In silico analysis of genes and pathways related to acute myeloid leukemia presenting leukopenia

Ajeet Kumar

Banaras Hindu University

Ravi Bhushan

Banaras Hindu University

Pawan K. Dubey

Banaras Hindu University

Vijai Tilak

Banaras Hindu University

Vineeta Gupta

Banaras Hindu University

Nilesh Kumar

Banaras Hindu University

S. V.S Raju

Banaras Hindu University

Akhtar Ali (✉ akhtar_genetics@yahoo.co.in)

Banaras Hindu University

Research Article

Abstract

Acute myeloid leukemia (AML) is the progenitor and hematopoietic stem cell blood cancer. Chromosomal abnormalities include balanced translocations between two chromosomes like t[8;21] and t[15;17]) in malignant cells. The current study aimed to investigate AML with leukopenia and gene expression changes in high-white and low-white counts B-cell. The total number of samples used was ten. The raw gene expression profiles (ID: GSE20482) of bone marrow obtained from AML patients showed five high-white counts B-cell and five low-white counts B-cell expressed genes differentially expressed. Genes that corresponded to official gene symbols were selected for protein-protein interaction (PPI) and sub-network construction (score > 0.4). In addition, functional annotation of gene ontology (GO) and pathway analysis were performed for the genes involved in networking.

Results

A total of 846 genes were identified as differentially expressed, and 406 genes were upregulated; and another 440 genes were downregulated. The remaining 14 genes that interacted with each other were significantly identified. The hub genes GNB4, LAMTOR2, ACTN4, HGSNAT, and TMED1 were upregulated while the downregulated DEGs forming hub nodes were UBR4, FBXO30, KLHL21, DCTN6, RNF123, RNF114. AML significantly affects the expression of genes involved in cell differentiation, apoptosis, cell signaling, and protein modification. AML cells enter the blood quickly and spread to the liver, spleen, and central nervous system. These are a total of thirteen pathways. They were enriched, and AML was found to significantly impact genes involved in oxidative phosphorylation, actin cytoskeleton regulation, endocytosis, phagocytosis, shigellosis, and epithelial cell signaling in helicobacter, adherent junction, pertussis, bile secretion, malaria, and African trypanosomiasis.

Conclusions

Hub genes like GNB4 and UBR4 provide a novel biomarker in AML.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous malignant disorder of the hematopoietic stem cells these are characterized by abnormal proliferation of the immature cell, blast cells, and disabling of the production of normal blood cells [1]. AML is a genetic disease first established by Janet Rowley, who discovered the somatic chromosomal abnormalities in the leukemic cell of the patient sample, and they show the balanced translocations of chromosome (i.e., t[8;21] and t[15;17]) [2]. AML affects people worldwide, such as leukemia, lymphoma, and multiple myelomas. It is the most common type of leukemia in adults and has the lowest survival rate of all different kinds of leukemia. AML is a rare disease. They are highly malignant neoplasms and supervise a large number of cancer-related deaths. Approximately 25% of all leukemia in adults in the West constitutes the most frequent form of leukemia. In AML the chromosomal abnormalities are found in t(8;21), t(15;17), inv (16), and 11q23. These aberrations produce PML-RARA, RUNX1-RUNX1T1, CBF-MYH11, and MLL-fusion genes. Mutations

generally occur in transcription factors, signaling of molecules, tumor suppressor genes, epigenetic regulators, RNA splicing factors, and cohesion complexes, with FLT3, NPM1, and DNMT3A the most frequently mutated genes in AML.

Cytogenetics provides robust diagnostic information and provides the framework for risk stratification in AML patients, and it has several limitations[3]. Some technical failures, like cytogenetics, cannot relate the gene fusions, for example, NUP98-NSD1, CBFA2T3-GLIS2, and MNX1-ETV6, which forecast the poor outcome in pediatric patients of AML[4] [5].AML was initially subdivided based on morphology (French-American-British system). Those are helpful for the categorization of pathology. Later, some genomics and clinical factors are initiated to coordinate with chemotherapy's reaction and overall survival. In 2017 the European Leukemia Net (ELN) separated into three prognostic groups based on genetic mutation and abnormalities in acute myeloid leukemia. These are categorized as favorable, intermediate, and adverse[6].

Material And Methodology

2.1. Differentially expressed genes and Microarray data Screening in AML

The raw data procured from the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) datasets (http://www.ncbi.nlm.nih.gov/geo/) and gene expression profile datasets obtained from (GENE ID: GSE20482). Hematological stem cells, bone marrow, and peripheral blood from AML patients with leukopenia were used in this study. The datasets derived from a study utilize the GPL6848 Agilent-012391 Whole Human Genome Oligo Microarray G4112A platform. These studies are only on genes of high and low blood count B-cell samples analyzed through bioinformatics techniques.

2.2. Preprocessing of data and screening of DEGs in AML

Gene expression datasets were obtained from NCBI and further these data were used for pre-processed. The expression values of probes communicate to a particular gene have to mean calculated to find the advantage of gene expression. Additionally, up and down-regulated genes were distinguished by using BiGGEsTs software analysis. The probe-level ideogram is transformed into a gene-level ideogram by using GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/). These DEGs pick out, calibrate p values < 0.5 and threshold logFC values are > 0.1 for up and < − 0.1 for down-regulated genes.

2.3 Principal component analysis and heat map generation in AML

Principal component analysis (PCA) was executed, and constructed the heat map using the ClustVis software available online and plotted the PCA graph on DEGs data. ClustVis support maximum file size is larger, so it is not feasible to assemble the PCA plot for all gene expression studies.

2.4 PPI network creation and generation of sub-network in AML

DEGs data were uploaded to the STRING v 10.5 (http://www. string-db.org/). It is an online software dataset for anticipating the functional exchange between the proteins and predicting a collaborated outcome for protein-protein interaction among gene positions, and the combined outcome N 0.4 was set as the standard. Prepare for modeling networking and sub-network; we used Cytoscape v 3.2.1. Cytoscape (http://www.cytoscape.org/) is a bioinformatics software for the investigation and visualization of biological networks with high throughput data. The assemble and edge coefficient were taken as standard for network construction.

