns with these histone modifiers, especially H3K9 demethylases.

Combined with our existing knowledge, we hypothesize a regulatory axis to exist along H3K9me3, hub genes in these pathways, and GEM resistance, as we observed the former two were varying in the similar pattern and might have a regulatory relationship, meanwhile the correlation of the latter ones could be summarized from previous studies. To find a potent gene that involved, the protein-protein network of the turquoise module was constructed and partitioned into functional clusters. As an example, we took a look at the cluster whose function is endocytosis, and RAB5B, which has the highest local network density inside the cluster, was chosen as a hub node for the following discussion.

RAB5B as a potential player in GEM resistance

RAB5B is a gene that involved in vesicular traffic. Analysis of PDAC dataset derived significantly overexpression of RAB5B in PDAC and such overexpression is considerably correlated with poor survival rate. Moreover, the RAB5B expression is positively correlated with multiple H3K9 demethylases and public ChIP-Seq data shows RAB5B is silenced directly by H3K9me3 in normal pancreatic tissue but not in PDAC. The relationship between RAB5B expression and GEM resistant phenotype request further validation, however, some existing evidence might encourage this hypothesis.

A recent study by Igarashi and his colleagues concluded that RAB5 expression is correlated to EMT in various cancers (Igarashi et al., 2017). They confirmed that high RAB5 expression was correlated with low E-cadherin level, a marker for EMT (Arumugam et al., 2009), and by suppressing RAB5 in PDAC cells E-cadherin expression was enhanced, leading to impeded proliferation, invasion, and migration of tumor cells. On the other hand, GEM resistant cells were observed to have a decreased E-cadherin expression and the tendency of EMT in multiple studies on PDAC (Arumugam et al., 2009; Samulitis et al., 2015). Our RNA-Seq results showed E-cadherin was almost eliminated in GEM resist cells at the starting stage and only accumulated slightly. In contrast, it was always massive in GEM sensitive cells. However, its concentration reduced at 48 hours but then salvaged at 72 hours, and this pattern was in accordance with that of KDMs and RAB5B but not inversely correlated. Therefore, further justification would be necessary to validate our prediction.

Another potent relationship might lie in the interaction between RAB5B and other genes in its cluster. The cluster was found to be mainly enriched in vesicle-mediated transcript pathways, and endocytosis and exosome secretion have become a promising target to overcome GEM resistance (Richards et al., 2017). At the same time, the other core genes, such as EGFR (Hessmann et al., 2017; Holohan, Van Schaeybroeck, Longley, & Johnston, 2013) and PIK3C2A (Binenbaum et al., 2015; Borazanci et al., 2017; Holohan et al., 2013; Hu & Zhang, 2016), have been proved to have a strong correlation with GEM sensitivity, therefore it would be reasonable to expect RAB5B to incorporate with them and play its role in the tumor plasticity and GEM adaptation.

Justification, Limitations, and Alternations of Current Workflow

In this study, our goal of RNA-Seq analysis is to identify genes or pathways whose expressions are modulated by H3K9 and are closely correlated to GEM resistance. We observed an apparent dissimilarity between the responses of GEM sensitive and resistant cells through sample clustering, thus we expected a group of genes that expressed significantly different between cell types. To identify those genes, we applied both standard statistical approach using DESeq2 and consensus network analysis based on WGCNA to select genes that were altered differently in such distinct drug response.

The statistical approach would not give convincing results due to the lack of replicates. This approach assumes a specific model to describe the underlying distribution of count data, in DESeq2 the negative binomial distribution, so that genes can be selected if their expressions between conditions are significantly different based on a test statistic predicted by the model. However, since we only have one sample for each test condition, the statistical test could not estimate the variability of each gene and provide proper inference on differential gene expression. Therefore, we used this method only for the purpose to explore the data.

The other method for gene screening is WGCNA, which is a systematic method to group genes with highly correlated expression into modules and relate them to external traits. In our experiment, since there was a huge difference between the two samples at 48 hours, we were expecting to build up the module – sample correlation matrix, and identify candidate genes whose patterns had the greatest correlation with these two samples. This method is more suitable for our data than the statistical-based approach. Although it also requires biological replicates to build up the network, it does not require an arbitrary threshold for dividing expression data into significant and non-significant pools but uses the gene expression data in order to detect

coordinated changes in the same pathway. Moreover, it accounts for correlations between genes and all samples to give a broader impression on the gene expression landscape.

In the analysis part, the functional annotation and pathway analysis could also be conducted in other methods instead of the current simple enrichment analysis. For example, Gene Set Enrichment Analysis interprets the gene profiles into common biological functions at gene set level and robust to the noises from a single gene.

Although the results presented herein are very promising, this study contains three main limitations. The largest limitation of our experiment is the lack of replicates, and consequently, all the results could only be considered as interpretation but not as solid conclusion. Another major weakness of our experiment design was the missing control groups. To make our RNA-Seq results more convincing, RNA samples of S and R cells without treatment should also be included. What's more, since the expression profiles at 24 hours were not measured, it became risky to draw the expression curve for each gene during GEM treatment. Lastly, the analysis was not completed since the genes expressed along with H3K9me3 methyltransferases had not been analyzed.

To overcome these limitations, our following RNA-Seq experiment had already included three replicates for each condition and 24-hour into the time course. Together with our ChIP-Seq data, the upcoming results should be more valuable to address our experimental needs.

CONCLUSION

In this paper, we presented the results from our RNA-Seq analysis and revealed an epigenetic-transcriptomic regulatory axis that might closely related to GEM adaptation and PDAC plasticity. We used weighted gene co-expression analysis to construct a gene co-expression network, identify and validate network hub genes associated with both global H3K9me3 level and GEM resistance. RAB5B were discussed in detail as one of the hub genes and its implicit association with GEM adaptation and H3K9me3 modulation was supported using public patient data. However, a more comprehensive analysis is in demand to identify other hub genes and further studies are required to characterize the hub genes functionally for potential therapeutic targets.

MATERIALS AND METHODS

Differential Expression Analysis

In our experiment, each sample were categorized based on the cell type (sensitive or resistant, labeled in S or R) and GEM treatment duration (0hr/48hr/72hr), and then listed as independent samples without replicates. The raw sequencing data were uploaded to the Galaxy web platform (Afgan et al., 2016), and we used the public server at usegalaxy.org to analyze the data. The sequencing results were first aligned to reference genome retrieved from UCSC database using HISAT (D. Kim, Langmead, & Salzberg, 2015)and then annotated by featureCounts (Liao, Smyth, & Shi, 2014) to generate the expression matrix, and all the setting was as default. Next, the count matrix was filtered in R (Ihaka & Gentleman, 1996) by keeping entries without zero read, and the resulted matrix had 12839 entries in total. The expression matrix was then normalized based on the library size of each sample and then log2-transformed to stabilized the variance.

The R package hclust and pheatmap was used for sample clustering and heat maps. Euclidean distances between the transcriptomic profile of each sample pairs were calculated and visualized. Lower and higher correlations were depicted in blue and red respectively. Each row/column indicates one sample as labeled.

The R package DESeq2 was used to screen the DEGs between the sensitive and resistant samples before treatment (Love, Huber, & Anders, 2014). For each pairwise comparison, the fold change was calculated for all genes as a ratio of their expression level between the resistant sample and sensitive sample. Since no replicates was made for each condition, absolute log2 fold change greater than 1.5 were chosen as the only cut-off criterion to select genes to be further considered for our following analysis.

