Degree project

UNIVERSITY OF SKÖVDE

STUDY OF UP-REGULATED GENES IN GENE CLUSTERS DURING FORMATION OF MATURE HEPATOCYTES FROM HUMAN INDUCED PLURIPOTENT STEM CELLS TO IDENTIFY TRANSCRIPTION FACTORS AND MIRNAS

Master Degree Project (120 credits) in Systems Biology

Second Cycle 45 credits

Autumn term 2021

Student: Suraj Thomas Alexander

Supervisor: Patric Nilsson, University of Skövde

Examiner: Zelmina Lubovac, University of Skövde

Abstract

The multifunctional purpose of hepatocytes, the functional liver cells within the metabolic, endocrine and secretory functions highlights key importance in emphasizing the research and treatment methods that utilize these cells. Forming 80% of the liver's cells, hepatocytes are involved in many of the primary functions of the liver including the delivery of immune response against pathogens and aiding in the detoxification of drugs. As a result, it provides a valuable basis for medical research. Through the findings of Ghosheh et al. (2017), a method of generating mature hepatocytes was achieved through the human pluripotent stem cells (HPSC), but the generation of hepatocytes in which all the genes are expressed at the right amount through this method proves to be a difficult endeavor. The primary goal of this project is to utilize the established findings to enhance and improve the efficacy of the process that goes behind the generation of mature hepatocytes. The approach towards the current project was initiated with culturing and differentiating three human embryonic stem cell lines and three human-induced pluripotent stem cell lines into mature hepatocytes. In the study mentioned, k-means clustering along with Pearson correlation as the distant measure was run in R to subdivide the top 2000 genes with the highest differential expression into 10 clusters. The cluster data from this paper was used to do the current study, in which the up-regulated and down-regulated gene were first identified for clusters 2, 4 & 6 and 9. The interactions of up-regulated genes in these clusters were further analyzed using Enrichr to identify the different miRNAs for various genes from the clusters. Within cluster 2, a total of 8 genes showed the possibility of being regulated using 4 miRNAs. Transcription factors were also identified for cluster 2 and a combination of HNF1A, EP300, AHR, NFKB1 and HIF1A could repress 8 genes that were not repressed by miRNAs. In cluster 4 & 6, most of the up-regulated genes showcased tumorigenicity and all 20 genes identified could be regulated with the combination of 7 miRNAs. In cluster 9, a combination of 11 miRNAs could be used to regulate 26 out of 27 genes that were analyzed. Ensuring that stem cell products do not turn cancerous is a priority in the medical field. Conducting the analyses of the other clusters aside from 2, 4 & 6 and 9 will prove highly beneficial in reducing the risks pertaining to stem cell mutation due to overexpression of genes.

Popular scientific summary

The prospects of stem cell therapy is heard very often across society. It is conveyed as a treatment of the future, but the fact remains that there is still quite a lot the medical field has to uncover with the subject of stem cells. Aside from difficulties of culturing stem cells and generally ascertaining the optimal way of implementing a treatment, there's a certain risk factor that comes with the cultured cells either promoting or developing the growth of benign or malignant tumors. Although there are clear benefits that have been shown, it has to be evaluated from the perspective of risks to society in order to ensure that the ratio of those two elements is favorable to society. The risk factors at the moment come from the potency of stem cells, the application, and through the manipulation process during vitro culturing; conducting a study in order to understand which methods lead to stem cells turning cancerous, becomes a crucial goal within the medical field.

One of the methods used to study this factor is through the transformation of stem cells (HPSCs) into mature liver cells (hepatocytes). Hepatocytes comprise 80% of the liver's cell mass and are involved in the primary functions of the liver. They are a crucial part of the metabolic, endocrine, and secretory functions. A few of the purposes that hepatocytes serve within the liver include the generation of immunologic response against pathogens, conversion of carbohydrates into fatty acids, detoxification of drugs in the system, and aiding in the homeostasis of glucose. As a result, there are large prospects for the purpose of conducting research and providing treatment. With that being said, the process of converting HPSCs into mature hepatocytes is a difficult one, and specifically with context to tumorigenicity and the genes that are affected. This paper aims to take up the research established by Ghosheh et al. (2017) where 2000 genes were identified and subsequently divided into ten sections known as clusters which contain groups of genes which are similar in function. The focus is on establishing the up-regulated and down-regulated genes, at which point the interaction of the up-regulated genes is analyzed. One of the primary goals of this paper is to ascertain the miRNAs for these for the genes that are within their respective clusters. The miRNAs are non-coding RNA molecule that once identified, can help to regulate gene expression levels.

Throughout the analysis, it was shown how the genes had high expression levels during the process of generation of the mature hepatocytes which act as a trigger to enhance tumorigenicity. Conversely, PAG1 that was highly expressed but favors tumor suppression. Additionally, the paper connects the miRNA to the genes from the four clusters. As a result, it becomes very important to make these distinctions in order to bridge the gap between where stem cell therapy is at the moment, and the potentiality it holds. Currently, stem cells are not at a point where they can be extensively applied into the public sector because of the possibility of causing harm to an individual is larger compared to the benefits in the present moment. This method can also apply to any list of up-regulated genes in general.

Table of Contents

Abbreviations

ESC	Embryonic stem cell
HEP	Hepatocyte
hPSC-HEP	Hepatocyte derived from human pluripotent stem cells
iPSC	Induced pluripotent stem cell
iPSC-HEP	Hepatocyte derived from human induced pluripotent stem cells
NANOG	Nanog Homeobox
OCT3	Organic cation transporter 3

Introduction

Hepatocytes

Hepatocytes are functional liver cells that are involved in metabolic, endocrine and secretory functions. Hepatocytes form 80% of the liver's cell mass and they synthesize proteins like lipoproteins, transferrin and glycoproteins. Hepatocytes are also critical for the homeostasis of glucose (Klover and Mooney, 2004). They help in carbohydrate metabolism by converting carbohydrates into fatty acid (Ali et al., 1863). They are also involved in detoxification of drugs like converting ammonia to urea for excretion. They are considered as a highly predictive in-vitro model for preclinical drug metabolism studies (Lübberstedt et al., 2011). They are also used for cell modelling for drug induced liver injury studies (Goldring et al., 2017). A lot of human pathogens infect hepatocytes but since they interact directly with T cells, they deliver cell-autonomous innate immune responses that can result in host defense (Crispe 2016). There is high demand for hepatocytes because of their use in research and for treatment in the medical field. Alternate methods have been investigated because of this need. Human pluripotent stem cells (HPSC) can be transformed into mature hepatocytes by controlled in-vitro hepatic differentiation (Ghosheh et al., 2017). Some methods to differentiate pluripotent stem cells can be seen in Figure 1 (Rashid & Alexander, 2013). Generating mature hepatocytes derived from human pluripotent stem cells (hPSC-HEP) remains a challenge and this project aims to improve the efficacy of the process. Human embryonic stem cells can divide indefinitely and differentiate into all mature cell types of the human body (Jensen et al., 2009). The first embryonic stem cells were derived from mouse embryos in 1981 in which mouse embryonic stem cells were cultured from embryos in the uterus for increased cell count and then derived from these embryos (Evans & Kaufman, 1981; Martin GR, 1981). By 1998, a technique was developed to isolate and grow human embryonic stem cells in cell culture (James et al., 1998). Induced pluripotent stem cells (iPSC) are pluripotent stem cells that are artificially procured from an adult differentiated somatic cell that is non-pluripotent and is induced by the forced expression of specific genes (Yu et al., 2007; Hockemeyer et al., 2008). Fibroblasts are mostly used for iPSC generation, but they can also be obtained from liver, pancreas β cells and mature B cells (Yu and Thomson, 2008). Even though ESC and iPSC differ in origin, the have highly similar growth characteristics, gene expression profiles, epigenetic modifications and developmental potential (Šarić & Hescheler, 2008; Xu et al., 2009).

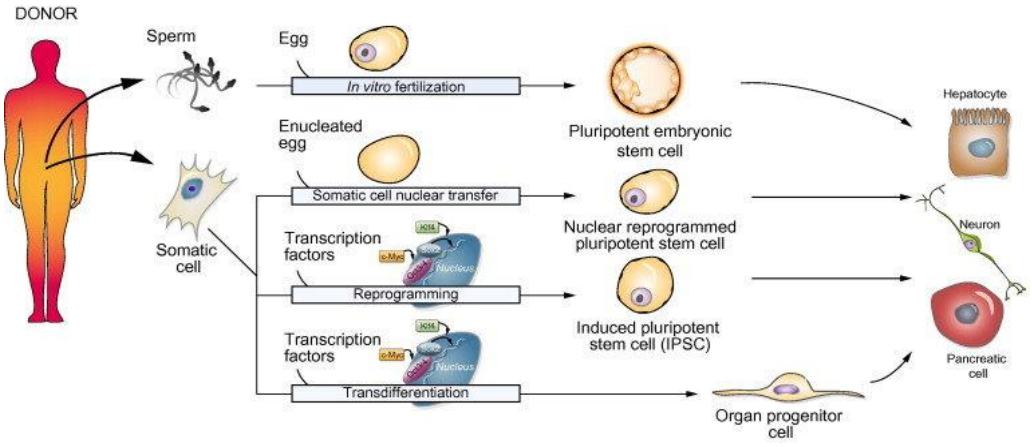

Figure 1: The in-vitro transformation of a pluripotent stem cell to a hepatocyte can be seen in this figure. Other stem cell types and some methods to obtain them can also be seen (Rashid & Alexander, 2013).

Micro RNAs (miRNA)

Micro RNAs or miRNAs were first discovered in nematodes and have been identified to play a key regulatory role in biological processes in animals (Gebert & MacRae, 2019). miRNAs are short non-coding RNAs with a length of 21 to 23 nucleotides that regulate gene expression post transcriptionally by binding to the untranslated region of their target mRNAs which in turn destabilizes the mRNA and induces translational silencing (Cannell et al., 2008). They contribute to a major role in cell growth, proliferation, differentiation, immune response, and apoptosis (Cai et al., 2009). miRNAs can cause diseases like cancer when the gene complimentary to the miRNA is mutated (Farazi et al., 2011). In diseased conditions, miRNA expression levels could be changed due to alterations in the transcriptional or posttranscriptional regulation of miRNA expression (Gommans & Berezikov, 2012). The lack of miRNA in a specific location will lead to genes being up-regulated which has been seen in multiple cancers. miRNAs have the ability to target multiple genes within a signaling pathway which makes them a promising target for drug development (Tiwari et al., 2018). There are multiple websites and online tools available to identify miRNAs for a given list of genes. In this study multiple miRNAs were identified for some up-regulated genes in mature iPSC-HEPs using miRTarBase 2017 in the Enrichr online tool (Hsu et al., 2011; Chen et al., 2013).

Transcription factors

Transcription factors are proteins that have a precise structure that allows them to bind to particular DNA sequences in gene regulatory regions and control their transcription (David Latchman, 1997). These specific sequences of DNA are called enhancer or promoter sequences. Some transcription factors bind to a DNA promoter sequence near the transcription start site and form a transcription initiation complex which also contains RNA polymerase. The RNA polymerase is triggered by the transcription initiation complex and starts mRNA synthesis which leads to transcription of the targeted gene. Other transcription factors bind to regulatory sequences like enhancer sequences and can either stimulate or repress transcription of the related gene (Athanasios G. Papavassiliou, 1995). Transcription factors with their ability to activate and repress genes play a vital role in controlling gene expression but only the transcription factors that repress genes are looked into in this paper. The online tool TRRUST v2 was used to find transcription factors in this study (Han et al., 2018).