2.5 Analysis of differentially functional expressed genes in AML

The database for annotation, visualization, and integrated discovery (DAVID, http://david.abcc.ncifcrf.gov/) draft of action is a database combined with a comprising set of functional interpretations and a huge genes list. Gene ontology (GO) amplification performed as well as biological, molecular, and cellular modules were implemented by using DAVID v 6.8 and STRING v 10.5. Based on hypergeometric distribution, DAVID clutches the genes with exactly similar functions or related as a whole set. DEGs related pathways analysis and completed with the Panther Classification System (http://pantherdb.org/)

2.6 Analysis of Interaction network

Biological General Repository for Interaction Datasets (BioGRID) tools are used to identify (https://wiki.thebiogrid.org/doku.php/ORCS:tools) the possible interconnection for acute myeloid leukemia. BioGRID is curated for biological databases like protein-protein interactions and acute myeloid leukemia's genetic or chemical interactions.

2.7 Survival analysis

Overall survival analysis for selected hub genes was performed using GEPIA V2.0. The association between mRNA prognosis and expression of acute myeloid leukemia.

Results

3.1. Differentially expressed genes (DEGs) in AML

846 DEGs were allowed with their formal gene ideogram and up and down-regulated DEGs 406 & 440. Their p-value is < .05. **(Supplementary file 1).** Fourteen genes only were recognized to remarkably interconnect with each other. Mean of AML B cell high and low B cell shown in **Fig. 1.** DEGs were affected based on their common gene expression value.

3.2 Principal component and hierarchical clustering analysis of DEGs in AML

PCA plot shows a distributed plot with axes communicated to two independent principal components, 1 and 2, shown in **Fig. 2A.** Another shows a Heat-map data matrix in which shade summarizes the arithmetic divergence. A heat map was fabricated for differentially expressed genes, shown in **Fig. 2B.** Conceive of screened DEGs along with assembling of Volcano plot shows in **Fig. 3A.** MD plot releases the up and down-regulated genes shown in **Fig. 3B.** Venn diagram showing the overhang between GO terms down-regulated and up-regulated in the AML shows in **Fig. 3C.** Total 198 genes were set up, the novel up-regulated and 235 genes were set up the novel down regulated. DEGs were determined based upon standard gene expression value shows in **Fig. 3D.**

3.3 Protein-protein interaction meshwork in AML

Protein-protein interaction (PPI) of DEGs shown in **Fig. 4.** Pink and blue circle represents up and down-regulated genes. Based on the combined score estimated through STRING. Seventy-six gene pairs (integrated score N 0.4) were elevated to link together and constitute one prime network with 69 nodes and 381 edges. Total five sub-networks will eliminate independently. The up-regulated DEGs and down-regulated DEGs forming hub nodes were GNB4, LAMTOR2, ACTN4, HGSNAT, TMED1, and UBR4 FBXO30, KLHL21, DCTN6, RNF123, RNF114. Clustering coefficient and edge were appropriated as the basis of the selection criteria of hub nodes.

3.4 Construction of Sub-network in AML

Five sub-networks (3 nodes and three edges in network1; 2 nodes and one boundary in networks 2 and 3 both) were pulled out from the primary network beside using Cytoscape shown in **Fig. 5.** Each gene in network **A** was up-regulated, and network **B** was down-regulated. Another network, **C, D**, and **E** lines, shows the correlation between up and downregulated genes.

3.5 Analysis of functional enrichment in AML

Analysis of functional enrichment was carried out and remarkably improved their processes, functions, and components of DEGs (FDR b 0.05) were recorded in Tables 1 and 2 results of GO shown in **Fig. 6.**

Table 1
Gene Ontology (GO) analysis of up regulated DEGs.

Category	Term	PValue	FDR
GOTERM_BP_FAT	GO:0006796 ~ phosphate-containing compound metabolic process	2.98E-06	0.007334
GOTERM_BP_FAT	GO:0006793 ~ phosphorus metabolic process	3.25E-06	0.007334
GOTERM_BP_FAT	GO:0032268 ~ regulation of cellular protein metabolic process	5.35E-05	0.063377
GOTERM_BP_FAT	GO:0051246 ~ regulation of protein metabolic process	5.62E-05	0.063377
GOTERM_BP_FAT	GO:0030029 ~ actin filament-based process	1.48E-04	0.096366
GOTERM_BP_FAT	GO:0044403 ~ symbiosis, encompassing mutualism through parasitism	1.93E-04	0.096366
GOTERM_BP_FAT	GO:0044419 ~ interspecies interaction between organisms	1.93E-04	0.096366
GOTERM_BP_FAT	GO:0030198 ~ extracellular matrix organization	2.07E-04	0.096366
GOTERM_BP_FAT	GO:0016032 ~ viral process	2.07E-04	0.096366
GOTERM_BP_FAT	GO:0043062 ~ extracellular structure organization	2.14E-04	0.096366
GOTERM_BP_FAT	GO:0044764 ~ multi-organism cellular process	2.37E-04	0.097261
GOTERM_BP_FAT	GO:0080134 ~ regulation of response to stress	3.18E-04	0.118647
GOTERM_BP_FAT	GO:0035336 ~ long-chain fatty-acyl-CoA metabolic process	3.42E-04	0.118647
GOTERM_BP_FAT	GO:0007010 ~ cytoskeleton organization	3.75E-04	0.120822
GOTERM_BP_FAT	GO:0035337 ~ fatty-acyl-CoA metabolic process	4.85E-04	0.142612
GOTERM_BP_FAT	GO:0030036 ~ actin cytoskeleton organization	5.06E-04	0.142612
GOTERM_BP_FAT	GO:0007015 ~ actin filament organization	6.11E-04	0.162057
GOTERM_BP_FAT	GO:1901135 ~ carbohydrate derivative metabolic process	0.001101	0.274455
GOTERM_BP_FAT	GO:0032956 ~ regulation of actin cytoskeleton organization	0.001196	0.274455
GOTERM_BP_FAT	GO:0008610 ~ lipid biosynthetic process	0.001217	0.274455
GOTERM_BP_FAT	GO:0033036 ~ macromolecule localization	0.001309	0.281035
GOTERM_CC_FAT	GO:0070062 ~ extracellular exosome	4.89E-12	1.23E-09
GOTERM_CC_FAT	GO:1903561 ~ extracellular vesicle	6.82E-12	1.23E-09

