Iran. J. Biotechnol. July 2022;20(3): e3245 DOI: 10.30498/ijb.2022.323464.3245

Research Article

Meta-Analysis of EGF-Stimulated Normal and Cancer Cell Lines to Discover EGF-Associated Oncogenic Signaling Pathways and Prognostic Biomarkers

Shahrokh Garousi[1], Sodabeh Jahanbakhsh Godehkahriz[1*], Kasra Esfahani[2], Tahmineh Lohrasebi[2], Amir Mousavi[3], Ali Hatef Salmanian[2], Mahsa Rezvani[4], Maryam Moein[4]

[1]Department of plant genetics and production engineering, Faculty of agriculture and natural resources, University of Mohaghegh Ardabili, Ardabil, Iran.
[2]Plant Bioproducts Department, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
[3]Plant Molecular Biotechnology Department, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.
[4]Institute of Agricultural Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.

[*]*Corresponding author*: Sodabeh Jahanbakhsh Godehkahriz, Department of plant genetics and production engineering, Faculty of agriculture and natural resources, University of Mohaghegh Ardabili, Ardabil, Iran. Tel: +98-4533512204, Fax: +98-4533512204, E-mail: jahanbakhsh@uma.ac.ir

Background: Although epidermal growth factor (EGF) controls many crucial processes in the human body, it can increase the risk of developing cancer when overexpresses.

Objectives: This study focused on detecting cancer-associated genes that are dysregulated by EGF overexpression.

Materials and Methods: To identify differentially expressed genes (DEGs), two independent meta-analyses with normal and cancer RNA-Seq samples treated by EGF were conducted. The new DEGs detected only via two meta-analyses were used in all downstream analyses. To reach count data, the tools of FastQC, Trimmomatic, HISAT2, SAMtools, and HTSeq-count were employed. DEGs in each individual RNA-Seq study and the meta-analysis of RNA-Seq studies were identified using DESeq2 and metaSeq R package, respectively. MCODE detected densely interconnected top clusters in the protein-protein interaction (PPI) network of DEGs obtained from normal and cancer datasets. The DEGs were then introduced to Enrichr and ClueGO/CluePedia, and terms, pathways, and hub genes enriched in Gene Ontology (GO) and KEGG and Reactome were detected.

Results: The meta-analysis of normal and cancer datasets revealed 990 and 541 new DEGs, all upregulated. A number of DEGs were enriched in protein K48-linked deubiquitination, ncRNA processing, ribosomal large subunit binding, and protein processing in endoplasmic reticulum. Hub genes overexpression (DHX33, INTS8, NMD3, OTUD4, P4HB, RPS3A, SEC13, SKP1, USP34, USP9X, and YOD1) in tumor samples were validated by TCGA and GTEx databases. Overall survival and disease-free survival analysis also confirmed worse survival in patients with hub genes overexpression.

Conclusions: The detected hub genes could be used as cancer biomarkers when EGF overexpresses.

Keywords: Biomarker, Cancer, EGF, Meta-analysis, RNA-Seq

1. Background

Invasive cancer results from various dysregulated genes (1). Oncogenic mutations drive targeted genes to up and down-regulate or generate dysfunctional proteins (2). Since genes and pathways in tumors undergo massive changes, extensive research is crucial for detecting involved biomarkers. EGF plays a vital role in numerous essential processes, it proliferates many types of cells and

cures injuries (3), nevertheless, EGF is known to be implicated in the development of several cancers. In a study, EGF-treated various cancer cells showed EGF receptor overexpression, dephosphorization, and under-expression of focal adhesion kinase, and also an increase in invasion, migration, and metastasis of cancer cells (4). Likewise, when breast cancer cells were treated with EGF, cell migration and drug resistance were raised (5). Following EGF overexpression, the expression profiles of core genes and pathways change, complicated irregular patterns form and some cancers emerge. Mining DEGs involved in cancer is a perfectly valid method to discover new biomarkers (6). Meta-analysis provides researchers with an invaluable tool to uncover novel DEGs (7). It enlarges the sample size and increases statistical power (8).

2. Objectives

Nowadays EGF cosmetics and other products are available in the market, therefore the current study aimed at discovering novel EGF-stimulated genes and pathways associated with cancer using normal and cancer cell lines, obtained from several independent and related studies. Two different meta-analyses with normal and cancer datasets were conducted, then each of which was independently analyzed, and their shared dysfunctional DEGs and pathways were discerned. Malfunctioned oncogenes in cancer cell lines were also detected. To the best of our knowledge, the present study is the first to carry out and report the meta-analysis of EGF-stimulated normal and cancer datasets at once.

3. Materials and Methods

3.1. RNA-Seq Data Selection and Analysis

The NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and European Nucleotide Archive (https://www.ebi.ac.uk/ena) was mined for finding the publicly available RNA-Seq datasets. Two independent trials were conducted, the meta-analysis of RNA-Seq data obtained from EGF-stimulated normal and cancer cell lines. The main features of each study are summarized in **Table 1**. Since studies one and three

Table 1. Details of RNA-Seq studies. RNA-Seq data of normal and cancer cell lines were analyzed through two independent meta-analyses, studies 1 and 3 in the meta-analysis of cancer datasets with three and two subpopulations, respectively, treated as different datasets.

Group	Study	Accession	Cell line	Sample	size	Illumina platform
Normal	1	GSE121697	Mesenchymal stromal cells	Control	3	HiSeq 3000
				EGF	3	
	2	GSE156089	Skin-derived epidermal stem cells	Control	3	NovaSeq 6000
				EGF	3	
	3	GSE124586	HTR-8/SVneo cells	Control	2	HiSeq 2000
				EGF	2	
Cancer	1	PRJNA322427	MDA468 Breast Cancer	Control	9	HiSeq 2000
				EGF	7	
			PMC42-ET Breast Cancer	Control	8	
				EGF	8	
			PMC42-LA Breast Cancer	Control	9	
				EGF	8	
	2	GSE40632	HeLa cells	Control	2	HiSeq 2000
				EGF	2	
	3	GSE85089	A549- Lung - KRAS mutation	Control	2	NextSeq 500
				EGF	2	
			HeLa cervix	Control	4	
				EGF	4	
	4	GSE94374	HCT116_WT	Control	2	HiSeq 2500
				EGF	2	