Aim

This project aims to extend on the findings proved by Ghosheh et al. (2017) who identified sets of genes which induced expression during various developmental stages during the differentiation of human pluripotent stem cells into mature hepatocytes. In the previous study, the top 2000 differentially expressed genes were clustered with k-means clustering using Pearson correlation as the distant measure in R which resulted in 10 clusters. Genes in cluster 2 were typical for an immature phenotype. CYP1A1 and CYP1B1, for example, are expressed in fetal liver and subsequently down-regulated or silenced in adult liver. These results imply that some maturation genes can be induced, in addition to either incomplete differentiation of human pluripotent stem cells toward mature hepatocytes or failure in turning off transcription of fetal genes, or both. These deviations are going to be checked to improve the functionality of hPSC-HEPs. The top 5 differentially expressed genes in cluster 2 were LRRC19, ISX, AREG, SLC51B, and CYP1A1. Cluster 4 and cluster 6 had very similar expression profiles and the top differentially expressed genes were CXCR4, RP4-559A3.6, MIXL1, EOMES, and LGR5. Cluster 9 included genes that were

expressed in the later stages of hepatic differentiation, and these genes were slightly up-regulated in hPSC-HEPs when compared with liver tissue controls. The top five differentially expressed genes from cluster 9 are AFP, TTR, FGA, FGB, and APOB.

The overall aim with this project was to perform extensive bioinformatics analysis on cluster 2, 4, 6 and 9 using Enrichr to find interactions of genes that were up-regulated in hPSC-HEPs when compared to the control liver tissues in these clusters (Chen et al., 2013). MicroRNAs and transcription factors involved in the regulation of these genes were investigated so that the expression of up-regulated genes in the mentioned clusters can be controlled. Therefore, only up-regulated genes had been studied because miRNAs mainly regulate gene expression and the transcription factors that have been investigated are only the ones that down-regulate gene expression.

Materials and Methods

Experimental background described by Ghosheh et al. (2017)

In the previous work by Ghosheh, three human embryonic stem cell lines Cellartis AS034, SA121, and SA181 and three human induced pluripotent stem cell lines (human iPS cell line ChiPSC6b, P11012, and P11025) were cultured and differentiated into mature hepatocytes using DEF-CS Culture System. The RNA processing and extraction was performed using the MagMAX-96 Total RNA Isolation Kit and quantified by using GeneQuantpro spectrophotometer (Ghosheh et al., 2016). Totally 33720 gene transcripts were obtained when the microarray data was normalized using the robust multiple average normalization method to reduce nonbiological variation (Irizarry et al., 2003). The reproducibility of the differentiation and the microarray experiments were confirmed by clustering the dataset using Pearson correlation as distance measurement and average linkage method in R using the genefilter package (Gentleman et al., 2021). Genes with high differential expression were then identified by applying Significance Analysis of Microarray data using the siggenes package in R (Holger Schwender, 2020).

The top 2000 genes with the most significant differential expression were then clustered using k-means clustering and Pearson correlation as the distance measure using the amap package in R to get 10 clusters (Caussinus et al., 2003). The number of clusters was set to 10 with an arbitrary seed of 1987 after calculating the within-clusters sum of squares and the between-clusters sum of squares and considering a compromise between these two measurements. The datasets have been provided in the supplemental data in the online version of the article mentioned and were used for this project.

Gene expression data analysis

It is to be noted that every time clustering is performed using k-means, the genes in the cluster can differ from the previous iteration. The gene expression levels from Ghoseh's study had been looked into in the current project from the clusters that have been previously established in order to find out which genes were up-regulated in the mature hPSC-HEPs when compared to the control liver tissue. The ClueGo app in Cytoscape was used for the functional annotation of cluster 2, 4, 6 and 9 to show pathways (Bindea et al., 2009). Expression data of the mature hPSC-HEPs from 6 different cell lines and 2 control liver tissue samples was used in this experiment. The mean

of the expression data from the 6 cell lines corresponds to the variable "day30 mean" and the mean of the 2 liver tissue samples was labelled "hlt mean". Gene expression data of mature hPSC-HEPs versus the control liver tissue was plotted in a barplot using Microsoft Excel because the bar plot was more interactive than the one in R. When the averages were compared, it helped identify which gene was up-regulated or down-regulated in the mature hPSC-HEPs when compared to the control liver tissue.

Identification of microRNAs

Enrichr was used to find the miRNAs for the up-regulated genes in the different clusters so that they could be regulated (Chen et al., 2013). Enrichr gave a set of results with different databases and miRTarBase 2017 was chosen from among them. In 2010, the first version of miRTarBase was created to integrate miRNA-target interaction (MTI) studies and their functional roles in different biological processes (Hsu et al., 2011). Since then, the database has had some updates including the latest one in 2020 in which a text mining system was integrated to increase the recognition of MTI related articles by using a scoring system. A high score corresponded to an article that was highly related to MTIs. A couple of biological databases were also incorporated to provide the regulatory network of miRNAs and its expression in blood. Currently, the database has compiled more than13,404 validated MTIs from 11,021 articles from manual curations. With these latest revisions, miRTarBase is one of the most extensively annotated and experimentally validated MTI databases that currently exist (Huang et al., 2020). In Enrichr, under the section miRTarBase 2017, the results were first shown in a bar column and the miRNA with the highest p-value was shown on top. The p-value was calculated from a Fisher exact test which assumes a binomial distribution and independence for the probability of any gene belonging to any set. The "Clustergram" tab was chosen to view the interaction of the miRNAs with the genes inputted in the list. For the parameters, "Row Order" was set to "Cluster" and "Column order" was set to "Sum". For the score, the "Combined Score" was selected as it was a combined score of the p-value from the Fisher exact test and a z-score. The "Top Enriched Terms" parameter represented the miRNAs and could be set between 10, 20 and 30 and this was set to 30. The final parameter was "Top rows sum" which represented the genes that corresponded to the miRNAs, and this was set to "all rows" so that all the maximum number of genes could be displayed. The following parameters were set for all the gene lists inputted so that the miRNAs were arranged from the highest to the lowest combined score which was shown using a red bar over the miRNA. The results can be recreated by using the cluster data in the Enrichr website (Chen et al., 2013).

Identification of transcription factors

Transcription factors (TFs) were investigated for the up-regulated genes in cluster 2 using TRRUST v2 because most of the miRNA that were identified were *Mus musculus* miRNA (Han et al., 2018). The latest version of TRRUST includes 8,444 and 6,552 TF-target regulatory relationships of 800 human TFs and 828 mouse TFs, respectively. These TF-target regulatory relationships were formulated from 11,237 Pubmed articles which report small-scale experimental studies of transcriptional regulations. This was done using a sentence-based text mining approach to search for regulatory interactions from over 20 million Pubmed articles. Data for human genes and mouse genes are included in TRRUST (Han et al., 2018). In the search tab of the TRRUST website, there were two options. The second option is a tool that finds TFs for a list of genes. In this section, the "Species" parameter was set to "human", and the list of up-regulated genes were inputted in the query bar and submitted. In the output, the number of valid genes, invalid genes and the query genes included in TRRUST were shown. The output also included a

table with a list of TFs, their description, the number genes interacting with the TF, the p-value and the false discovery rate. The TFs were arranged in this table depending on the highest p-value. The interaction of the TFs with the genes is opened in a new tab automatically and it contains a table with the TF, the target gene and the mode of regulation (up-regulation or down-regulation), a Pubmed ID reference for the proof of regulation and a final column containing Gene Ontology biological process of the target genes. With the help of this last table, a few TFs were found that could down-regulate the up-regulated genes in cluster 2.

Pathway analysis

The online tool ConsensusPathDB was used to identify networks between the genes that were upregulated in the mature hPSC-HEPs (Kamburov et al., 2011). This was applied on one cluster at a time. ConsensusPathDB uses over 31 public repositories to get information on interactions (Herwig et al., 2016). It contains 20499 unique physical entities and 859848 unique interactions with 5578 different pathways. The tool describes interaction network modules, biochemical pathways and functional information that are significantly enriched by the user's input by using computational methods for enrichment, statistical over-representation, and graph analysis. In the left panel of the website the enrichment analysis option under gene set analysis is chosen. The list of upregulated genes is inputted for each cluster separately and then submitted. In the parameters section, "pathways as defined by pathway databases" was checked and all the databases available were selected, the minimum number of measured genes was set to 4 and the p-value cut off was set to 0.01 before enriching the genes. The output gives the pathway, the number of genes in the pathway, the p-value, and the q-value. There was also an option to visualize the selected pathways once they were selected. All the pathways were eventually derived from the Reactome database (Griss et al., 2020; Jassal et al., 2020). The results from the pathway analysis were discussed briefly with the discussion of this project.

This study could give a broader explanation on which TFs and miRNA interact in the development process of mature hepatocytes from human embryonic stem cells to give mature cells of higher quality. A brief description of the genes (and their implications when up-regulated) that could be regulated by the miRNAs and TFs were also included in the discussion.

Results

Gene expression levels

Excel was used to create a bar plot for each cluster showing the genes that were up-regulated and down-regulated in mature hPSC-HEPs compared to the control liver tissues in cluster 2, 4 & 6, and 9. The blue bars represent the gene expression level of the mature hPSC-HEPs and the red bars show the gene expression levels of the control liver tissues. The genes up-regulated in the hPSC-HEPs were shown before the genes down-regulated in the hPSC-HEPs when compared to the control liver tissue. Figures 2, 3 and 4 represent the average of gene expression levels in mature hPSC-HEPs against the average gene expression levels in the control liver tissues in cluster 2, 4 & 6 and 9 respectively. From the data of the bar plot, the genes that were up-regulated in mature hPSC-HEPs compared to the control tissue were selected so that regulatory factors could be investigated for them. With this, each cluster studied provided a list of genes that could be used as input to different enrichment tools. A brief description of the number of genes that were up-

regulated in the mature hPSC-HEPs compared to the control tissue was provided after Figure 2, 3 and 4.

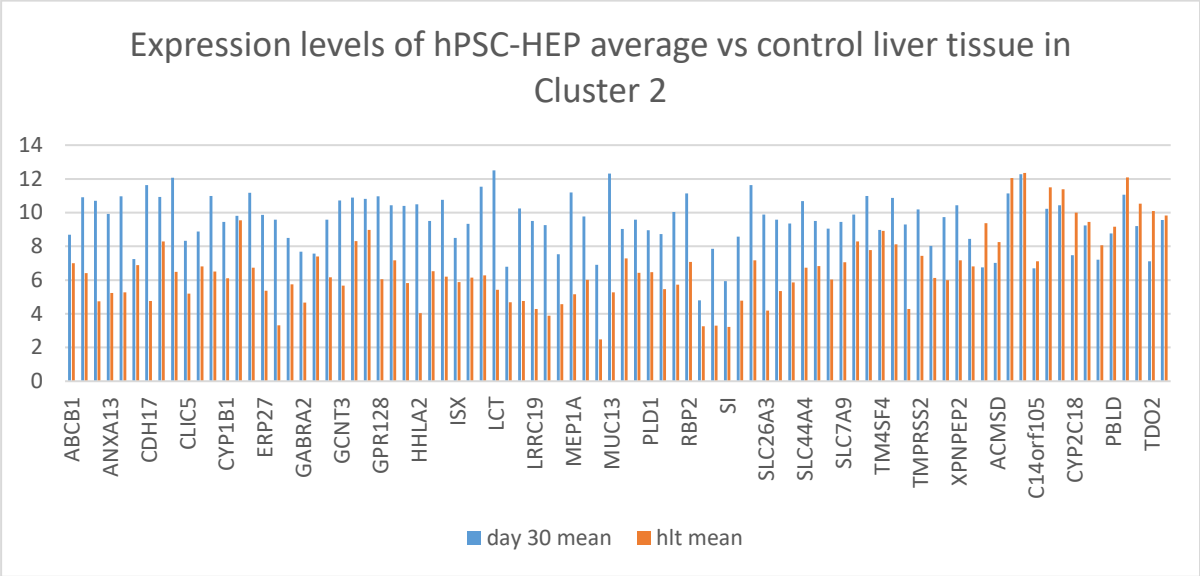

Figure 2: The barplot shows the gene expression levels of the average of the mature hPSC-HEPs (blue) and the control liver tissues (red) for the different genes in cluster 2.