Category	Term	PValue	FDR
GOTERM_CC_FAT	GO:0043230 ~ extracellular organelle	6.98E-12	1.23E-09
GOTERM_CC_FAT	GO:0031988 ~ membrane-bounded vesicle	1.59E-10	2.10E-08
GOTERM_CC_FAT	GO:0044421 ~ extracellular region part	1.61E-06	1.71E-04
GOTERM_CC_FAT	GO:0005925 ~ focal adhesion	6.53E-06	5.63E-04
GOTERM_CC_FAT	GO:0005924 ~ cell-substrate adherens junction	7.45E-06	5.63E-04
GOTERM_CC_FAT	GO:0030055 ~ cell-substrate junction	9.21E-06	6.09E-04
GOTERM_CC_FAT	GO:0005912 ~ adherens junction	2.85E-05	0.001435
GOTERM_CC_FAT	GO:0000323 ~ lytic vacuole	2.98E-05	0.001435
GOTERM_CC_FAT	GO:0005764 ~ lysosome	2.98E-05	0.001435
GOTERM_CC_FAT	GO:0070161 ~ anchoring junction	4.61E-05	0.002033
GOTERM_CC_FAT	GO:0005576 ~ extracellular region	1.56E-04	0.006356
GOTERM_CC_FAT	GO:0005773 ~ vacuole	2.02E-04	0.007627
GOTERM_MF_FAT	GO:0008092 ~ cytoskeletal protein binding	8.72E-05	0.076204
GOTERM_MF_FAT	GO:0019899 ~ enzyme binding	7.15E-04	0.196724
GOTERM_MF_FAT	GO:0016818 ~ hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	0.00115	0.196724
GOTERM_MF_FAT	GO:0016817 ~ hydrolase activity, acting on acid anhydrides	0.001195	0.196724
GOTERM_MF_FAT	GO:0003924 ~ GTPase activity	0.001305	0.196724
GOTERM_MF_FAT	GO:0019901 ~ protein kinase binding	0.001807	0.196724
GOTERM_MF_FAT	GO:0017111 ~ nucleoside-triphosphatase activity	0.001865	0.196724
GOTERM_MF_FAT	GO:0016779 ~ nucleotidyltransferase activity	0.002001	0.196724
GOTERM_MF_FAT	GO:0016462 ~ pyrophosphatase activity	0.002194	0.196724
GOTERM_MF_FAT	GO:0019001 ~ guanyl nucleotide binding	0.002251	0.196724
GOTERM_MF_FAT	GO:0044325 ~ ion channel binding	0.002533	0.201247
GOTERM_MF_FAT	GO:0005525 ~ GTP binding	0.002974	0.216623
GOTERM_MF_FAT	GO:0003779 ~ actin binding	0.004198	0.282254
GOTERM_MF_FAT	GO:0032561 ~ guanyl ribonucleotide binding	0.00513	0.305882

Category	Term	PValue	FDR
GOTERM_MF_FAT	GO:0044769 ~ ATPase activity, coupled to transmembrane movement of ions, rotational mechanism	0.00527	0.305882
GOTERM_MF_FAT	GO:0032403 ~ protein complex binding	0.0056	0.305882
GOTERM_MF_FAT	GO:0051015 ~ actin filament binding	0.006499	0.328758
GOTERM_MF_FAT	GO:0019900 ~ kinase binding	0.006975	0.328758
GOTERM_MF_FAT	GO:0045182 ~ translation regulator activity	0.007147	0.328758
GOTERM_MF_FAT	GO:0005200 ~ structural constituent of cytoskeleton	0.008236	0.359914
GOTERM_MF_FAT	GO:0030371 ~ translation repressor activity	0.009062	0.374378
GOTERM_MF_FAT	GO:0015078 ~ hydrogen ion transmembrane transporter activity	0.009485	0.374378
GOTERM_MF_FAT	GO:0003684 ~ damaged DNA binding	0.010045	0.374378

Table 2
Gene Ontology (GO) analysis of down regulated DEGs.

Category	Term	PValue	FDR
GOTERM_BP_FAT	GO:0046501 ~ protoporphyrinogen IX metabolic process	3.44E-08	6.86E-05
GOTERM_BP_FAT	GO:0006779 ~ porphyrin-containing compound biosynthetic process	3.87E-08	6.86E-05
GOTERM_BP_FAT	GO:0006778 ~ porphyrin-containing compound metabolic process	6.72E-08	6.86E-05
GOTERM_BP_FAT	GO:0015669 ~ gas transport	6.79E-08	6.86E-05
GOTERM_BP_FAT	GO:0033014 ~ tetrapyrrole biosynthetic process	1.01E-07	8.12E-05
GOTERM_BP_FAT	GO:0033013 ~ tetrapyrrole metabolic process	3.52E-06	0.002366
GOTERM_BP_FAT	GO:0006783 ~ heme biosynthetic process	4.56E-06	0.002633
GOTERM_BP_FAT	GO:0015671 ~ oxygen transport	9.93E-06	0.005014
GOTERM_BP_FAT	GO:0006782 ~ protoporphyrinogen IX biosynthetic process	2.20E-05	0.009871
GOTERM_BP_FAT	GO:0042168 ~ heme metabolic process	3.84E-05	0.015509
GOTERM_BP_FAT	GO:0051188 ~ cofactor biosynthetic process	1.32E-04	0.048324
GOTERM_BP_FAT	GO:0030218 ~ erythrocyte differentiation	1.84E-04	0.061892
GOTERM_BP_FAT	GO:0048534 ~ hematopoietic or lymphoid organ development	2.22E-04	0.068828
GOTERM_BP_FAT	GO:0021953 ~ central nervous system neuron differentiation	2.96E-04	0.085258
GOTERM_BP_FAT	GO:0034101 ~ erythrocyte homeostasis	3.62E-04	0.09735
GOTERM_BP_FAT	GO:0030097 ~ hemopoiesis	4.36E-04	0.110089
GOTERM_BP_FAT	GO:0002520 ~ immune system development	5.65E-04	0.13421
GOTERM_BP_FAT	GO:0055072 ~ iron ion homeostasis	7.66E-04	0.171931
GOTERM_BP_FAT	GO:0021515 ~ cell differentiation in spinal cord	8.13E-04	0.172802
GOTERM_BP_FAT	GO:0046148 ~ pigment biosynthetic process	8.99E-04	0.181474
GOTERM_BP_FAT	GO:0002262 ~ myeloid cell homeostasis	0.001208	0.232225
GOTERM_BP_FAT	GO:0018130 ~ heterocycle biosynthetic process	0.001338	0.235364
GOTERM_BP_FAT	GO:0051186 ~ cofactor metabolic process	0.001341	0.235364
GOTERM_CC_FAT	GO:0005833 ~ hemoglobin complex	1.30E-09	6.64E-07

