



OPEN ACCESS

EDITED BY

Chunquan Li,
University of South China, China

REVIEWED BY

Yonglong Zhang,
Shanghai Jiao Tong University, China
Farhan Abu,
University of Texas Southwestern
Medical Center, United States

*CORRESPONDENCE

Mohammad Mobashir,
m.mobashir@cslifesciences.com
Esam Ibraheem Azhar,
eazhar@kau.edu.sa

SPECIALTY SECTION

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 21 February 2022

ACCEPTED 02 November 2022

PUBLISHED 21 November 2022

CITATION

El-Kafrawy SA, El-Daly MM, Bajrai LH, Alandijany TA, Faizo AA, Mobashir M, Ahmed SS, Ahmed S, Alam S, Jeet R, Kamal MA, Anwer ST, Khan B, Tashkandi M, Rizvi MA and Azhar EI (2022), Genomic profiling and network-level understanding uncover the potential genes and the pathways in hepatocellular carcinoma. *Front. Genet.* 13:880440. doi: 10.3389/fgene.2022.880440

COPYRIGHT

© 2022 El-Kafrawy, El-Daly, Bajrai, Alandijany, Faizo, Mobashir, Ahmed, Ahmed, Alam, Jeet, Kamal, Anwer, Khan, Tashkandi, Rizvi and Azhar. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genomic profiling and network-level understanding uncover the potential genes and the pathways in hepatocellular carcinoma

Sherif A. El-Kafrawy^{1,2}, Mai M. El-Daly^{1,2}, Leena H. Bajrai^{1,3},
Thamir A. Alandijany^{1,2}, Arwa A. Faizo^{1,2},
Mohammad Mobashir^{4,5*}, Sunbul S. Ahmed⁵, Sarfraz Ahmed⁶,
Shoaib Alam⁷, Raja Jeet⁸, Mohammad Amjad Kamal^{9,10,11,12,13},
Syed Tauqeer Anwer⁵, Bushra Khan⁵, Manal Tashkandi¹⁴,
Moshahid A. Rizvi⁵ and Esam Ibraheem Azhar^{1,2*}

¹Special Infectious Agents Unit-BSL3, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia, ²Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, ³Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, ⁴Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institute, Stockholm, Sweden, ⁵Genome Biology Lab, Department of Biosciences, Jamia Millia Islamia, New Delhi, India, ⁶Department of Biosciences, Jamia Millia Islamia, New Delhi, India, ⁷Department of Biotechnology, Jamia Millia Islamia, New Delhi, India, ⁸Botany Department, Ganesh Dutt College, Begusarai, Bihar, India, ⁹Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, China, ¹⁰King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia, ¹¹Department of Pharmacy, Faculty of Allied Health Sciences, Daffodil International University, Dhaka, Bangladesh, ¹²Enzymoics, Hebersham, NSW, Australia, ¹³Novel Global Community Educational Foundation, Hebersham, NSW, Australia, ¹⁴Department of Biochemistry, College of Science, University of Jeddah, Jeddah, Saudi Arabia

Data integration with phenotypes such as gene expression, pathways or function, and protein-protein interactions data has proven to be a highly promising technique for improving human complex diseases, particularly cancer patient outcome prediction. Hepatocellular carcinoma is one of the most prevalent cancers, and the most common cause is chronic HBV and HCV infection, which is linked to the majority of cases, and HBV and HCV play a role in multistep carcinogenesis progression. We examined the list of known hepatocellular carcinoma biomarkers with the publicly available expression profile dataset of hepatocellular carcinoma infected with HCV from day 1 to day 10 in this study. The study covers an overexpression pattern for the selected biomarkers in clinical hepatocellular carcinoma patients, a combined investigation of these biomarkers with the gathered temporal dataset, temporal expression profiling changes, and temporal pathway enrichment following HCV infection. Following a temporal analysis, it was discovered that the early stages of HCV infection tend to be more harmful in terms of expression shifting patterns, and that there is no significant change after that, followed by a set of genes that are consistently altered. PI3K, cAMP, TGF, TNF, Rap1, NF- κ B, Apoptosis, Longevity regulating pathway, signaling pathways regulating pluripotency of stem cells, Cytokine-cytokine receptor

interaction, p53 signaling, Wnt signaling, Toll-like receptor signaling, and Hippo signaling pathways are just a few of the most commonly enriched pathways. The majority of these pathways are well-known for their roles in the immune system, infection and inflammation, and human illnesses like cancer. We also find that ADCY8, MYC, PTK2, CTNNB1, TP53, RB1, PRKCA, TCF7L2, PAK1, ITPR2, CYP3A4, UGT1A6, GCK, and FGFR2/3 appear to be among the prominent genes based on the networks of genes and pathways based on the copy number alterations, mutations, and structural variants study.

KEYWORDS

HCV and HCC, biomarkers, gene expression/mutational profiling, co-expression, network-level understanding