in the meta-analysis of cancer datasets had three and two different cell lines as subpopulations, respectively, they were considered different datasets. The meta-analysis of normal datasets contained three studies with 8 EGF-stimulated and 8 control samples, whereas the meta-analysis of cancer datasets comprised four studies with seven subpopulations, and 33 EGF-stimulated and 36 control samples. If the samples of eligible datasets had been stimulated only with EGF and not EGF in combination with any other chemicals, they were allowed to include in the final selection. **Figure 1A** illustrates the workflow of the present study. Any low-quality bases, adapters, and other Illumina-specific sequences were trimmed off the FASTQ data with Trimmomatic v0.39. All reads with an average quality of less than 20 were dropped. The quality of reads was tested before and after trimming with FASTQC v0.11.9. The trimmed reads were aligned to the human reference genome, GRCh38.101, by HISAT2 v2-2.2.1. The aligned sequences were then post-processed with SAMtools v1.11. The uniquely overlapped reads with GRCh38.101.gtf were counted with HTSeq v0.12.4. All preceding analyses were carried out using the Linux terminal, Ubuntu 20.04. After pre-filtering low count genes, DEGs between EGF- stimulated and control read counts were detected in every single RNA-Seq study by DESeq2 v1.31.7 package of R software. Meta-analysis of normal and cancer count data was independently conducted by the metaSeq v1.3.0 package of R. To screen statistically significant DEGs in DESeq2 and metaSeq analyses, p-values < 0.05 adjusted by Benjamini- Hochberg false discovery rate (FDR) was chosen. Detected DEGs, in this study, can be classified into four groups, 1. Three sets of DEGs were obtained from the analysis of three RNA-Seq studies of normal cell lines, 2. Seven sets of DEGs were identified by the analysis of four (with seven subgroups) RNA-Seq studies of cancer cell lines, 3. One set of DEGs was detected by the meta-analysis of three RNA-Seq studies of normal cell lines, and 4. One set of DEGs was extracted by the meta-analysis of four (with seven subgroups) RNA-Seq studies of cancer cell lines. The number of common and different DEGs among various comparisons of RNA-Seq studies and the meta-analysis of RNA-Seq studies were illustrated by Venn diagram, VennDiagram v1.6.20 R package, and chord diagram, circlize v0.4.12 R package, for the meta-analysis of normal and cancer datasets, respectively. Shared DEGs between two meta-analyses were also visualized by chordDiagram. In addition, the pheatmap v1.0.12 R package was used to plot a heatmap, with hierarchical clustering, for the top 250 DEGs identified by metaSeq in normal and cancer datasets.

3.2. Constructing PPI Network for DEGs and Detecting Densely Interconnected Clusters

Cytoscape v3.8.2 was used for constructing the PPI network. Using stringApp, a Cytoscape app, the network of DEGs was imported from Search Tool for the Retrieval of Interacting Genes (STRING v11.0b) database into Cytoscape. The minimum confidence (score) cut-off was adjusted to 0.4, Molecular Complex Detection plugin (MCODE v2.0.0), a Cytoscape app, was hired to identify thickly interlinked areas.

3.3. Enrichment Analysis of DEGs via Enrichr and ClueGO and Identifying Hub Genes

Using Enrichr, term and pathway enrichment in GO, KEGG and Reactome were inspected, and hub genes were identified. The top terms and pathways, based on p-value, enriched in biological process (Bp), molecular function (Mf), and cellular component (Cc) of GO, and KEGG and Reactome were plotted by ggplot2 v3.3.2 R package. To get highly accurate results in enrichment analysis and hub genes detection, DEGs found in the meta-analysis of normal, and cancer datasets were also analyzed with ClueGO v2.5.8/CluePedia v1.5.8, a Cytoscape app, with FDR ≤ 0.05.

3.4. Validating Higher Expression of Cancer-Associated Hub Genes

To validate the clinical and prognostic reliability of discovered hub genes, GEPIA2 (http://gepia2.cancer-pku.cn/) was employed. This web server illustrates statistically significant differences in the expression of normal and tumor samples using The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databanks. Box plot and Kaplan–Meier survival plots (overall survival, OS, and disease free survival, DFS) were hired for verifying higher expression of hub genes in various cancers stimulated by EGF overexpression and recognizing the connection between hub genes expression levels and prognostic results. The cutoff of $|\log_2 FC| > 1$ and p-value < 0.05 for box plots and median 50-50 and logrank p-value < 0.05 for survival plots was set.