Gene Ontology and KEGG Pathway Analysis

Gene sets were uploaded to Metascape (Tripathi et al., 2015) for functional enrichment using ontology terms collected from Gene Ontology for biological processes and pathway annotations from KEGG Pathway (Ashburner et al., 2000; Kanehisa & Goto, 2000). Default parameters were used (Min Overlap:3, P-Value Cutoff: 0.01, Min Enrichment:1.5).

Weighted Gene Correlation Network Analysis

The scale-free co-expression network for the genes was defined using the R package WGCNA (Langfelder & Horvath, 2008). 5000 entries in the expression matrix with highest standard deviation across six samples were extracted for this analysis. Firstly, power value was screened out: the average connectivity degree of different modules with power values ranging from 1 to 30 was calculated and the smallest power value with the score above 0.9 was chosen. The Pearson's correlation matrices for all genes was by raising all values to this power value and transformed into a topological overlap matrix (TOM). Thus, network connectivity of a gene was defined as the sum of its adjacency with all other genes and the corresponding dissimilarity (1-TOM) was calculated. To classify genes into gene modules by hierarchical clustering according to the dissimilarity. The dendrogram was simplified by merging similar modules and every module was assigned to a specific color.

Next, Module Eigengenes (MEs) were derived as the major component in the principal components for each gene module, whose expression profile is a representative of all gene expression patterns within the given module. To identify the relevant module, module-sample associations were estimated using the correlation between the ME and each sample. Modules that highly correlated to the sample were picked for downstream analysis.

Gene Network Construction and Hub Gene Selection

To find hub genes that comprised highly interconnected nodes within the turquoise module, we imported the gene list of this cluster into the STRING database, which is an extensive database of previously discovered networks and screened for significant gene-gene interactions (http://www.string-db.org/). We chose the default confidence threshold of 0.4 to construct protein-protein interaction (PPI). The PPI network was further partitioned into clustered using MCODE algorithms and densely connected gene clusters were identified (Bader & Hogue, 2003). In each cluster, genes with the highest degree or the largest local network connection score were regarded as hub genes and selected for further analyses. Default parameters were used for MCODE (Node Score Cutoff: 0.02, K-core: 2, Max.Depth = 100).

Hub Gene Validation

The RNA-Seq dataset of PDAC from TCGA was accessed and analyzed for using GEPIA(Tang et al., 2017). Survival plot was based on overall survival rate using the median as the group cutoff. Boxplots on differentially expressed genes in PDAC and normal pancreatic tissue was generated using default cutoff. Gene expression correlations were calculated as the Pearson correlation coefficients.

Public ChIP-Seq Data Visualization

The raw ChIP-Seq data of PDAC was retrieved from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/gds/) under the Super-Series accession number GSE64560. Samples of H3K9me3 in seven PDAC subtypes (PT45, PANC-1, Mia Paca-2, HPAF-II, CFPAC-1, Capan-2 and Capan-1). Two sets of ChIP-Seq data of normal pancreatic tissue were obtained from ENCODE for comparison (Birney et al., 2007). All data was queried

through the Cistrome Data Browser database (Liu et al., 2011) and visualized in WashU Epigenome Browser (Zhou et al., 2013).

Table 1: Modules of WGCNA

Module Color	Number of Genes	Most Enriched GO Term
Purple	87	GO:0044283: small molecule biosynthetic process
Tan	83	GO:0044283: small molecule biosynthetic process
Brown	800	GO:0051960: regulation of nervous system development
Turquoise	1636	GO:0043087: regulation of GTPase activity
Salmon	55	GO:0006270: DNA replication initiation
Black	99	GO:0007162: negative regulation of cell adhesion
Greenyellow	87	GO:0032101: regulation of response to external stimulus
Pink	95	GO:0098609: cell-cell adhesion
Blue	1374	GO:0022613: ribonucleoprotein complex biogenesis
Green	115	GO:0043588: skin development
Magenta	91	GO:0097435: supramolecular fiber organization
Red	107	GO:0006260: DNA synthesis
Yellow	355	GO:0097190: apoptotic signaling pathway
Grey	16	N/A

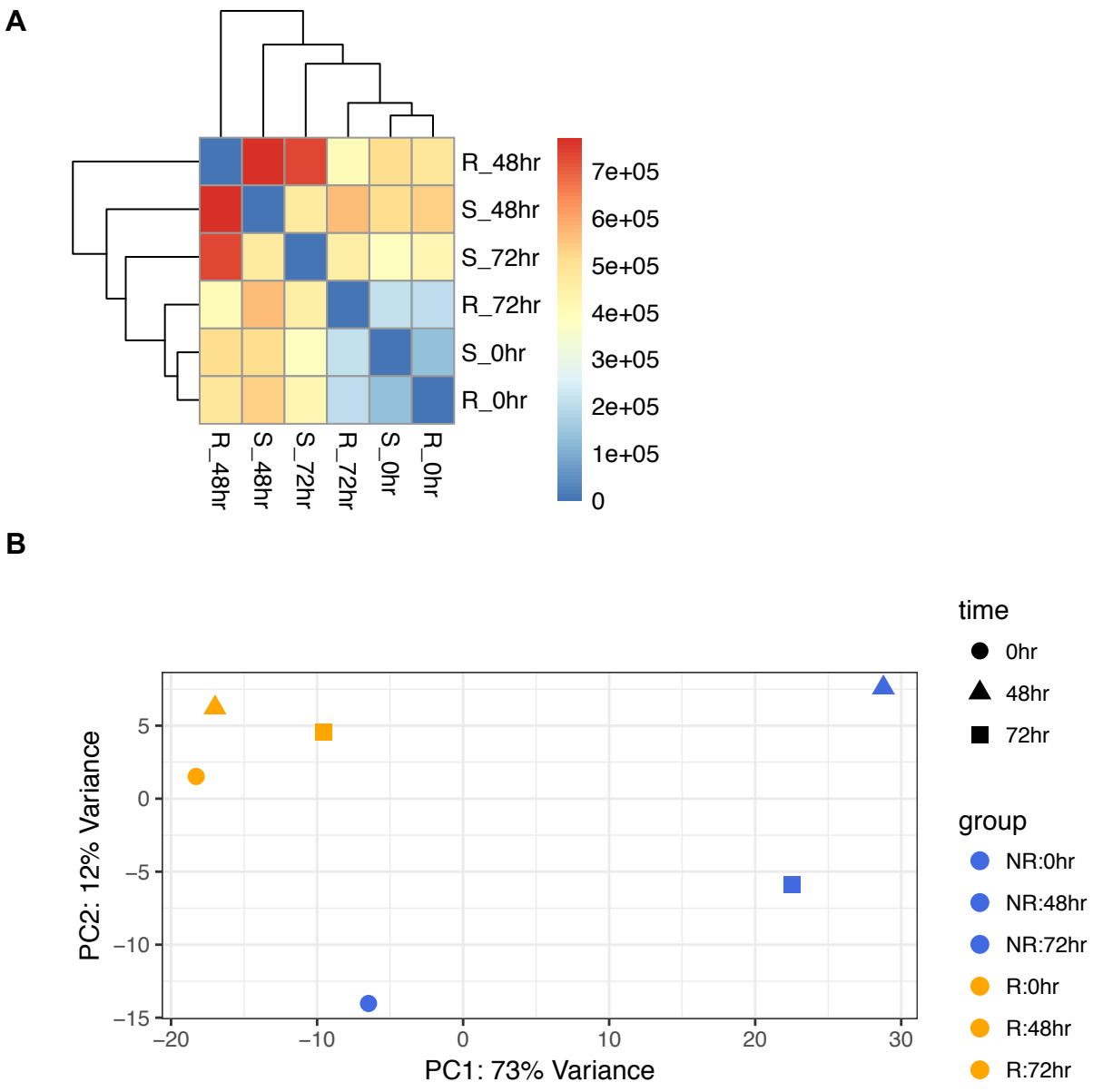


Figure 1: Sample clustering according to whole transcriptome profiles. (**A**) Hierachical clustering. (**B**) Principal component analysis