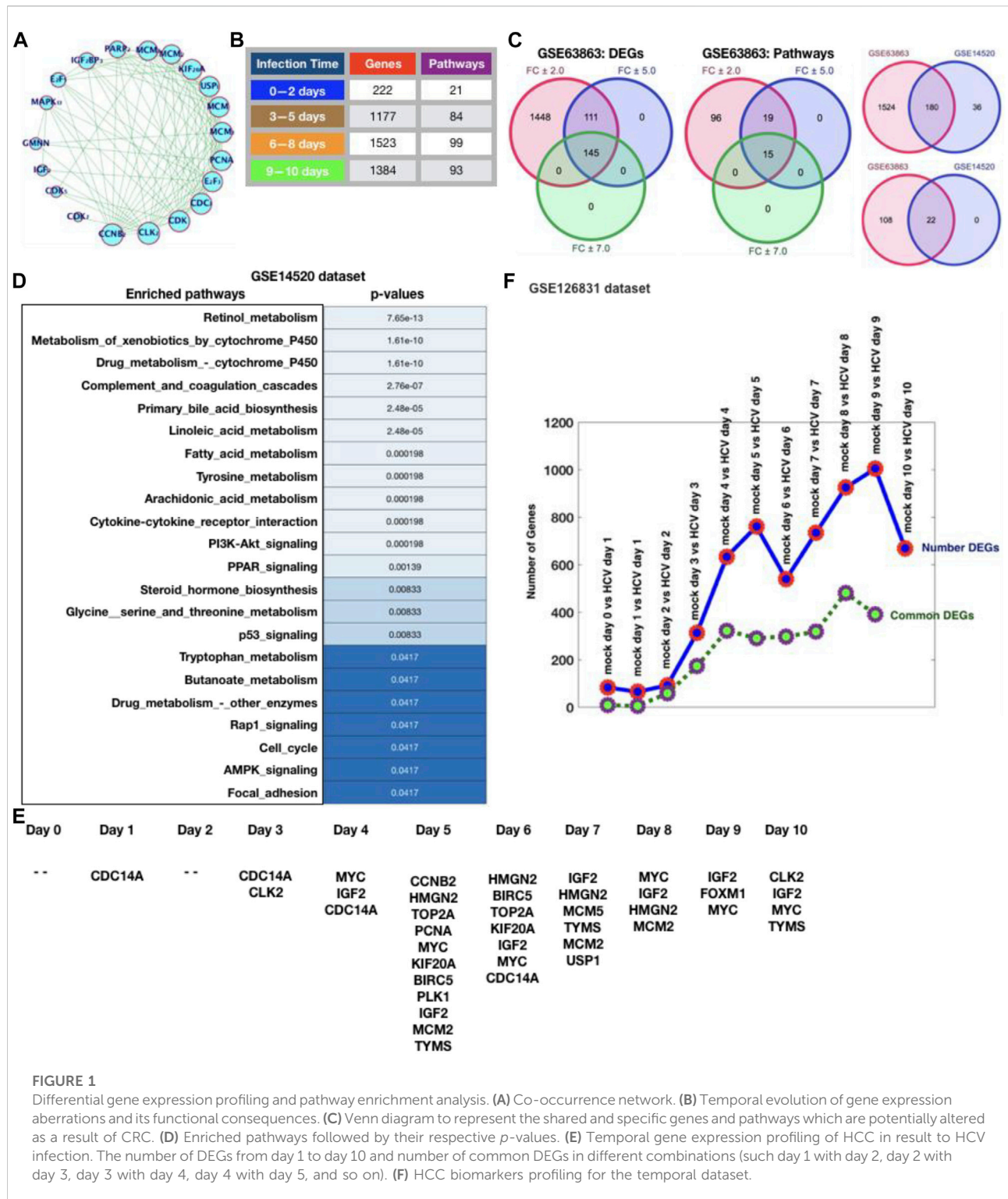
Introduction

Acquired genomic aberrations of various sorts and sizes, ranging from single nucleotide variants to structural abnormalities, are a common feature of cancer. Cancer genomes have a wide range of genomic abnormalities of various sorts and sizes. Single nucleotide variants (SNVs) to bigger structural variants (SVs) all have an impact on genome organization (Bruin et al., 2013; Dienstmann et al., 2014; Prandi et al., 2014). Different types of mutations are seen in cancer cells, and they are linked to the cell's ability to reproduce uncontrollably. Certain modifications to the genetic code only affect one or a few letters (Futreal et al., 2004; Zarrei et al., 2015; Yizhak et al., 2019). Others, referred to as copy number changes (CNA), involve bigger segments of the genome that can be deleted (deletions) or duplicated (duplications) (amplifications) (Pelham et al., 2006; Grubor et al., 2009; Agell et al., 2012; Li and Li, 2014). Various patients' tumors have different quantities of these deletions or amplifications, which are collectively known as the CNAs burden. Scientists can now scan the genomes of cancers and assess the types of mutations present in each patient thanks to new technologies. The outcomes can assist in determining the best course of action. Patients with a high CNAs burden in their tumors, for example, have a higher chance of relapse after treatment. However, it is unclear whether these persons have shorter survival rates as well, or whether CNAs levels might predict the prognosis of other cancers. Over a hundred samples from prostate cancer patients who were not treated with surgery or radiation were analyzed by Hieronymus et al. The findings revealed that a higher CNA burden in tumors is linked to more disease-related mortality (Rigail et al., 2012; Beerenwinkel et al., 2014; Li and Li, 2014; Cooper et al., 2015). The findings in prostate cancer were also true in other cancer types. When Hieronymus et al. looked at genomic data from individuals with various tumors using a different DNA sequencing assay that is authorized for clinical use, they came to the same conclusions. This suggests that CNA load could be a valuable clinical measure for assessing risk in cancer patients. Structural variation, in which rearrangements remove, increase, or reorganize genomic regions ranging in size

from kilobases (kb) to whole chromosomes, is a crucial mutational mechanism in cancer. Somatic acquired big structural variations (SVs) are a type of abnormality that can cause cancer by deactivating tumor suppressor genes and upregulating oncogenes, among other things. Detecting and characterizing these variations could lead to better cancer medicines and diagnostics (Lim and van Oudenaarden, 2007; Barbosa-Morais et al., 2010; Biesecker and Spinner, 2013; Gerstberger et al., 2014; Moncunill et al., 2014; Zarrei et al., 2015).

Cancer is caused by beginning cells that undergo a lot of evolutionary selection as the disease progresses and can change dramatically throughout treatment. Tumor cell evolution may result in subclonal divergence, leading in genetic and molecular heterogeneity. Computational approaches for creating maps of cancer evolution could help clinical risk classification and therapy techniques. There is still a gap in the study of slightly aberrant or extremely varied malignancies, despite the development of tools for assessing tumor DNA purity and cancer cell ploidy (Bardwell et al., 2001; Thomas et al., 2004; Cui et al., 2007; Carja and Feldman, 2012; Klinke, 2013; Murugaesu et al., 2013; Paguirigan et al., 2015).

The most common type of cancer in the world, hepatocellular carcinoma (HCC), is the leading cause of cancer-related fatalities (Ieta et al., 2007; Consortium et al., 2010; El-Serag, 2011; Repana and Ross, 2015; HASS et al., 2016). A high number of HCC patients show signs of vascular invasion with intrahepatic metastases, which tend to invade portal vein branches and create portal vein tumor thrombus (PVTT), which can obstruct the portal vein and cause portal hypertension (ROBINSON, 1994; Jhunjhunwala et al., 2014; Llovet et al., 2015). HCC advancement can be linked to a variety of causes, the most common of which being HBV and HCV. Aflatoxin B1, alcohol consumption, cigarette smoking, hepatotoxic chemical agents, and host co-factors such as elevated serum androgen levels, genetic polymorphisms, and DNA repair enzymes may all be linked to the progressive accumulation of a number of genomic aberrations within the hepatocytes, with TP53 and CTNNB1 being two well-known cancer drivers (Fujikawa et al., 2001; Ichikawa et al., 2008; Attari et al., 2019).



HCV is a single-stranded RNA virus with four structural proteins: capsid protein C, envelope glycoproteins E1 and E2, and protein P7, as well as six non-structural proteins: NS2, NS3, NS4A, NS4B, NS5A, and NS5B. Chronic inflammation, immune-

mediated hepatocyte death and disorder, fibrosis, and multilayer diseases (cellular pathways such as proliferation, apoptosis, and DNA repair) are all possible outcomes of HCV infection (core and structural proteins) (Ahmad et al., 2012; Jost and Altfeld,

2013; Roberts and Gordenin, 2014; Schwarzenbach et al., 2014; Sacerdote and Ricceri, 2018; Lupberger et al., 2019b).

As previously noted, HCV infection appears to be a potential cause of liver disorders such as liver cancer, steatosis, and fibrosis, and the mechanisms behind infection, liver disease development, and carcinogenesis are not fully or well understood. There are also a number of factors associated with HCC. So, in order to learn more about the leading cause of liver cancer/hepatocellular carcinoma, we used a method in which we gathered and studied previously identified biomarkers, a publicly available dataset for hepatocellular carcinoma (temporal data) with and without HCV infection, a combined study, clinical relevance, and functional impact. We examined changes in gene expression patterns, mutation mutations, CNAs, and SVs using publically available information from Gene Expression Omnibus (GEO) and TCGA, followed by cBioPortal. Furthermore, we investigated the enriched pathways for their overall functional implications and used network-level understanding to determine the impact of changed genes on other genes.

Results

As noted in the preceding section, we compiled a list of known HCC biomarkers before working with the GEO and TCGA datasets. The GEO dataset contains HCV-infected data that spans 10 days. So, in the first section of the results, we focused on data related to HCC biomarkers, followed by temporal gene expression profiling and functional significance, and finally, CNAs, mutations, and SVs analyses.

HCC biomarkers and its clinical relevance

Using cBioPortal in HCC, we were able to map out the proportion of over-expression (both individually and overall) and co-occurrence for the selected genes (biomarkers picked from previously published work) inside the TCGA database. We provided the co-occurrence in Figure 1A, and for co-occurrence, we also presented the network with the relevant connectivity in terms of co-occurrence. CCNB2, CLK2, CDK4, CDC7, E2F3, PCNA, MCM3, MCM4, USP1, KIF20A, MCM2, and MCM7 are shown to be dominantly controlling a large number of genes, or to put it another way, most of the genes are interdependent. The majority of the genes here are involved in the cell cycle, however there are a few that are specifically involved in infection and inflammatory processes (E2F5, MAPK13, IGF2BP3, IGF2). We investigated the temporal gene expression profiling for HCV infection acquired from GEO after assessing the biomarkers association. First, as shown in Figure 1B, we projected DEGs for each day of infection by combining the genes into four groups (0–2 days, 3–5 days, 6–8 days, and 9–10 days). Figure 1B shows that

increased infection duration causes significant changes in gene expression patterns until a certain time point, after which there are few changes in gene expression patterns and a slight decrease in the number of DEGs between 9 and 10 days, as well as enriched pathways or biological functions affected by changes in gene expression patterns. Figure 1B shows an exponential growth in the number of DEGs up to day eight, after which there is volatility, leading to the conclusion that there is a greater level of distribution in gene expression pattern during early HCV infection in HCC.