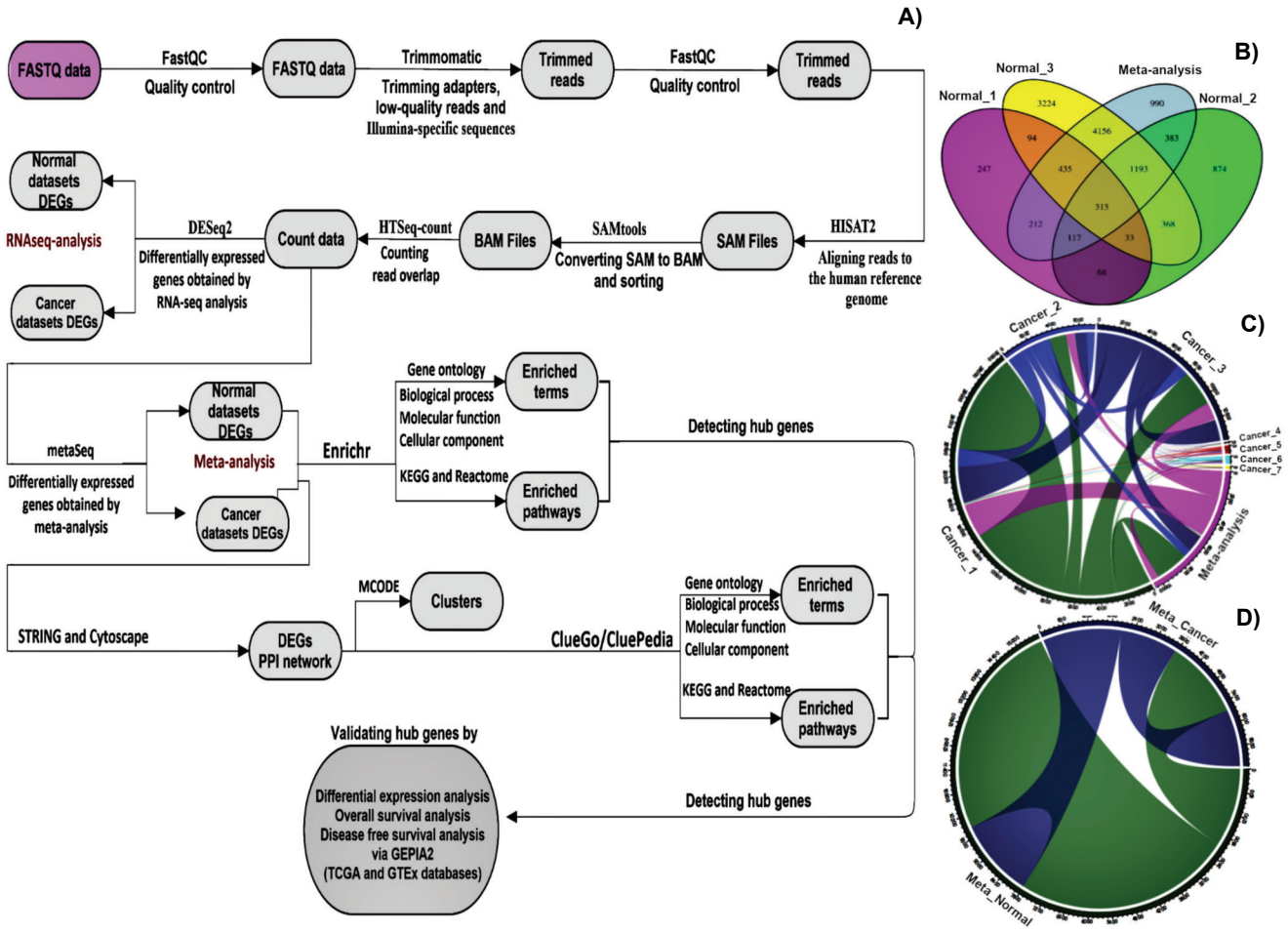

Figure 1. A) Workflow of the current study. Various comparisons of common and different DEGs among different groups are summarized for normal cell lines by Venn diagram, in **B)**, and cancer cell lines by chord diagram in **C)**, **D)** Chord diagram illustrates common and different DEGs between two meta-analyses.

4. Results

In the current study, expression variations of EGF-stimulated normal and cancer cell lines were analyzed. In the meta-analysis of normal datasets, 7799 upregulated, and only 2 downregulated DEGs (MT-ND6 and AC011603.3) were detected. A total of 315 common DEGs between all individual RNA-Seq studies and the meta-analysis of normal datasets were recognized. Besides, this meta-analysis identified 990 gained DEGs, the DEGs that none of the RNA-Seq studies could detect; all gained DEGs were upregulated (**Fig. 1B**). The meta-analysis of cancer datasets identified 3751 DEGs, with 541 gained DEGs; all DEGs were upregulated. There were no common DEGs between RNA-Seq studies and the meta-analysis of cancer datasets (**Fig. 1C**). Further DEGs information is detailed in **supplementary Table 1**. Statistically significant DEGs were selected at FDR < 0.05. Only gained DEGs identified by the meta-analysis of normal and cancer datasets were included in all downstream analyses. To cluster comparable expression patterns, a heatmap of $\log_2$-transformed counts for the top 250 DEGs, ranked by adjusted p-value and identified by metaSeq in the normal and cancer datasets, was plotted. Plotting was done with the pheatmap R package, based on hierarchical clustering, Euclidean distance (**Fig. S1A-B**).

4.1. Constructing PPI Network of DEGs

The PPI subnetworks, composed of 127 nodes and 244 edges, and 149 nodes and 208 edges, for the DEGs detected by the meta-analysis of normal and cancer datasets, respectively, were constructed with Cytoscape v3.8.2. (**Fig. S2A-B**).

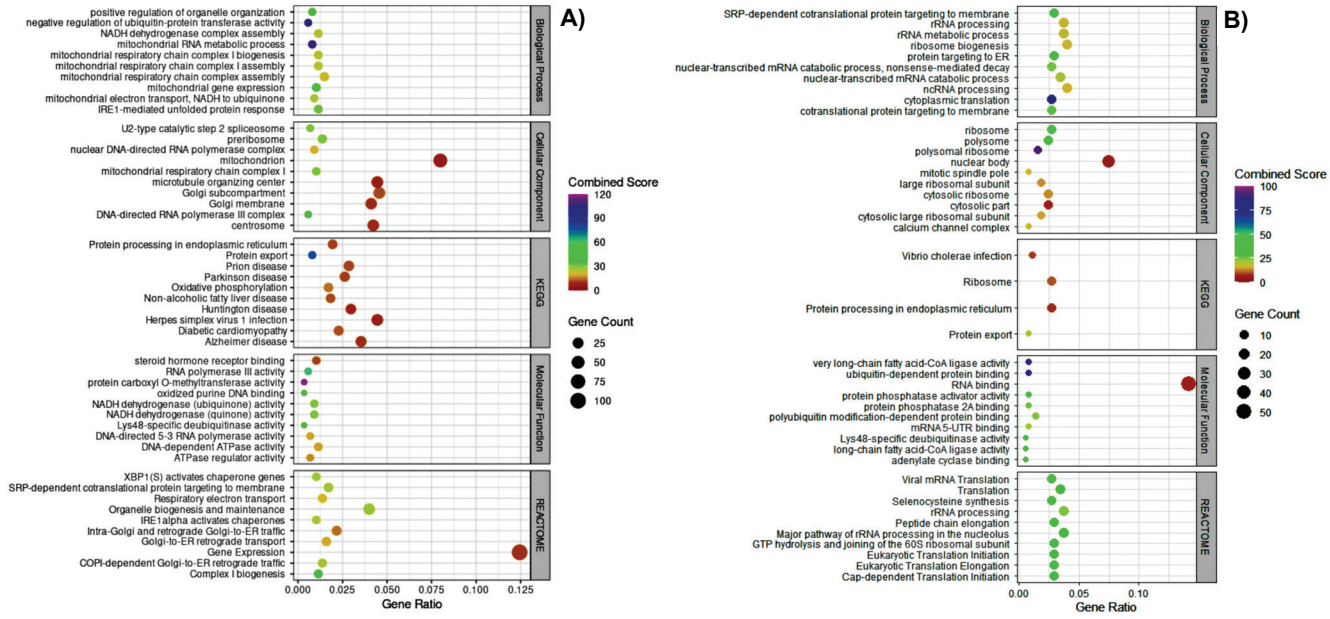

Figure 2. The top ten enriched GO terms, and KEGG and Reactome pathways in the meta-analysis of **A)** normal cell lines and **B)** cancer cell lines obtained by Enrichr. The Gene count represents the number of DEGs included in a term or a pathway, Gene ratio = Gene count / the number of DEGs introduced to Enrichr.

4.2. Identifying Seed Nodes

To detect very interconnected modules in the PPI networks of normal and cancer DEGs via MCODE, significant-top seven and four clusters, respectively, with scores four and more than four were selected (**Figs. S3 - 4**). In the normal PPI network, ASB7, AKAP9, LYRM4, ATR, OTUD7B, and ATAD1 received the highest score in clusters of one, two, three four, six, and seven, respectively. However, RPL10, EXOSC6, and ASPM in clusters one, two, and four were the nodes from which the clusters originated in the cancer PPI network. The calculated scores were 18.758, 11.93, 6.8, 5.545, 5, 4.5, and 4 for the normal PPI network, and 11.5, 6, 4, and 4 for the cancer PPI network.