Figure 2 showed that out of a total of 86 genes, 71 genes were up-regulated and 15 genes down-regulated in mature hPSC-HEPs when compared to the gene expression level in the control liver tissue in cluster 2.

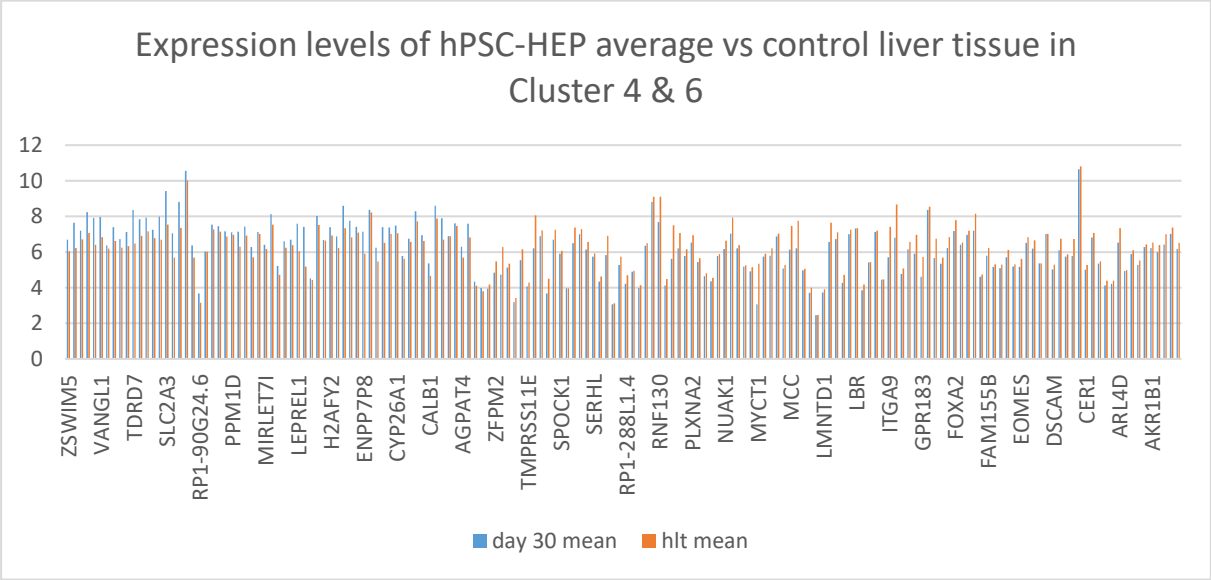

Figure 3: The barplot shows the gene expression levels of the average of the mature hPSC-HEPs (blue) and the control liver tissues (red) for the different genes in cluster 4 & 6.

Figure 3 showed that out of a total of 170 genes, 64 genes were up-regulated, and 106 genes were down-regulated in mature hPSC-HEPs when compared to the gene expression level in the control liver tissue in cluster 4 & 6.

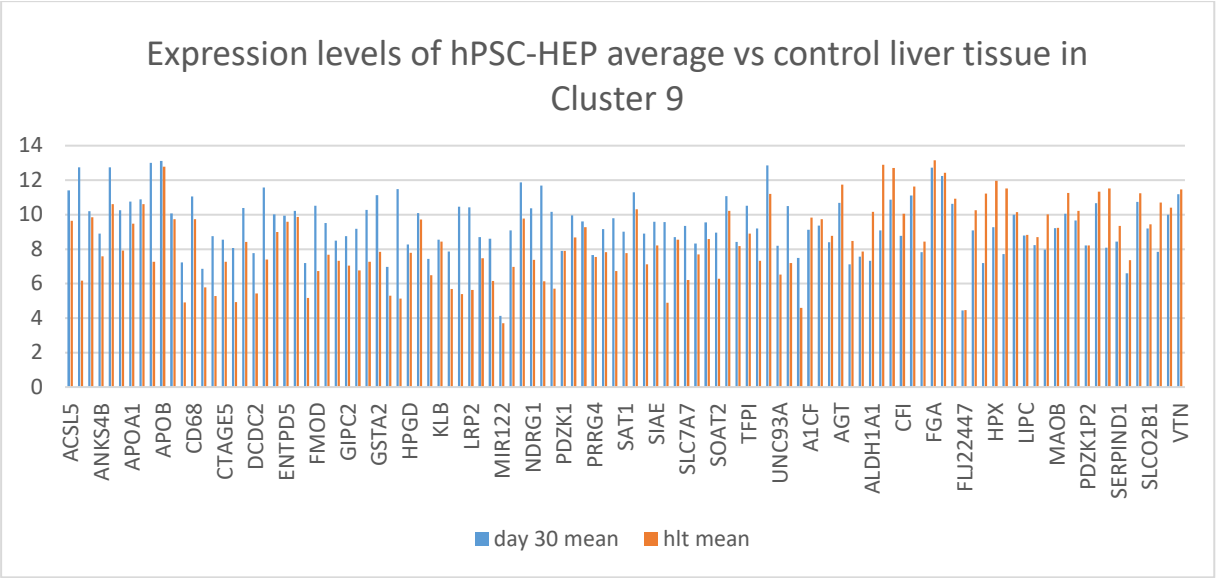

Figure 4: The barplot shows the gene expression levels of the average of the mature hPSC-HEPs (blue) and the control liver tissues (red) for the different genes in cluster 9.

Figure 4 showed that out of a total of 109 genes, 72 genes are up-regulated and 37 genes were down-regulated in hPSC-HEPs when compared to the gene expression level in the control liver tissue in cluster 4 & 6.

Identification of miRNAs

Once the gene list is inputted in Enrichr, miRTarBase 2017 was selected and the view was changed to "Clustergram" for visualization (Chen et al., 2013; Hsu et al., 2011). The results showed *Mus musculus* miRNAs (mmu-mir-*) and also *Homo sapiens* miRNAs (hsa-mir-*). For the rest of this paper, hsa-mir-* will be referred to as miR-* because the only miRNAs that were mentioned in this paper were human miRNAs, as they were relevant to the current study. The hsa-mir-* and their interactions with the genes were noted before each figure. Figures 5, 6 and 7 show the results of the up-regulated genes hPSC-HEPs and the miRNA that can regulate them for clusters 2, 4 & 6 and 9 respectively. Since the figures itself were difficult to interpret, figures were also made using BioRender to visualize the Enrichr results.

Cluster 2

The miRNAs found for the up-regulated genes in cluster 2 had the following interactions. MiR-4539, miR-181a-2-3p, miR-105-p and miR6826-3p regulated 8 genes (PAG1, ABCG2, CYP1B1, CLIC5, TSPAN1, DUOX2, TMEM45B and PRR15L) from the list of 71 genes that was inputted. Since only 4 miRs were found for the up-regulated genes in mature hPSC-HEPs, TFs were also investigated for these genes.

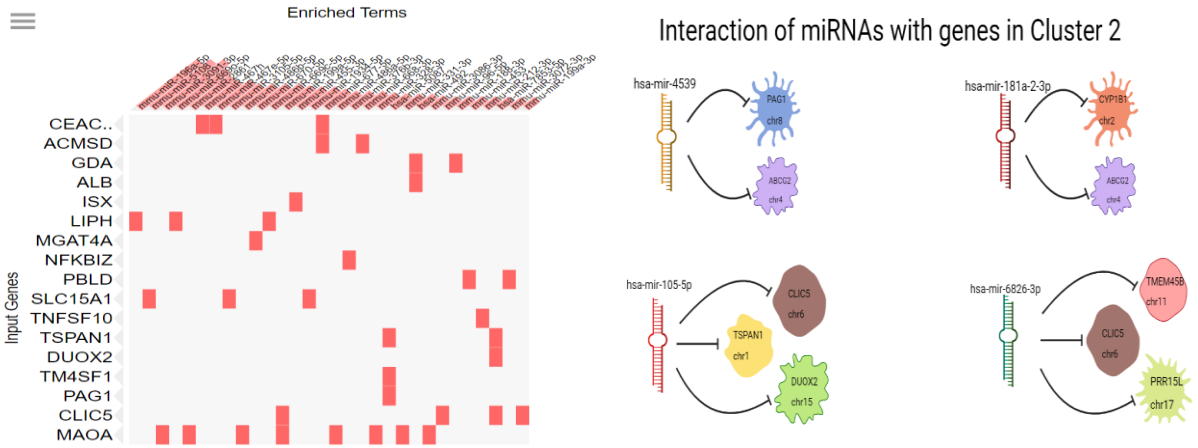

Figure 5: The image on the left is the result from Enrichr using the genes from cluster 2 without down-regulated genes in mature HEP-iPSCs. miRNAs that can regulate genes are shown under "Enriched Terms" and the genes that the miRNAs regulate are shown on the left as "Input Genes" The image on the right is a representation of the miRNA interaction with the up-regulated genes (Reprinted from "Cluster 2 gene-miRNA interaction", by BioRender.com). The chromosome number for each gene is also included as "chr (chromosome number)".

Transcription Factors for cluster 2

In cluster 2, most of the miRNAs that were identified were mouse miRNAs and because of this, transcription factors for this cluster were investigated. A gene list of the up-regulated genes excluding the genes that were in the Enrichr results was used to get transcription factors for the genes using TRRUST v2 (Han et al., 2018). Transcription factors can regulate genes or increase their expression but only regulatory transcription factors and the genes they interact with were investigated in this study. Table 1 shows the list of transcription factors and the different genes that they regulate.

Table 1. This table shows the action of the transcription factors on the query genes included in TRRUST v2 (Han et al., 2018). Transcription factors that could repress the genes of cluster 2 were compiled.

Transcription Factor	**Repression of gene(s)**
HNF1A	TM4SF1
EP300	LAMA3, NR1H4
AHR	CYP1A1
RELA, NFKB1	SERPINA3, TNFSF10
NFIL3	TNFSF10
ESR1	CYP1A1
USF1	CYP1A1
HIF1A	ABCB1, ACE2
EGR1	TNFSF10

Out of the 10 transcription factors provided, HNF1A, EP300, AHR, NFKB1 and HIF1A in combination could repress 8 genes (TM4SF1, LAMA3, NR1H4, CYP1A1, SERPINA3, TNFSF10, ABCB1 and ACE2) in cluster 2. All the genes that are up-regulated in mature hIPSC- HEPs in cluster 2 that can be regulated using miRNAs and transcription factors were shown to favor either the development or progression of various tumors except for PAG1.

Cluster 4 and 6

To regulate 20 genes (BMPR2, FRZB, DCBLD2, GATA6, MYOCD, SERINC3, RHOBTB3, LRIG3, COLEC12, LRP12, PLXNA2, DDHD1, AJAP1, LRIG3, GNAL, RAP2A, DCBLD2, LIX1L, ZEB1, ZFPM2 and DLC1), a combination of miR-125b-1-3p, miR-4760-3p, miR-520f-3p, miR-200b-3p, miR-770-5p, miR-204-3p and miR-335-3p could be used. Most of the genes that have been up-regulated in hPSC-HEPs in cluster 4 & 6 were associated with tumors when up-regulated. The exceptions are BMPR2, GATA6, RHOBTB3 and AJAP1 out of which high expression of RHOBTB3 and AJAP1 did not favor tumorigenicity. These were investigated in the discussion.