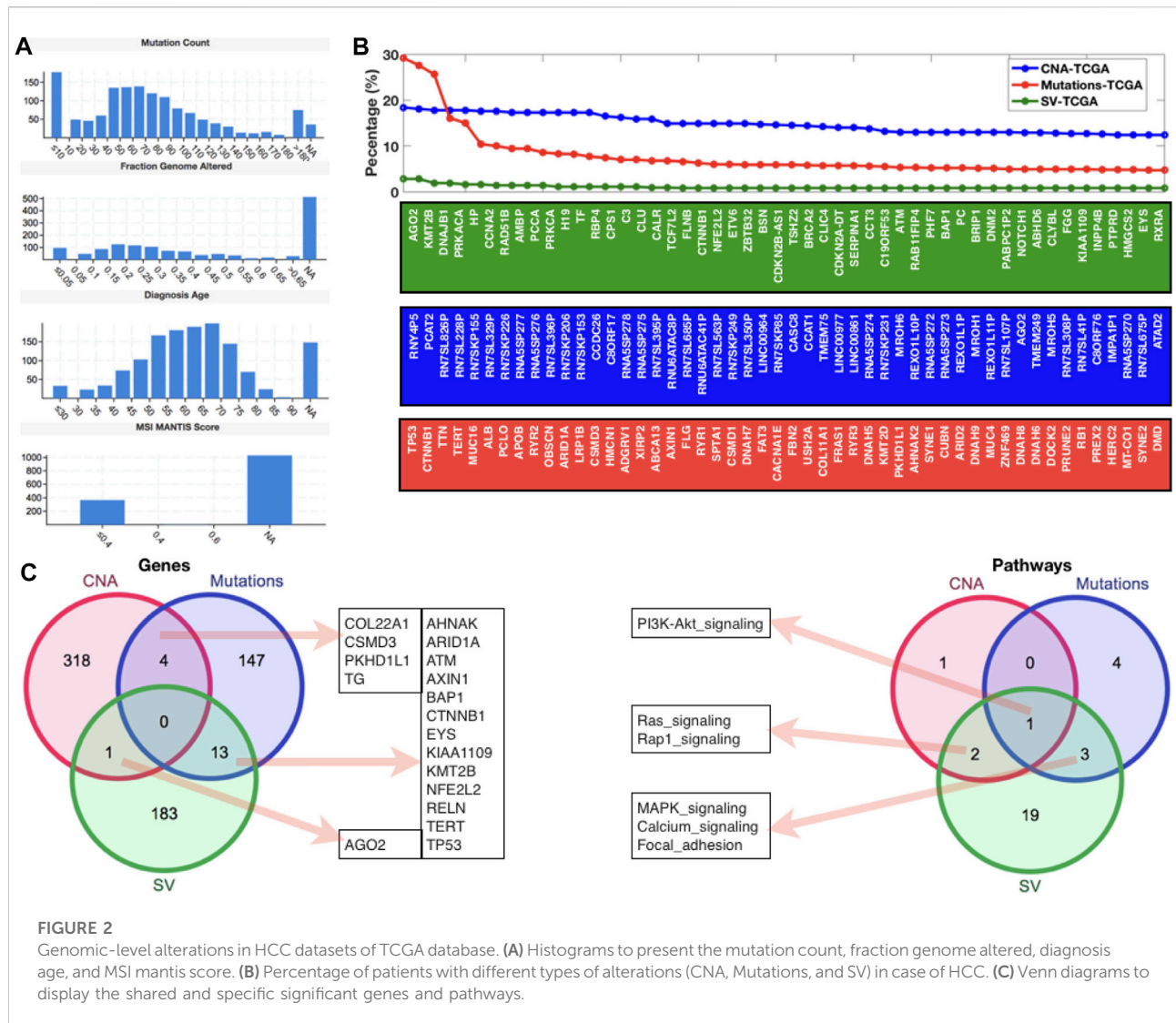
PI3K, cAMP, TGF, TNF, Rap1, NF- κ B, Apoptosis, Longevity regulating pathway, signaling pathways regulating pluripotency of stem cells, Cytokine-cytokine receptor interaction, p53 signaling, Wnt signaling, Toll-like receptor signaling, and Hippo signaling pathways are just a few of the most commonly enriched pathways. The majority of these pathways are well-known for their roles in the immune system, infection and inflammation, and human illnesses like cancer.

In addition, we conducted a comparison analysis of HCC gene expression datasets that were not infected with HCV. We observed that there are a large number of DEGs, so we prepared lists of DEGs for these three different fold changes, 2.0, 5.0, and 7.0, and analyzed the enriched pathways for all three datasets, finding that 145 DEGs and 15 enriched pathways were shared across all the three fold changes (2.0, 5.0, and 7.0), 111 DEGs and 19 enriched pathways shared between fold changes 2.0 and 5.0, and 1448 DEGs and 96 enriched pathways were unique to fold change 2.0. We compared this dataset to another dataset for the same after evaluating it at different fold changes. 180 DEGs and 22 enriched pathways were shared between the two datasets, and GSE63863 had its own set of DEGs and enriched pathways. The majority of these 22 pathways are well-known and acknowledged as the most important pathways linked to various malignancies, including HCC (Figures 1C,D; Supplementary Data S1). Furthermore, we have also presented the HCV-infected HCC temporal data in Figure 1E which contains temporal gene expression profiling of HCC in result to HCV infection. The number of DEGs from day 1 to day 10 and number of common DEGs in different combinations (such day 1 with day 2, day 2 with day 3, day 3 with day 4, day 4 with day 5, and so on).

Moreover, we have also performed the mapping of known HCC biomarkers with temporal gene expression dataset and observe that day 0 and day 2 have no HCC biomarkers as DEGs while day 5 contains the maximum number (11) of HCC biomarkers in the predicted DEGs list (Figure 1E).

Analysis of CNAs, mutations, and SVs from TCGA database

After examining gene expression profiling from the GEO database, we went on to look at global genomic aberrations using TCGA and cBioPortal, as well as all of the HCC



datasets to look at overall CNAs, mutations, and SVs in the case of HCC. Figure 2A shows the MANTIS Score distributions for mutation count, fraction genome altered, diagnostic age, and microsatellite instability (MSI) (which predicts the MSI status of tumors). For this study, all of the HCC samples from TCGA were chosen. In terms of mutation count, we can see that 10 samples have the most (>150), while 40–70 samples have a similar number of mutations (>120 and 150), and the fraction of genome altered has similar histogram patterns. The majority of the diagnosed patients were between the ages of 50 and 75, with an MSI MANTIS score of 0.4 for almost 400 patients and an MSI MANTIS score of unknown for over 1000 samples. Figure 2B shows the top 50 genes after giving the fundamental data of mutations, CNAs, and SVs. Most of the top 50 genes are specific, although AGN2 (which plays a vital function in

RNA interference) was found in both CNAs and SVs lists, and CTNNB1 (a putative component of the adherens junction) was found in both mutations and SVs lists. After mapping the top 50 genes, we applied a threshold level to all three scenarios (CNAs (10.0), mutations (3.0), and SVs (0.5)) and used a venn diagram to compare these gene lists to the enriched pathways lists (Figure 2C). We can see that none of these three lists have a gene in common. There were four genes shared by CNAs and the mutations list, thirteen genes shared by mutations and SVs, and one gene shared by SVs and the CNAs list. In terms of gene set comparison, one pathway (PI3K-AKT) was shared by all three lists, three pathways (MAPK, calcium, and focal adhesion signaling) were shared by mutations and SVs, and two routes (Ras and Rap1 signaling) were shared by SVs and CNAs (Figure 2C) (Table 1).

TABLE 1 Temporal enriched pathways.

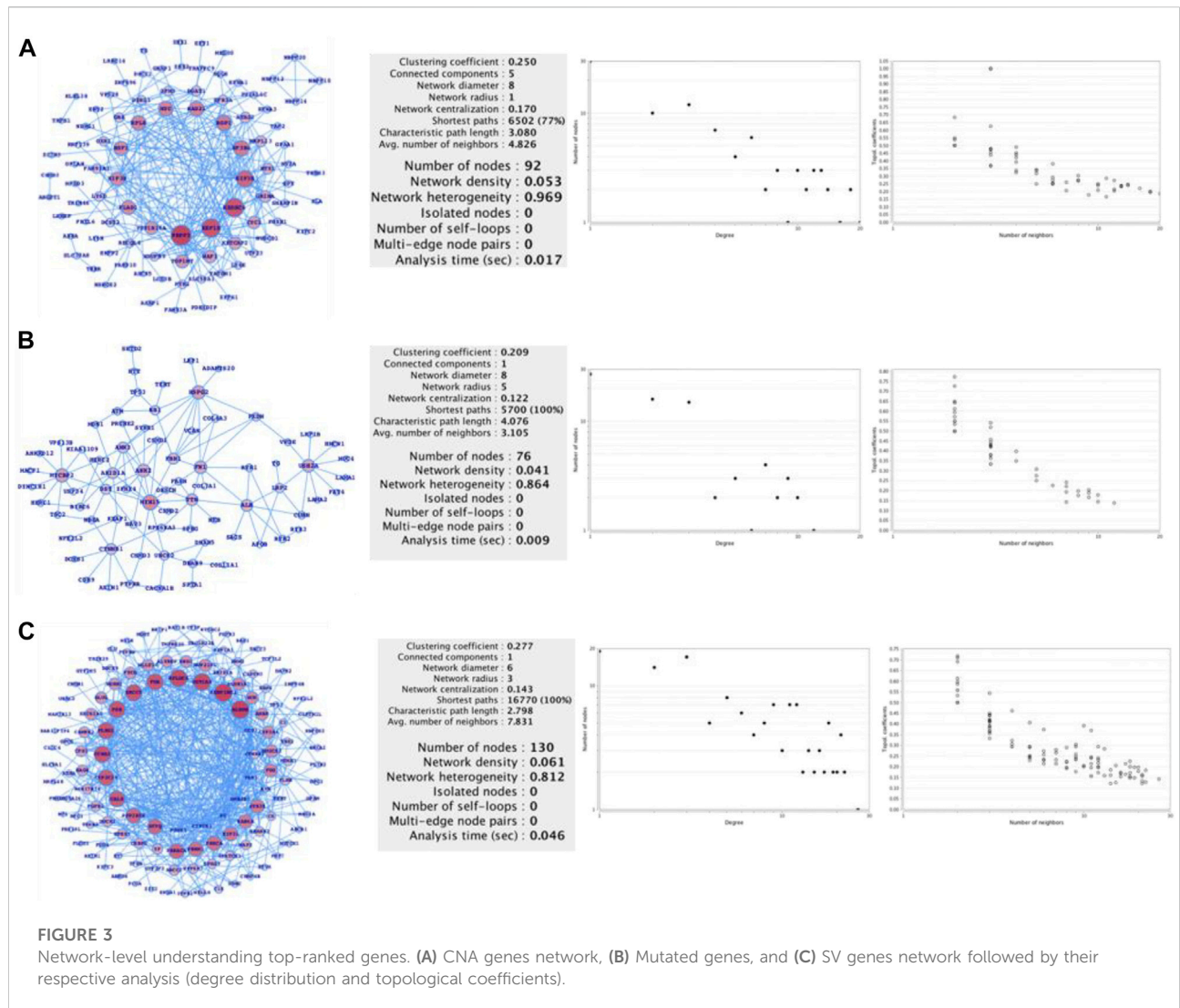
Enriched pathways (CNA genes)