4.3. Enrichment Analysis of GO, KEGG, and Reactome and Identifying Hub Genes via Enrichr and ClueGO/CluePedia

To detect dysregulated hub genes in GO terms and KEGG and Reactome pathways, involved in cancer, Enrichr, an excellent tool with updated libraries, and ClueGO/CluePedia were used. The top Bp, Mf, Cc terms of GO, and KEGG and Reactome pathways with enriched DEGs in Enrichr and ClueGo/CluePedia are detailed in **Figure 2**, and **supplementary Table 2** and **Figure 3 A-D**, respectively.

4.4. Finding Common DEGs between Two Meta-Analyses and Common Upregulated DEGs between All RNA-Seq Datasets and Two Meta-Analyses

By comparing the DEGs identified only by two meta-analyses, 29 DEGs, all of which upregulated, were matched (**Fig. 1D**). These DEGs are as follows: RAB21, GSS, KIF2A, TOP2B, ATP6V0B, RPS3A, INTS8, CEP295, KLHL15, YIF1A, MAPK6, HOMER1, SCOC, BTG3, SEC13, OTUD4, RAD1, RARS1, NMD3, POMGNT1, PHIP, H1-10, CEP290, RESF1, YOD1, FANCM, AC092614.1, NBPF26, and SKP1. A total of 12 DEGs, including HSPA5, NSF, ATP6V1D, GSS, KNOP1, PRDM4, PFDN1, EIF2B2, EMC4, SERPINE2, LETM2, and SMIM29, with upregulating expression profiles, between DEGs of two meta-analyses and DEGs of all RNA-Seq studies were also detected. The 29 and 12 gene sets had the GSS gene in common.

4.5. Validating Higher Expression of Cancer-Associated Hub Genes

With the help of TCGA and GTEx databases, differential expression analysis of hub genes via boxplot function of GEPIA2 validated the higher expression of hub genes (DHX33, INTS8, NMD3, OTUD4, P4HB,

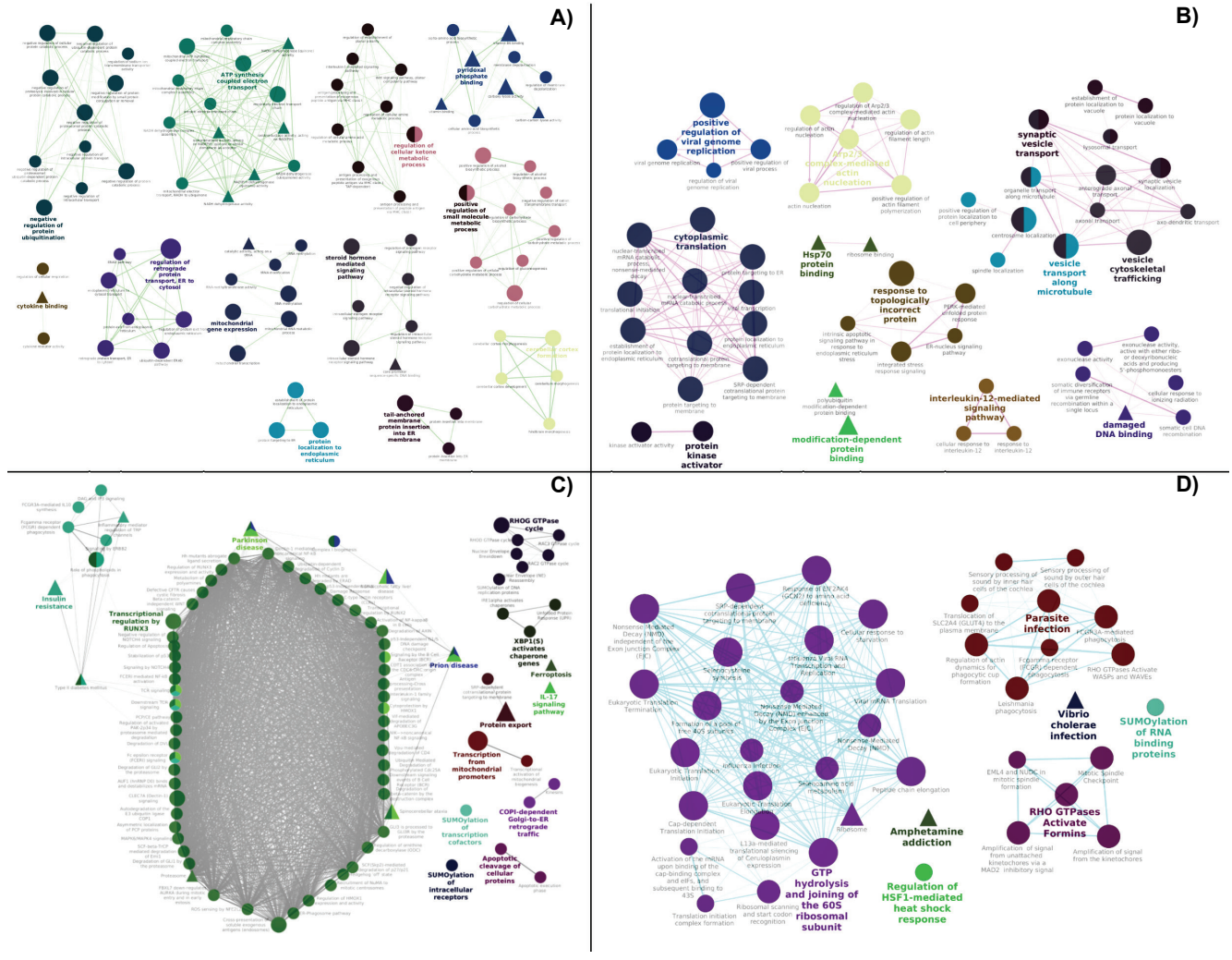

Figure 3. Enrichment network of GO terms **(A-B)** and KEGG and Reactome pathways **(C-D)** for top DEGs in the meta-analysis of normal and cancer cell lines, respectively, by ClueGO/CluePedia app of Cytoscape. Biological processes and Reactome are circular in shape, and molecular functions and KEGG are triangular. The color and size of nodes represent category and enrichment significance (p-values < 0.05, adjusted by Benjamini- Hochberg). Edge width represents the kappa score.

