Title: Identification of Genes and Pathways Implicated in Treatment of Intraheptatic Cholangiocarcinoma with GC and ICB (CLTA-4)

Author: Victoria Palecek (2024)

Introduction:

Intrahepatic cholangiocarcinoma, or ICC, is a rare type of cancer that makes up approximately 10% of cholangiocarinomas. ICC can be incredibly hard to treat and is often associated with a poor prognosis due to limited treatment options (Dana-Farber Cancer Institute, 2024). One of the chemotherapies used to treat ICC is gemcitabine/cisplatin (GC) which aids in lymphocyte infiltration into the tumor. Immunecheckpoint blockades (ICBs) ensure the elimination of cancer cells by negatively regulating elements involved in immune system suppression. In a paper by Chen et al., mice with ICC were treated with GC and ICB both separately and in tandem to determine their effects on prognosis. In this study, we aim to use bioinformatics methods to analyze RNA-seq data from a study completed in early 2024 aligned to the mm10 mouse genome as well as an alignment to the most recent mouse genome (GRCm39) to identify new genes and pathways associated with ICC treatment response to GC chemotherapy and ICB (Chen et al.,2024).

Background

ICC is not only difficult to treat because of the limited treatment options available. The most successful treatment for ICC is surgical removal of the tumor. Unfortunately, most patients with ICC are not diagnosed before the cancer has metastasized due to non-specific symptoms. Some of the risk factors for ICC include chronic inflammation, sclerosing cholangitis, fibropolycystic liver disease, gall stones, biliary tract diseases, and genetic predisposition (Tanaka et al.,2010). This type of cancer is much more prevalent in individuals 65 years of age and older. A majority of patients will require a hemihepatectomy and some may require a partial bile duct removal and reconstruction (Bridgewater et al.,2014).

Significance

Our study holds significant clinical implications for ICC management by attempting to identify novel therapeutic targets and pathways associated with GC chemotherapy (gemcitabine/cisplatin) and ICB. By aligning RNA-seq data to the most recent mouse genome (GRCm39), we aim to identify differentially expressed genes (DEGs), whose discovery was limited by the previous genome alignment, relevant to ICC treatment response. This approach may offer insights into the molecular mechanisms driving ICC response to therapy, which may aid in the future development of targeted therapies.

Current Knowledge of GC/ICB in ICC

The work of Chen et al. provides an important resource of information regarding the outcome of combining GC chemotherapy and CTLA-4 blockade (ICB) and its use in reprogramming the immune microenvironment of ICC. Our study was based on the data provided by this paper. The study's results indicate that the combination of GC chemotherapy and CTLA-4 blockade enhances the immune response against tumors by increasing the number of tumor-infiltrating lymphocytes (TILs) and upregulating immune checkpoint molecules like PD-L1. The authors also observed that the combination of GC chemotherapy, CTLA-4 blockade, and anti-PD-1 (both ICBs) therapy had an even greater anti-tumor effects in preclinical ICC models. These observations indicate potential in combination therapies involving ICBs in treating ICC and improving outcomes.

GEO Accession ID	SRR ID	Treatment Type	Test/Control
GSM8017532	SRR27553696	N/A	Control
GSM8017533	SRR27553695	N/A	Control
GSM8017534	SRR27553694	N/A	Control
GSM8017535	SRR27553693	GC/ICB	Test

GSM8017536	SRR27553692	GC/ICB	Test
GSM8017537	SRR27553691	GC/ICB	Test
GSM8017538	SRR27553690	GC	Test
GSM8017539	SRR27553689	GC	Test
GSM8017540	SRR27553688	GC	Test
GSM8017541	SRR27553687	ICB	Test
GSM8017542	SRR27553686	ICB	Test
GSM8017543	SRR27553685	ICB	Test

Figure 1: Table indicates each treatment replicate with its associated GEO accession ID, SRR ID, treatment method and grouping.