KEGG_04151_P13K-Akt_signaling_pathway_-_Homo_sapiens_(human)	1.984127e-04
KEGG_04015_Rap1_signaling_pathway_-_Homo_sapiens_(human)	8.333333e-03
KEGG_04014_Ras_signaling_pathway_-_Homo_sapiens_(human)	4.166667e-02
KEGG_04360_Axon_guidance	4.166667e-02
Enriched pathways (Mutated genes)	
KEGG_04151_P13K-Akt_signaling_pathway_-_Homo_sapiens_(human)	2.480159e-05
KEGG_04510_Focal_adhesion	1.984127e-04
KEGG_04512_ECM-receptor_interaction	1.984127e-04
KEGG_04020_Calcium_signaling_pathway	8.333333e-03
KEGG_04010_MAPK_signaling_pathway	4.166667e-02
KEGG_04110_Cell_cycle	4.166667e-02
KEGG_04120_Ubiquitin_mediated_proteolysis	4.166667e-02
KEGG_04918_Thyroid_hormone_synthesis	4.166667e-02
Enriched pathways (SV genes)	
KEGG_04010_MAPK_signaling_pathway	2.755732e-06
KEGG_04610_Complement_and_coagulation_cascades	2.480159e-05
KEGG_04151_P13K-Akt_signaling_pathway_-_Homo_sapiens_(human)	1.984127e-04
KEGG_04310_Wnt_signaling_pathway	1.984127e-04
KEGG_04510_Focal_adhesion	1.984127e-04
KEGG_00980_Metabolism_of_xenobiotics_by_cytochrome_P450	1.388889e-03
KEGG_04014_Ras_signaling_pathway_-_Homo_sapiens_(human)	1.388889e-03
KEGG_04611_Platelet_activation	1.388889e-03
KEGG_04810_Regulation_of_actin_cytoskeleton	1.388889e-03
KEGG_04919_Thyroid_hormone_signaling_pathway	1.388889e-03
KEGG_00830_Retinol_metabolism	8.333333e-03
KEGG_00982_Drug_metabolism_-_cytochrome_P450	8.333333e-03
KEGG_04020_Calcium_signaling_pathway	8.333333e-03
KEGG_04371_Apelin_signaling_pathway_-_Homo_sapiens_(human)	8.333333e-03
KEGG_04921_Oxytocin_signaling_pathway	8.333333e-03
KEGG_00020_Citrate_cycle_(TCA_cycle)	4.166667e-02
KEGG_04015_Rap1_signaling_pathway_-_Homo_sapiens_(human)	4.166667e-02
KEGG_04022_cGMP-PKG_signaling_pathway_-_Homo_sapiens_(human)	4.166667e-02
KEGG_04141_Protein_processing_in_endoplasmic_reticulum	4.166667e-02
KEGG_04144_Endocytosis	4.166667e-02
KEGG_04261_Adrenergic_signaling_in_cardiomyocytes	4.166667e-02
KEGG_04392_Hippo_Signaling_Pathway	4.166667e-02
KEGG_04550_Signaling_pathways_regulating_pluripotency_of_stem_cells	4.166667e-02
KEGG_04723_Retrograde_endocannabinoid_signaling	4.166667e-02
KEGG_04916_Melanogenesis	4.166667e-02

Network-level understanding potential HCC genes

Finally, we used the FunCoup network database of CNAs, mutations, and SVs genes list to map out the networks, which we then processed in cytoscape using network analyzer (Figures 3A–C). The statistics, degree distribution, and topological coefficients of the networks were shown in Figure 3. The

degree distribution, topological coefficients, and statistical features all show that the SVs network is densely connected, followed by the CNAs network and mutations network (thinly connected). PRPF3, EEF1D, EXOSC4, EIF3E, SF3B4, BOP1, RAD21, MYC, RPL8, HSF1, HIF3E, FLAD1, PPP1R16A, TOP1MT, MAF1, KRTCAP2, CYC1, and GRINA were shown to be substantially related in the CNAs genes network. MYH15, MYCBP2, HSPG2, USH2A, FN1, FBN1, CTNBN1, ARID1A, and

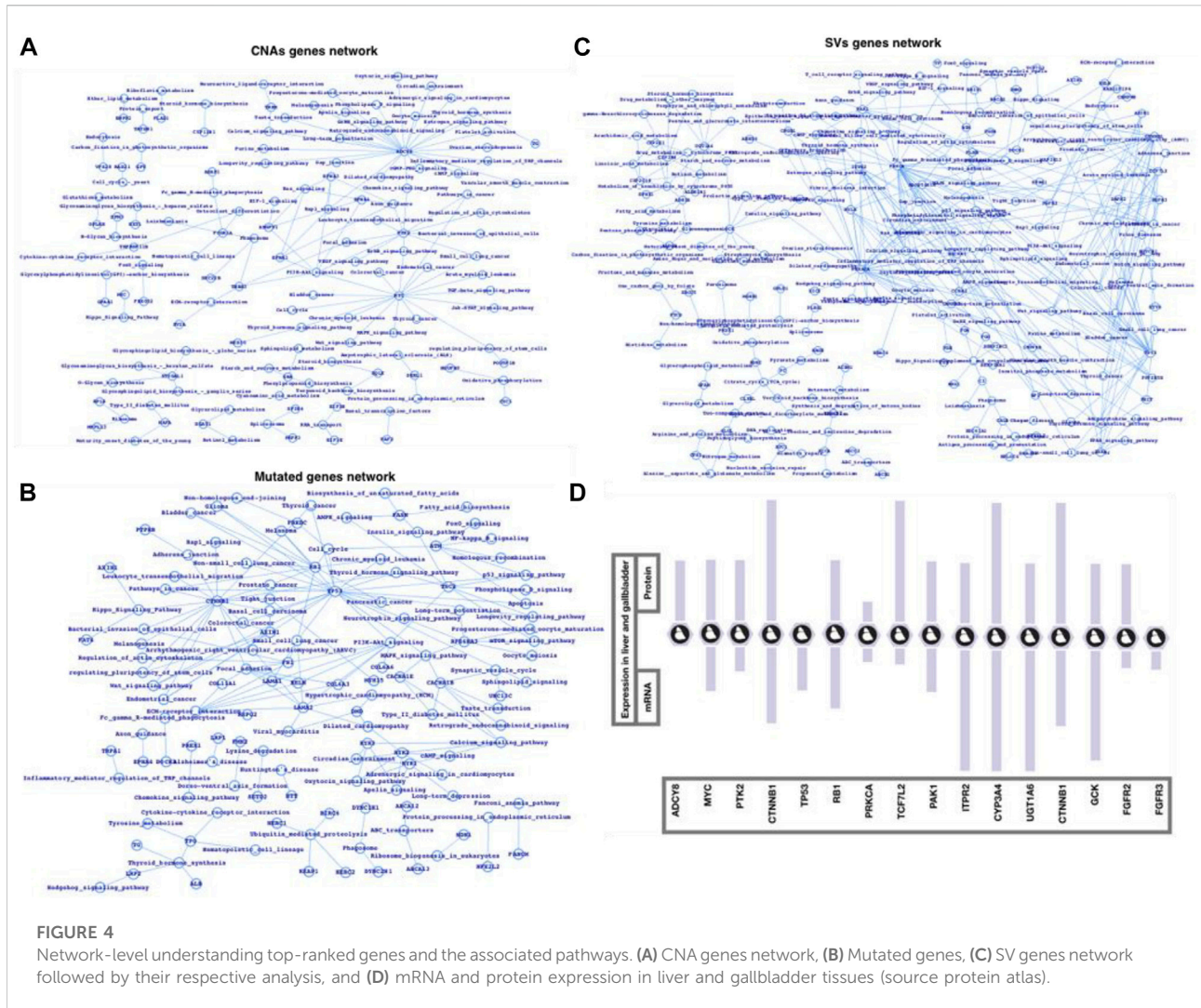


TTN were shown to be substantially related in the mutant genes network. The strongly related genes in the SVs genes network were ALDOB, SERPINC1, UGT1A6, NPLOC4, FGA, KRCC5, FGB, PLRG1, CCNA2, CYP2C18, CALR, PPP2R5E, SFPQ, PRKACA, PBRM1, PRKCA, EIF3L, RAB6A, and STK38. Based on the general network notion, it might be concluded that genes that appear to be heavily connected within the network are more significant than genes that appear to be less connected. Similarly, the more coupled genes have the potential to change more genes, and as a result, more biological activities. Furthermore, we plotted the gene networks and associated pathways for CNAs genes, mutant genes, and SVs genes (Figures 4A–C), where ADCY8, MYC, and PTK2 appear to be part of a large number of essential signaling pathways in the case of the CNAs genes network. CTNNB1, TP53, and RB1 have all been linked to cancer or cancer-related signaling pathways, primarily in HCC. PRKCA, TP53, TCF7L2, PAK1, ITPR2, CYP3A4, UGT1A6,