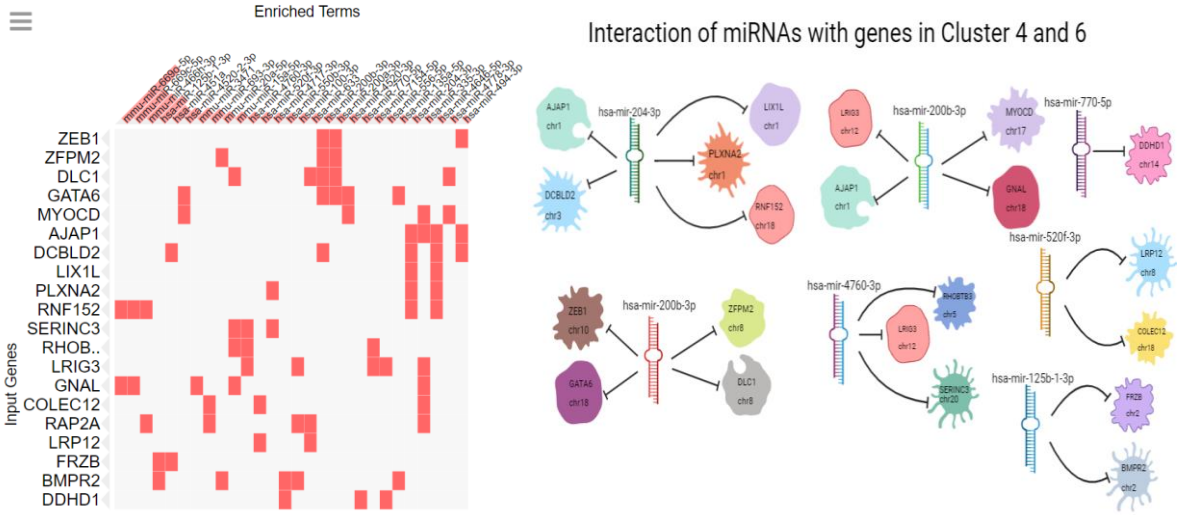

Figure 6: The image on the left is the result from Enrichr using the genes from cluster 4 & 6 without down-regulated genes in mature HEP-iPSCs. miRNAs that can regulate genes are shown under "Enriched Terms" and the genes that the miRNAs regulate are shown on the left as "Input Genes" The image on the right is a representation of miRNA interaction with the up-regulated genes using BioRender (Reprinted from "Cluster 4 & 6 gene-miRNA interaction", by BioRender.com). The chromosome number for each gene is also included as "chr (chromosome number)".

Cluster 9

All of the genes in cluster 9 that were up-regulated in hPSC-HEPs either promoted or caused tumorigenicity in different cells except for TFPI and APOB which when up-regulated repressed cancers. These were investigated in the discussion. MiR-3689f , miR-500b-5p, miR-494b-3p, miR-2681, miR-146a-5p, miR-6514-5p, miR-1291-3p, miR-200a-3p, miR-292a-5p, miR-1248 and miR-4503 could be used to regulate 26 out of 27 genes (THRB, SLC7A7, SGMS2, CLIC6, PRRG4, DCDC2, IL1RAP, PROS1, CFTR, TFPI, APOB, TTR, APOA1, ENTPD5, LRP2, EGFR, SPTLC3, ANKS4B, MAF, KLB, SLC35D1, XKRX, FRK, RNF128, UNC93A and NDRG1) in Figure 5 except for AFP for which a mouse miRNA is the only interaction from the given list.

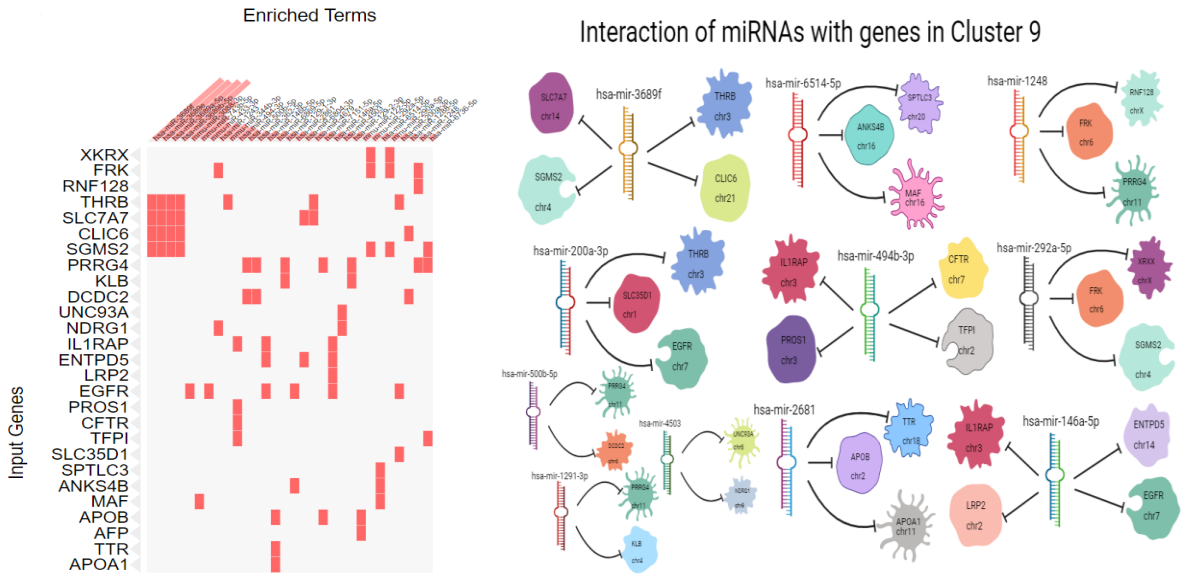

Figure 7: The figure on the left is the result from Enrichr using the genes from cluster 9 without down-regulated genes in mature HEP-iPSCs. miRNAs that can regulate genes are shown under “Enriched Terms” and the genes that the miRNAs regulate are shown on the left as “Input Genes” The image on the right is a representation of miRNA interaction with the up-regulated genes using BioRender (Reprinted from “Cluster 9 gene-miRNA interaction”, by BioRender.com). The chromosome number for each gene is also included as “chr (chromosome number)”.

Discussion

Genes that can be regulated in cluster 2 using miRNAs

The up-regulated genes from cluster 2 that could be regulated with the miRNAs found were discussed in this section. The ubiquitously expressed transmembrane adaptor protein phosphoprotein membrane anchor with glycosphingolipid microdomains 1 (PAG1) acts as a tumor suppressing gene (Svec, 2009). It was up-regulated in mature hPSC-HEP and can be regulated using miR-4539. Another gene that can be regulated using miR-4539 was ABCG2. The ATP-binding cassette transporter G2 (ABCG2) gene had been involved in clinical multi drug resistance in cancer. It excretes endogenous and exogenous substrates on its localized cellular plasma membrane in cancer cells and in normal tissues. It also protects the tissue against xenobiotics (Todoya et al., 2019). MiR-181a-2-3p also regulates the activity of ABCG2. The metabolism of various xenobiotics was carried out by the gene CY1PB1, and it was seen to be expressed in and extrahepatic tissues. High frequency of CYP1B1 in multiple cancers suggest that its modulation can decrease tumorigenesis and prevent cancer (Li et al., 2017). In this study CYP1B1 had to be regulated because it was up-regulated in the mature hPSC-HEPs, and this can be done using miR-181a-2-3p. The genes CLIC5, TSPAN1 and DUOX2 can be regulated by miR-105-p. In hepatocellular carcinoma, chloride intracellular channel 5 (CLIC5) had been found to be up-regulated and its inhibition resulted in decreased migration and invasion (Flores-Tellez et al., 2015). The protein coding gene, tetraspanin 1 (TSPAN1), when up-regulated caused cancer in pancreas and its downregulation contributed to reduce migration and invasion (Wang et al., 2021). In several types of cancers including pancreatic and colorectal cases, the gene dual oxidase 2 (DUOX2) had been shown to be highly expressed (Cao et al., 2021; Burgueño et al., 2021). Upregulation of transmembrane 45b (TMEM45B) caused multiple types of cancer including prostate, gastric and also tumorigenesis in osteogenic carcinoma and downregulation can be associated with alleviated conditions (Luo et al., 2018; Shen et al., 2018; Li et al., 2017). Also referred to as ATAD4 Proline-rich protein 15-like protein (PRR15L) and not much was known about this gene, but it had been associated with breast carcinoma and osteosarcoma (Sheils et al., 2021).

Genes that can be regulated in cluster 2 using transcription factors

The up-regulated genes from cluster 2 that could be regulated with the transcription factors found were discussed in this section. Belonging to the tetraspanin family, transmembrane 4 L six family member 1(TM4SF1) and its high expression levels were associated with ovarian cancer and other epithelial cancers (Gao et al., 2019; Tang et al., 2020). The TF HNF1A can repress TM4SF1. The gene laminin subunit α3 (LAMA3) had been associated with ovarian cancer when mutated and down-regulated (Feng et al., 2021). In the data provided, since LAMA3 was up-regulated, it can be regulated with the transcription factor EP300. The TF EP300 can also regulate the gene NR1H4 which encodes a nuclear receptor called farsenoid X receptor which controls bile acid homeostasis and regulates several metabolic pathways that are essential to liver energy balance (Preidis et al., 2017). The production of reactive oxidative species is modulated by the gene cytochrome P450 1A1 (CYP1A1) and it should be regulated as the oxidative stress plays an important role as hepatic liver peroxidation progresses with its overexpression (Huang et al., 2018). The transcription factor ESR1 can regulate the expression of CYP1A1. When up-regulated, the gene serpin peptidase inhibitor, clade A member 3 (SERPINA3) promotes hepatocellular carcinoma and the factors

associated with it (Ko et al., 2019). A compound called ELFDO can be used to sensitize liver cancer cells to tumor necrosis factor superfamily member 10 (TNFSF10) which help induce apoptosis in the liver cancer cells (Qu et al., 2019). The TFs RELA and NFKB1 regulate the expression of SERPINA3 and TNFSF10. When up-regulated, ATP binding cassette subfamily B member 1 (ABCB1) had been seen to promote multidrug resistance in cancers (Feng et al., 2020). The gene angiotensin-converting enzyme 2 (ACE2) was involved in the modulation of the renin-angiotensin system and blood pressure and was a molecular receptor for SARS-CoV (Wu et al., 2020). The genes ABCB1 and ACE2 can be down-regulated by the TF EGR1.