Analysis Methods:

For our analysis, we used a multi-step approach to extract meaningful results from RNA-seq data obtained from different experimental groups. Initially, we uploaded RNA-seq data from three replicates of each experimental group (control, ICB treatment, GC treatment, and ICB/GC treatment) into the Galaxy platform. Subsequently, we conducted quality control checks using FASTQC to assess the overall sequencing quality. Following this, we utilized trimmomatic software to remove low-quality reads and any remaining primers, ensuring the integrity of our dataset, before conducting a second round of FASTQC analysis. Next, we aligned all 12 samples to both the mm10 and GRCm39 genomes to account for potential discrepancies between genome assemblies using HISAT2. Counts for these alignments were then obtained by running the aligned reads through htseq-count. To identify DEGs from the count data, we employed the DESeq2 package in R, comparing expression levels between the control group and each treatment group using a corrected t-test. Then, we conducted a comparative analysis between the significant genes identified in both genome alignments, utilizing an anti-join method to isolate DEGs uniquely identified by the GRCm39 alignment. Finally, we utilized DAVID, WebGestalt, REViGO and STRING bioinformatics tools to elucidate functional annotations, pathways and connections associated with the identified DEGs, enabling us to infer the molecular mechanisms underlying the therapeutic effects of the drugs in each treatment group.

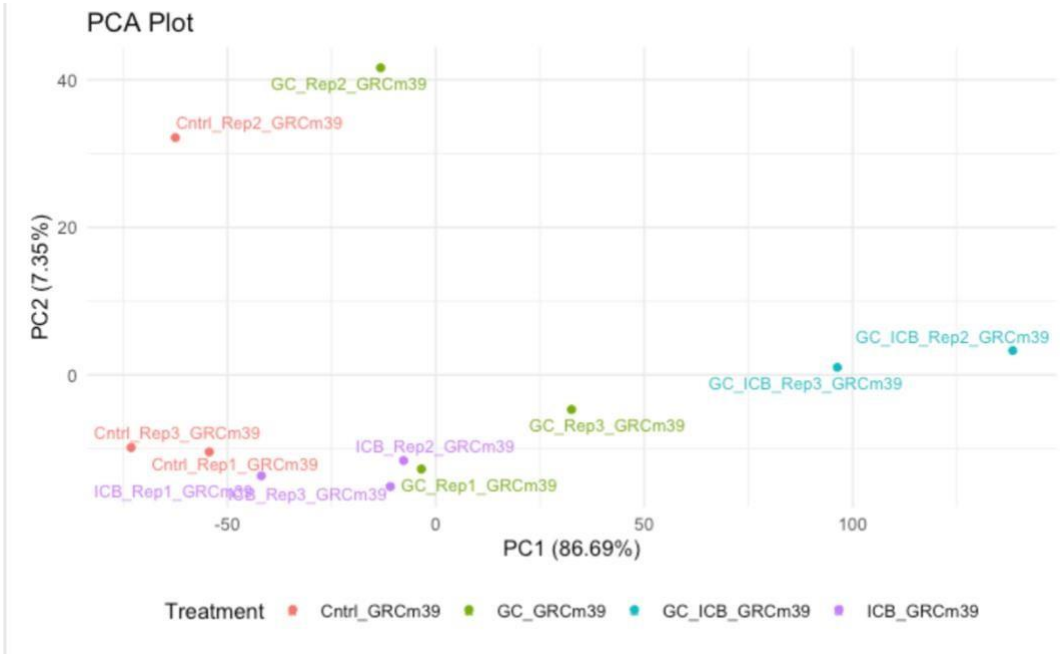


Figure 2: PCA plot of the GRCm39 treatment and control groups. The ICB test group replicate 3 was removed as it did not align with the rest of the ICB replicates or any other replicate aligned to the GRCm39 genome.