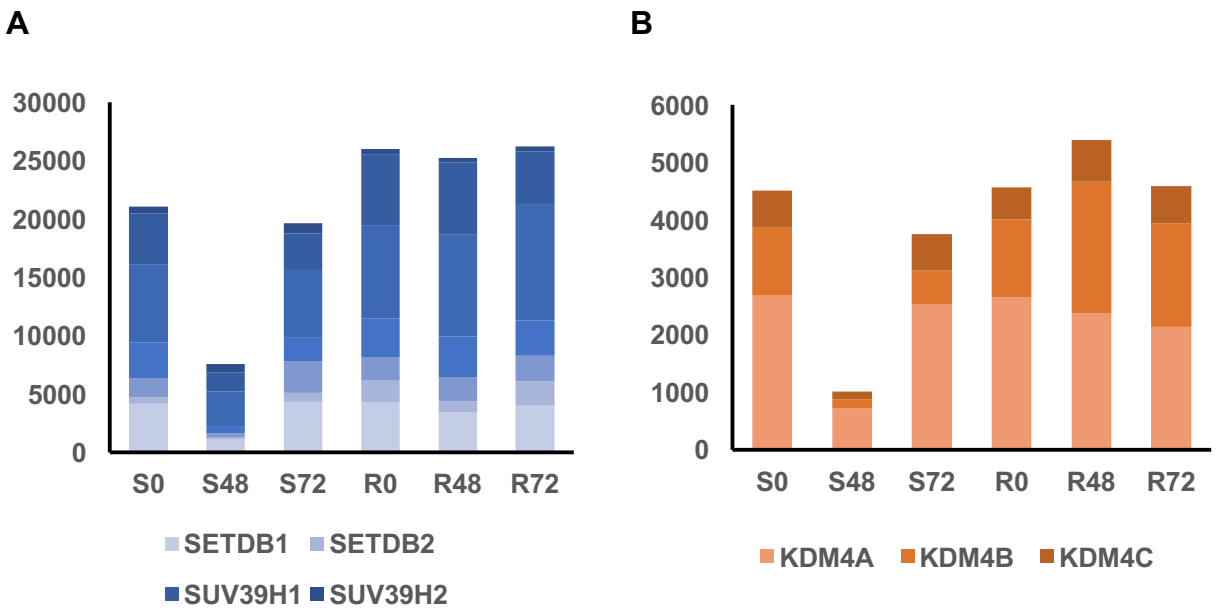


Figure 2: Expression levels of H3K9 modifiers during GEM treatment. (**A**) H3K9me3 methyltransferases. (**B**) H3K9me3 demethylases

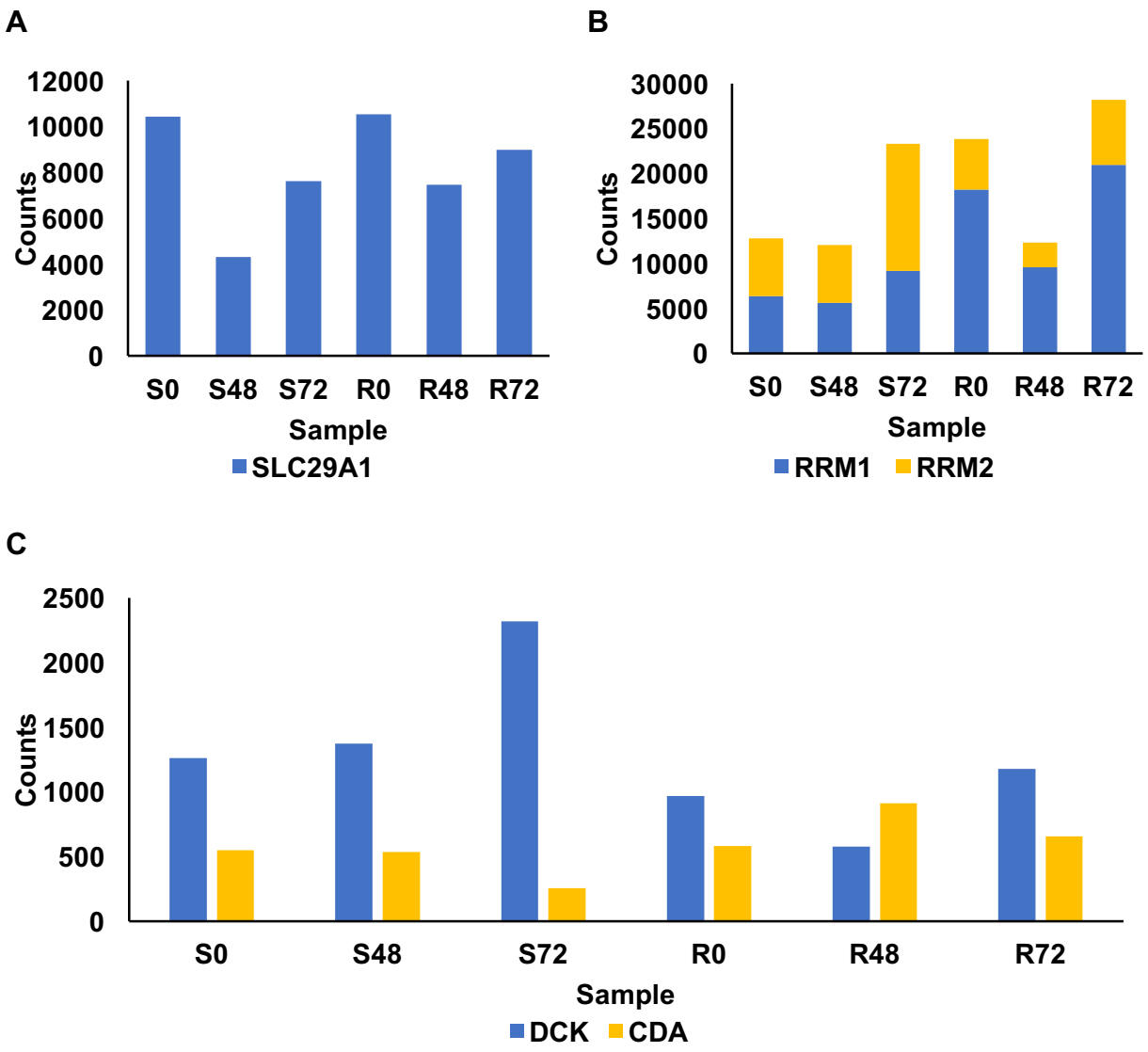


Figure 3: Expression levels of GEM metabolism enzymes during GEM treatment. (**A**) GEM transporters SLC families. (**B**) RRMs. (**C**) DCK and CDA.