CTNNB1, GSK3, and FGFR2/3 are among the genes in the SVs genes network that connect a vast number of signaling pathways. We conclude that the top-ranked CNAs, mutant, and SVs genes have the ability to change at a higher-scale at the functional level based on these three genes and pathways association networks.

Discussion

Using GEO and TCGA datasets, we adopted an interdisciplinary strategy to investigate gene expression profiles, somatic mutations, CNAs, and SVs analyses. The gene expression datasets were divided into two categories: temporal datasets infected with HCV and non-temporal datasets clear of HCV infection. This study took into account all of the HCC datasets in the TCGA database. Furthermore, we used a network biology technique (Barabasi and Oltvai, 2004; Emmert-Streib and Glazko,



2010; Hu et al., 2016) to better understand the relationship between top-ranked genes in terms of linkage while they were altered. The SVs genes network appears to be the most densely connected, followed by the CNAs and mutant gene networks. Moreover, we have also used those data where the infection is associated with HBV to evaluate the broad spectrum of the impact of infection in addition to HCC at gene expression and functional levels.

The assessment of the clonality of each somatic aberration enables the deconvolution of the sequence of oncogenic events that occur during tumor initiation or progression. Assuming that clonal alterations originated prior to subclonal alterations within the same tumor, we examined pairs of genes that are aberrant in the same sample and across multiple tumors to determine the directionality of the clonal-subclonal hierarchy (Cibulskis et al., 2012; Klijn et al., 2013; Li and Li, 2014; Swanton, 2014). HCC subtypes are classified by gene clustering of tumor specific genes which resolve the HCC pathogenesis according to their etiological factor, clinical stage, recurrence rate, and prognosis. The

expression in genes regulating cell proliferation and anti-apoptotic pathways such as PNCA and cell cycle regulators CDK4, CCNB1, CCNA2, and CKS2 and ubiquitination mechanisms were studied previously. In addition to that several molecular markers of tumor progression like HSP70, CAP2, GPC3, and GS were also expressed in expression profiling. The expression profiling by time course analysis has identified several genes as a progression marker in HCC such as GPC3, CXCL12, SPINK1, GLUL, UBD, TM4SF5, DPT, SCD, MAL2, TRIM55, and COL4A2. Meanwhile the specific alteration of HCC signals transduction pathways and protein expression have given the opportunities for new therapies targeting new molecular factors. High-throughput data (genomic and proteomic) are frequently generated with the goal to understand the genotype-phenotype relationship in the complex diseases (Emilsson et al., 2008; Gonzalez-Perez et al., 2013; van't Veer et al., 2002).

Among the most common enriched pathways are PI3K, cAMP, TGF, TNF, Rap1, NF- κ B, Apoptosis, Longevity regulating pathway,

signaling pathways regulating pluripotency of stem cells, Cytokine-cytokine receptor interaction, p53 signaling, Wnt signaling, Toll-like receptor signaling, and Hippo signaling pathways. Majority of these pathways well characterized for immune controlling system, infection and inflammation, and human diseases such as cancer. PRPF3, EEF1D, EXOSC4, EIF3E, SF3B4, BOP1, RAD21, MYC, RPL8, HSF1, HIF3E, FLAD1, PPP1R16A, TOP1MT, MAF1, KRTPCAP2, CYC1, and GRINA were highly connected in case of CNAs network, in mutated genes network, MYH15, MYCBP2, HSPG2, USH2A, FN1, FBN1, CTNNB1, ARID1A, and TTN were highly connected, and ALDOB, SERPINC1, UGT1A6, NPLOC4, FGA, KRCC5, FGB, PLRG1, CCNA2, CYP2C18, CALR, PPP2R5E, SFPQ, PRKACA, PBRM1, PRKCA, EIF3L, RAB6A, and STK38 were among the highly connected genes in SVs genes network. CTNNB1, TP53, RB1, ADCY8, MYC, PTK2, PRKCA, TP53, TCF7L2, PAK1, ITPR2, CYP3A4, UGT1A6, CTNNB1, GCK, and FGFR2/3 were among the genes whose alterations could possibly alter a large number of critical biological functions including those which directly infer the cancer mainly the HCC pathways. Moreover, we have also presented the expression (mRNA and protein) (Figure 4D) of some of the potential genes in case of human liver and gallbladder tissues by using the Protein Atlas database (Uhlén et al., 2005, 2017, 2019; Cancer Genome Atlas Research Network, 2008). This study could be an example to apply the integrative approach for a number of complex diseases such as cancers, type-2 diabetes, cardiovascular diseases, and neurological disorders (Varambally et al., 2005; Taylor et al., 2010; Van Herle et al., 2012; Zhang et al., 2015; Huwait and Mobashir, 2022).

Conclusions

According to our findings, only a few genes, such as CLK2, E2F5, CDK5, E2F3, MCM3, PCNA, and CDK4, are highly overexpressed among HCC patients, and the overall expression of all the selected biomarkers appears in more than 60% of the patients, and in terms of co-occurrence, CCNB2, CLK2, CDK4, CDC7, E2F3, PCNA, and MCM3 appear to be the dominantly c Following a temporal analysis, it was discovered that the early stages of HCV infection tend to be more harmful in terms of expression shifting patterns, and that there is no significant change after that, followed by a set of genes that are consistently altered. In contrast to our expression data profile, following 4 days of HCV infection, a group of pathways is always affected. PI3K, cAMP, TGF, TNF, Rap1, NF- κ B, Apoptosis, Longevity regulating pathway, signaling pathways regulating pluripotency of stem cells, Cytokine-cytokine receptor interaction, p53 signaling, Wnt signaling, Toll-like receptor signaling, and Hippo signaling pathways are all highly altered pathways in HCC infected with HCV, according to our findings. The majority of these pathways are well-known for their roles in the immune system, infection and inflammation, and human illnesses like cancer. PI3K, cAMP, TGF, TNF, Rap1, NF- κ B, Apoptosis, Longevity regulating pathway, signaling pathways

regulating pluripotency of stem cells, Cytokine-cytokine receptor interaction, p53 signaling, Wnt signaling, Toll-like receptor signaling, and Hippo signaling pathways are just a few of the most commonly enriched pathways. Most of these pathways are well-known for their functions in the immune system, infection and inflammation, and human diseases such as cancer. According to the networks of genes and pathways based on CNAs, mutations, and SVs, ADCY8, MYC, PTK2, CTNNB1, TP53, RB1, PRKCA, TCF7L2, PAK1, ITPR2, CYP3A4, UGT1A6, GCK, and FGFR2/3 appear to be among the prominent genes.

Methods

We have selected genome-wide expression and mutational data for HCC with HCV infection and without HCV infection samples. By applying computational approach and integrating experimental data, we have unraveled the critical genes and the pathways which appear to be associated with human HCC. We have selected different datasets and the dataset details are as follows: In GSE63863 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63863>), using the Mass Array EpiTyper, they have looked at a TERT methylation assay that included the UTSS region in 125 matched HCC samples and then analyzed a validation set of 12 matched HCC samples and obtained the TERT gene's FPKM value to determine the association between TERT promoter methylation status and TERT expression level. In case of GSE14520 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse14520>), tumors and the associated non-tumor tissues were analyzed independently using a single channel array technology for gene expression profiling. On Affymetrix GeneChip HG-U133A 2.0 arrays, tumor and paired non-tumor samples from 22 patients in cohort 1 and the normal liver pool were analyzed according to the manufacturer's methodology. An Affymetrix GeneChip Scanner 3000 was used to measure fluorescence intensities, which was controlled by GCOS Affymetrix software. The 96 HT HG-U133A microarray platform was used to process all samples from cohort 2 as well as 42 tumor and non-tumor samples. An Affymetrix GeneChip HT Array Plate Scanner was used to determine the fluorescence intensities, which was controlled by GCOS Affymetrix software. We have also used HCV specific dataset GSE126831 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE126831>) where integrated genomic analysis was used to investigate time-resolved HCV infection of hepatocyte-like cells and they discovered pathways relevant for liver disease pathogenesis that have verified in the livers of 216 cirrhotic patients with HCV using differential expression, gene set enrichment analysis, and protein-protein interaction mapping.