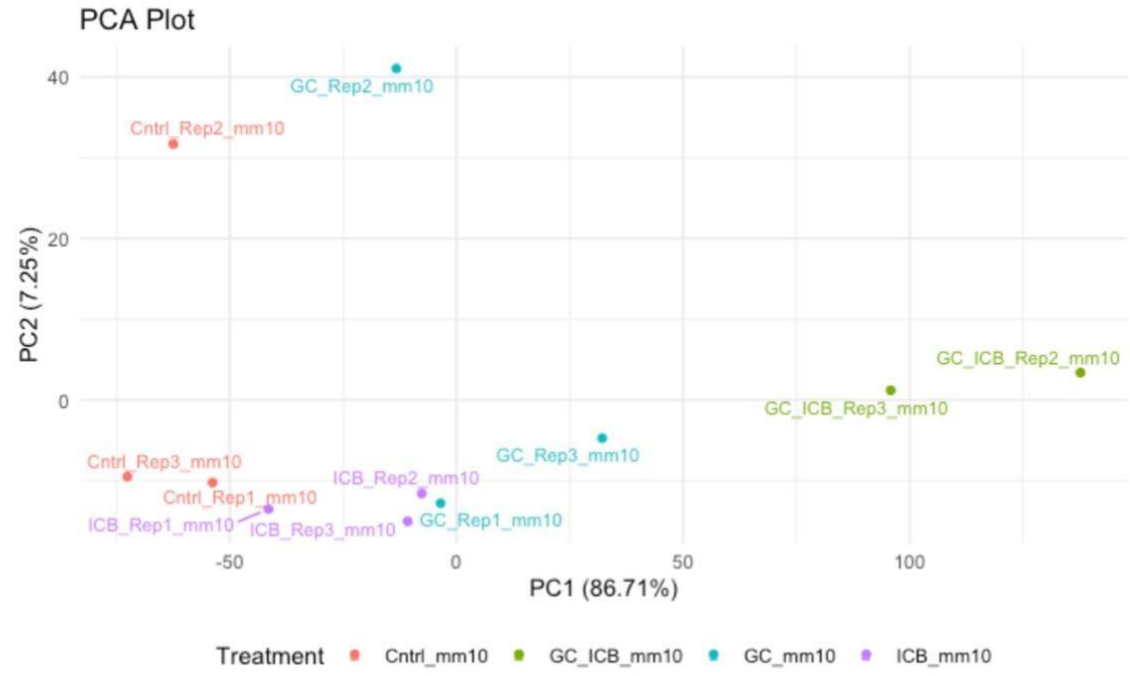


Figure 3: PCA plot of the mm10 treatment and control groups. The ICB test group replicate 3 was removed as it did not align with the rest of the ICB replicates or any other replicate aligned to the mm10 genome.

Analysis Results:

In our study, we compared the test groups (GC, GC/ICB, and ICB) to the control group and identified 24,377 DEGs using the DESeq2 package. To focus on the most biologically relevant genes, we applied a parameter to filter the DEGs. We excluded genes with an adjusted p-value greater than 0.005 and an absolute log2 fold change of less than 5. This helped us identify genes that exhibited substantial changes in expression levels across the treatment groups. After applying the filtering criteria, we observed distinct distributions of significantly expressed genes in each treatment group. The GC/ICB group showed 962 genes that met our criteria, indicating substantial transcriptional changes induced by the combination of GC chemotherapy and ICB. Similarly, the GC group exhibited 220 significantly expressed genes, highlighting the impact of GC chemotherapy alone on gene expression profiles. In contrast, the ICB group showed a modest response, with only 39 genes meeting our filtering criteria. We further examined the magnitude of gene expression changes by calculating the maximum fold change observed in each treatment group. In the GC/ICB group the highest fold change was 22.82, the GC group exhibited a maximum fold change of 19.03 and the ICB group had a maximum fold change of 18.82 indicating a pronounced increase in gene expression in the top genes in all treatment groups. Additionally, we found 68 unique genes in the GC/ICB treatment group, 8 in the GC treatment group, and only 1 in the ICB treatment group that met our criteria of significance and were found in the GRCm39 alignment but not in the mm10 alignment.

GC/ICB Most Highly Expressed Genes:

Ensembl Gene ID	Gene name	Gene type
ENSMUSG00000001155	Ftcd	protein_coding
ENSMUSG00000005547	Cyp2a5	protein_coding
ENSMUSG00000022227	Mcpt1	protein_coding
ENSMUSG00000022347	A1bg	protein_coding
ENSMUSG00000025003	Cyp2c39	protein_coding
ENSMUSG00000026272	Agxt	protein_coding

ENSMUSG00000027761	Aadac	protein_coding
ENSMUSG00000032808	Cyp2c38	protein_coding
ENSMUSG00000035780	Ugt2a3	protein_coding
ENSMUSG00000035836	Ugt2b1	protein_coding
ENSMUSG00000040660	Cyp2b9	protein_coding
ENSMUSG00000052595	A1cf	protein_coding
ENSMUSG00000053303	Slc22a26	protein_coding
ENSMUSG00000054757	Akr1c20	protein_coding
ENSMUSG00000060407	Cyp2a12	protein_coding
ENSMUSG00000066072	Cyp4a10	protein_coding
ENSMUSG00000070704	Ugt2b36	protein_coding
ENSMUSG00000074254	Cyp2a4	protein_coding
ENSMUSG00000074639	Rdh16f2	protein_coding
ENSMUSG00000092008	Cyp2c69	protein_coding

Figure 4: Table indicates the top 20 genes from the GC/ICB treatment group when compared to the control group and their associated Ensembl gene ID and gene type.