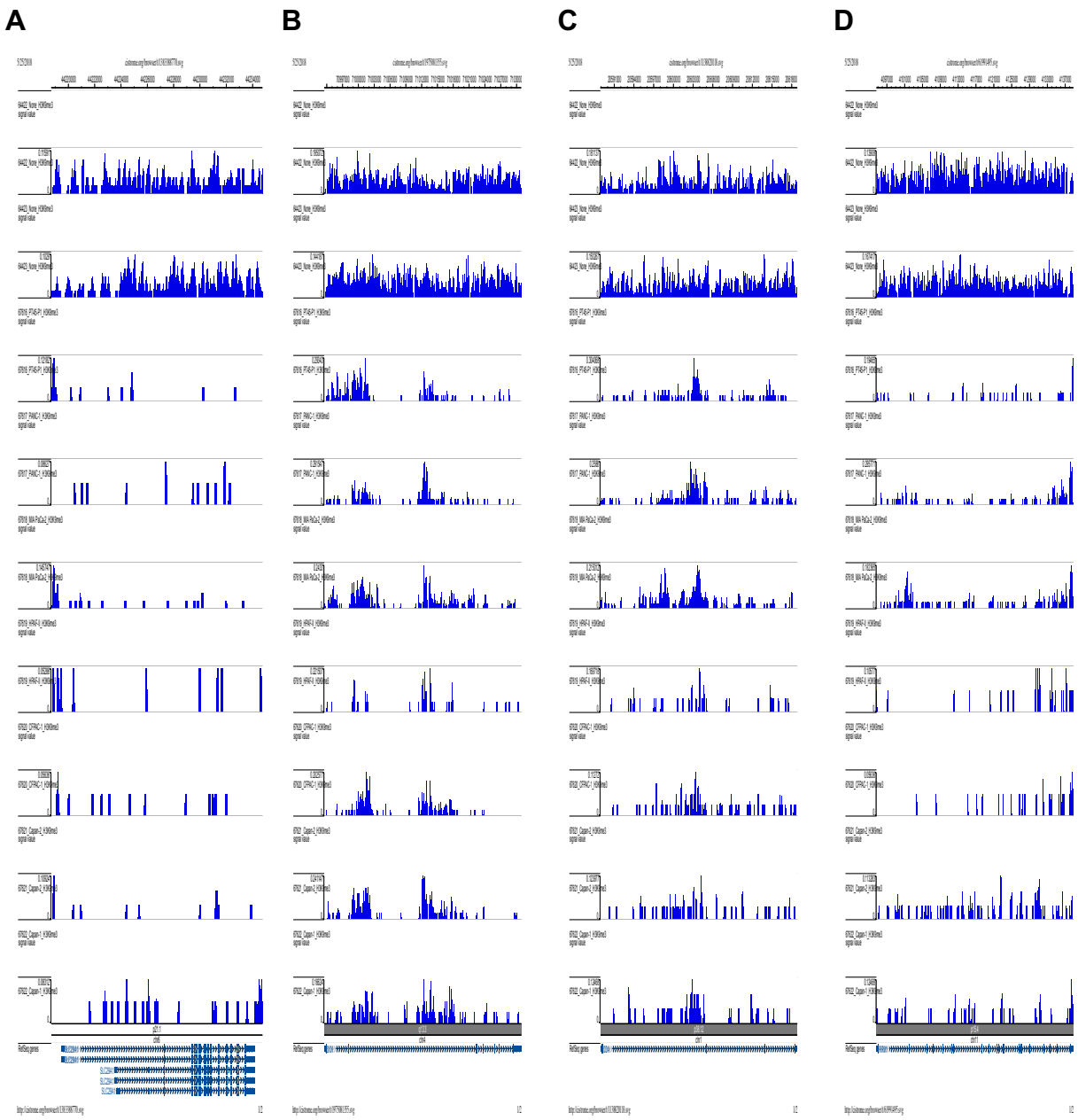


Figure 4: Chip-seq of H3K9me3 in normal pancreatic tissues and different subtypes of PDAC. The top two rows are from normal pancreatic tissues, and the five rows below are from subtypes of PDAC. (**A**) SLC29A1. (**B**) DCKs. (**C**) CDA. (D) RRM1

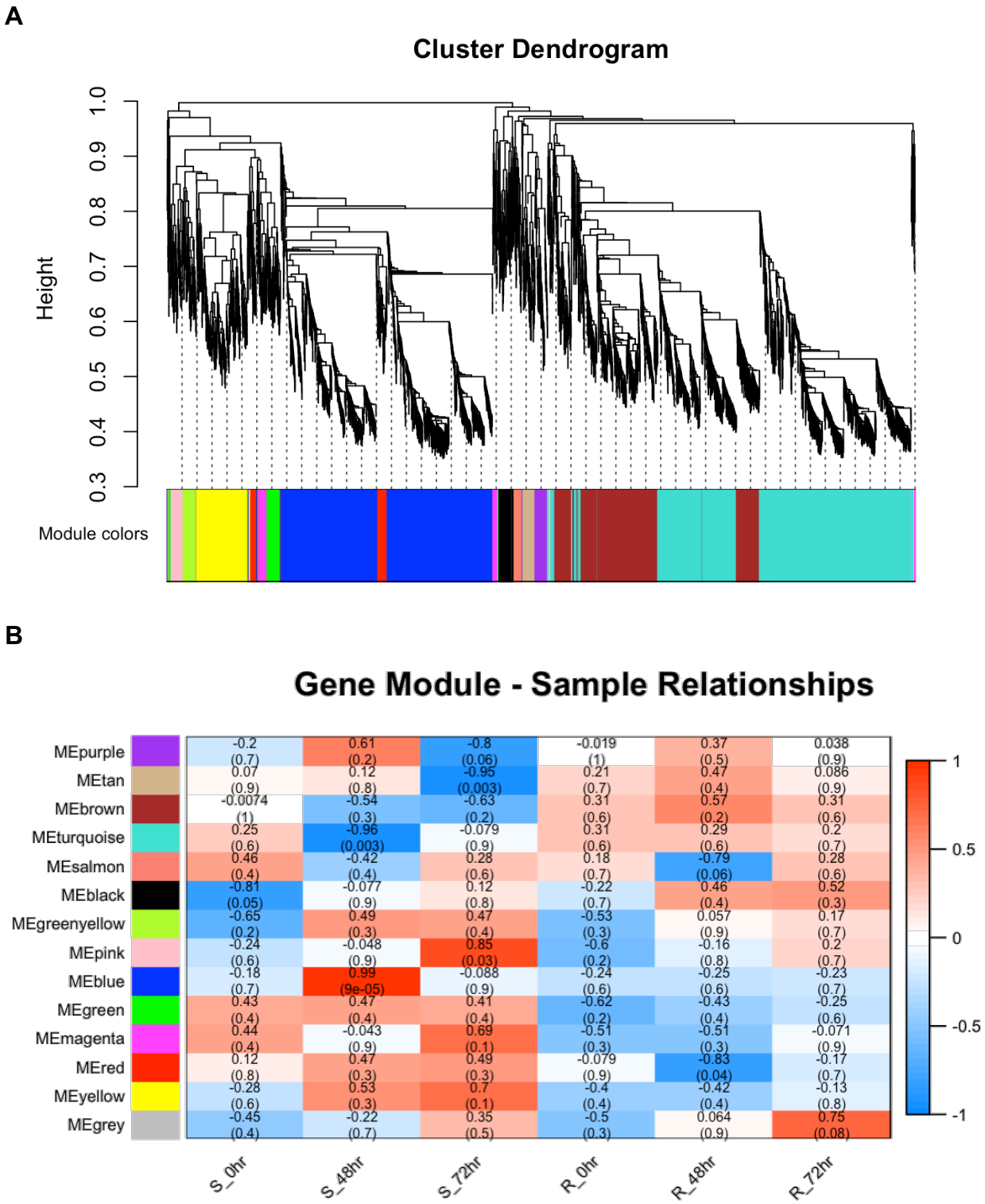
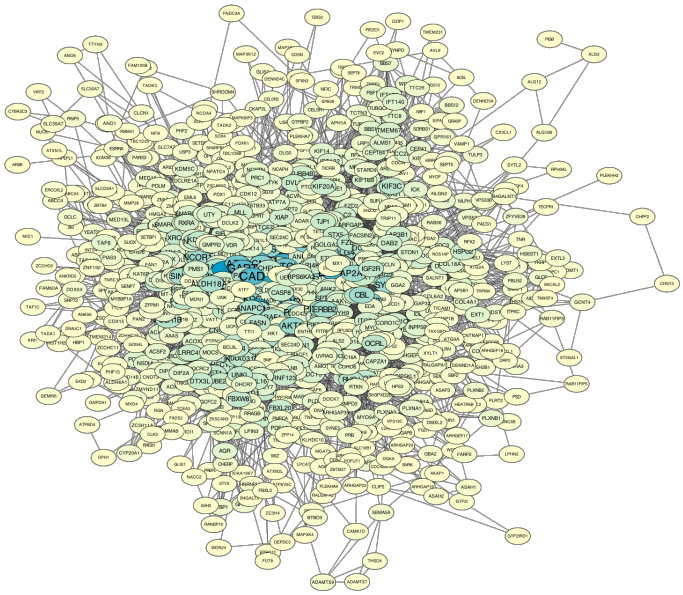