In this study, from on previous study, we have collected the genes as biomarkers in case of HCC and studied their clinical relevance and have also studied the publicly available dataset (GSE126831 (Lupberger et al., 2019a)) related to gene expression profiling. In comparison from the previous work, we have applied different approach where we have started our work by mapping the

known association (publicly available network database) FunCoup (Alexeyenko and Sonnhammer, 2009), investigated the clinical significance of the overexpression of HCC biomarkers, and finally studied DEGs and the enriched pathways from the gene expression data (obtained from Gene Expression Omnibus). Further, we have utilized the HCC datasets from TCGA database and by using cBioPortal explored all possible mutations, CNAs, and SVs (Koboldt et al., 2012; Werner et al., 2014).

Initially, we have selected the dataset (raw expression dataset) GSE126831 (Lupberger et al., 2019a) for HCC and processed it for normalization and log₂ values of all the mapped genes. GSE126831 comprises 63 samples ranging from day 0–10 (temporal samples infected with HCV and mocked samples), with three mocked RNA samples for day 0 and three mocked and three RNA infected with HCV samples for days 1–10. We compared faked samples to HCV infected samples at the respective day of infection for differential gene expression profiling. mRNA profiles of sham or HCV-infected Huh7.5.1dif cells, obtained every day between days 0 and 10 after infection in triplicate. At 7 days after infection, the HCV infection had reached a halt (pi). Unspecific effects cannot be ruled out after day 7 pi.

The paired-end reads from all 63 samples were aligned to the human hg19 UCSC reference using TopHat software for transcriptome profiling at Illumina NextSeq 500 (Homo sapiens) RNA sequencing (v2.0.14). The Cufflinks package's cuffquant and cuffnorm were used to calculate gene expression levels (FPKM values) (v2.2.1). By creating analytical groups, proteins and transcripts were mapped. Supplementary Files format and content: hg19 Genome build: hg19 The RPKM values for each sample and the results of a differential expression analysis of mapped transcripts are stored in tab-delimited text files. Now, we proceed for our major goal which is to understand the gene expression patterns (Lapointe et al., 2004; Subramanian et al., 2005) and its inferred functions (Subramanian et al., 2005; Mi et al., 2016) and also the impact of HCC biomarker genes. We used MATLAB tools (e.g., matest) for differential gene expression prediction and statistical analysis, and for pathway analysis, we used the KEGG database (Kanehisa et al., 2007, 2009) and in-house code created for pathway and network research (Bajrai L. et al., 2021; Kamal et al., 2020; Khouja et al., 2022a; Kumar et al., 2020; Warsi et al., 2020). Furthermore, we took all of the HCC samples from the TCGA database and used cBioPortal to look for mutations, CAN, and SV, as well as prepare a list of genes using a threshold cutoff. The CNA threshold was set at 10.0, the mutation threshold at 3.0, and the SV threshold was set at 0.5. As previously stated, this collection of genes has been processed for pathway enrichment analysis. For the GEO datasets, GEO2R was applied for the calculation of *p*-values and fold changes. GEO2R is a web-based tool that allows users to compare two or more groups of Samples in a GEO Series to find genes that are differentially expressed under different experimental settings. The results are supplied as a table of genes ordered by significance, as well as a set of graphic graphs to help visualise differentially expressed genes and assess data set quality (Bajrai L. et al., 2021; Bajrai et al., 2021 L. H.; Eldakhkhny et al., 2021; Khouja et al., 2022b). FunCoup

(Reynolds et al., 2010) was used to generate DEGs networks for all of the networks in this study, and cytoscape was utilized to visualize the networks. Protein complexes, protein-protein physical interactions, metabolic, and signaling pathways are among the four types of functional coupling or linkages predicted by FunCoup. MATLAB has been used for the majority of our code and calculations. Cytoscape (Shannon et al., 2003; Skov et al., 2012), network database (PPI), ProgeneV2, and other fundamental tools are among the extra applications and resources used (Krishnamoorthy et al., 2020; Bajrai L. et al., 2021; Bajrai et al., 2021 L. H.; Eldakhkhny et al., 2021; Ahmed et al., 2022; Anwer et al., 2022).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

Conceptualization, SAE, MME, LHB, AAF, SSA, SaA, ShA, RJ, MAK, BK, MT, STA, MAR, and EIA; methodology, SAE, MME, LHB, AAF, SSA, SaA, ShA, RJ, MAK, BK, MT, STA, MAR, and EIA; software, MM and EIA; validation, LHB, TAA and EIA; formal analysis, LHB, MM, SaA, MAK, EIA; investigation, LHB, TAA, AAF, MAK, and EIA; resources, MAK, and EIA; data curation, MM, and EIA; writing—original draft preparation, SAE, MME, LHB, AAF, SSA, SaA, ShA, RJ, MAK, BK, MT, STA, MAR, and EIA; writing—review and editing, SAE, MME, LHB, AAF, SSA, SaA, ShA, RJ, MAK, BK, MT, STA, MAR, and EIA; visualization, LHB, EIA; supervision, MAK, MM, MAR, and EIA; project administration, EIA; funding acquisition, MAK, MAR, and EIA. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, grant number FP-5-42 and The APC was funded by FP-5-42. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Acknowledgments

We are thankful to DSR, KAU and for providing us the resources and the facility to carry out the work to Special Infectious Agents Unit, King Fahd Medical Research Centre,

King Abdulaziz University, Jeddah, Saudi Arabia, Medical Laboratory Sciences Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia, and Enzymoics, Hebersham, NSW; Novel Global Community Educational Foundation, Australia.