GC Most Highly Expressed Genes:

Ensembl Gene ID	Gene name	Gene type
ENSMUSG00000001155	Ftcd	protein_coding
ENSMUSG00000004038	Gstm3	protein_coding
ENSMUSG00000021208	Ifi27l2b	protein_coding
ENSMUSG00000024331	Dsc2	protein_coding
ENSMUSG00000025405	Inhbc	protein_coding
ENSMUSG00000026272	Agxt	protein_coding
ENSMUSG00000027761	Aadac	protein_coding
ENSMUSG00000028003	Lrat	protein_coding
ENSMUSG00000029273	Sult1d1	protein_coding
ENSMUSG00000034785	Dio1	protein_coding
ENSMUSG00000035811	Ugt2b35	protein_coding
ENSMUSG00000035836	Ugt2b1	protein_coding
ENSMUSG00000041660	Bbox1	protein_coding
ENSMUSG00000043681	Fam25a	protein_coding
ENSMUSG00000054757	Akr1c20	protein_coding
ENSMUSG00000062624	Cyp2c67	protein_coding
ENSMUSG00000070704	Ugt2b36	protein_coding

ENSMUSG00000074639	Rdh16f2	protein_coding
ENSMUSG00000096546	Smlr1	protein_coding

Figure 5: Table indicates the top 20 genes from the GC treatment group when compared to the control group and their associated Ensembl gene ID and gene type.

ICB Most Highly Expressed Genes:

Ensembl Gene ID	Gene name	Gene type
ENSMUSG00000001155	Ftcd	protein_coding
ENSMUSG00000006522	Itih3	protein_coding
ENSMUSG00000006529	Itih1	protein_coding
ENSMUSG00000010064	Slc38a3	protein_coding
ENSMUSG00000010175	Prox1	protein_coding
ENSMUSG00000017453	Pipox	protein_coding
ENSMUSG00000019880	Rspo3	protein_coding
ENSMUSG00000019989	Enpp3	protein_coding
ENSMUSG00000020609	Apob	protein_coding
ENSMUSG00000020884	Asgr1	protein_coding
ENSMUSG00000021620	Acot12	protein_coding
ENSMUSG00000021999	Cpb2	protein_coding
ENSMUSG00000022057	Adamdec1	protein_coding
ENSMUSG00000022227	Mcpt1	protein_coding
ENSMUSG00000022809	Nr1i2	protein_coding
ENSMUSG00000022868	Ahsg	protein_coding
ENSMUSG00000022871	Fetub	protein_coding
ENSMUSG00000022875	Kng1	protein_coding
ENSMUSG00000024331	Dsc2	protein_coding
ENSMUSG00000024391	Apom	protein_coding

Figure 6: Table indicates the top 20 genes from the ICB treatment group when compared to the control group and their associated Ensembl gene ID and gene type.

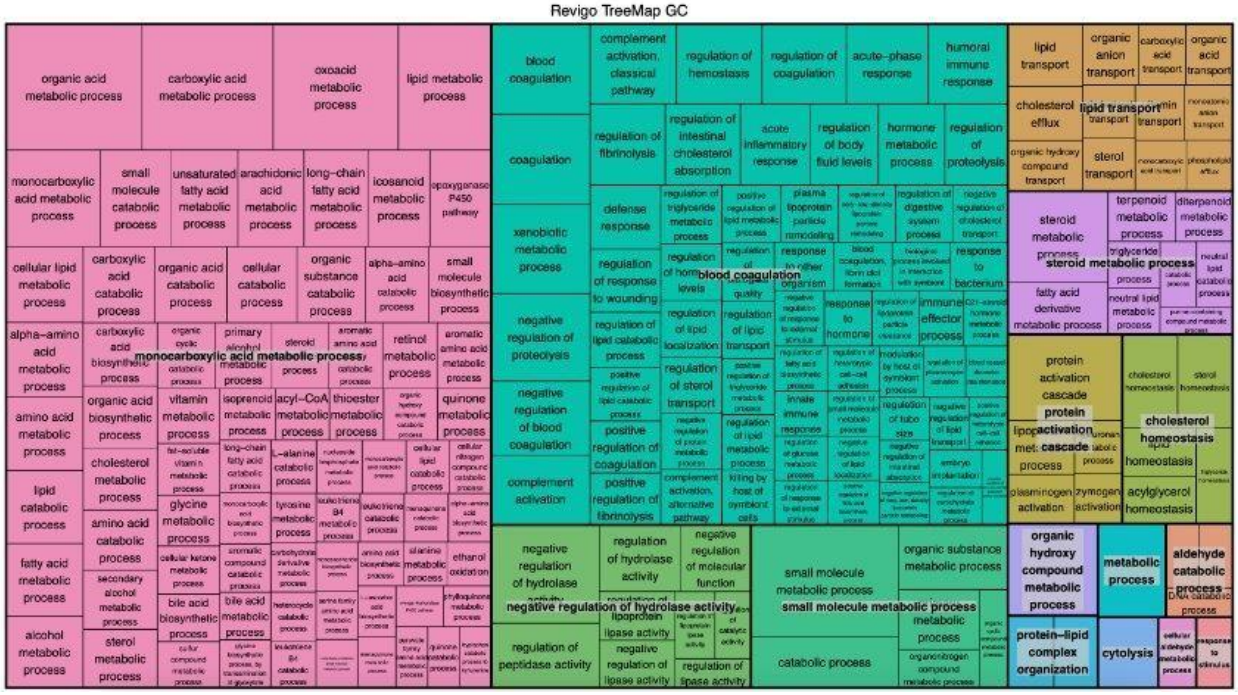
Discussion:

The GC/ICB treatment group showed a significantly low false discover rate (FDR) in Gene Ontology (GO) terms related to cellular metabolism, stress response, and drug detoxification processes in the context of ICC. Chemotherapy drugs like gemcitabine and cisplatin affect cancer cells by inducing cellular stress, DNA damage, and metabolic pathway disruption. In response, cells activate stress response pathways and

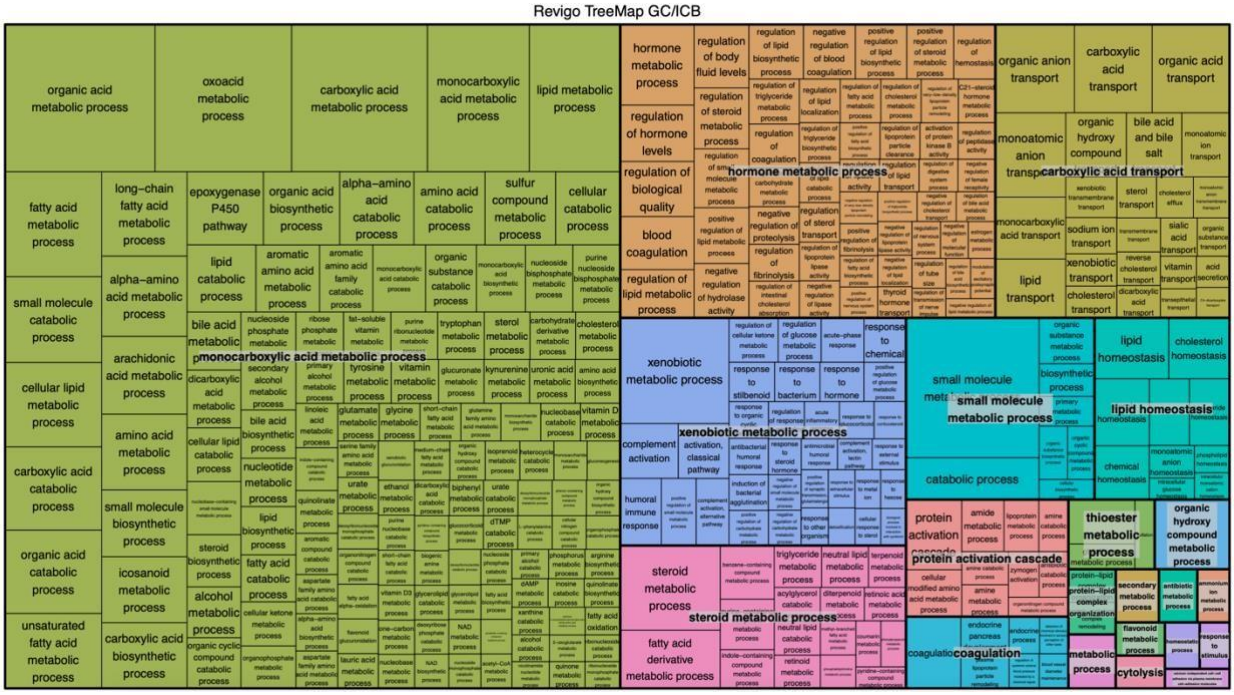
upregulate metabolic processes to cope with increased energy demands and maintain cellular homeostasis (Herr and Debatin, 2001). Enrichment of GO terms associated with lipid metabolism, drug metabolism, and organic compound metabolism indicate complex interaction between chemotherapy-induced cellular damage and compensatory mechanisms aimed at detoxifying and eliminating cytotoxic agents (Van Den Boogaard et al.,2022). Interestingly, the enriched GO terms in the GC treatment group were nearly identical to those observed in the GC/ICB group, indicating that GC chemotherapy alone also triggers similar metabolic and stress response pathways in ICC cells.