Genes that can be regulated in cluster 4 & 6 using miRNAs

The up-regulated genes from cluster 4 & 6 that could be regulated with the miRNAs were discussed in this section. The variants of bone morphogenetic protein receptor 2 (BMPR2) when down-regulated plays a critical role in the development of pulmonary artery hypertension (Song et al., 2020). The gene frizzled-related protein (FRZB) was up-regulated in patients with hepatocellular carcinoma (Huang et al., 2015). Since BMPR2 and FRZB was up-regulated in the mature HEP-iPSCs they can be regulated with miR-125b-1-3p. The protein coding gene discoidin, CUB and LCCL domain containing 2 (DCBLD2) was up-regulated in colorectal cancer and lung cancer (Pagnotta et al., 2013; Koshikawa et al., 2002). The genes DCBLD2 and FRZB can be regulated by miR-451-a. The gene GATA binding protein 6 (GATA6) belongs to a family of zinc finger transcription regulators and was involved in various stages of development of the liver (Zhang & He, 2018). The gene MYOCD encodes for the protein myocardin, and its increased expression can lead to the activation of non-small cell lung cancer cells (Tong et al., 2020). The genes GATA6 and MYOCD can be regulated by miR-4520-2-3p since they are up-regulated in the mature HEP-iPSCs. The genes serine incorporator 3 (SERINC3) along with SERINC5 was highly expressed in human HIV-1 target cells (Usami et al., 2015). The gene rho-related BTB domain-containing protein 3 (RHOBTB3) belongs to the RHOBTB subfamily of the rho family GTPases and high expression of RHOBTB3 favored the overall survival of patients with non M3 acute myeloid leukemia (Yang et al., 2021). The gene leucine rich repeats and immunoglobulin like domains 3 (LRIG3) interacts with circular RNA to form circ-LRIG which when highly expressed, promotes hepatocellular carcinoma (Sun et al., 2020). The genes SERINC3, RHOBTB3 and LRIG3 can be regulated with miR-4760-3p. High expression of COLEC12 in osteosarcoma patients increased tumor weight, migration and invasion of the tumor (Li et al., 2020). The gene LDL Receptor related protein 12 (LRP12) was up-regulated in oral squamous cell carcinomas (Garnis et al., 2004). The genes COLEC12 and LRP12 can down-regulated by miR-520f-3p. The protein coding gene plexin A2 (PLXNA2) was up-regulated in the cerebellum of patients with schizophrenia (Mah et al., 2006). The genes SERINC3 and PLXNA2 can be down-regulated by miR-4717-3p. The protein coding gene DDHD domain containing 1 (DDHD1) was up-regulated in colorectal cancers (Raimondo et al., 2018). The genes BMPR2 and DDHD1can be down-regulated using miR-550b-3p. The gene ras-related protein rap-2a (RAP2A) was highly expressed in breast cancer cells and tissues (Liu et al., 2020). The genes LRIG3, RAP2A and BMPR2 can be down-regulated by miR-100-3p. The gene deleted in liver cancer-1 (DLC1) was up-regulated in multiple cancers and acts as a potential tumor suppressor (Zhang & Li, 2020). The genes DLC1, RAP2A and LRP12 can be down-regulated with miR-663. The gene zinc finger E-box binding homeobox 1 (ZEB1) was up-regulated in multiple cancers and favored their migration, invasion and metastasis (Caramel et al., 2018). The gene zinc finger protein, FOG family member 2 (ZFPM2) when mutated in its long coding RNA forms ZFPM2 antisense RNA 1 (ZFPM2-AS1) which when up-regulated, promoted gastric carcinogenesis and hepatocellular carcinoma (Kong et al., 2018; Liu et al., 2020). The genes

ZEB1, ZFPM2, DLC1, GATA6 and DCBLD2 can be down-regulated using miR-200a-3p. The gene DDHD1 can be down-regulated by miR-770-5p. In hepatocellular carcinoma and glioblastoma multiforme the gene adherens junctions-associated protein 1 (AJAP1) was down-regulated (Han et al., 2017; Di et al., 2018). The gene limb and CNS expressed 1 like (LIX1L) was up-regulated in multiple cancers and promoted the progression of hepatocellular carcinoma (Zou et al., 2021). When up-regulated, the gene ring finger protein 152 (RNF152) inhibited colorectal cancer cell proliferation (Cui et al., 2018). The genes AJAP1, DCBLD2, LIX1L, PLXNA2 and RNF152 can be down-regulated using miR-204-3p. The gene g protein subunit alpha L (GNAL) when mutated, caused various forms of dystonia (Adam et al., 2019). The genes MYOCD, AJAP1, LRIG3, GNAL, COLEC12 and RAP2A could be down-regulated with miR-335-3p.

Genes that can be regulated in cluster 9 using miRNAs

The up-regulated genes from cluster 9 that could be regulated with the miRNAs found were discussed in this section. The gene thyroid hormone receptor β (THRB) primarily produces thyroid hormones but was also important for the normal functioning of the adult liver (Ortiga-Carvalho et al., 2014). The gene, mutation in solute carrier family 7 member 7 (SLC7A7), caused lysinuric protein intolerance and its overexpression played a key role in the development of T-cell acute lymphoblastic leukemia in patients (Sperandeo et al., 2008; Ji et al., 2018). When the gene sphingomyelin synthase 2 (SGMS2) was up-regulated in esophageal squamous cell carcinoma it counteracted the effect of THAP9 antisense RNA 1 (THAP9-AS1) (Pan et al., 2021). The gene chloride intracellular channel 6 (CLIC6) was up-regulated in breast cancer in a specific group of patients studied (Li et al., 2020). The genes THRB, SLC7A7, SGMS2 and CLIC6 can be down-regulated with miR-3689f. When up-regulated in breast cancer patients, the gene proline rich γ-carboxyglutamic acid protein 4 (PRRG4) promoted metastasis (Zhang et al., 2020). Biallelic mutations in doublecortin domain containing 2 (DCDC2) caused neonatal sclerosing cholangitis in a proband who had the condition (Lin et al., 2020). miR-500b-5p down-regulates the genes PRRG4 and DCDC2. The gene interleukin 1 receptor accessory protein (IL1RAP) was up-regulated in hematopoietic stem cells of patients with acute myeloid leukemia (Barreyro et al., 2012). When up-regulated in oral squamous cell carcinoma, the gene protein S (PROS1) regulated AXL receptor tyrosine kinase (AXL) and increased tumorigenicity (Abboud-Jarrous et al., 2017). When mutated, CF transmembrane conductance regulator (CFTR) gene caused cystic fibrosis (Lopes-Pacheco et al., 2016). When the tissue factor pathway inhibitor (TFPI) gene was inhibited in patients with haemophilia, functional hemostasis was restored (Peterson et al., 2016). The genes IL1RAP, PROS1, CFTR and TFPI can be down-regulated with miR-494b-3p. When inactivated the apolipoprotein B (APOB) gene favored the tumorigenicity of hepatocellular carcinoma (Lee et al., 2018). When mutated, the gene transthyretin (TTR) was expressed in large amounts in the liver that led to familial amyloid polyneuropathy (Niemietz et al., 2015). When regulated by tripartite motif (TRIM) family member TRIM15, the gene apolipoprotein A1 (APOA1) increased the invasion and metastasis of pancreatic cancer cells (Sun et al., 2021). The genes APOB, TTR and APOA1 can be down-regulated using miR-2681. The overexpression of the ectonucleoside triphosphate diphosphohydrolase 5 (ENTPD5) gene favored the progression of lung cancer by regulating Caspase expression (Xue et al., 2015). The gene LDL receptor related protein 2 (LRP2) was down-regulated to block the c-Jun N-terminal kinases signaling pathway to induce apoptosis in thyroid cancer cells (He et al., 2020). The overexpression of epidermal growth factor receptor (EGFR) was associated with poor prognosis in patients with non-small cell lung cancer (Ohsake et al., 2000, Selvaggi et al., 2004). The genes IL1RAP, ENTPD5, LRP2 and EGFR can be down-regulates using miR-146a-5p. High expression of serine palmitoyltransferase long chain base subunit 3 (SPTLC3)

mRNA in a nonalcoholic steatohepatitis mouse model was associated with development of hepatocellular carcinoma (Yoshimine et al., 2015). The gene ankyrin repeat and sterile alpha motif domain containing 4B (ANKS4B) was identified to be one of the genes that were significantly expressed in ovarian carcinoma (Wang et al., 2019). In tissues with esophageal cancer MAF BZIP transcription factor (MAF) expression was significantly increased in macrophages due to interferon-γ (Takeya et al., 2019). The genes SPTLC3, ANKS4B and MAF can be down-regulated with miR-6514-5p. The klotho beta (KLB) gene regulates multiple metabolic systems in the liver and treatment with exogenous protein of KLB with human bladder cancer cell lines increased their proliferation, migration and growth (Hori et al., 2016). The genes PRRG4 and KLB can be down-regulated using miR-1291-3p. The solute carrier family 35 member D1 (SLC35D1) gene was found to be a key gene in the development of ulcerative colitis-associated colorectal cancer along with the transcription factor TEF3 (Zhang et al., 2021). The genes SLC35D1, EGFR and THRB can be down-regulated with miR-200a-3p. The expression of the XK related x-linked (XKRX) gene was significantly higher in colorectal cancer tissues than in normal tissues (Pan et al., 2018). The activity of fyn-related kinase (FRK) in malignant tumors still needs to be researched but it was highly expressed in non-small cell lung cancer and provides stemness by induction of metabolic reprogramming (Zhang et al., 2020). The genes XKRX, FRK and SGMS2 can be down-regulated using miR-292a-5p. The overexpression of ring finger protein 128 (RNF128) in hepatoma cells enhanced their proliferation, migration and invasion (Bai et al., 2020). The genes FRK, RNF128 and PRRG4 can be down-regulated with miR-1248. Elevated expression of unc-93 homolog A (UNC93A) was associated with worse overall survival of patients with high-risk neuroblastoma (Ognibene et al., 2020). The n-myc downstream regulated 1 (NDRG1) gene was highly expressed in tissues of patients with inflammatory breast cancer (Villodre et al., 2020). The genes UNC93A and NDRG1 can be down-regulated using miR-4503.

Pathway analysis

The online tool ConsensusPathDB was used to find networks between the genes that were upregulated in the mature hPSC-HEPs (Kamburov et al., 2011). The top two pathways, the number of genes and the genes in common were noted for the up-regulated genes of each cluster. In cluster 2, there were 14 hits for metabolism pathway and 10 hits for the pathway of transport of small molecules with the genes SLC44A3, SLC44A4 and ABCB1 in common. In cluster 4&6 there was 10 hits for the signal transduction pathway and the metabolism pathway with the gene CYP26A1 in common. In cluster 9, there was 23 hits for the metabolism and 12 hits for the pathway of metabolism of proteins with the genes APOA1, APOA2 and APOB in common. It was noteworthy that the metabolism pathway was in the top two hits for clusters 2, 4 & 6 and 9.

Conclusion

During this study, 244 genes that were up-regulated in mature hPSC-HEPs when compared to the control liver tissue were identified. From these 244 up-regulated genes, 10 TFs were found that could regulate 8 genes and 22 miRNAs were identified that could regulate 54 genes. Out of the 62 genes that could be regulated, 57 genes promoted tumorigenicity when they were up-regulated and 5 genes (PAG1, RHOBTB3, AJAP1, TFPI and APOB) had no effect or aided in the reduction of tumorigenicity on up-regulation. The up-regulation of genes that led to tumors was a major motivation to find regulatory factors for these genes and could help in preventing hPSC-HEPs from turning cancerous. A pathway analysis on the upregulated genes showed that in all the clusters studied, the metabolism pathway was in the top two hits. In the aim of this experiment the top 5 differentially expressed genes in each cluster that was investigated in Ghosheh's study was

mentioned. In cluster 2, a miRNA was identified for the gene ISX but it was not mentioned because it was a *Mus musculus* miRNA and 3 different transcription factors (AHR, ESR1 and USF1) that interact with CYP1A1 were identified. In cluster 4 & 6, no miRNAs could be identified for the top differentially expressed genes mentioned. In cluster 9 however, the miRNA miR-2681, was identified for TTR and APOB. Since all the results obtained in this paper were purely analytical, they need to be validated with wet lab experiments for their efficacy in an experimental setup.

Ethical aspects, gender perspectives, and impact on the society

Stem cell based medicinal products are used in the differentiation status, proliferation capacity, in-vitro culture, and many other aspects, all of which involve risk factors which question the safe application of stem cell therapy. A vast majority of clinical trials focused on mesenchymal stem cells (MSC) in regenerative medicine applications has not specified any major health concerns which suggest it could be an alternative (Herberts et al., 2011). A risk factor is defined as a potential source of harm (ISO 31000:2018). Since the variety of risk factors in stem cell based medicinal products is large, the risks can differ widely as well. Because of the same, all important identified risks and potential risks should be determined extensively (European Medicines Agency). The risk factors include risks of tumor formation, risk of immune response in allogeneic stem cell transplantation, risk of transmission of a human pathogen and external agents and potential risk factors that have not been accounted because they have not been found so far (Herberts et al., 2011). Stem cells and tumor cells are similar in lifespan, apoptosis resistance and the ability to proliferate over a prolonged time period (Werbowetski et al., 2009). Identical growth regulators and control mechanisms are associated with both cancer and stem cell maintenance (Li et al., 2006).