Category	Term	PValue	FDR
GOTERM_CC_FAT	GO:0014731 ~ spectrin-associated cytoskeleton	4.88E-04	0.124537
GOTERM_CC_FAT	GO:0030863 ~ cortical cytoskeleton	7.42E-04	0.126063
GOTERM_CC_FAT	GO:0044448 ~ cell cortex part	0.001803	0.229872
GOTERM_CC_FAT	GO:0044445 ~ cytosolic part	0.005077	0.517882
GOTERM_CC_FAT	GO:0008091 ~ spectrin	0.014618	1
GOTERM_CC_FAT	GO:0005829 ~ cytosol	0.035669	1
GOTERM_CC_FAT	GO:0005938 ~ cell cortex	0.038142	1
GOTERM_CC_FAT	GO:0072562 ~ blood microparticle	0.045891	1
GOTERM_CC_FAT	GO:0031672 ~ A band	0.04615	1
GOTERM_CC_FAT	GO:0030864 ~ cortical actin cytoskeleton	0.053691	1
GOTERM_CC_FAT	GO:0099568 ~ cytoplasmic region	0.054347	1
GOTERM_MF_FAT	GO:0005344 ~ oxygen transporter activity	7.41E-06	0.006183
GOTERM_MF_FAT	GO:0019825 ~ oxygen binding	6.57E-04	0.27416
GOTERM_MF_FAT	GO:0003700 ~ transcription factor activity, sequence-specific DNA binding	0.00173	0.366809
GOTERM_MF_FAT	GO:0001071 ~ nucleic acid binding transcription factor activity	0.001759	0.366809
GOTERM_MF_FAT	GO:0008047 ~ enzyme activator activity	0.002244	0.374321
GOTERM_MF_FAT	GO:0020037 ~ heme binding	0.002743	0.381242
GOTERM_MF_FAT	GO:0046906 ~ tetrapyrrole binding	0.004189	0.47989
GOTERM_MF_FAT	GO:0015399 ~ primary active transmembrane transporter activity	0.005179	0.47989
GOTERM_MF_FAT	GO:0015405 ~ P-P-bond-hydrolysis-driven transmembrane transporter activity	0.005179	0.47989
GOTERM_MF_FAT	GO:0042910 ~ xenobiotic transporter activity	0.006461	0.538849
GOTERM_MF_FAT	GO:0030506 ~ ankyrin binding	0.008435	0.639508
GOTERM_MF_FAT	GO:0042626 ~ ATPase activity, coupled to transmembrane movement of substances	0.011255	0.663746
GOTERM_MF_FAT	GO:0016820 ~ hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	0.012306	0.663746

Category	Term	PValue	FDR
GOTERM_MF_FAT	GO:0022804 ~ active transmembrane transporter activity	0.012817	0.663746
GOTERM_MF_FAT	GO:0003810 ~ protein-glutamine gamma-glutamyltransferase activity	0.014862	0.663746
GOTERM_MF_FAT	GO:0097159 ~ organic cyclic compound binding	0.015401	0.663746
GOTERM_MF_FAT	GO:0051537 ~ 2 iron, 2 sulfur cluster binding	0.015733	0.663746
GOTERM_MF_FAT	GO:0044212 ~ transcription regulatory region DNA binding	0.01618	0.663746
GOTERM_MF_FAT	GO:0000975 ~ regulatory region DNA binding	0.016386	0.663746
GOTERM_MF_FAT	GO:0019209 ~ kinase activator activity	0.016927	0.663746
GOTERM_MF_FAT	GO:0001067 ~ regulatory region nucleic acid binding	0.017304	0.663746

3.6 Enrichment of KEGG pathway for DEGs in AML

KEGG pathway enrichment examination was carried out, and remarkably up-regulated KEGG pathway enrichment for DEGs are African trypanosomiasis, malaria, bile secretion, porphyrin, and chlorophyll. Down regulated KEGG pathway enrichment for DEGs is Pertussis, adherent junction, epithelial cell signaling, shigellosis, γ R mediated phagocytosis, endocytosis, regulation of actin cytoskeleton, oxidative phosphorylation, infection of Escherichia coli shown in **Fig. 7.**

3.7 Survival analysis in AML

For hub genes like LAMTOR2, KLHL21 and UBR4 overall survival was found to be low for their increased expression shown in **Fig. 8.** and **Fig. 9.** shown the flowchart illustrating the process of our research.