RPS3A, SEC13, SKP1, USP34, USP9X, and YOD1) in tumor samples versus normal ones (**Fig. 4**). These hub genes were identified by enrichment analysis of GO, KEGG, and Reactome using Enrichr and ClueGO/CluePedia. Surprisingly, nearly all hub genes (except for USP9X and YOD1) overexpressed in Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC) and Thymoma (THYM). Similarly, in Cholangiocarcinoma (CHOL), DHX33, INTS8, OTUD4, RPS3A, SKP1, and USP34 had higher expression levels, whereas NMD3, SEC13, SKP1, USP34, USP9X, and YOD1 upregulated in Pancreatic adenocarcinoma (PAAD). In addition, NMD3, P4HB, RPS3A, and SEC13 in Glioblastoma multiforme (GBM), NMD3, P4HB, RPS3A, and SEC13 in Brain Lower Grade Glioma (LGG), and DHX33, P4HB, and RPS3A in Testicular Germ Cell Tumors (TGCT) highly expressed. In **supplementary Table 3** relationship between the overexpression of hub genes and cancer type is described in detail. In addition, to evaluate the prognostic and clinical value of hub genes, Kaplan–Meier curves (OS and DFS analyses) were plotted. OS function of GEPIA2 according to cancer type confirmed that P4HB gene for Bladder Urothelial Carcinoma (BLCA), GBM, Kidney renal papillary cell carcinoma (KIRP) and LGG, and DHX33, NMD3, and SEC13 genes for Skin Cutaneous Melano-

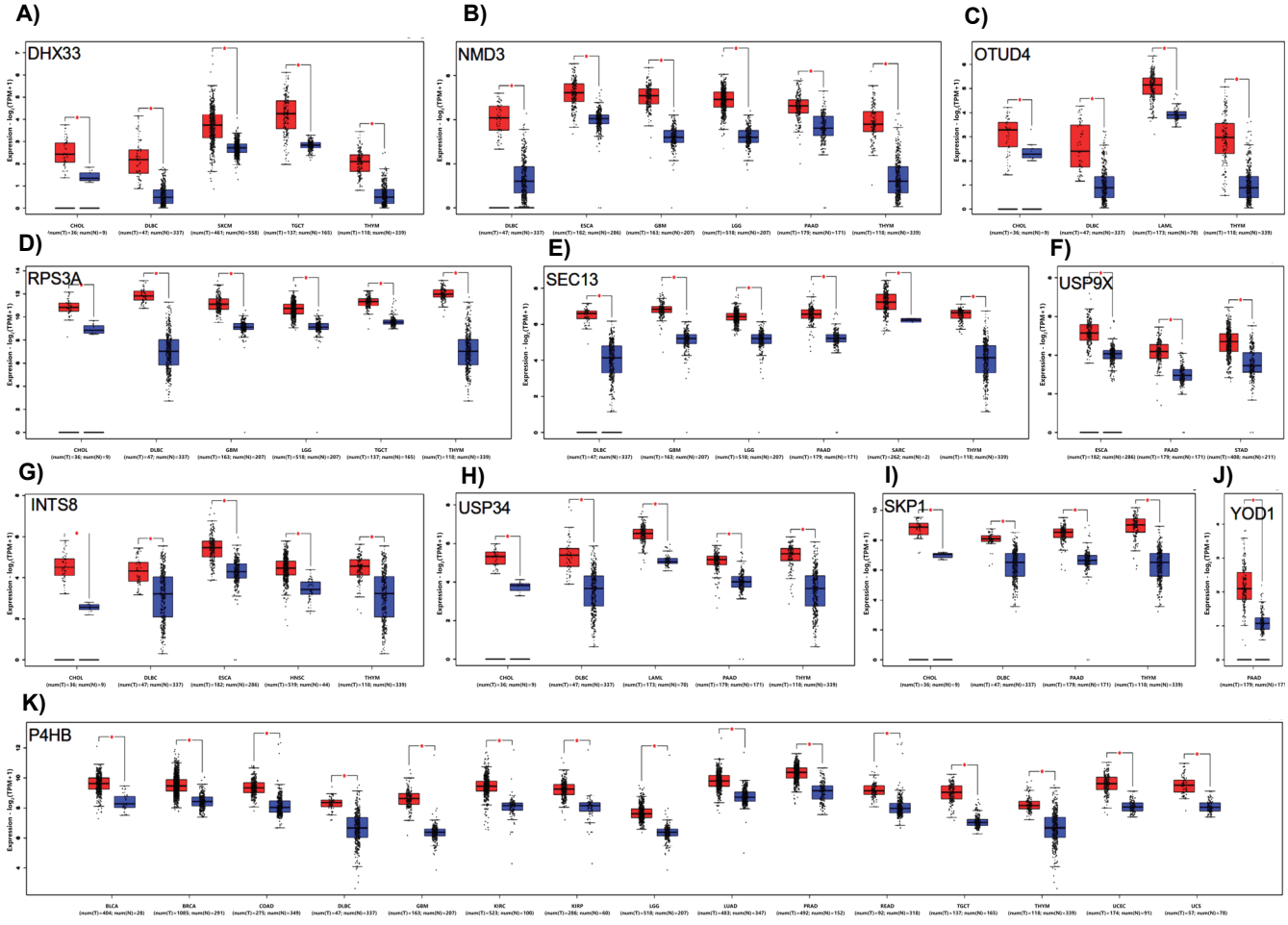

Figure 4. Overexpression validation of hub genes in tumor samples **A)** DHX33, **B)** NMD3, **C)** OTUD4, **D)** RPS3A, **E)** SEC13, **F)** USP9X, **G)** INTS8, **H)** UPS34, **I)** SKP1, **J)** YOD1, and **K)** P4HB.

ma (SKCM), PAAD and Sarcoma (SARC), respectively, are reliable prognostic biomarkers, and those genes can identify patients with significantly worse survival (**Fig. 5**). DFS analysis also demonstrated that the overexpression of the P4HB gene in BLCA, GBM, Kidney renal clear cell carcinoma (KIRC), KIRP and LGG, and NMD3, SEC13, and YOD1 genes in PAAD, SARC, and PADD, respectively, were significantly connected to poorer survival (**Fig. S5**).

5. Discussion

In the present study, protein K48-linked deubiquitination (GO:0071108) term of Bp enriched in Enrichr and ClueGO. When the systems regulating ubiquitination are impaired, associated biological processes change, and life-threatening illnesses, especially cancer, manifest (9). K48-linked polyubiquitin chains target ubiquitinated proteins for proteasomal degradation, and they mass where DNA has been damaged for degrading reacting proteins to damage (10). The results of Enrichr and ClueGO revealed that the two meta-analyses had OTUD4 and YOD1 DEGs in common. OTUD4, encoding a specific DUB to K48, seems to frustrate proteasomal degradation proteins that repair DNA. OTUD4 overexpression profoundly influences many tumor types such as breast and glioblastoma through the transforming growth factor-β (TGFβ) pathway (11). It was reported that OTUD4 overexpression suppressed motility, invasion, and overproduction of breast, liver, and lung cancer in apoptosis/AKT signaling pathway, and so it could be used as a biomarker for many cancers (12).