The potency of a stem cell is one of the fundamental factors for the risk of tumor formation. Intrinsic factors such as site of administrations and extrinsic factors like manipulation of the stem cells during in vitro culturing can also trigger tumor formation. Benign teratomas and malignant teratocarcinomas have been observed after treatment with human ESCs or mouse iPSCs (Shih et al., 2007). Many genes that are used to synthesize iPSC are either indisputable oncogenes such as Myc (MYC proto-oncogene) and KLF4 (Kruppel like factor 4) (PS Knoepfler, 2008; Wei et al., 2005) or are in various ways linked to tumorigenesis such as Sox2, Nanog, and Oct3 (Chen et al., 2008; Chiou et al., 2008; Palma et al., 2008). Myc expression is not only shared between stem and tumor cells, but distinct groups of Myc regulated target genes that are co-expressed in both malignant tumors and ESC (Ben et al., 2008). In a major paper about iPSCs, it had been reported that the creation of an iPSC without the presence of Myc was not possible (Takahashi and Yamanaka, 2006). Another master stem cell regulator KLF4 was studied, and it could be seen that lowered levels affected ESC pluripotency and self-renewal which forced ESC to differentiate (Jiang J et al., 2008).

When it comes to the gender perspective, there is a lot of controversy and abuse relating to the donation of female body tissues for stem cell research (SCR). There have been acts of cross border in-vitro fertilization between Eastern and Western Europe in which poor women were manipulated or bribed for in-vitro fertilization, extraction of research ova and/or to generate research embryos (Catherine Waldby, 2008). There are many social and environmental factors that impact the public opinion on SCR. Many countries have their own rules and regulations on what can be donated and how much research is acceptable. In a study conducted in the

Netherlands, the Society for Gender and Technology asserted that there was a severe psychological impact on women who donated their egg cells when compared to men donating their sperm. Research into the prevention of infertility could be conducted instead of pursuing embryo research as women´s health interests could be supported (Marta Kirejczyk, 2008). In a country like the UK where ESR is promoted, it is allowed to create an embryo just for research purposes and the opinion of scientists from two different laboratories was taken as an example. A socio-ethical study was conducted in the UK in which 15 scientists who work with hESCs were interviewed about their ethical opinion regarding their work. All of them were prepared to use spare embryos were used for research whereas most of them were against the creating an embryo just for research (Wainwright et al., 2006).

Because of the diversity of the risks involved with dealing with ESCs for medicinal use, methods have been made so that the process has reduced tumorigenicity. One of the ways to slow or even eliminate the tumorigenicity in normal stem cells prior to transplantation may be to use their pluripotency by partially differentiating them into progenitors. Stem cells can be used to produce progenitor or precursor cells of the desired lineage and then the transplant progenitors could be purified by sorting. The sorting could be either positive (sorting for the progenitors based on markers) or negative (sorting based on stem cell markers for their elimination). On attaining adequate purity by weeding out through differentiation combined with sorting possibly all contaminating stem cells that remain, the progenitor transplant should be both dependable and potent. Since differentiation is a dynamic process, the presence of residual stem cells can be observed in differentiated cultures which makes sorting imperfect (PS Knoepfler, 2009).

Thymidine kinase (tk) which acts as a suicide gene when genetically integrated into stem cells has been reported to be effective in combination with Ganciclovir (Gan) treatment (Schuldiner et al., 2003). The treatment in this study was not stem cell specific and would have also killed any differentiated progeny from those stem cells in a hypothetical treatment situation causing it to fail. Differentiated teratoma cells were readily killed by Gan treatment. Using oct3 or nanog promoter driven expression of tk, would make the system ideally kill only those hESC that have escaped differentiation. One of the concerns was that if small, but functionally relevant subpopulations of hESC may not express key stem cell factors like oct3 and nanog even though most major hESC do so. Even with efficient stem cell killing, the possibility of patients requiring life-long treatment with Gan or other agents to contain the growth of residual stem cells questions the possibility of the re-emergence of proliferating, drug-resistant stem cells which could in turn lead to tumors at later dates. A study was conducted on the safety of viral transduction of human hematopoietic stem cells in which and MSCs in which the progress in the animals was followed for 18 months and no evidence of tumorigenesis was found which implied that introduction of a single suicide gene with limited genetic modification may be safe. This addresses a major concern which takes into consideration the risk of tumorigenicity when genetically modifying stem cells (Bauer et al., 2008). It is possible to design viral vectors so that they integrate with a high frequency at specific non-coding genomic sites at regions distant from the genes so that there is enhanced safety.

Therefore, the most practical approach to safe regenerative medicine could be an incorporation of differentiation and sorting with the introduction of a stem cell specific suicide gene or using a non-genetic method directed at killing residual stem cells.

A proline rich nuclear protein named undifferentiated embryonic stem cell transcription factor 1 (UTF1) has been shown to have a role in cell differentiation in pluripotent and germ cells and also

acts as either an oncogene or a tumor suppressor gene in cancer cells. It can also additionally as a biomarker to identify iPSCs (Raina et al., 2021).

Future perspectives

Out of the 10 clusters that were provided in Ghosheh et al., 2017, only 2, 4 & 6 and 9 were looked into in this paper. The genes in cluster 10 seemed to have a huge difference in expression levels between the mature hPSC-HEPs and control liver tissues and could be looked into using the same methodology in this paper. An alternate way to do the same methodology provided would be to run Enrichr multiple times with the same gene list. With every iteration, the genes for which the miRNAs have been found can be removed. This would make the resulting iteration have a whole new set of genes and in this way the entire cluster can be covered. Identification of miRNAs and transcription factors will prove vital in regulating the right genes which could help prevent stem cell products from turning cancerous.

Acknowledgments

I would like to thank my supervisor Patric Nilsson for all his help and guidance throughout the duration of my work. Thanks to Jane Synnergren for all her help and input which helped shape the outline for this project. My examiner Zelmina Lubovac also introduced me to some interesting concepts about clustering which I could not implement but made me understand the concept much better. A special thanks to Nidal Ghosheh for clearing doubts about the paper and providing assistance with the same.

References

Abboud-Jarrous, G., Priya, S., Maimon, A., Fischman, S., Cohen-Elisha, M., Czerninski, R., & Burstyn-Cohen, T. (2017). Protein S drives oral squamous cell carcinoma tumorigenicity through regulation of AXL. Oncotarget, 8(8), 13986.

Adam, M. P., Ardinger, H. H., Pagon, R. A., Wallace, S. E., Bean, L. J. H., Stephens, K., & Amemiya, A. DYT-GNAL--GeneReviews®.

Ali, E. S., Hua, J., Wilson, C. H., Tallis, G. A., Zhou, F. H., Rychkov, G. Y., & Barritt, G. J. (2016). The glucagon-like peptide-1 analogue exendin-4 reverses impaired intracellular Ca2+ signalling in steatotic hepatocytes. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1863(9), 2135-2146.

Bai, X. S., Zhang, C., Peng, R., Jiang, G. Q., Jin, S. J., Wang, Q., ... & Bai, D. S. (2020). RNF128 Promotes Malignant Behaviors via EGFR/MEK/ERK Pathway in Hepatocellular Carcinoma. OncoTargets and therapy, 13, 10129.

Barreyro, L., Will, B., Bartholdy, B., Zhou, L., Todorova, T. I., Stanley, R. F., ... & Steidl, U. (2012). Overexpression of IL-1 receptor accessory protein in stem and progenitor cells and outcome correlation in AML and MDS. Blood, The Journal of the American Society of Hematology, 120(6), 1290-1298.

Bauer, G., Dao, M. A., Case, S. S., Meyerrose, T., Wirthlin, L., Zhou, P., ... & Nolta, J. A. (2008). In vivo biosafety model to assess the risk of adverse events from retroviral and lentiviral vectors. Molecular therapy, 16(7), 1308-1315.

Ben-Porath, I., Thomson, M. W., Carey, V. J., Ge, R., Bell, G. W., Regev, A., & Weinberg, R. A. (2008). An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. Nature genetics, 40(5), 499.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., ... & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics, 25(8), 1091-1093.

Burgueño, J. F., Fritsch, J., González, E. E., Landau, K. S., Santander, A. M., Fernández, I., ... & Abreu, M. T. (2021). Epithelial TLR4 signaling activates DUOX2 to induce microbiota-driven tumorigenesis. Gastroenterology, 160(3), 797-808.

Cai, Y., Yu, X., Hu, S., & Yu, J. (2009). A brief review on the mechanisms of miRNA regulation. Genomics, proteomics & bioinformatics, 7(4), 147-154.

Cannell, I. G., Kong, Y. W., & Bushell, M. (2008). How do microRNAs regulate gene expression?. Biochemical Society Transactions, 36(6), 1224-1231.

Cao, M., Zhang, P. B., Wu, P. F., Chen, Q., Ge, W. L., Shi, G. D., ... & Jiang, K. R. (2021). DUOX2 As a Potential Prognostic Marker which Promotes Cell Motility and Proliferation in Pancreatic Cancer. BioMed Research International, 2021.

Caramel, J., Ligier, M., & Puisieux, A. (2018). Pleiotropic roles for ZEB1 in cancer. Cancer research, 78(1), 30-35.

Caussinus, H., Hakam, S., & Ruiz-Gazen, A. (2003). Projections révélatrices contrôlées: Groupements et structures diverses. Revue de statistique appliquée, 51(1), 37-58.

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., ... & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC bioinformatics, 14(1), 1-14.

Chen, Y., Shi, L., Zhang, L., Li, R., Liang, J., Yu, W., ... & Shang, Y. (2008). The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. Journal of Biological Chemistry, 283(26), 17969-17978.

Chiou, S. H., Yu, C. C., Huang, C. Y., Lin, S. C., Liu, C. J., Tsai, T. H., ... & Lo, J. F. (2008). Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. Clinical cancer research, 14(13), 4085-4095.

Crispe, I. N. (2016). Hepatocytes as immunological agents. The Journal of Immunology, 196(1), 17-21.

Cui, X., Shen, W., Wang, G., Huang, Z., Wen, D., Yang, Y., ... & Cui, L. (2018). Ring finger protein 152 inhibits colorectal cancer cell growth and was a novel prognostic biomarker. American journal of translational research, 10(11), 3701.

Di, C., Mladkova, N., Lin, J., Fee, B., Rivas, M., Chunsheng, K., ... & Adamson, D. C. (2018). AJAP1 expression modulates glioma cell motility and correlates with tumor growth and survival. International journal of oncology, 52(1), 47-54.

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. nature, 292(5819), 154-156.

Farazi, T. A., Spitzer, J. I., Morozov, P., & Tuschl, T. (2011). miRNAs in human cancer. The Journal of pathology, 223(2), 102-115.

Feng, L. Y., Huang, Y. Z., Zhang, W., & Li, L. (2021). LAMA3 DNA methylation and transcriptome changes associated with chemotherapy resistance in ovarian cancer. Journal of Ovarian Research, 14(1), 1-10.

Feng, W., Zhang, M., Wu, Z. X., Wang, J. Q., Dong, X. D., Yang, Y., ... & Yang, D. H. (2020). Erdafitinib antagonizes ABCB1-mediated multidrug resistance in cancer cells. Frontiers in Oncology, 10, 955.