Discussion

The present study is a little effort to understand the molecular mechanism and its complexity involved in the AML patient. In this study, we used different bioinformatic tools to analyze high throughput gene expression datasets exposed to 846 differentially expressed genes (DEGs). The total number of the up-regulated gene was 406, while 440 genes were down-regulated. Based on GO cluster analysis, major biological processes related to DEGs were the homeostasis process, apoptosis, and protein modification. LAMTOR2, ACTN4, HGSNAT, TMED10 are up-regulated, and UBR4, FBXO30, KLHL21, DCTN6, RNF123, RNF114 were down-regulated respectively. GNB4gene is located on chromosome 3q26.33. Guanine nucleotide-binding proteins β-4, which produce signals to communicate receptors and effector molecules. Which is made up three subunits α, β and γ. Subunits of G protein are encoded by mammalian cells [7].

GNB4 gene mainly encodes through beta subunit. The β subunits are dominant regulators and as well as it also regulates signal transduction in various signaling systems. The LAMTOR2 was normally acknowledged in a yeast two-hybrid on a specific binding partner of MEK1 [8] which assemble in late

endosomes beside the adaptor protein LAMTOR2 (p14)[9]. MP1 and p14 are nearly identical on structurally and very stable heterodimeric complex. Which is essential for ERK stimulation on endosomes[10],[11]. Which depends on gene disruption of p14 and p14/MP1-MEK1 signaling complex modulates the endosomal traffic, EGFR degradation and cellular proliferation [12]. These action are determining for early embryogenesis and throughout tissue homeostasis as released by specific deletion of p14 gene in epidermis [13]. ACTN4 gene is actin cross-linking protein which is encoded by human alpha-actin-4 protein. These are correlated with cell motility, invasion and metastasis in cancer[14]. Excessive expression and massive copy number extension of ACTN4 in different cancer tissue has been also reported and they are associated with the imperfect prognosis in diverse type of cancer [15]. The spectrin genes are superfamily which belongs to the alpha actin and it represents a various group of cytoskeletal proteins. Diverse roles of alpha actin are an actin-binding protein in various types of cell. Gene scramble an isoform of actinin non-muscle, which is condensed in the cytoplasm, and elaborate in metastatic processes. HGSNAT gene encodes acetyl-CoA:alpha-glucosaminide N-acetyltransferase is an enzyme that catalyzes acetylation of the terminal glucosamine residues of sulfate before its hydrolysis by alpha-N-acetyl glucosaminidase [16]. HGSNAT is located in the lysosomal membrane, and it catalyzes transmembrane acetylation in which the terminal glucosamine residue of heparan sulfate acquires an acetyl group, so it converts N-acetylglucosamine. TMED10 genes are trans-membrane proteins that can alter distinct proteins of different segments. TMED10 can form more advance oligomeric assemble in cross linking and depend upon the assemble of mIL-1β in THP-1 cells. These cells are containing a unique transmembrane domain, a luminal signal peptide (SS), a Golgi dynamics (GOLD) domain, and a C-terminal tail (CT) facing the cytoplasm.

Downregulation gene

The UBR4 gene encodes Ubiquitin-protein ligase UBR4 enzyme (UBR4) in humans, found in chromosome number one. It encodes the protein that appears to be a component of cytoplasm in cytoskeletal [17]. UBR4 gene mark to human papillomavirus. These related pathways are the PI3K-Akt signaling pathway and the Innate Immune System[18]. The function of UBR4 gene calmodulin, ubiquitinase ligase activity, transferase activity, binding of zinc ion, and monitoring the integrin-mediated signaling (https://www.genecards.org/cgi-bin/carddisp.pl?gene=UBR4). The therapeutically consistent protein interconnection organized by N-terminal acetylation involves assembling an E2−E3 ubiquitin-like protein−ligation complex, nucleosome binding by an epigenetic regulator, cytoskeletal organization, and integrity of the anaphase-promoting complex [18]. F-Box Protein 30 (FBXO30) is a Protein Coding gene, and it is paralogous of the FBXO40. It associated with FBXO30 include Nasopharyngeal Carcinoma disease. The related pathways are MHC Class I mediated antigen presentation and Immune System. FBX030 also regulates the chromosome segregation of oocyte meiosis. (https://www.genecards.org/cgi-bin/carddisp.pl?gene=FBXO30). The gene family of Kelch-like protein 21 (KLHL) is an interchangeable signal transduction mechanism and protein degeneration along with ligase and cell-cycle regulation[19]. KLHLs attached and stain as a particular membrane for degradation and significance of sequence dissimilarity in the BTB (Broad Complex, Tramtrack, and BricàBrac) domain20. They transpire through the E3 ligase complex, which KLHLs connect to the Cullin-3 (CUL3)[20]. BACK domains and namesake Kelch

motifs distinguish KLHLs from other BTB proteins[21]. The Kelch motif connects with actin, which recognizes organelles, plasma membrane, and cytoskeleton [22]. These genes are adapters for BCR (BTB-CUL3-RBX1). Ubiquitin-protein ligase complexes are essential for well-organized chromosome arrangement or cytokinesis. The ligase complex control and restrain the chromosomal passenger complex (CPC) from chromosomes and moderate the ubiquitinated AURKB[23]. DCTN6 Dynactin (DCTN) is a diverse subunit protein that drives retrograde transport in cells [24]. DCTN6 (Dynactin Subunit 6) is a Protein Coding gene which is RGD (Arg-Gly-Asp) motif in the N-terminal region encoded by DCTN6. Adherent effect of macromolecular proteins such as fibronectin and identical biological function is still not recognized in some pathways relevant to the cell cycle, chromosome separation and transport to the Golgi, and consecutive modification. Total six subunits of DCTN specify to as dynactin 1 to 6 (DCTN1−6), and these subunits of DCTN are severe to the structural and function of DCTN[25],[26]. RNF123, the protein encoded by the RNF gene, contains a C-terminal ring finger domain, a motif present in various functionally distinct proteins involved in protein-protein or protein-DNA interactions. This protein shows the E3 ligase activity regarding the cyclin-dependent kinase inhibitor, also known as p27 or KIP1.