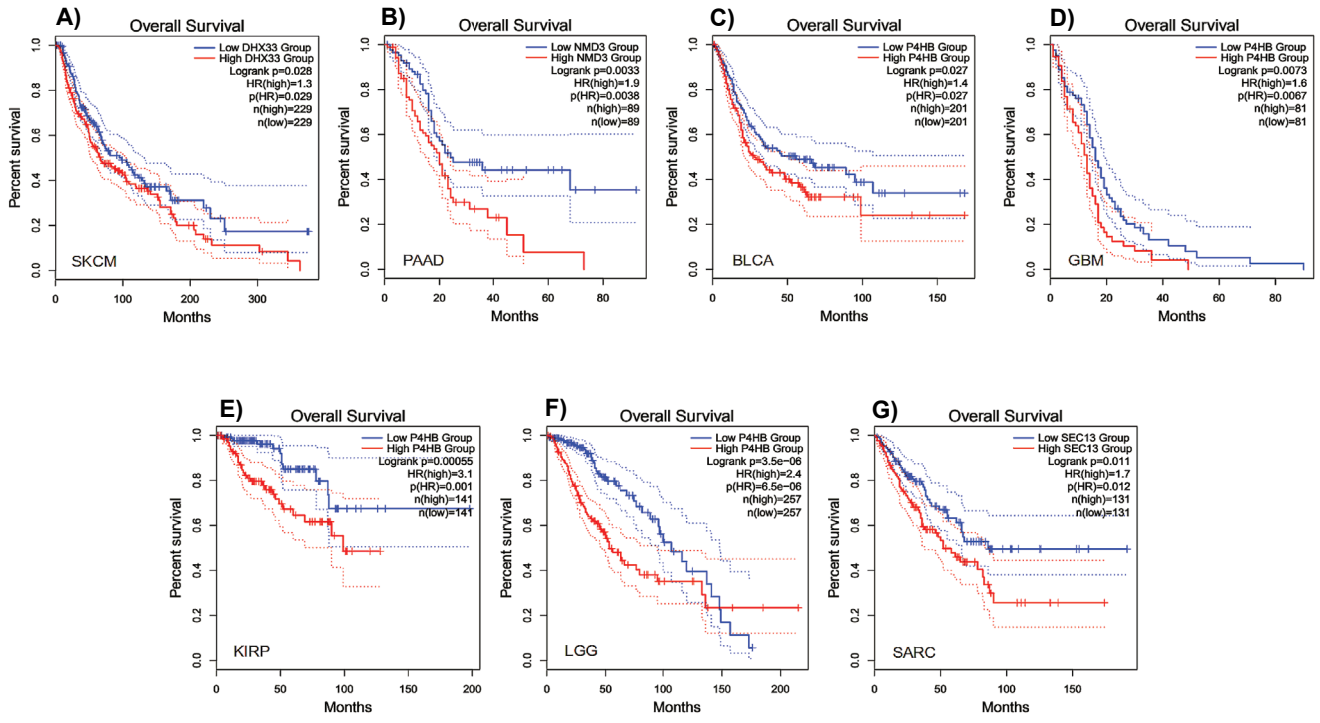

Figure 5. Overall survival (OS) analysis during hub genes overexpression. OS analysis of **A)** DHX3 in SKCM, **B)** NMD3 in PAAD, **C)** P4HB in BLCA, **D)** P4HB in GBM, **E)** P4HB in KIRP, **F)** P4HB in LGG, and **G)** SEC13 in SARC. OS in patients with hub genes overexpression is statistically poor.

YOD1 is a DUB of the ovarian tumor family (13). YOD1 binds to miR-4429 and improves ovarian cancer and cancer cells motility (14). In search of valid miRNA biomarkers for discriminating prostate cancer from benign prostatic hyperplasia, miR-93 and miR-375 were found suitable; they significantly targeted YOD1 and five other genes (15).

Concerning protein K48-linked deubiquitination (GO:0071108) term, except for OTUD4 and YOD1, only USP34 by Enrichr and USP34 and USP9X by ClueGO were found to be differentially expressed in the cancer datasets. If USPs undergo genetic modification or malfunction, in many cases cause cancer. USPs are also believed to be connected with p53 regulation in carcinogenesis (16).

USP9X reduction or inhibition causes the gathering of centrosomal MARCH7, K48/K63 ubiquitination of NPHP5, protein degradation, and deprivation of cilia, the microtubule-based protrusions. This deprivation is responsible for numerous cancers (17). USP9X also improves cell growth of high-grade glioma by β-catenin K48-linked deubiquitination (18).

Searching common DEGs in two meta-analyses involved in ncRNA processing (GO:0034470) in Enrichr eventuated in finding RPS3A and INTS8.

INTS8 is a constituent of the RNA polymerase II-mediated transcription machinery. The dysfunctional INTS8 is the root of many cancers. Investigating 31 different tumors showed a high mutation in INTS8 (19). In search of accurate biomarkers for the timely diagnosis of gastric cancer, INTS8 was found to be reliable (20). To find biomarkers associated with the Barcelona Clinic Liver Cancer (BCLC) staging and hepatocellular carcinoma (HCC) in a study, INTS8 under-expression prolonged overall survival (21). There are also reports of the neuro-development fault and UsnRNA abnormal processing due to mutated INTS8 (22).

RPS3A expression, with boosting NF-κB signaling, intensifies HBx liquefying and develops HCC (23).

Surveying DEGs involved in the term of ribosomal large subunit binding (GO:0043023) in the molecular function in Enrichr indicated the enrichment of DHX33 and NMD3 in the meta-analysis of cancer datasets, sharing the latter DEG with the me-

ta-analysis of normal datasets.

The NMD3, in a study, was discovered to be implicated in resistance to breast anti-cancer drugs. It was reported that under-expression and overexpression of both NMD3 and 4F2 cell-surface antigen heavy chain (SLC3A2) took place in apoptotic signaling pathway and cell redox homeostasis in McF-7 breast cancer cells, respectively (24).

Apart from regulating cell growth and proliferation, the DHX33 boosts the initiation of mRNA translation, as its decreased levels inhibit mRNA translation (25). In addition, the nucleus and cytosol localizing of DHX33 in various cancer cell lines have been reported. What is more, DHX33 strongly induces tumorigenesis as its overexpressed level was observed in Ras-mutated human lung cancer cell lines (26). DHX33 also lifts cancer cell movement (27). It has been shown that DHX33 cooperates with the AP-2 transcription factor incites Bcl-2 gene expression and survives cancer cells (28). Overexpressing DHX33 in human Glioblastoma Multiforme and preserving movement and multiplying glioblastoma cells is also reported (29). Additionally, it has been discovered that DHX33 regulates vital genes included in the cell cycle, apoptosis, and movement, and improves colon cancer (30). Cancer cells trigger cells to proliferate via stimulating deregulated genes involved in glycolysis. DHX33 promotes these genes through the Warburg effect (31).