In contrast, our analysis of significant DEGs from the ICB test group revealed enriched GO term descriptions that align with the expected effects of ICB therapy such as negative regulation of hydrolase activity and negative regulation of endopeptidase activity. The observed negative regulation of enzymatic activities indicates immunesuppressive mechanisms. This inhibition may prevent the degradation of antitumor immune effectors and enhance their effector functions, which may contribute to the effectiveness of the combination of GC chemotherapy and ICB in ICC tumor cells as ween in the Chen et al. study (Wei et al.,2018). Additionally, the enrichment of GO terms related to small molecule metabolic processes and organic hydroxy compound metabolic processes suggests metabolic adaptations induced by ICB therapy.

Revigo TreeMap GC



Revigo TreeMap GC/ICB



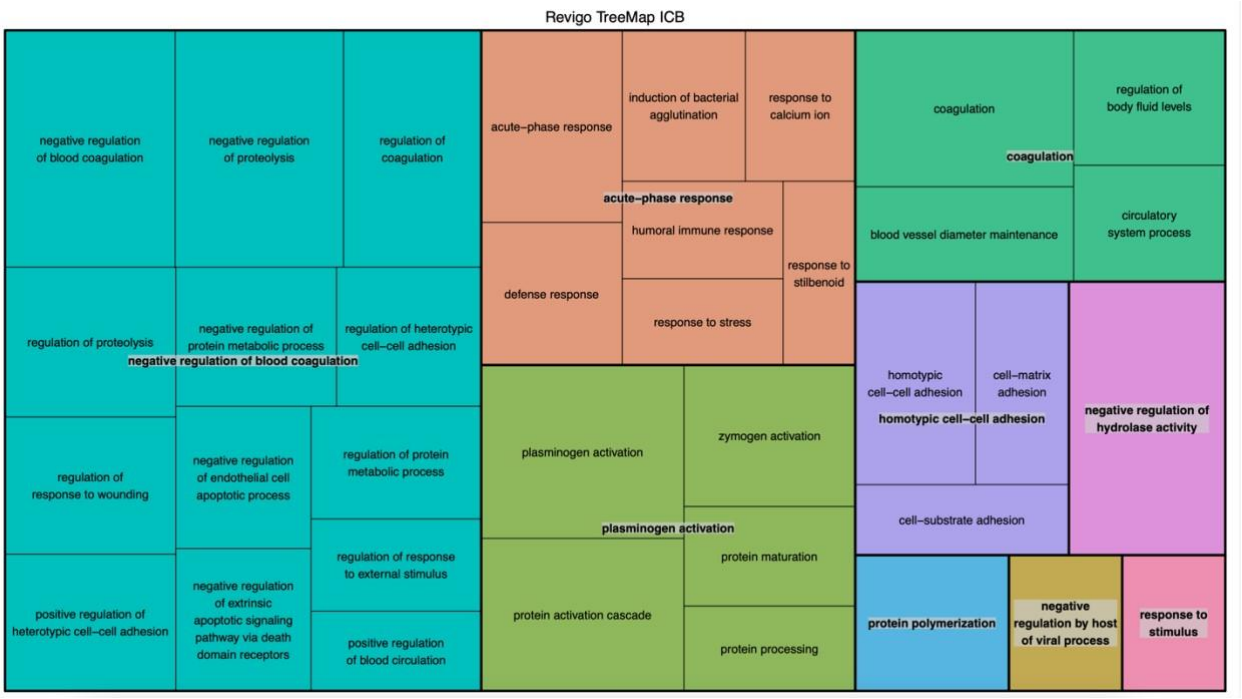


Figure 7: Image indicates REViGO tree maps for each of the three treatment groups and their enriched GO terms.