Conclusion

The present study is a miniature attempt to understand the molecular mechanisms and complexity of the AML patient sample. In this study, we analyzed high throughput gene expression datasets. However, additional research and experimental authentication are still required to certify the results. The crucial biological processes associated with DEGs, found on GO cluster exploration, were metabolic process, signal transduction, apoptosis, and protein purifying. DEGs that extensively initiate the hub nodes are LAMTOR2, KLHL21, and UBR4. For hub genes like LAMTOR2, KLHL21, and UBR4 overall survival was low for their increased expression. The hallmark of cancer is an irregular cellular metabolism. Besides, it promotes glycolysis and lipid biosynthesis and plays a crucial role in the growth of tumors in AML. Most of the carbon sources synthesize fatty acids. It is a form of glucose in mammalian cells, and it conducts de novo lipid synthesis and building blocks for a tumor cell. Thirteen pathways were enriched, and genes related to oxidative phosphorylation, regulation of actin cytoskeleton, endocytosis, phagocytosis, shigellosis, epithelial cell signaling, adherent junction, pertussis, bile secretion, malaria, African trypanosomiasis were found significantly affected by AML.

Declarations

Ethical Approval- Ethical approval was obtained from the institutional ethical committee (Ref. No: I.Sc./ECM-XII/2021-22) Institute of Science, Banaras Hindu University.

Competing interests - can be non-financial in nature.

Authors' contributions

Ajeet Kumar, M.Sc.: Ph.D. Scholar, Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, Varanasi, U.P., India

Contributions in the manuscript: Data analysis, manuscript writing.

Ravi Bhushan, M.Sc.: Ph.D. Scholar, Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, Varanasi, U.P., India

Contributions in the manuscript: Data analysis, manuscript writing.

Pawan K. Dubey, Ph.D.: Professor and head, Department of pathology, Institute of Medical Sciences, Banaras Hindu University, Varanasi, U.P., India.

Contributions in the manuscript: Data analysis, reviewing the literatures, manuscript writing.

Vijai Tilak, M.D.: Professor and head, Department of pathology, Institute of Medical Sciences, Banaras Hindu University, Varanasi, U.P., India.

Contributions in the manuscript: data analysis, manuscript writing.

Vineeta Gupta, Professor, Department of Pediatrics, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India.

Contributions in the manuscript: manuscript writing.

Nilesh Kumar, Assistant Professor Department of Pediatrics, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India.

Contributions in the manuscript: manuscript writing.

S.V.S Raju Professor, Department of Entomology and Agricultural Zoology, Institute of Agriculture Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India.

Contributions in the manuscript: data analysis, manuscript writing.

Akhtar Ali, Ph.D.: Assistant Professor, Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, Varanasi, U.P., India

Contributions in the manuscript: Supervised the study, conceived and designed the experiments; data analysis, manuscript writing.

Availability of data and materials

Data derived from public domain resources. The data that support the findings of this study are available in NCBI GEO dataset at SE20482. These data were derived from the following resources available in the public domain: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20482

References

1. Tamamyan G, Kadia T, Ravandi F, Borthakur G, Cortes J, Jabbour E, Daver N, Ohanian M, Kantarjian H, Konopleva M. Frontline treatment of acute myeloid leukemia in adults. Critical Reviews in Oncology/Hematology, 2017 110: 20–34.

2. Pescia G, Jotterand M. POSSIBLE EVIDENCE OF X-Y INTERCHANGE IN AN XX MALE. The Lancet, 1977 309: 550.

3. Grimwade D, Mrózek K. Diagnostic and Prognostic Value of Cytogenetics in Acute Myeloid Leukemia. Hematology/Oncology Clinics of North America, 2011 25: 1135–1161.

4. Gruber TA, Larson Gedman A, Zhang J, Koss CS, Marada S, Ta HQ, Chen S-C, Su X, Ogden SK, Dang J, Wu G, Gupta V, Andersson AK, Pounds S, Shi L, Easton J, Barbato MI, Mulder HL, Manne J, Wang J, Rusch M, Ranade S, Ganti R, Parker M, Ma J, Radtke I, Ding L, Cazzaniga G, Biondi A, Kornblau SM, Ravandi F, Kantarjian H, Nimer SD, Döhner K, Döhner H, Ley TJ, Ballerini P, Shurtleff S, Tomizawa D, Adachi S, Hayashi Y, Tawa A, Shih L-Y, Liang D-C, Rubnitz JE, Pui C-H, Mardis ER, Wilson RK, Downing JR. An Inv(16)(p13.3q24.3)-Encoded CBFA2T3-GLIS2 Fusion Protein Defines an Aggressive Subtype of Pediatric Acute Megakaryoblastic Leukemia. Cancer Cell, 2012 22: 683–697.

5. Beverloo HB, Panagopoulos I, Isaksson M, van Wering E, van Drunen E, de Klein A, Johansson B, Slater R. Fusion of the Homeobox Gene HLXB9 and the ETV6 Gene in Infant Acute Myeloid Leukemias with the t(7;12)(q36;p13)5.

6. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, Dombret H, Ebert BL, Fenaux P, Larson RA, Levine RL, Lo-Coco F, Naoe T, Niederwieser D, Ossenkoppele GJ, Sanz M, Sierra J, Tallman MS, Tien H-F, Wei AH, Löwenberg B, Bloomfield CD. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood, 2017 129: 424–447.

7. Simon M, Strathmann M, Gautam N. Diversity of G proteins in signal transduction. Science, 1991 252: 802–808.

8. Schaeffer HJ. MP1: A MEK Binding Partner That Enhances Enzymatic Activation of the MAP Kinase Cascade. Science, 1998 281: 1668–1671.

9. Wunderlich W, Fialka I, Teis D, Alpi A, Pfeifer A, Parton RG, Lottspeich F, Huber LA. A Novel 14-Kilodalton Protein Interacts with the Mitogen-Activated Protein Kinase Scaffold Mp1 on a Late Endosomal/Lysosomal Compartment. Journal of Cell Biology, 2001 152: 765–776.