Flores-Tellez, T. N., Lopez, T. V., Vasquez Garzon, V. R., & Villa-Trevino, S. (2015). Co-expression of ezrin-CLIC5-podocalyxin was associated with migration and invasiveness in hepatocellular carcinoma. PLoS One, 10(7), e0131605.

Gao, C., Yao, H., Liu, H., Feng, Y., & Yang, Z. (2019). TM4SF1 was a potential target for anti-invasion and metastasis in ovarian cancer. BMC cancer, 19(1), 1-12.

Garnis, C., Coe, B. P., Zhang, L., Rosin, M. P., & Lam, W. L. (2004). Overexpression of LRP12, a gene contained within an 8q22 amplicon identified by high-resolution array CGH analysis of oral squamous cell carcinomas. Oncogene, 23(14), 2582-2586.

Gebert, L. F., & MacRae, I. J. (2019). Regulation of microRNA function in animals. Nature reviews Molecular cell biology, 20(1), 21-37.

Gentleman, V. Carey, W. Huber and F. Hahne (2021). genefilter: genefilter: methods for filtering genes from high-throughput experiments. R package version 1.72.1

Ghosheh, N., Küppers-Munther, B., Asplund, A., Edsbagge, J., Ulfenborg, B., Andersson, T. B., ... & Synnergren, J. (2017). Comparative transcriptomics of hepatic differentiation of human pluripotent stem cells and adult human liver tissue. Physiological genomics, 49(8), 430-446.

Ghosheh, N., Olsson, B., Edsbagge, J., Küppers-Munther, B., Van Giezen, M., Asplund, A., ... & Synnergren, J. (2016). Highly synchronized expression of lineage-specific genes during in vitro hepatic differentiation of human pluripotent stem cell lines. Stem cells international, 2016.

Goldring, C., Antoine, D. J., Bonner, F., Crozier, J., Denning, C., Fontana, R. J., ... & Park, B. K. (2017). Stem cell–derived models to improve mechanistic understanding and prediction of human drug-induced liver injury. Hepatology, 65(2), 710-721.

Gommans, W. M., & Berezikov, E. (2012). Controlling miRNA regulation in disease. Next-Generation MicroRNA Expression Profiling Technology, 1-18.

Griss, J., Viteri, G., Sidiropoulos, K., Nguyen, V., Fabregat, A., & Hermjakob, H. (2020). Reactomegsa-efficient multi-omics comparative pathway analysis. Molecular & Cellular Proteomics, 19(12), 2115-2125.

Han, J., Xie, C., Pei, T., Wang, J., Lan, Y., Huang, K., ... & Liu, L. (2017). Deregulated AJAP1/β-catenin/ZEB1 signaling promotes hepatocellular carcinoma carcinogenesis and metastasis. Cell death & disease, 8(4), e2736-e2736.

Hattori, F., Chen, H., Yamashita, H., Tohyama, S., Satoh, Y. S., Yuasa, S., ... & Fukuda, K. (2010). Nongenetic method for purifying stem cell–derived cardiomyocytes. Nature methods, 7(1), 61-66.

He, Y., Cao, L., Wang, L., Liu, L., Huang, Y., & Gong, X. (2020). Metformin inhibits proliferation of human thyroid cancer TPC-1 cells by decreasing LRP2 to suppress the JNK pathway. OncoTargets and therapy, 13, 45.

Herberts, C. A., Kwa, M. S., & Hermsen, H. P. (2011). Risk factors in the development of stem cell therapy. Journal of translational medicine, 9(1), 1-14.

Herwig, R., Hardt, C., Lienhard, M., & Kamburov, A. (2016). Analyzing and interpreting genome data at the network level with ConsensusPathDB. Nature protocols, 11(10), 1889-1907.

Hockemeyer, Dirk, Frank Soldner, Elizabeth G. Cook, Qing Gao, Maisam Mitalipova, and Rudolf Jaenisch. "A drug-inducible system for direct reprogramming of human somatic cells to pluripotency." Cell stem cell 3, no. 3 (2008): 346-353.

Holger Schwender (2020). siggenes: Multiple Testing using SAM and Efron's Empirical Bayes Approaches. R package version 1.64.0.

Hori, S., Miyake, M., Onishi, S., Tatsumi, Y., Morizawa, Y., Nakai, Y., ... & Fujimoto, K. (2016). Clinical significance of α-and β-Klotho in urothelial carcinoma of the bladder. Oncology reports, 36(4), 2117-2125.

Hsu, S. D., Lin, F. M., Wu, W. Y., Liang, C., Huang, W. C., Chan, W. L., ... & Huang, H. D. (2011). miRTarBase: a database curates experimentally validated microRNA–target interactions. Nucleic acids research, 39(suppl_1), D163-D169.

Huang, B., Bao, J., Cao, Y. R., Gao, H. F., & Jin, Y. (2018). Cytochrome P450 1A1 (CYP1A1) catalyzes lipid peroxidation of oleic acid-induced HepG2 cells. Biochemistry (Moscow), 83(5), 595-602.

Huang, H. Y., Lin, Y. C. D., Li, J., Huang, K. Y., Shrestha, S., Hong, H. C., ... & Huang, H. D. (2020). miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. Nucleic acids research, 48(D1), D148-D154.

Huang, J., Hu, W., Lin, X., Wang, X., & Jin, K. (2015). FRZB up-regulated in hepatocellular carcinoma bone metastasis. International journal of clinical and experimental pathology, 8(10), 13353.

Irizarry, R. A. (2003). Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics, 4, 249-264.

James, A. (1998). Thomson, Joseph Itskovitz-Eldor SSS, Michelle A. Waknitz, Jennifer J. Swiergiel VSM, Jones JM. Embryonic Stem Cell Lines Derived from Human Blastocysts. Science, 282(5391), 1145.

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., ... & D'Eustachio, P. (2020). The reactome pathway knowledgebase. Nucleic acids research, 48(D1), D498-D503.

Jensen, J., Hyllner, J., & Björquist, P. (2009). Human embryonic stem cell technologies and drug discovery. Journal of cellular physiology, 219(3), 513-519.

Ji, X., Yang, X., Wang, N., Kang, M., Wang, Y., Rong, L., ... & Xue, Y. (2018). Function of SLC7A7 in T-Cell Acute Lymphoblastic Leukemia. Cellular Physiology and Biochemistry, 48(2), 731-740.

Jiang, J., Chan, Y. S., Loh, Y. H., Cai, J., Tong, G. Q., Lim, C. A., ... & Ng, H. H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nature cell biology, 10(3), 353-360.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., & Herwig, R. (2011). ConsensusPathDB: toward a more complete picture of cell biology. Nucleic acids research, 39(suppl_1), D712-D717.

Kirejczyk, M. (2008). On women, egg cells and embryos: gender in the regulatory debates on embryonic research in the Netherlands. European journal of women's studies, 15(4), 377-391.

Klover, P. J., & Mooney, R. A. (2004). Hepatocytes: critical for glucose homeostasis. The international journal of biochemistry & cell biology, 36(5), 753-758.

Knoepfler, P. S. (2008). Why myc? An unexpected ingredient in the stem cell cocktail. Cell Stem Cell, 2(1), 18-21.

Ko, E., Kim, J. S., Bae, J. W., Kim, J., Park, S. G., & Jung, G. (2019). SERPINA3 was a key modulator of HNRNP-K transcriptional activity against oxidative stress in HCC. Redox biology, 24, 101217.

Kong, F., Deng, X., Kong, X., Du, Y., Li, L., Zhu, H., ... & Xie, K. (2018). ZFPM2-AS1, a novel lncRNA, attenuates the p53 pathway and promotes gastric carcinogenesis by stabilizing MIF. Oncogene, 37(45), 5982-5996.

Koshikawa, K., Osada, H., Kozaki, K. I., Konishi, H., Masuda, A., Tatematsu, Y., ... & Takahashi, T. (2002). Significant up-regulation of a novel gene, CLCP1, in a highly metastatic lung cancer subline as well as in lung cancers in vivo. Oncogene, 21(18), 2822-2828.

Latchman, D. S. (1997). Transcription factors: an overview. The international journal of biochemistry & cell biology, 29(12), 1305-1312.

Lee, G., Jeong, Y. S., Kwak, M. J., Koh, J., Joo, E. W., Lee, J. S., ... & Yim, S. Y. (2018). Clinical significance of APOB inactivation in hepatocellular carcinoma. Experimental & molecular medicine, 50(11), 1-12.

Li, F., Zhu, W., & Gonzalez, F. J. (2017). Potential role of CYP1B1 in the development and treatment of metabolic diseases. Pharmacology & therapeutics, 178, 18-30.

Li, G. Z., Deng, J. F., Qi, Y. Z., Liu, R., & Liu, Z. X. (2020). COLEC12 regulates apoptosis of osteosarcoma through Toll-like receptor 4–activated inflammation. Journal of Clinical Laboratory Analysis, 34(11), e23469.

Li, H. C., Stoicov, C., Rogers, A. B., & Houghton, J. (2006). Stem cells and cancer: evidence for bone marrow stem cells in epithelial cancers. World journal of gastroenterology: WJG, 12(3), 363.

Li, S., Jiang, L., Tang, J., Gao, N., & Guo, F. (2020). Kernel fusion method for detecting cancer subtypes via selecting relevant expression data. Frontiers in Genetics, 11.

Li, Y., Guo, W., Liu, S., Zhang, B., Yu, B. B., Yang, B., ... & Feng, S. Q. (2017). Silencing transmembrane protein 45B (TNEM45B) inhibits proliferation, invasion, and tumorigenesis in osteosarcoma cells. Oncology research, 25(6), 1021.

Lin, Y., Zhang, J., Li, X., Zheng, D., Yu, X., Liu, Y., ... & Wang, Z. (2020). Biallelic mutations in DCDC2 cause neonatal sclerosing cholangitis in a Chinese family. Clinics and research in hepatology and gastroenterology, 44(5), e103-e108.

Liu, J. Q., Zhao, L. P., Liu, S. X., Sun, W., & Qi, H. Y. (2020). Abnormal expression of Rap2A as a prognostic marker for human breast cancer. European Review for Medical and Pharmacological Sciences, 24(18), 9541-9548.

Liu, W., Zhang, G. Q., Zhu, D. Y., Wang, L. J., Li, G. T., Xu, J. G., ... & Yang, X. Y. (2020). Long noncoding RNA ZFPM2-AS1 regulates ITGB1 by miR-1226-3p to promote cell proliferation and invasion in hepatocellular carcinoma. Eur Rev Med Pharmacol Sci, 24(14), 7612-7620.

Lopes-Pacheco, M. (2016). CFTR modulators: shedding light on precision medicine for cystic fibrosis. Frontiers in pharmacology, 7, 275.

Lübberstedt, M., Müller-Vieira, U., Mayer, M., Biemel, K. M., Knöspel, F., Knobeloch, D., ... & Zeilinger, K. (2011). HepaRG human hepatic cell line utility as a surrogate for primary human hepatocytes in drug metabolism assessment in vitro. Journal of pharmacological and toxicological methods, 63(1), 59-68.

Luo, F., Yang, K., Wang, Y. Z., & Lin, D. (2018). Tmem45b was a novel predictive biomarker for prostate cancer progression and metastasis. Neoplasma, 65(5), 815-821.

Mah, S., Nelson, M. R., Delisi, L. E., Reneland, R. H., Markward, N., James, M. R., ... & Braun, A. (2006). Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. Molecular psychiatry, 11(5), 471-478.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proceedings of the National Academy of Sciences, 78(12), 7634-7638.

Niemietz, C., Chandhok, G., & Schmidt, H. (2015). Therapeutic oligonucleotides targeting liver disease: TTR amyloidosis. Molecules, 20(10), 17944-17975.