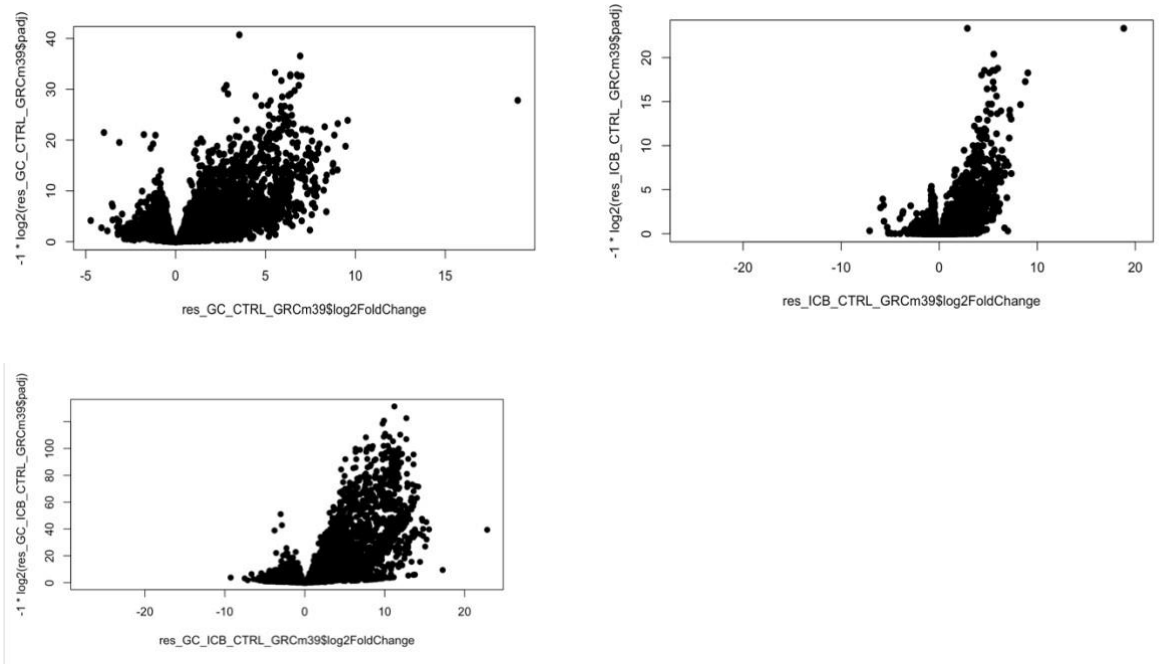


Figure 8: Image indicates volcano plots for all three treatment groups aligned to the GRCm39 genome. The log2 transformed adjusted p-value and fold change in each treatment group plotted against each other when compared show an important visual representation of why there are many more significant genes in the GC/ICB test group than either the GC or ICB test groups alone. There are few outliers indicated by these plots.

In our analysis, we found a group of unique genes that showed significant upregulation following treatment in each experimental group. The GC/ICB treatment group had 68 unique genes, of which 61 were long non-coding RNAs (lncRNAs). Interestingly, only two genes among these have been previously annotated, indicating the potential novelty of the remaining lncRNAs and their relevance to ICC treatment. The GC treatment group had eight unique genes, and four of them were lncRNAs. One of the unique genes in this group belonged to the TEC gene family, indicating that it may play a role in the T cell response to GC chemotherapy (Berg et al.,2005). The ICB treatment group had a single unique gene, the protein-coding gene Pbld1, which was also included in all treatment groups, while the protein coding genes Rasgef1b and Cyp2d13 were identified in the GC/ICB group and Cyp2d13 was also identified in the GC treatment group's list of unique genes. The identification of highly upregulated lncRNAs is important in the context of cancer treatment, as they are believed to be involved in many cellular processes, including proliferation, apoptosis, and metastasis (Jiang et al.,2019). Further investigation into the functions of these lncRNAs can provide insights into novel regulatory mechanisms involved in the treatment response and resistance in ICC. Additionally, understanding the roles of annotated genes such as Pbld1 and uncharacterized genes such as those in the TEC gene family can contribute to our understanding of the molecular mechanisms driving therapeutic effects of cancer drugs such as GC and ICB on ICC.

Conclusion:

Our research used bioinformatics techniques to analyze RNA-seq data mapped to both the latest mouse genome assembly (GRCm39) and the mm10 assembly, revealing new genes linked to response to GC chemotherapy and ICB in ICC. Following analysis, we found several DEGs and unique genes specific to each treatment group in the GRCm39 alignment, highlighting the specific response of ICC to therapeutic interventions such as GC and ICB. Our findings also reveal the potential roles of

lncRNAs and the annotated genes Pbld1, Cyp2d13, Rasgef1b, and members of the TEC gene family in regulating the immune microenvironment. These results suggest that further experimental validation and functional annotation of the identified genes and pathways will be important in further ICC research.