10. Teis D, Wunderlich W, Huber LA. Localization of the MP1-MAPK Scaffold Complex to Endosomes Is Mediated by p14 and Required for Signal Transduction. Developmental Cell, 2002 3: 803–814.

11. Kurzbauer R, Teis D, de Araujo MEG, Maurer-Stroh S, Eisenhaber F, Bourenkov GP, Bartunik HD, Hekman M, Rapp UR, Huber LA, Clausen T. Crystal structure of the p14/MP1 scaffolding complex: How a twin couple attaches mitogen-activated protein kinase signaling to late endosomes. Proceedings of the National Academy of Sciences, 2004 101: 10984–10989.

12. Teis D, Taub N, Kurzbauer R, Hilber D, de Araujo ME, Erlacher M, Offterdinger M, Villunger A, Geley S, Bohn G, Klein C, Hess MW, Huber LA. p14–MP1-MEK1 signaling regulates endosomal traffic and cellular proliferation during tissue homeostasis. Journal of Cell Biology, 2006 175: 861–868.

13. Honda K, Yamada T, Endo R, Ino Y, Gotoh M, Tsuda H, Yamada Y, Chiba H, Hirohashi S. Actinin-4, a Novel Actin-bundling Protein Associated with Cell Motility and Cancer Invasion. Journal of Cell Biology, 1998 140: 1383–1393.

14. Yamamoto S, Tsuda H, Honda K, Kita T, Takano M, Tamai S, Inazawa J, Yamada T, Matsubara O. Actinin-4 expression in ovarian cancer: a novel prognostic indicator independent of clinical stage and histological type. Mod Pathol, 2007 20: 1278–1285.

15. Wang M-C, Chang Y-H, Wu C-C, Tyan Y-C, Chang H-C, Goan Y-G, Lai W-W, Cheng P-N, Liao P-C. Alpha-Actinin 4 Is Associated with Cancer Cell Motility and Is a Potential Biomarker in Non–Small Cell Lung Cancer. Journal of Thoracic Oncology, 2015 10: 286–301.

16. Klein U, KRESSEt H. Sanfilippo syndrome type C: Deficiency of acetyl-CoA:a-glucosaminide N-acetyltransferase in skin fibroblasts. Proc Natl Acad Sci USA, 1978 5.

17. Lin R, Tao R, Gao X, Li T, Zhou X, Guan K-L, Xiong Y, Lei Q-Y. Acetylation Stabilizes ATP-Citrate Lyase to Promote Lipid Biosynthesis and Tumor Growth. Molecular Cell, 2013 51: 506–518.

18. Huh K-W, DeMasi J, Ogawa H, Nakatani Y, Howley PM, Munger K. Association of the human papillomavirus type 16 E7 oncoprotein with the 600-kDa retinoblastoma protein-associated factor, p600. Proceedings of the National Academy of Sciences, 2005 102: 11492–11497.

19. Scott DC, Hammill JT, Min J, Rhee DY, Connelly M, Sviderskiy VO, Bhasin D, Chen Y, Ong S-S, Chai SC, Goktug AN, Huang G, Monda JK, Low J, Kim HS, Paulo JA, Cannon JR, Shelat AA, Chen T, Kelsall IR, Alpi AF, Pagala V, Wang X, Peng J, Singh B, Harper JW, Schulman BA, Guy RK. Blocking an N-terminal acetylation–dependent protein interaction inhibits an E3 ligase. Nat Chem Biol, 2017 13: 850–857.

20. Stogios P, Prive G. The BACK domain in BTB-kelch proteins. Trends in Biochemical Sciences, 2004 29: 634–637.

21. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. Hum Genomics, 2013 7: 13.

22. Schleifer RJ, Li S, Nechtman W, Miller E, Bai S, Sharma A, She J-X. KLHL5 knockdown increases cellular sensitivity to anticancer drugs. Oncotarget, 2018 9: 37429–37438.

23. Schroer TA. DYNACTIN. Annu Rev Cell Dev Biol, 2004 20: 759–779.

24. Eckley DM, Gill SR, Melkonian KA, Bingham JB, Goodson HV, Heuser JE, Schroer TA. Analysis of Dynactin Subcomplexes Reveals a Novel Actin-Related Protein Associated with the Arp1 Minifilament Pointed End. Journal of Cell Biology, 1999 147: 307–320.

25. Karki S, Tokito MK, Holzbaur ELF. A Dynactin Subunit with a Highly Conserved Cysteine-rich Motif Interacts Directly with Arp1. Journal of Biological Chemistry, 2000 275: 4834–4839.