Ohsaki, Y. O. S. H. I. N. O. B. U., Tanno, S. A. C. H. I. E., Fujita, Y., Toyoshima, E., Fujiuchi, S., Nishigaki, Y., ... & Kikuchi, K. (2000). Epidermal growth factor receptor expression correlates with poor prognosis in non-small cell lung cancer patients with p53 overexpression. Oncology reports, 7(3), 603-610.

Ognibene, M., Morini, M., Garaventa, A., Podestà, M., & Pezzolo, A. (2020). Identification of a minimal region of loss on chromosome 6q27 associated with poor survival of high-risk neuroblastoma patients. Cancer biology & therapy, 21(5), 391-399.

Ortiga-Carvalho, T. M., Sidhaye, A. R., & Wondisford, F. E. (2014). Thyroid hormone receptors and resistance to thyroid hormone disorders. Nature Reviews Endocrinology, 10(10), 582-591.

Pagnotta, S. M., Laudanna, C., Pancione, M., Sabatino, L., Votino, C., Remo, A., ... & Ceccarelli, M. (2013). Ensemble of gene signatures identifies novel biomarkers in colorectal cancer activated through PPARγ and TNFα signaling. PloS one, 8(8), e72638.

Palma, I., Peña, R. Y., Contreras, A., Ceballos-Reyes, G., Coyote, N., Eraña, L., ... & Queipo, G. (2008). Participation of OCT3/4 and β-catenin during dysgenetic gonadal malignant transformation. Cancer letters, 263(2), 204-211.

Pan Yangjian, Gao Yuan, Liu Xiaolong, Sun Jingbo, Huang Xiaoping, & Liu Lixin. (2018). Expression and clinical significance of XKRX in colorectal cancer. Chinese Journal of Clinical Anatomy, 36(2), 202-205.

Papavassiliou, A. G. (1995). Transcription factors. New England Journal of Medicine, 332(1), 45-47.

Peterson, J. A., Maroney, S. A., & Mast, A. E. (2016). Targeting TFPI for hemophilia treatment. Thrombosis research, 141, S28-S30.

Preidis, G. A., Kim, K. H., & Moore, D. D. (2017). Nutrient-sensing nuclear receptors PPARα and FXR control liver energy balance. The Journal of clinical investigation, 127(4), 1193-1201.

Qu, Y., Liao, Z., Wang, X., Zhang, J., & Liu, C. (2019). EFLDO sensitizes liver cancer cells to TNFSF10 induced apoptosis in a p53 dependent manner. Molecular medicine reports, 19(5), 3799-3806.

Raimondo, S., Cristaldi, M., Fontana, S., Saieva, L., Monteleone, F., Calabrese, G., ... & Alessandro, R. (2018). The phospholipase DDHD1 as a new target in colorectal cancer therapy. Journal of Experimental & Clinical Cancer Research, 37(1), 1-12.

Raina, K., Dey, C., Thool, M., Sudhagar, S., & Thummer, R. P. (2021). An Insight into the Role of UTF1 in Development, Stem Cells, and Cancer. Stem Cell Reviews and Reports, 1-14.

Rashid, S. T., & Alexander, G. J. (2013). Induced pluripotent stem cells: from Nobel Prizes to clinical applications. Journal of hepatology, 58(3), 625-629.

Šarić, T., & Hescheler, J. (2008). Stem cells and nuclear reprogramming. Minimally Invasive Therapy & Allied Technologies, 17(2), 64-78.

Schuldiner, M., Itskovitz-Eldor, J., & Benvenisty, N. (2003). Selective ablation of human embryonic stem cells expressing a “suicide” gene. Stem cells, 21(3), 257-265.

Selvaggi, G., Novello, S., Torri, V., Leonardo, E., De Giuli, P., Borasio, P., ... & Scagliotti, G. V. (2004). Epidermal growth factor receptor overexpression correlates with a poor prognosis in completely resected non-small-cell lung cancer. Annals of oncology, 15(1), 28-32.

Sheils, T. K., Mathias, S. L., Kelleher, K. J., Siramshetty, V. B., Nguyen, D. T., Bologa, C. G., ... & Oprea, T. I. (2021). TCRD and Pharos 2021: mining the human proteome for disease biology. Nucleic acids research, 49(D1), D1334-D1346.

Shen, K., Yu, W., Yu, Y., Liu, X., & Cui, X. (2018). Knockdown of TMEM45B inhibits cell proliferation and invasion in gastric cancer. Biomedicine & Pharmacotherapy, 104, 576-581.

Shih, C. C., Forman, S. J., Chu, P., & Slovak, M. (2007). Human embryonic stem cells are prone to generate primitive, undifferentiated tumors in engrafted human fetal tissues in severe combined immunodeficient mice. Stem cells and development, 16(6), 893-902.

Song, J., Hinderhofer, K., Kaufmann, L. T., Benjamin, N., Fischer, C., Grünig, E., & Eichstaedt, C. A. (2020). BMPR2 Promoter Variants Effect Gene Expression in Pulmonary Arterial Hypertension Patients. Genes, 11(10), 1168.

Sperandeo, M. P., Andria, G., & Sebastio, G. (2008). Lysinuric protein intolerance: update and extended mutation analysis of the SLC7A7 gene. Human mutation, 29(1), 14-21.

Sun, S., Gao, J., Zhou, S., Li, Y., Wang, Y., Jin, L., ... & Li, X. (2020). A novel circular RNA circ-LRIG3 facilitates the malignant progression of hepatocellular carcinoma by modulating the EZH2/STAT3 signaling. Journal of Experimental & Clinical Cancer Research, 39(1), 1-13.

Sun, Y., Ren, D., Yang, C., Yang, W., Zhao, J., Zhou, Y., ... & Wu, H. (2021). TRIM15 promotes the invasion and metastasis of pancreatic cancer cells by mediating APOA1 ubiquitination and degradation. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 166213.

Svec, A. (2009). PAG: a potential tumour suppressor and how it all started. From immune signalling to neoplastic transformation. Ceskoslovenska patologie, 45(2), 35-39.

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. cell, 126(4), 663-676.

Takeya, H., Ohnishi, K., Shiota, T., Saito, Y., Fujiwara, Y., Yagi, T., ... & Komohara, Y. (2019). Maf expression in human macrophages and lymph node sinus macrophages in patients with esophageal cancer. Journal of Clinical and Experimental Hematopathology, 59(3), 112-118.

Tang, Q., Chen, J., Di, Z., Yuan, W., Zhou, Z., Liu, Z., ... & Di, M. (2020). TM4SF1 promotes EMT and cancer stemness via the Wnt/β-catenin/SOX2 pathway in colorectal cancer. Journal of Experimental & Clinical Cancer Research, 39(1), 1-17.

Tiwari, A., Mukherjee, B., & Dixit, M. (2018). MicroRNA key to angiogenesis regulation: miRNA biology and therapy. Current cancer drug targets, 18(3), 266-277.

Tong, X., Wang, S., Lei, Z., Li, C., Zhang, C., Su, Z., ... & Zhang, H. T. (2020). MYOCD and SMAD3/SMAD4 form a positive feedback loop and drive TGF-β-induced epithelial–mesenchymal transition in non-small cell lung cancer. Oncogene, 39(14), 2890-2904.

Toyoda, Y., Takada, T., & Suzuki, H. (2019). Inhibitors of human ABCG2: from technical background to recent updates with clinical implications. Frontiers in pharmacology, 10, 208.

Usami, Y., Wu, Y., & Göttlinger, H. G. (2015). SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. Nature, 526(7572), 218-223.

Villodre, E. S., Gong, Y., Hu, X., Huo, L., Yoon, E. C., Ueno, N. T., ... & Debeb, B. G. (2020). NDRG1 expression was an independent prognostic factor in inflammatory breast cancer. Cancers, 12(12), 3711.

Xue, Y., Wu, L., Liu, Y., Ma, Y., Zhang, L., Ma, X., ... & Chen, J. (2015). ENTPD5 induces apoptosis in lung cancer cells via regulating caspase 3 expression. PLoS One, 10(3), e0120046.

Waldby, C. (2008). Oocyte markets: women's reproductive work in embryonic stem cell research. New genetics and society, 27(1), 19-31.

Wang, L., Gao, P., Yuan, P., Zhou, P., Fan, H., Lin, X., ... & Wang, Z. (2021). miR-573 suppresses pancreatic cancer cell proliferation, migration, and invasion through targeting TSPAN1. Strahlentherapie und Onkologie, 197(5), 438-448.

Wang, Y., Lei, L., Chi, Y. G., Liu, L. B., & Yang, B. P. (2019). A comprehensive understanding of ovarian carcinoma survival prognosis by novel biomarkers. Eur Rev Med Pharmacol Sci, 23(19), 8257-8264.

Wei, D., Kanai, M., Huang, S., & Xie, K. (2005). Emerging role of KLF4 in human gastrointestinal cancer. Carcinogenesis, 27(1), 23-31.

Werbowetski-Ogilvie, T. E., Bossé, M., Stewart, M., Schnerch, A., Ramos-Mejia, V., Rouleau, A., ... & Bhatia, M. (2009). Characterization of human embryonic stem cells with features of neoplastic progression. Nature biotechnology, 27(1), 91-97.

Wu, J., Deng, W., Li, S., & Yang, X. (2020). Advances in research on ACE2 as a receptor for 2019-nCoV. Cellular and Molecular Life Sciences, 1-14.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J. A., & Kosik, K. S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell, 137(4), 647-658.

Yang, S. H., Liu, W., Peng, J., Xu, Y. J., Liu, Y. F., Li, Y., ... & Liu, E. Y. (2021). High Expression of RhoBTB3 Predicts Favorable Chemothrapy Outcomes in non-M3 Acute Myeloid Leukemia. Journal of Cancer, 12(14), 4229.

Yoshimine, Y., Uto, H., Kumagai, K., Mawatari, S., Arima, S., Ibusuki, R., ... & Ido, A. (2015). Hepatic expression of the Sptlc3 subunit of serine palmitoyltransferase was associated with the development of hepatocellular carcinoma in a mouse model of nonalcoholic steatohepatitis. Oncology reports, 33(4), 1657-1666.

Yu, J., & Thomson, J. A. (2008). Pluripotent stem cell lines. Genes & development, 22(15), 1987-1997.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., ... & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. science, 318(5858), 1917-1920.

Zhang, D., Yan, P., Han, T., Cheng, X., & Li, J. (2021). Identification of key genes and biological processes contributing to colitis associated dysplasia in ulcerative colitis. PeerJ, 9, e11321.

Zhang, L., & He, J. B. (2018). Progress of GATA6 in liver development. Yi Chuan= Hereditas, 40(1), 22-32.

Zhang, Y., & Li, G. (2020). A tumor suppressor DLC1: The functions and signal pathways. Journal of cellular physiology, 235(6), 4999-5007.

Zhang, L., Qin, Y., Wu, G., Wang, J., Cao, J., Wang, Y., ... & Gu, H. (2020). PRRG4 promotes breast cancer metastasis through the recruitment of NEDD4 and downregulation of Robo1. Oncogene, 39(49), 7196-7208.

Zhang, L., Yang, Y., Chai, L., Bu, H., Yang, Y., Huang, H., ... & Li, W. (2020). FRK plays an oncogenic role in non-small cell lung cancer by enhancing the stemness phenotype via induction of metabolic reprogramming. International journal of cancer, 146(1), 208-222.

Zou, J., Zhu, X., Xiang, D., Zhang, Y., Li, J., Su, Z., ... & Zhang, H. (2021). LIX1-like protein promotes liver cancer progression via miR-21-3p-mediated inhibition of fructose-1, 6-bisphosphatase. Acta Pharmaceutica Sinica B.

Guideline on Risk Management Systems for Medicinal Products for Human Use - https://www.emwa.org/Documents/Freelancer/riskmanagement/rmp%20guidelines.pdf

https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en