Analyzing protein processing in endoplasmic reticulum term in KEGG pathway in Enrichr and ClueGO revealed SEC13, SKP1, and YOD1 upregulation in both meta-analyses, and P4HB in the meta-analysis of cancer datasets.

A high expression of SKP1 in non-small cell lung tumors, leading to late diagnosis, was reported (32). So to cure cancer with the conventional drugs, SKP1 is targeted. Research on colorectal cancer (CRC) stem cells (CRC-SCs) disclosed SKP1 overexpression in CRC-SCs with a poor diagnosis (33).

SEC13 involvement in cancer has been reported. With implicating in the signaling cascade of transforming growth factor-β (TGF-β), SEC13 involvement in carcinogenesis is believed. To detect upregulated genes in gastric adenocarcinoma, suppression subtractive hybridization was applied on this cancer tissue, and eight overexpressed genes, including SEC13, were screened (6).

P4HB is connected to many cancers. This chaperon collaborates in a handful of enzymatic activities in the endoplasmic reticulum (34). Based on a report, abnormally expressed P4HB managed to deteriorate glioma cancer phenotypically. In addition, overexpressed P4HB enforced cancer attributes such as developed tumorigenesis, migration, invasion through the MAPK signaling pathway (35). Similarly, P4HB was introduced as a biomarker for diffuse gliomas (35). Western blot analysis has also confirmed P4HB overexpression in bladder carcinoma (36). In a report, by using western blot, RT-qPCR, and bioinformatics analyses, high expression of P4HB in bladder cancer with feeble overall survival was also shown (37)

6. Conclusion

Finding reliable cancer-associated biomarkers for the timely diagnosis of cancers is of paramount importance. EGF, modulating many crucial processes in the human body, when upregulates can cause cancer. The present study by conducting two meta-analyses detected DEGs, biological processes, molecular functions, and pathways malfunction by EGF-stimulation. Using Enrichr and ClueGO, DEGs involvement in specific GO terms and pathways were clarified. The hub genes (OTUD4, YOD1, USP34, USP9X, RPS3A, INTS8, SEC13, SKP1, and P4HB), which their overexpression by EGF-stimulation was confirmed by conducting two meta-analyses, showed to participate fully in tumorigenesis.

Acknowledgments

The authors would like to appreciate Dr. Nemat Hedayat Ivrig, a faculty member of the University of Mohaghegh Ardabili, for cloud computing support. This study has been funded by no organizations.

Author Contributions

Conceptualization, G.S; Methodology, G.S, J.G.S and L.T.; Software, G.S; Validation, G.S, L.T, J.G.S., R.M and M.M., Formal Analysis, G.S; Investigation, R.M, and M.M.; Resources, G.S, J.G.S, L.T, E.K, M.A, S.A.H, M.M and R.M., Writing–Original Draft Preparation, G.S.; Writing–Review and Editing, L.T, J.G.S, E.K, S.A.H, and M.A.; Visualization, G.S, R.M, and M.M.; Supervision, L.T, and J.G.S; Project Administration, J.G.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The results of this study are based on the analysis of datasets are as follows: GSE121697, GSE156089, GSE124586, PRJNA322427, GSE40632, GSE85089 and GSE94374.