Figure 5: Identification of modules associated with clinical traits. (**A**) Clustering dendrogram of 5000 genes with largest variances. (**B**) Heatmap of the correlation between module eigengenes and samples. Each cell contained the corresponding correlation and p-value

A



B

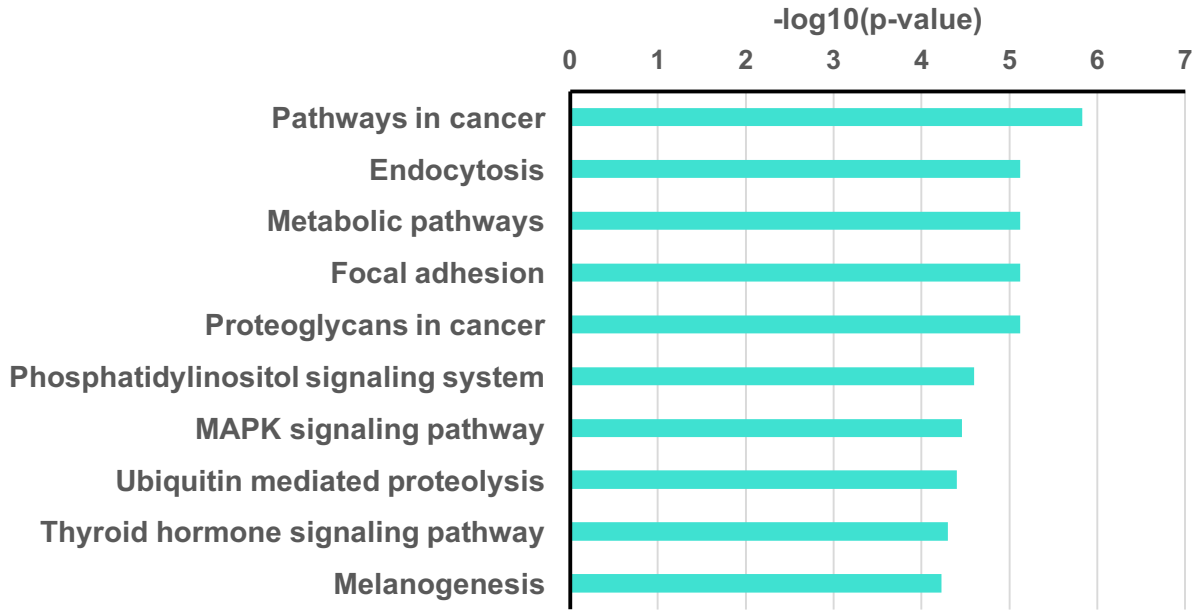


Figure 6: Protein Interaction Network of Module Turquoise.(**A**) Illustration of the network in Cytoscape (**B**) KEGG annotation of the gene set of module turquiose.

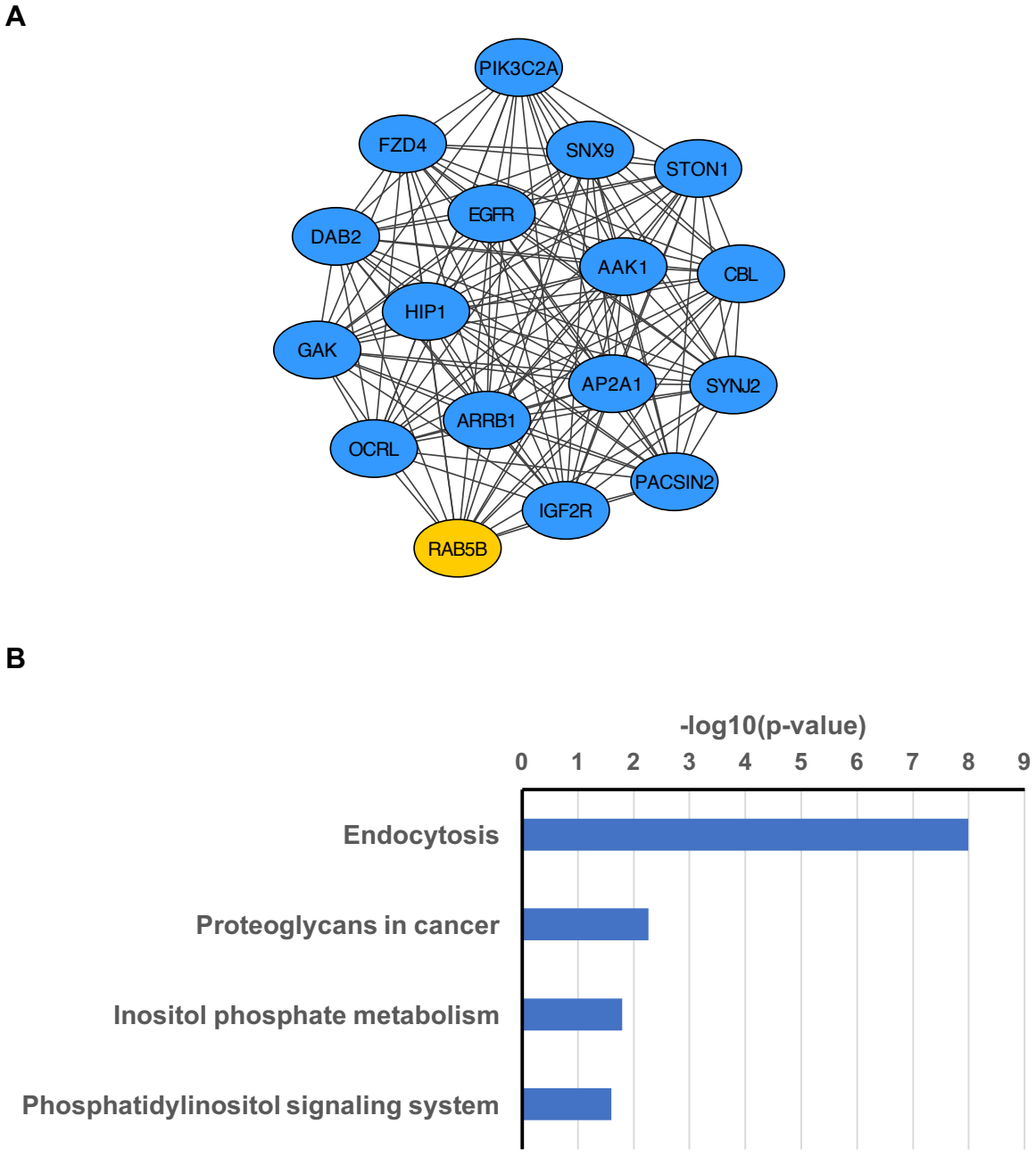


Figure 7: RAB5B leads gene cluster that functions in Golgi vesicle transport.(**A**) Gene expression levels of RAB5B between normal pancreatic and PDAC samples. (**B**) KEGG annotation of the cluster