Conflict of interest

Author MK was employed by the company Enzymoics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Agell, L., Hernández, S., Nonell, L., Lorenzo, M., Puigdecamet, E., de Muga, S., et al. (2012). A 12-gene expression signature is associated with aggressive histological in prostate cancer: SEC14L1 and TCEB1 genes are potential markers of progression. *Am. J. Pathol.* 181, 1585–1594. doi:10.1016/j.ajpath.2012.08.005
- Ahmad, W., Ijaz, B., and Hassan, S. (2012). Gene expression profiling of HCV genotype 3a initial liver fibrosis and cirrhosis patients using microarray. *J. Transl. Med.* 10, 41. doi:10.1186/1479-5876-10-41
- Ahmed, S., Mobashir, M., Al-Keridis, L. A., Alshammari, N., Adnan, M., Abid, M., et al. (2022). A network-guided approach to discover phytochemical-based anticancer therapy: Targeting MARK4 for hepatocellular carcinoma. *Front. Oncol.* 1, 15. doi:10.3389/fonc.2022.914032
- Alexeyenko, A., and Sonhammer, E. L. L. (2009). Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Res.* 19, 1107–1116. doi:10.1101/gr.087528.108
- Anwer, S. T., Mobashir, M., Fantoukh, O. I., Khan, B., Imtiyaz, K., Naqvi, I. H., et al. (2022). Synthesis of silver nano particles using myricetin and the *in-vitro* assessment of anti-colorectal cancer activity: *In-silico* integration. *Int. J. Mol. Sci.* 23, 11024–11118. doi:10.3390/ijms231911024
- Attari, M. M. A., Ostadian, C., Saei, A. A., Mihanfar, A., Darband, S. G., Sadighparvar, S., et al. (2019). DNA damage response and repair in ovarian cancer: Potential targets for therapeutic strategies. *DNA Repair* 80, 59–84. doi:10.1016/j.dnarep.2019.06.005
- Bajrai, L. H., Sohrab, S. S., Mobashir, M., Kamal, M. A., Rizvi, M. A., and Azhar, E. I. (2021a). Understanding the role of potential pathways and its components including hypoxia and immune system in case of oral cancer. *Sci. Rep.* 11, 19576–19610. doi:10.1038/s41598-021-98031-7
- Bajrai, L., Sohrab, S. S., Alandijany, T. A., Mobashir, M., Parveen, S., Kamal, M. A., et al. (2021b). Gene expression profiling of early acute febrile stage of dengue infection and its comparative analysis with *Streptococcus pneumoniae* infection. *Front. Cell. Infect. Microbiol.* 11, 707905–707930. doi:10.3389/fcimb.2021.707905
- Barabasi, A.-L., and Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi:10.1038/nrg1272
- Barbosa-Morais, N. L., Dunning, M. J., Samarajiwa, S. A., Darot, J. F. J., Ritchie, M. E., Lynch, A. G., et al. (2010). A re-annotation pipeline for Illumina BeadArrays: Improving the interpretation of gene expression data. *Nucleic Acids Res.* 38, e17. doi:10.1093/nar/gkp942
- Bardwell, A. J., Flatauer, L. J., Matsukuma, K., Thorner, J., and Bardwell, L. (2001). A conserved docking site in MEKs mediates high-affinity binding to MAP kinases and cooperates with a scaffold protein to enhance signal transmission. *J. Biol. Chem.* 276, 10374–10386. doi:10.1074/jbc.M010271200
- Bereenwinkel, N., Schwarz, R. F., Gerstung, M., and Markowetz, F. (2014). Cancer evolution: Mathematical models and computational inference. *Syst. Biol.* 64, e1–e25. doi:10.1093/sysbio/syu081
- Biesecker, L. G., and Spinner, N. B. (2013). A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* 14, 307–320. doi:10.1038/nrg3424
- Bruin, S. C., de Ronde, J. J., Wiering, B., Braaf, L. M., de Wilt, J. H. W., Vincent, A. D., et al. (2013). Selection of patients for hepatic surgery of colorectal cancer liver metastasis based on genomic aberrations. *Ann. Surg. Oncol.* 20, 560–569. doi:10.1245/s10434-013-2985-7
- Cancer Genome Atlas Research Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455, 1061–1068. doi:10.1038/nature07385
- Carja, O., and Feldman, M. W. (2012). An equilibrium for phenotypic variance in fluctuating environments owing to epigenetics. *J. R. Soc. Interface* 9, 613–623. doi:10.1098/rsif.2011.0390
- Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P. W., et al. (2012). Absolute quantification of somatic DnA alterations in human cancer. *Nat. Biotechnol.* 30, 413–421. doi:10.1038/nbt.2203
- Consortium, T. I. C. G., committee, E., committee, E. A. P., group, T. A. C. A. W., group, T. W., group, B. A. W., et al. (2010). International network of cancer genome projects. *Nature* 464, 993–998. doi:10.1038/nature08987
- Cooper, C. S., Eeles, R., Wedge, D. C., Van Loo, P., Gundem, G., Alexandrov, L. B., et al. (2015). Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* 47, 367–372. doi:10.1038/ng.3221
- Cui, Q., Ma, Y., Jaramillo, M., Bari, H., Awan, A., Yang, S., et al. (2007). A map of human cancer signaling. *Mol. Syst. Biol.* 3, 152. doi:10.1038/msb4100200
- Dienstmann, R., Rodon, J., Prat, A., Perez-Garcia, J., Adamo, B., Felip, E., et al. (2014). Genomic aberrations in the FGFR pathway: Opportunities for targeted therapies in solid tumors. *Ann. Oncol.* 25, 552–563. doi:10.1093/annonc/mdt419
- El-Serag, H. B. (2011). Hepatocellular carcinoma. *N. Engl. J. Med.* 365, 1118–1127. doi:10.1056/NEJMra1001683
- Eldakhkhny, B. M., Sadoun, AIH., Choudhry, H., and Mobashir, M. (2021). *In-silico* study of immune system Associated genes in case of type-2 diabetes with insulin action and resistance, and/or obesity. *Front. Endocrinol. (Lausanne)* 12, 1–10. doi:10.3389/fendo.2021.641888
- Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A. S., Zink, F., Zhu, J., et al. (2008). Genetics of gene expression and its effect on disease. *Nature* 452, 423–428. doi:10.1038/nature06758
- Emmert-Streib, F., and Glazko, G. V. (2010). Network biology: A direct approach to study biological function. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 3, 379–391. doi:10.1002/wsbm.134
- Fujikawa, K., Kamiya, H., Yakushiji, H., Nakabepu, Y., and Kasai, H. (2001). Human MTH1 protein hydrolyzes the oxidized ribonucleotide, 2-hydroxy-ATP. *Nucleic Acids Res.* 29, 449–454. doi:10.1093/nar/29.2.449

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.880440/full#supplementary-material>

- Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., et al. (2004). A census of human cancer genes. *Nat. Rev. Cancer* 4, 177–183. doi:10.1038/nrc1299
- Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15, 829–845. doi:10.1038/nrg3813
- Gonzalez-Perez, A., Mustonen, V., Reva, B., Ritchie, G. R. S., Creixell, P., Karchin, R., et al. (2013). Computational approaches to identify functional genetic variants in cancer genomes. *Nat. Methods* 10, 723–729. doi:10.1038/nmeth.2562
- Grubor, V., Krasnitz, A., Troge, J. E., Meth, J. L., Lakshmi, B., Kendall, J. T., et al. (2009). Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). *Blood* 113, 1294–1303. doi:10.1182/blood-2008-05-158865
- Hass, H. G., Vogel, U., Scheurlen, M., and Jobst, J. (2016). Gene-expression analysis identifies specific patterns of dysregulated molecular pathways and genetic subgroups of human hepatocellular carcinoma. *Anticancer Res.* 36, 5087–5095. doi:10.21873/anticancer.11078
- Hu, J. X., Thomas, C. E., and Brunak, S. (2016). Network biology concepts in complex disease comorbidities. *Nat. Rev. Genet.* 17, 615–629. doi:10.1038/nrg.2016.87
- Huwait, E., and Mobashir, M. (2022). Potential and therapeutic roles of diosmin in human diseases. *Biomedicines* 10, 1076. doi:10.3390/biomedicines10051076
- Ichikawa, J., Tsuchimoto, D., Oka, S., Ohno, M., Furuichi, M., Sakumi, K., et al. (2008). Oxidation of mitochondrial deoxynucleotide pools by exposure to sodium nitroprusside induces cell death. *DNA Repair* 7, 418–430. doi:10.1016/j.dnarep.2007.11.007
- Ieta, K., Ojima, E., Tanaka, F., Nakamura, Y., Haraguchi, N., Mimori, K., et al. (2007). Identification of overexpressed genes in hepatocellular carcinoma, with special reference to ubiquitin-conjugating enzyme E2C gene expression. *Int. J. Cancer* 121, 33–38. doi:10.1002/ijc.22605
- Jhunjhunwala, S., Jiang, Z., Stawiski, E. W., Gnani, F., Liu, J., Mayba, O., et al. (2014). Diverse modes of genomic alteration in hepatocellular carcinoma. *Genome Biol.* 15, 436–514. doi:10.1186/s13059-014-0436-9
- Jost, S., and Altfield, M. (2013). Control of human viral infections by natural killer cells. *Annu. Rev. Immunol.* 31, 163–194. doi:10.1146/annurev-immunol-032712-100001
- Kamal, M. A., Warsi, M. K., Alnajeebi, A., Ali, H. A., Helmi, N., Izhari, M. A., et al. (2020). Gene expression profiling and clinical relevance unravel the role of hypoxia and immune signaling genes and pathways in breast cancer: Role of hypoxia and immune signaling genes in breast cancer. *jimsa*. 1. doi:10.36013/jimsa.v1i1.3
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2007). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 35, D480–D484. doi:10.1093/nar/gkm882
- Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2009). KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* 37, D355–D360. doi:10.1093/nar/gkp896
- Khouja, H. I., Ashankyty, I. M., Bajrai, L. H., Kumar, P. K. P., Kamal, M. A., Firoz, A., et al. (2022a). Multi-staged gene expression profiling reveals potential genes and the critical pathways in kidney cancer. *Sci. Rep.* 12, 7240–7311. doi:10.1038/s41598-022-11143-6
- Khouja, H. I., Ashankyty, I. M., Bajrai, L. H., Kumar, P. K. P., Kamal, M. A., Firoz, A., et al. (2022b). Multi-staged gene expression profiling reveals potential genes and the critical pathways in kidney cancer. *Sci. Rep.* 12, 7240–7311. doi:10.1038/s41598-022-11143-6
- Klijn, C., Koudijs, M. J., Kool, J., Hoeweten, J., Boer, M., de Moes, J., et al. (2013). Analysis of tumor heterogeneity and cancer gene networks using deep sequencing of MMTV-induced mouse mammary tumors. *PLoS ONE* 8, e62113. doi:10.1371/journal.pone.0062113
- Klinke, D. J. (2013). An evolutionary perspective on anti-tumor immunity. *Front. Oncol.* 2, 1–13. doi:10.3389/fonc.2012.00202
- Koboldt, D. C., Fulton, R. S., McLellan, M. D., Schmidt, H., Kalicki-Verzier, J., McMichael, J. F., et al. (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70. doi:10.1038/nature11412
- Krishnamoorthy, P. K. P., Kamal, M. A., Warsi, M. K., Alnajeebi, A. M., Ali, H. A., Helmi, N., et al. (2020). *In-silico* study reveals immunological signaling pathways, their genes, and potential herbal drug targets in ovarian cancer. *Inf. Med. Unlocked* 20, 100422. doi:10.1016/j.imu.2020.100422
- Kumar, P. P., Kamal, M. A., Warsi, M. K., Alnajeebi, A., Ali, H. A., Helmi, N., et al. (2020). *In-silico* study reveals immunological signaling pathways, their genes, and potential herbal drug targets in ovarian cancer. *Inf. Med. Unlocked* 20, 100422. doi:10.1016/j.imu.2020.100422
- Lapointe, J., Li, C., Higgins, J. P., van de Rijn, M., Bair, E., Montgomery, K., et al. (2004). Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* 101, 811–816. doi:10.1073/pnas.0304146101
- Li, B., and Li, J. Z. (2014). A general framework for analyzing tumor subclonality using SNP array and DNA sequencing data. *Genome Biol.* 15, 473–523. doi:10.1186/s13059-014-0473-4
- Lim, H. N., and van Oudenaarden, A. (2007). A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat. Genet.* 39, 269–275. doi:10.1038/ng1956
- Llovet, J. M., Villanueva, A., Lachenmayer, A., and Finn, R. S. (2015). Advances in targeted therapies for hepatocellular carcinoma in the genomic era. *Nat. Rev. Clin. Oncol.* 12, 408–424. doi:10.1038/nrclinonc.2015.103
- Lupberger, J., Croonenborghs, T., Roca Suarez, A. A., Van Renne, N., Jühling, F., Oudot, M. A., et al. (2019a). Combined analysis of metabolomes, proteomes, and transcriptomes of hepatitis C virus-infected cells and liver to identify pathways associated with disease development. *Gastroenterology* 157, 537–551. e9. doi:10.1053/j.gastro.2019.04.003
- Lupberger, J., Croonenborghs, T., Suarez, A. A. R., Van Renne, N., Jühling, F., Oudot, M. A., et al. (2019b). Combined analysis of metabolomes, proteomes, and transcriptomes of hepatitis C virus-infected cells and liver to identify pathways associated with disease development. *Gastroenterology* 157, 537–551. e9. doi:10.1053/j.gastro.2019.04.003
- Mi, H., Poudel, S., Muruganujan, A., Casagrande, J. T., and Thomas, P. D. (2016). PANTHER version 10: Expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* 44, D336–D342. doi:10.1093/nar/gkv1194
- Moncunill, V. I., Gonzalez, S., agrave, S. I. L. B., Andrieux, L. O., Salaverria, I., Royo, C., et al. (2014). Comprehensive characterization of complex structural variations in cancer by directly comparing genome sequence reads. *Nat. Biotechnol.* 32, 1106–1112. doi:10.1038/nbt.3027
- Murugaesu, N., Chew, S. K., and Swanton, C. (2013). Adapting clinical paradigms to the challenges of cancer clonal evolution. *Am. J. Pathol.* 182, 1962–1971. doi:10.1016/j.ajpath.2013.02.026
- Paguirin, A. L., Smith, J., Meshinchi, S., Carroll, M., Maley, C., and Radich, J. P. (2015). Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Sci. Transl. Med.* 7, 281re2. doi:10.1126/scitranslmed.aaa0763
- Pelham, R. J., Rodgers, L., Hall, I., Lucito, R., Nguyen, K. C. Q., Navin, N., et al. (2006). Identification of alterations in DNA copy number in host stromal cells during tumor progression. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19848–19853. doi:10.1073/pnas.0609635104
- Prandi, D., Baca, S. C., Romanel, A., Barbieri, C. E., Mosquera, J.-M., Fontugne, J., et al. (2014). Unraveling the clonal hierarchy of somatic genomic aberrations. *Genome Biol.* 15, 1–16. doi:10.1186/s13059-014-0439-6
- Repana, D., and Ross, P. (2015). Targeting FGF19/FGFR4 pathway: A Novel therapeutic strategy for hepatocellular carcinoma. *Diseases* 3, 294–305. doi:10.3390/diseases3040294
- Reynolds, C. A., Hong, M. G., Eriksson, U. K., Blennow, K., Wiklund, F., Johansson, B., et al. (2010). Analysis of lipid pathway genes indicates association of sequence variation near SREBF1/TOM1L2/ATPAF2 with dementia risk. *Hum. Mol. Genet.* 19, 2068–2078. doi:10.1093/hmg/ddq079
- Rigaill, G. J., Cadot, S., Kluin, R. J. C., Xue, Z., Bernards, R., Majewski, I. J., et al. (2012). A regression model for estimating DNA copy number applied to capture sequencing data. *Bioinformatics* 28, 2357–2365. doi:10.1093/bioinformatics/bts448
- Roberts, S. A., and Gordenin, D. A. (2014). Hypermutation in human cancer genomes: Footprints and mechanisms. *Nat. Rev. Cancer* 14, 786–800. doi:10.1038/nrc3816
- Robinson, W. S. (1994). Molecular events in the pathogenesis of hepatitis B virus-associated hepatocellular carcinoma. *Annu. Rev. Med.* 45, 297–323. doi:10.1146/annurev.med.45.1.297
- Sacerdote, C., and Ricceri, F. (2018). Epidemiological dimensions of the association between type 2 diabetes and cancer: A review of observational studies. *Diabetes Res. Clin. Pract.* 143, 369–377. doi:10.1016/j.diabres.2018.03.002
- Schwarzenbach, H., Nishida, N., Calin, G. A., and Pantel, K. (2014). Clinical relevance of circulating cell-free microRNAs in cancer. *Nat. Rev. Clin. Oncol.* 11, 145–156. doi:10.1038/nrclinonc.2014.5
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303
- Skov, V., Knudsen, S., Olesen, M., Hansen, M. L., and Rasmussen, L. M. (2012). Global gene expression profiling displays a network of dysregulated genes in non-atherosclerotic arterial tissue from patients with type 2 diabetes. *Cardiovasc. Diabetol.* 11, 15. doi:10.1186/1475-2840-11-15

- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102
- Swanton, C. (2014). Cancer evolution: The final frontier of precision medicine? *Ann. Oncol.* 25, 549–551. doi:10.1093/annonc/mdu005
- Taylor, B. S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B. S., et al. (2010). Integrative genomic profiling of human prostate cancer. *Cancer Cell* 18, 11–22. doi:10.1016/j.ccr.2010.05.026
- Thomas, J. D., Lee, T., and Suh, N. P. (2004). A function-based framework for understanding biological systems. *Annu. Rev. Biophys. Biomol. Struct.* 33, 75–93. doi:10.1146/annurev.biophys.33.110502.132654
- Uhlén, M., Björling, E., Agaton, C., Szgyarto, C. A.-K., Amini, B., Andersen, E., et al. (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* 4, 1920–1932. doi:10.1074/mcp.M500279-MCP200
- Uhlén, M., Karlsson, M. J., Zhong, W., Tebani, A., Pou, C., Mikes, J., et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science* 366, eaax9198. doi:10.1126/science.aax9198
- Uhlén, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357, eaan2507. doi:10.1126/science.aan2507
- Van Herle, K., Behne, J. M., Van Herle, A., Blaschke, T. F., Smith, T. J., Yeaman, M. R., et al. (2012). Integrative continuum: Accelerating therapeutic advances in rare autoimmune diseases. Gene expression profiling predicts clinical outcome of breast cancer. *Annu. Rev. Pharmacol. Toxicol.* 52415, 523530–547536. doi:10.1146/annurev-pharmtox-010611-134628
- Varambally, S., Yu, J., Laxman, B., Rhodes, D. R., Mehra, R., Tomlins, S. A., et al. (2005). Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* 8, 393–406. doi:10.1016/j.ccr.2005.10.001
- Warsi, M. K., Kamal, M. A., Baeshen, M. N., Izhari, M. A., and Mobashir, A. F. A. M. (2020). Comparative study of gene expression profiling unravels functions associated with pathogenesis of dengue infection. *Curr. Pharm. Des.* 26, 5293–5299. doi:10.2174/1381612826666201106093148
- Werner, H. M. J., Mills, G. B., and Ram, P. T. (2014). Cancer systems biology: A peek into the future of patient care? *Nat. Rev. Clin. Oncol.* 11, 167–176. doi:10.1038/nrclinonc.2014.6
- Yizhak, K., Aguet, F., Kim, J., Hess, J. M., Kübler, K., Grimsby, J., et al. (2019). RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* 364, eaaw0726. doi:10.1126/science.aaw0726
- Zarrei, M., MacDonald, J. R., Merico, D., and Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat. Rev. Genet.* 16, 172–183. doi:10.1038/nrg3871
- Zhang, Y., Delahanty, R., Guo, X., Zheng, W., and Long, J. (2015). Integrative genomic analysis reveals functional diversification of APOBEC gene family in breast cancer. *Hum. Genomics* 34, 34–42. doi:10.1186/s40246-015-0056-9