References

1. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004;**10**(8):789-799. doi:10.1038/nm1087
2. Sever R, Brugge JS. Signal transduction in cancer. *Perspect Med*. 2015;**5**(4):a006098. doi:10.1101/cshperspect.a006098
3. Kim BW, Kim SK, Heo KW, Bae KB, Jeong KH, Lee SH, *et al*. Association between epidermal growth factor (EGF) and EGF receptor gene polymorphisms and end-stage renal disease and acute renal allograft rejection in a Korean population. *Ren Fail*. 2020;**42**(1):98–106. doi:10.1080/0886022x.2019.1710535
4. Hunter T. Epidermal Growth Factor-Induced Tumor Cell Invasion and Metastasis Initiated by Dephosphorylation and Downregula. *Mol Cell Biol*. 2001;**21**(12):4016-4031. doi:10.1128/MCB.21.12.4016-4031.2001
5. Garcia R, Franklin RA, McCubrey JA. EGF induces cell motility and multi-drug resistance gene expression in breast cancer cells. *Cell Cycle*. 2006;**5**(23):2820–2826. doi:10.4161/cc.5.23.3535
6. Mottaghi-Dastjerdi N, Soltany-Rezaee-Rad M, Sepehrizadeh Z, Roshandel G, Ebrahimifard F, Setayesh N. Identification of novel genes involved in gastric carcinogenesis by suppression subtractive hybridization. *Hum Exp Toxicol*. 2015;**34**(1):3-11. doi:10.1177/0960327114532386
7. Alimadadi A, Aryal S, Manandhar I, Joe B, Cheng X. Identification of Upstream Transcriptional Regulators of Ischemic Cardiomyopathy Using Cardiac RNA-Seq Meta-Analysis. *Int J Mol Sci*. 2020;**21**(10):3472. doi:10.3390%2Fijms21103472
8. Rau A, Marot G, Jaffrézic F. Differential meta-analysis of RNA-Seq data from multiple studies. *BMC Bioinform*. 2014;**15**(1):1-10. doi:10.1186/1471-2105-15-91
9. Morrow JK, Lin H-K, Sun S-C, Zhang S. Targeting ubiquitination for cancer therapies. *Future Med Chem*. 20157;**7**(17):2333-2350. doi:10.4155%2Ffmc.15.148
10. Liu Z, Dong X, Yi H-W, Yang J, Gong Z, Wang Y, *et al*. Structural basis for the recognition of K48-linked Ub chain by proteasomal receptor Rpn13. *Cell Discov*. 2019;**5**(1):19. doi:10.1038/s41421-019-0089-7
11. Jaynes PW, Iyengar PV, Lui SKL, Tan TZ, Vasilevski N, Wright SCE, *et al*. OTUD4 enhances TGFβ signaling through regulation of the TGFβ receptor complex. *Sci Rep*. 2020;**10**(1):1-13. doi:10.1038/s41598-020-72791-0
12. Zhao X, Su X, Cao L, Xie T, Chen Q, Li J, *et al*. OTUD4: A potential prognosis biomarker for multiple human cancers. *Cancer Manag Res*. 2020;**12**:1503-1512. doi:10.2147/CMAR.S233028
13. Ernst R, Mueller B, Ploegh HL, Schlieker C. The otubain YOD1 is a deubiquitinating enzyme that associates with p97 to facilitate protein dislocation from the ER. *Mol Cell*. 2009;**36**(1):28-38. doi:10.1016%2Fj.molcel.2009.09.016
14. Zhu Y, Chen P, Shi L, Zhu T, Chen X. MiR-4429 suppresses the malignant development of ovarian cancer by targeting YOD1. *Eur Rev Med Pharmacol Sci*. 2020;**24**(17):8722-8730. doi:10.26355/eurrev_202009_22809
15. Ciszkowicz E, Porzycki P, Semik M, Kaznowska E, Tyrka M. MiR-93/miR-375: Diagnostic potential, aggressiveness correlation and common target genes in prostate cancer. *Int J Mol Sci*. 2020;**21**(16):5567. doi:10.3390/ijms21165667
16. Deng L, Meng T, Chen L, Wei W, Wang P. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduct Target Ther*. 2020;**5**(1):1-28. doi:10.1038/s41392-020-0107-0
17. Das A, Qian J, Tsang WY. USP9X counteracts differential ubiquitination of NPHP5 by MARCH7 and BBS11 to regulate ciliogenesis. *PLoS Genet*. 2017;**13**(5):e1006791. doi:10.1371%2Fjournal.pgen.1006791
18. Yang B, Zhang S, Wang Z, Yang C, Ouyang W, Zhou F, *et al*. Deubiquitinase USP9X deubiquitinates β-catenin and promotes high grade glioma cell growth. *Oncotarget*. 2016;**7**(48):79515. doi:10.18632%2Foncotarget.12819
19. Federico A, Rienzo M, Abbondanza C, Costa V, Ciccodicola A, Casamassimi A. Pan-cancer mutational and transcriptional analysis of the integrator complex. *Int J Mol Sci*. 2017;**18**(5):936. doi:10.3390%2Fijms18050936
20. Cheng L, Zhang Q, Yang S, Yang Y, Zhang W, Gao H, *et al*. A 4-gene panel as a marker at chromosome 8q in Asian gastric cancer patients. *Genomics*. 2013;**102**(4):323-330. doi:10.1016/j.ygeno.2013.05.004
21. Xu W, Rao Q, An Y, Li M, Zhang Z. Identification of biomarkers for Barcelona Clinic Liver Cancer staging and overall survival of patients with hepatocellular carcinoma. *PLoS One*. 2018;**13**(8):e0202763. doi:10.1371/journal.pone.0202763
22. Kirstein N, Dos Santos HG, Blumenthal E, Shiekhattar R. The Integrator complex at the crossroad of coding and noncoding RNA. *Curr Opin Cell Biol*. 2021;**70**:37-43. doi:10.1016/j.ceb.2020.11.003
23. Lim K-H, Kim K-H, Choi SI, Park E-S, Park SH, Ryu K, *et al*. RPS3a over-expressed in HBV-associated hepatocellular carcinoma enhances the HBx-induced NF-κB signaling via its novel chaperoning function. *PloS One*. 2011;**6**(8):e22258. doi:10.1371/journal.pone.0022258
24. Sommer AK, Hermawan A, Ljepoja B, Fröhlich T, Arnold GJ, Wagner E, *et al*. A proteomic analysis of chemoresistance development via sequential treatment with doxorubicin reveals novel players in MCF-7 breast cancer cells. International journal of molecular medicine. *Mol Med*. 2018;**42**(4):1987-1997. doi:10.3892/ijmm.2018.3781
25. Zhang Y, You J, Wang X, Weber J. The DHX33 RNA helicase promotes mRNA translation initiation. *Mol Cell Biol*. 2015;**35**(17):2918–2931. doi:10.1128/mcb.00315-15
26. Yuan B, Wang X, Fan C, You J, Liu Y, Weber JD, *et al*. DHX33 transcriptionally controls genes involved in the cell cycle. *Mol Cell Biol*. 2016;**36**(23):2903-2917. doi:10.1128/MCB.00314-16
27. Fu J, Liu Y, Wang X, Yuan B, Zhang Y. Role of DHX33 in c-Myc-induced cancers. *Carcinogenesis*. 2017;**38**(6):649-660. doi:10.1093/carcin/bgx041
28. Wang J, Feng W, Yuan Z, Weber JD, Zhang Y. DHX33 interacts with AP-2β to regulate Bcl-2 gene expression and promote cancer cell survival. *Mol Cell Biol*. 2019;**39**(17):e00017-e00019. doi:10.1128%2FMCB.00017-19
29. Wang H, Yu J, Wang X, Zhang Y. The RNA helicase DHX33

is required for cancer cell proliferation in human glioblastoma and confers resistance to PI3K/mTOR inhibition. *Cell Signal*. 2019;**54**:170-178. doi:10.1016/j.cellsig.2018.12.005

30. Zhu Y, Du Y, Zhang Y. DHX33 promotes colon cancer development downstream of Wnt signaling. *Gene*. 2020;**735**: 144402. doi:10.1016/j.gene.2020.144402
31. Peng C, Hou ST, Deng CX, Zhang Y. Function of DHX33 in promoting Warburg effect via regulation of glycolytic genes. *J Cell Physiol*. 2021;**236**(2):981-996. doi:10.1002/jcp.29909
32. Liu Y-Q, Wang X-L, Cheng X, Lu Y-Z, Wang G-Z, Li X-C, *et al*. Skp1 in lung cancer: clinical significance and therapeutic efficacy of its small molecule inhibitors. *Oncotarget*. 2015;**6**(33):34953. doi:10.18632/oncotarget.5547
33. Tian C, Lang T, Qiu J, Han K, Zhou L, Min D, *et al*. SKP1 promotes YAP-mediated colorectal cancer stemness via suppressing RASSF1. *Cancer Cell Int.* 2020;**20**(1):579. doi:10.1186/s12935-020-01683-0
34. Zou H, Wen C, Peng Z, Shao Y-Y, Hu L, Li S, *et al*. P4HB and PDIA3 are associated with tumor progression and therapeutic outcome of diffuse gliomas. *Oncol Rep*. 2018;**39**(2):501-510. doi:10.3892%2For.2017.6134
35. Sun S, Kiang KM, Ho AS, Lee D, Poon M-W, Xu F-F, *et al*. Endoplasmic reticulum chaperone prolyl 4-hydroxylase, beta polypeptide (P4HB) promotes malignant phenotypes in glioma via MAPK signaling. *Oncotarget*. 2017;**8**(42):71911. doi:10.18632%2Foncotarget.18026
36. Wang X, Bai Y, Zhang F, Yang Y, Feng D, Li A, *et al*. Targeted inhibition of P4HB promotes cell sensitivity to gemcitabine in Urothelial carcinoma of the bladder. *Onco Targets Ther*. 2020;**13**:9543. doi:10.2147/ott.s267734
37. Wu Y, Peng Y, Guan B, He A, Yang K, He S, *et al*. P4HB: A novel diagnostic and prognostic biomarker for bladder carcinoma. *Oncol Lett*. 2021;**21**(2):1-1. doi:10. 3892/ol.2020.12356