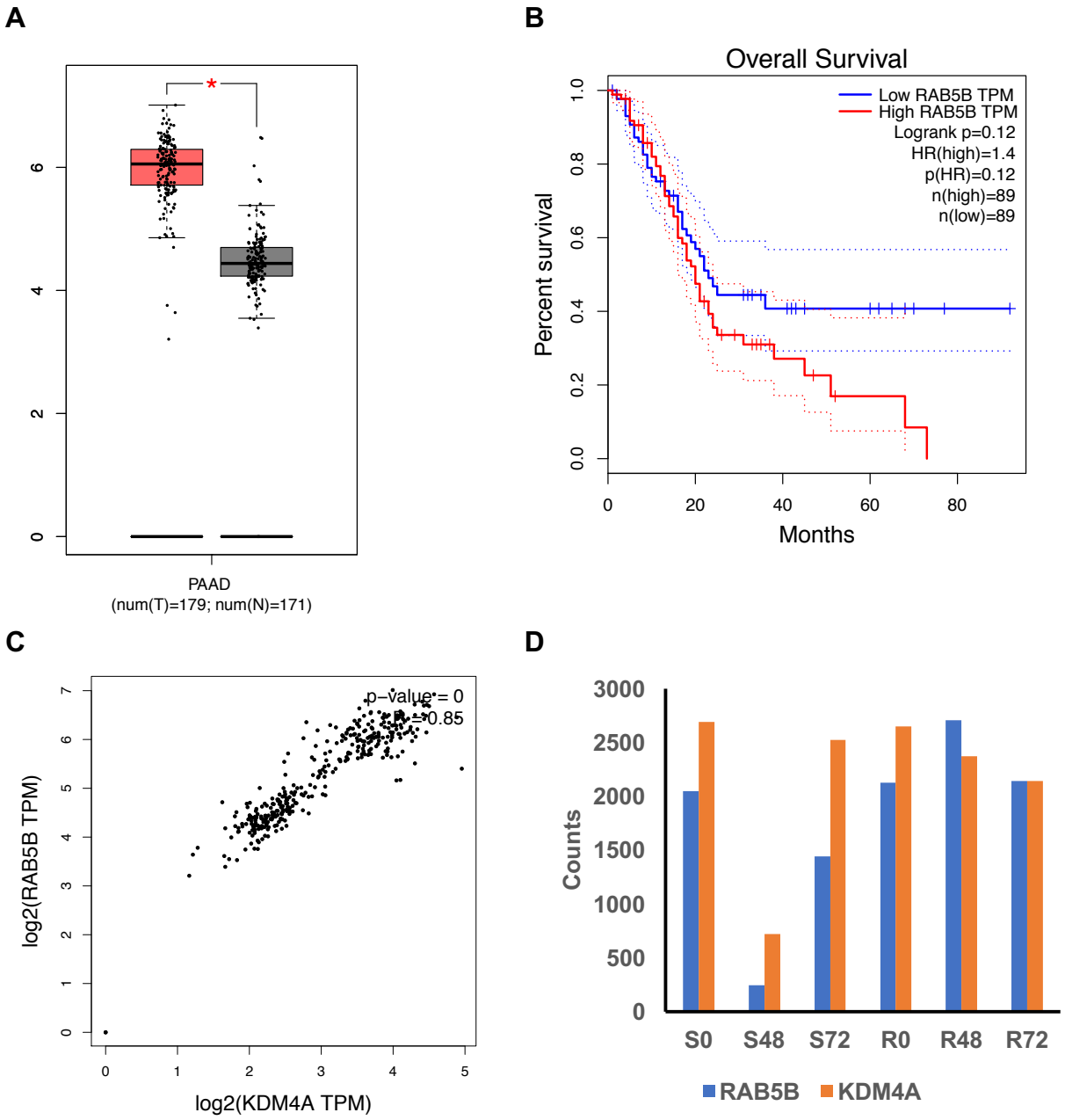


Figure 8: Validation of RAB5B associated with PDAC and its clinical outcome based on TCGA PAAD cohort (**A**) Gene expression levels of RAB5B between normal pancreatic and PDAC samples. (**B**) Kaplan-Meier curves for overall survival analysis based on the expression level of RAB5B. (**C**) Correlation between RAB5B and KDM4A. (**D**) Expression level of RAB5B and KDM4A in our RNA-Seq data