Experimental Validation:

To confirm the significance of the genes and pathways identified in our analysis and understand their importance in ICC, we suggest the following experimental validations. Firstly, we suggest using quantitative real-time polymerase chain reaction (qRT-PCR) to validate the DEGs identified in each treatment group. This will provide quantitative confirmation of the RNA-seq data and validate the significance of the identified genes. Additionally, we can perform in vivo studies using patient-derived tumor models to evaluate the impact of regulating indicated genes' expression on tumor growth, metastasis, and response to therapy. Finally, we can integrate clinical data from patients undergoing similar treatment regimens to validate the identified biomarkers/genes for treatment.

References:

Berg, Leslie J., Lisa D. Finkelstein, Julie A. Lucas, and Pamela L. Schwartzberg. "TEC FAMILY KINASES IN T LYMPHOCYTE DEVELOPMENT AND FUNCTION*." *Annual Review of Immunology* 23, no. Volume 23, 2005 (April 23, 2005): 549–600. https://doi.org/10.1146/annurev.immunol.22.012703.104743.

Boogaard, Winnie M. C. van den, Daphne S. J. Komninos, and Wilbert P. Vermeij. "Chemotherapy Side-Effects: Not All DNA Damage Is Equal." *Cancers* 14, no. 3 (January 2022): 627. https://doi.org/10.3390/cancers14030627.

Bridgewater, John, Peter R. Galle, Shahid A. Khan, Josep M. Llovet, Joong-Won Park, Tushar Patel, Timothy M. Pawlik, and Gregory J. Gores. “Guidelines for the Diagnosis and Management of Intrahepatic Cholangiocarcinoma.” *Journal of Hepatology* 60, no. 6 (June 1, 2014): 1268–89. https://doi.org/10.1016/j.jhep.2014.01.021.

Buettner, Stefan, Jeroen LA van Vugt, Jan NM IJzermans, and Bas Groot Koerkamp. “Intrahepatic Cholangiocarcinoma: Current Perspectives.” *OncoTargets and Therapy* 10 (February 22, 2017): 1131–42. https://doi.org/10.2147/OTT.S93629.

Chen, Jiang, Zohreh Amoozgar, Xin Liu, Shuichi Aoki, Zelong Liu, Sarah M. Shin, Aya Matsui, et al. “Reprogramming the Intrahepatic Cholangiocarcinoma Immune Microenvironment by Chemotherapy and CTLA-4 Blockade Enhances Anti-PD-1 Therapy.” *Cancer Immunology Research* 12, no. 4 (April 2, 2024): 400–412. https://doi.org/10.1158/2326-6066.CIR-23-0486.

Fulda, Simone, Adrienne M. Gorman, Osamu Hori, and Afshin Samali. “Cellular Stress Responses: Cell Survival and Cell Death.” *International Journal of Cell Biology* 2010 (2010): 214074. https://doi.org/10.1155/2010/214074.

Herr, Ingrid, and Klaus-Michael Debatin. “Cellular Stress Response and Apoptosis in Cancer Therapy.” *Blood* 98, no. 9 (November 1, 2001): 2603–14. https://doi.org/10.1182/blood.V98.9.2603.

“Intrahepatic Cholangiocarcinoma (ICC) | Dana-Farber Cancer Institute.” Accessed April 27, 2024. https://www.dana-farber.org/cancer-care/types/intrahepaticcholangiocarcinoma#.

Jiang, Ming-Chun, Jiao-Jiao Ni, Wen-Yu Cui, Bo-Ya Wang, and Wei Zhuo. “Emerging Roles of lncRNA in Cancer and Therapeutic Opportunities.” *American Journal of Cancer Research* 9, no. 7 (July 1, 2019): 1354–66.

Tanaka, M, H Tanaka, H Tsukuma, A Ioka, A Oshima, and T Nakahara. “Risk Factors for Intrahepatic Cholangiocarcinoma: A Possible Role of Hepatitis B Virus.” *Journal of Viral Hepatitis* 17, no. 10 (October 2010): 742–48. https://doi.org/10.1111/j.13652893.2009.01243.x.

Wei, Spencer C., Colm R. Duffy, and James P. Allison. “Fundamental Mechanisms of Immune Checkpoint Blockade Therapy.” *Cancer Discovery* 8, no. 9 (September 4, 2018): 1069–86. https://doi.org/10.1158/2159-8290.CD-18-0367.