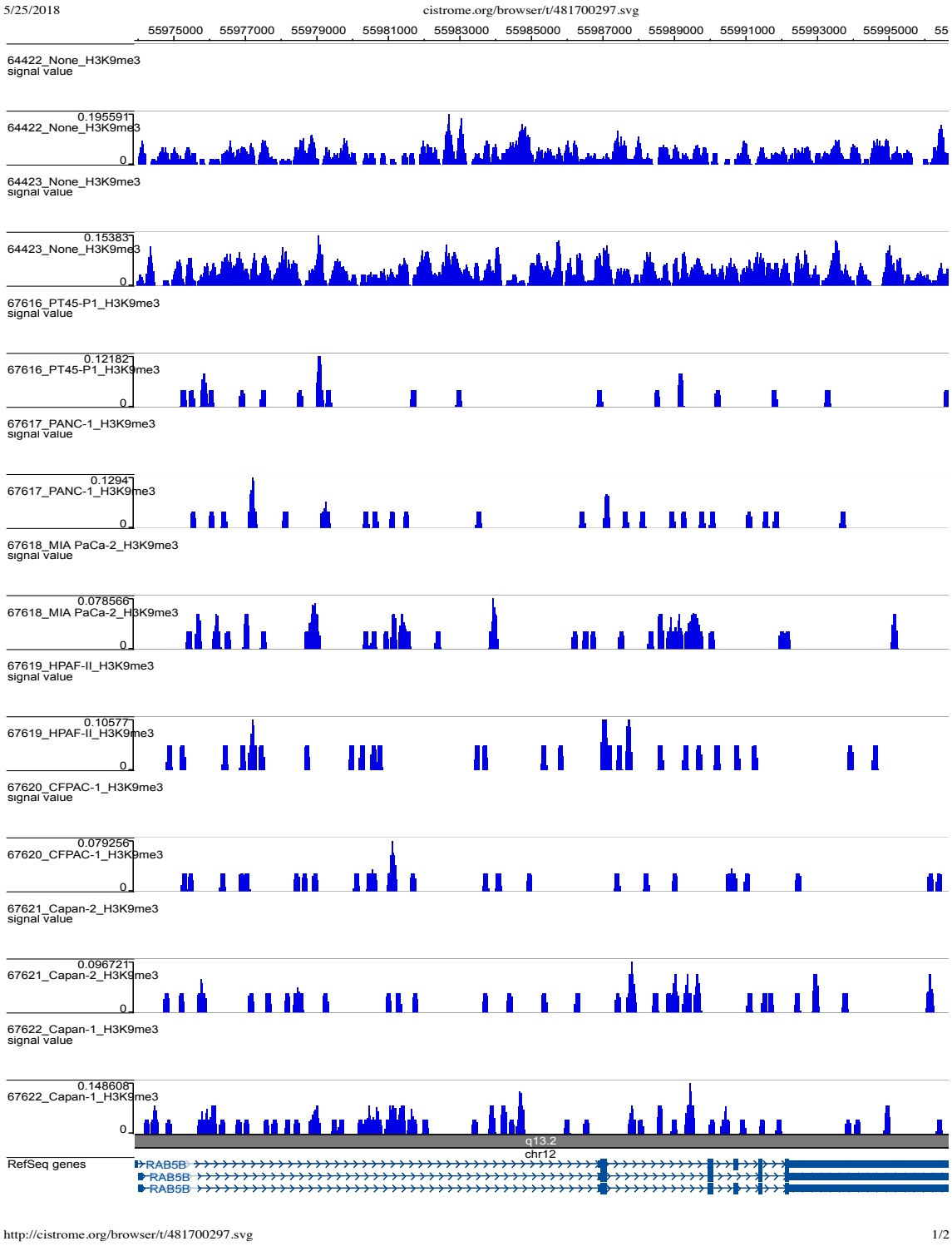


Figure 9: Chip-seq of H3K9me3 at RAB5B in normal pancreatic tissues and different subtypes of PDAC.

REFERENCE

Adamska, A., Domenichini, A., & Falasca, M. (2017). Pancreatic ductal adenocarcinoma: Current and evolving therapies. *International Journal of Molecular Sciences*, *18*. doi:10.3390/ijms18071338

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., … Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, *44*, W3–W10.

Aguirre, A. J., Hruban, R. H., & Raphael, B. J. (2017). Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*, *32*, 185–203.e13.

Alvarellos, M. L., Lamba, J., Sangkuhl, K., Thorn, C. F., Wang, L., Klein, D. J., … Klein, T. E. (2014). PharmGKB summary: gemcitabine pathway. *Pharmacogenetics and Genomics*, *24*, 564–74.

Arumugam, T., Ramachandran, V., Fournier, K. F., Wang, H., Marquis, L., Abbruzzese, J. L., … Choi, W. (2009). Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Research*, *69*, 5820–5828.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000, May). Gene ontology: Tool for the unification of biology. *Nature Genetics*.

Bader, G. D., & Hogue, C. W. V. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, *4*, 1–27.

Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., … Grimmond, S. M. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*, *531*, 47–52.

Biancur, D. E., & Kimmelman, A. C. (2018). The plasticity of pancreatic cancer metabolism in tumor progression and therapeutic resistance. *BBA - Reviews on Cancer*, 0–1.

Binenbaum, Y., Na'Ara, S., & Gil, Z. (2015). Gemcitabine resistance in pancreatic ductal adenocarcinoma. *Drug Resistance Updates*, *23*, 55–68.

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., … de Jong, P. J. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447*, 799–816.

Borazanci, E., Dang, C. V., Robey, R. W., Bates, S. E., Chabot, J. A., & Von Hoff, D. D. (2017). Pancreatic cancer: "A riddle wrapped in a mystery inside an enigma." *Clinical Cancer Research*, *23*, 1629–1637.

Brusslan, J. A., Bonora, G., Rus-Canterbury, A. M., Tariq, F., Jaroszewicz, A., & Pellegrini, M. (2015). A Genome-Wide Chronological Study of Gene Expression and Two Histone Modifications, H3K4me3 and H3K9ac, during Developmental Leaf Senescence. *Plant Physiology*, *168*, 1246–1261.

Ciccolini, J., Serdjebi, C., Peters, G. J., & Giovannetti, E. (2016). Pharmacokinetics and pharmacogenetics of Gemcitabine as a mainstay in adult and pediatric oncology: an EORTC-PAMM perspective. *Cancer Chemotherapy and Pharmacology*, *78*, 1–12.

Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., … Gray, J. W. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature Medicine*, *17*, 500–503.

De Sousa Cavalcante, L., & Monteiro, G. (2014). Gemcitabine: Metabolism and molecular mechanisms of action, sensitivity and chemoresistance in pancreatic cancer. *European Journal of Pharmacology*, *741*, 8–16.

Deer, E. L., González-Hernández, J., Coursen, J. D., Shea, J. E., Ngatia, J., Scaife, C. L., … Mulvihill, S. J. (2010). Phenotype and genotype of pancreatic cancer cell lines. *Pancreas*, *39*, 425–35.

Den, A. Van, Gremeaux, L., Topal, B., & Vankelecom, H. (2012). Human pancreatic adenocarcinoma contains a side population resistant to gemcitabine. *BMC Cancer*, *12*, 1.

Diaferia, G. R., Balestrieri, C., Prosperini, E., Nicoli, P., Spaggiari, P., Zerbi, A., & Natoli, G. (2016). Dissection of transcriptional and *cis* -regulatory control of differentiation in human pancreatic cancer. *The EMBO Journal*, *35*, 595–617.

Dreyer, S. B., Chang, D. K., Bailey, P., & Biankin, A. V. (2017). Pancreatic cancer genomes: Implications for clinical management and therapeutic development. *Clinical Cancer Research*, *23*, 1638–1646.

Fitzgerald, T. L., & McCubrey, J. A. (2014). Pancreatic cancer stem cells: Association with cell surface markers, prognosis, resistance, metastasis and treatment. *Advances in Biological Regulation*, *56*, 45–50.

Fryer, R. A., Barlett, B., Galustian, C., & Dalgleish, A. G. (2011). Mechanisms underlying gemcitabine resistance in pancreatic cancer and sensitisation by the iMiD??? lenalidomide. *Anticancer Research*, *31*, 3747–3756.

Genetics and Biology of Pancreatic Ductal Adenocarcinoma. (2016), *97*, 5421–5433.

Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics*, *13*, 343–357.

Grimison, P., Galettis, P., Manners, S., Jelinek, M., Metharom, E., de Souza, P. L., … Links, M. J. (2007). Randomized crossover study evaluating the effect of gemcitabine infusion dose rate: Evidence of auto-induction of gemcitabine accumulation. *Journal of Clinical Oncology*, *25*, 5704–5709.

Gusella, M., Pasini, F., Bolzonella, C., Meneghetti, S., Barile, C., Bononi, A., … Padrini, R. (2011). Equilibrative nucleoside transporter 1 genotype, cytidine deaminase activity and age predict gemcitabine plasma clearance in patients with solid tumours. *British Journal of Clinical Pharmacology*, *71*, 437–444.

Hessmann, E., Johnsen, S. A., Siveke, J. T., & Ellenrieder, V. (2017). Epigenetic treatment of pancreatic cancer: is there a therapeutic perspective on the horizon? *Gut*, *66*, 168–179.

Holohan, C., Van Schaeybroeck, S., Longley, D. B., & Johnston, P. G. (2013). Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer*, *13*, 714–726.

Hu, X., & Zhang, Z. (2016). Understanding the Genetic Mechanisms of Cancer Drug Resistance Using Genomic Approaches. *Trends in Genetics*, *32*, 127–137.

Igarashi, T., Araki, K., Yokobori, T., Altan, B., Yamanaka, T., Ishii, N., … Kuwano, H. (2017). Association of RAB5 overexpression in pancreatic cancer with cancer progression and poor prognosis via E-cadherin suppression. *Oncotarget*, *8*, 12290–12300.

Ihaka, R., & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, *5*, 299–314.

Jia, Y., & Xie, J. (2015). Promising molecular mechanisms responsible for gemcitabine resistance in cancer. *Genes and Diseases*, *2*, 299–306.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*, 27–30.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, *12*, 357–360.

Kim, M. P., & Gallick, G. E. (2008). Gemcitabine resistance in pancreatic cancer: Picking the key players. *Clinical Cancer Research*, *14*, 1284–1285.

Kroep, J. R., Loves, W. J. P., van der Wilt, C. L., Alvarez, E., Talianidis, I., Boven, E., … Peters, G. J. (2002). Pretreatment deoxycytidine kinase levels predict in vivo gemcitabine sensitivity. *Molecular Cancer Therapeutics*, *1*, 371–376.

Lane, A. A., & Chabner, B. A. (2009). Histone deacetylase inhibitors in cancer therapy. *Journal of Clinical Oncology*, *27*, 5459–5468.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*. doi:10.1186/1471-2105-9-559

Lee, C. J., Dosch, J., & Simeone, D. M. (2008). Pancreatic cancer stem cells. *Journal of Clinical Oncology*, *26*, 2806–2812.

Liang, C., Shi, S., Meng, Q., Liang, D., Ji, S., Zhang, B., … Yu, X. (2017). Complex roles of the stroma in the intrinsic resistance to gemcitabine in pancreatic cancer: Where we are and where we are going. *Experimental and Molecular Medicine*, *49*. doi:10.1038/emm.2017.255

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930.

Liu, T., Ortiz, J. A., Taing, L., Meyer, C. A., Lee, B., Zhang, Y., … Liu, X. S. (2011). Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biology*, *12*, R83.

López-Casas, P. P., & López-Fernández, L. A. (2010). Gene-expression profiling in pancreatic cancer. *Expert Review of Molecular Diagnostics*, *10*, 591–601.

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

Lu, C., Yang, D., Sabbatini, M. E., Colby, A. H., Grinstaff, M. W., Oberlies, N. H., … Liu, K. (2018). Contrasting roles of H3K4me3 and H3K9me3 in regulation of apoptosis and gemcitabine resistance in human pancreatic cancer cells. *BMC Cancer*, *18*, 1–12.

Manji, G. A., Olive, K. P., Saenger, Y. M., & Oberstein, P. (2017). Current and emerging therapies in metastatic pancreatic cancer. *Clinical Cancer Research*, *23*, 1670–1678.

Manuyakorn, A., Paulus, R., Farrell, J., Dawson, N. A., Tze, S., Cheung-Lau, G., … Dawson, D.

W. (2010). Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: Results from RTOG 9704. *Journal of Clinical Oncology*, *28*, 1358–1365.

McDonald, O. G., Li, X., Saunders, T., Tryggvadottir, R., Mentch, S. J., Warmoes, M. O., … Feinberg, A. P. (2017). Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nature Genetics*, *49*, 367–376.

McDonald, O. G., Wu, H., Timp, W., Doi, A., & Feinberg, A. P. (2011). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nature Structural and Molecular Biology*, *18*, 867–874.

Minucci, S., & Pelicci, P. G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nature Reviews Cancer*, *6*, 38–51.

Mori-Iwamoto, S., Kuramitsu, Y., Ryozawa, S., Taba, K., Fujimoto, M., Okita, K., … Sakaida, I. (2008). A proteomic profiling of gemcitabine resistance in pancreatic cancer cell lines. *Molecular Medicine Reports*, *1*, 429–434.

Nordh, S., Ansari, D., & Andersson, R. (2014). hENT1 expression is predictive of gemcitabine outcome in pancreatic cancer: A systematic review. *World Journal of Gastroenterology*.

Park, Y. S., Jin, M. Y., Kim, Y. J., Yook, J. H., Kim, B. S., & Jang, S. J. (2008). The Global Histone Modification Pattern Correlates with Cancer Recurrence and Overall Survival in Gastric Adenocarcinoma. *Annals of Surgical Oncology*, *15*, 1968–1976.

Richards, K. E., Zeleniak, A. E., Fishel, M. L., Wu, J., Littlepage, L. E., & Hill, R. (2017). Cancer-associated fibroblast exosomes regulate survival and proliferation of pancreatic cancer cells. *Oncogene*, *36*, 1770–1778.

Ryan, D. P., Hong, T. S., & Bardeesy, N. (2014). Pancreatic Adenocarcinoma. *New England Journal of Medicine*, *371*, 1039–1049.

Saiki, Y., Yoshino, Y., Fujimura, H., Manabe, T., Kudo, Y., Shimada, M., … Horii, A. (2012). DCK is frequently inactivated in acquired gemcitabine-resistant human cancer cells. *Biochemical and Biophysical Research Communications*, *421*, 98–104.

Samulitis, B. K., Pond, K. W., Pond, E., Cress, A. E., Patel, H., Wisner, L., … Landowski, T. H. (2015). Gemcitabine resistant pancreatic cancer cell lines acquire an invasive phenotype with collateral hypersensitivity to histone deacetylase inhibitors. *Cancer Biology & Therapy*, *16*, 43–51.

Santa Pau, E. C. de, Real, F. X., & Valencia, A. (2014). *Bioinformatics Analysis of Pancreas Cancer Genome in High-Throughput Genomic Technologies*. *Molecular Diagnostics and Treatment of Pancreatic Cancer*. Elsevier. doi:10.1016/B978-0-12-408103-1.00005-4

Sherman, M. H., Yu, R. T., Tseng, T. W., Sousa, C. M., Liu, S., Truitt, M. L., … Evans, R. M. (2017). Stromal cues regulate the pancreatic cancer epigenome and metabolome. *Proceedings of the National Academy of Sciences*, *114*, 1129–1134.

Shields, M. A., Dangi-Garimella, S., Krantz, S. B., Bentrem, D. J., & Munshi, H. G. (2011). Pancreatic cancer cells respond to type I collagen by inducing snail expression to promote membrane type 1 matrix metalloproteinase-dependent collagen invasion. *Journal of Biological Chemistry*, *286*, 10495–10504.

Shukla, S. K., Purohit, V., Mehla, K., Gunda, V., Chaika, N. V., Vernucci, E., … Singh, P. K. (2017). MUC1 and HIF-1alpha Signaling Crosstalk Induces Anabolic Glucose Metabolism to Impart Gemcitabine Resistance to Pancreatic Cancer. *Cancer Cell*, *32*, 392.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians*, *68*, 7–30.

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., … von Mering, C. (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research*, *45*, D362–D368.

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, *45*, W98–W102.

Tripathi, S., Pohl, M. O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D. A., … Chanda, S. K. (2015). Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. *Cell Host & Microbe*, *18*, 723–735.

Wei, Y., Xia, W., Zhang, Z., Liu, J., Wang, H., Adsay, N. V., … Hung, M. C. (2008). Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. *Molecular Carcinogenesis*, *47*, 701–706.

Wood, L. D., & Hruban, R. H. (2018). Molecular Understanding of Development of Ductal Pancreatic Cancer. In *The Pancreas: An Integrated Textbook of Basic Science, Medicine, and Surgery* (pp. 679–687).

Zhou, X., Lowdon, R. F., Li, D., Lawson, H. A., Madden, P. A. F., Costello, J. F., & Wang, T. (2013). Exploring long-range genome interactions using the WashU Epigenome Browser. *Nature Methods*, *10*, 375–376.