26. Garces JA, Clark IB, Meyer DI, Vallee RB. Interaction of the p62 subunit of dynactin with Arp1 and the cortical actin cytoskeleton. Current Biology, 1999 9: 1497–1502.
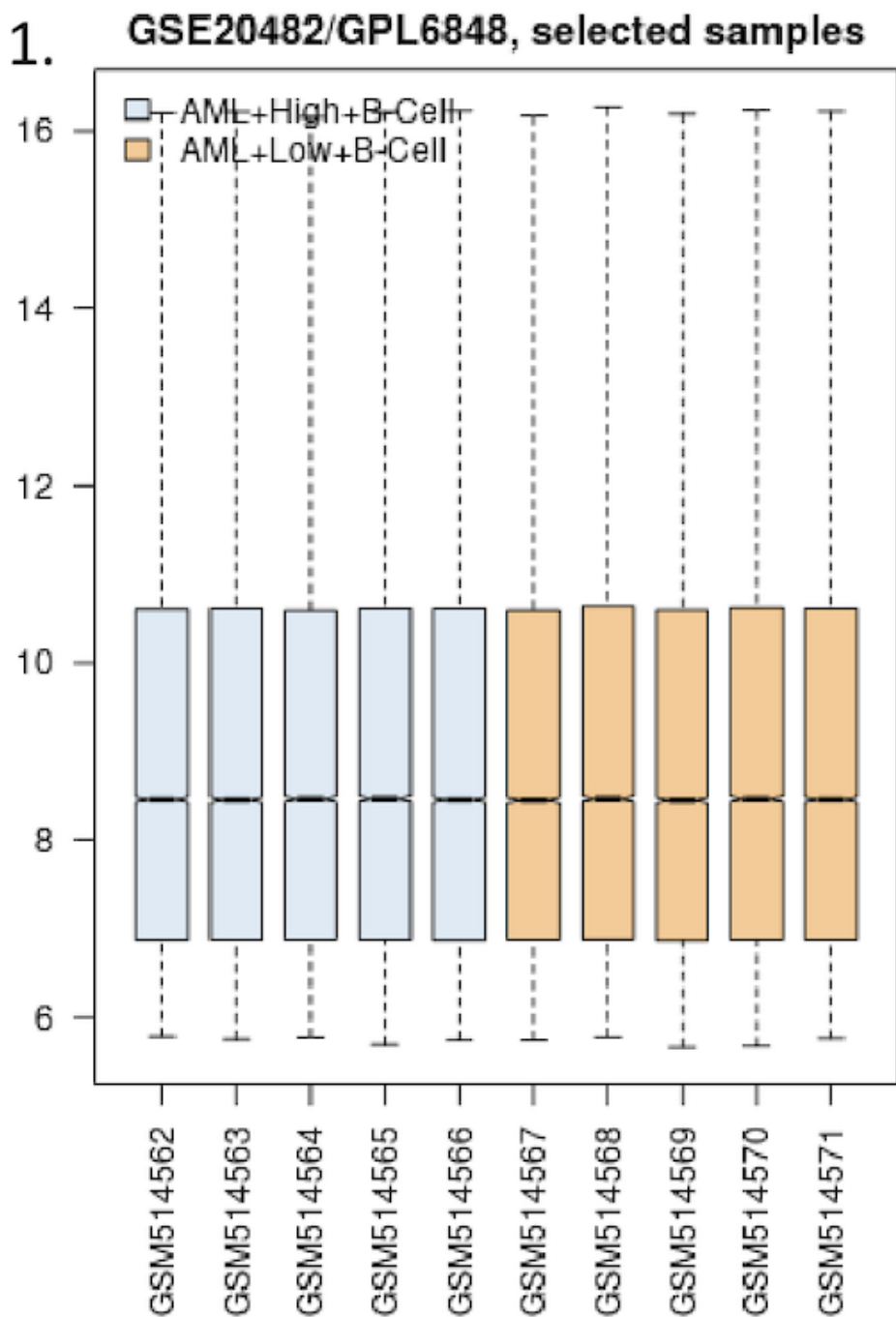
Figures

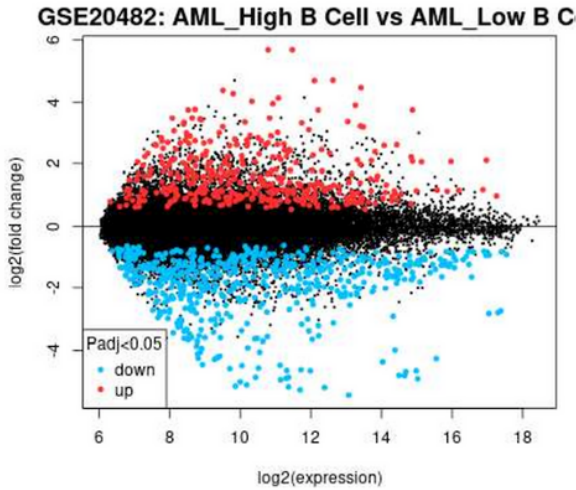**Fig. 1. A.** Mean of AML high and low B cell.
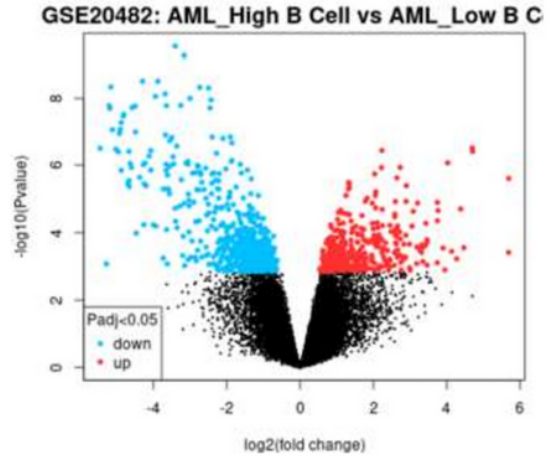
Figure 1

See image above for figure legend.
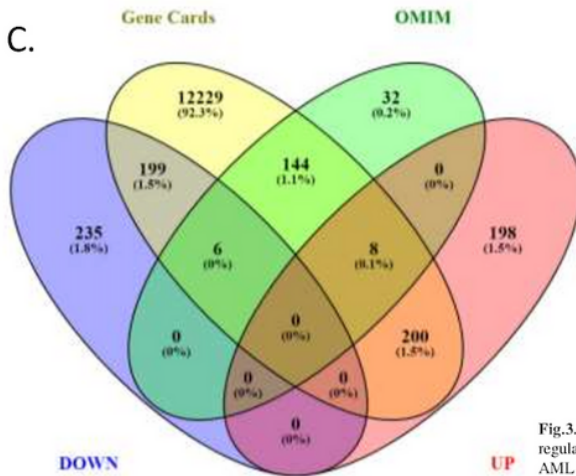
Figure 2

See image above for figure legend.

3. **A.**

B.

C.

D.

Total DEGs	846
UP DEGs	406
DOWN DEGs	440
Novel UP DEGs	198
Novel DOWN DEGs	235

Fig.3. **A.** Visualization of screened DEGs through constructed Volcano plot. **B.** MD plot reveals the up and down regulated genes (Figure 3A and 3B). **C.** Venn diagram showing the overlap between GO terms down-regulated in AML and up-regulated in the AML. **D.** 198 gene were found in novel up regulated and 235 gene were found in novel down regulated. DEGs were determined based upon average gene expression value

Figure 3

See image above for figure legend.

Figure 4

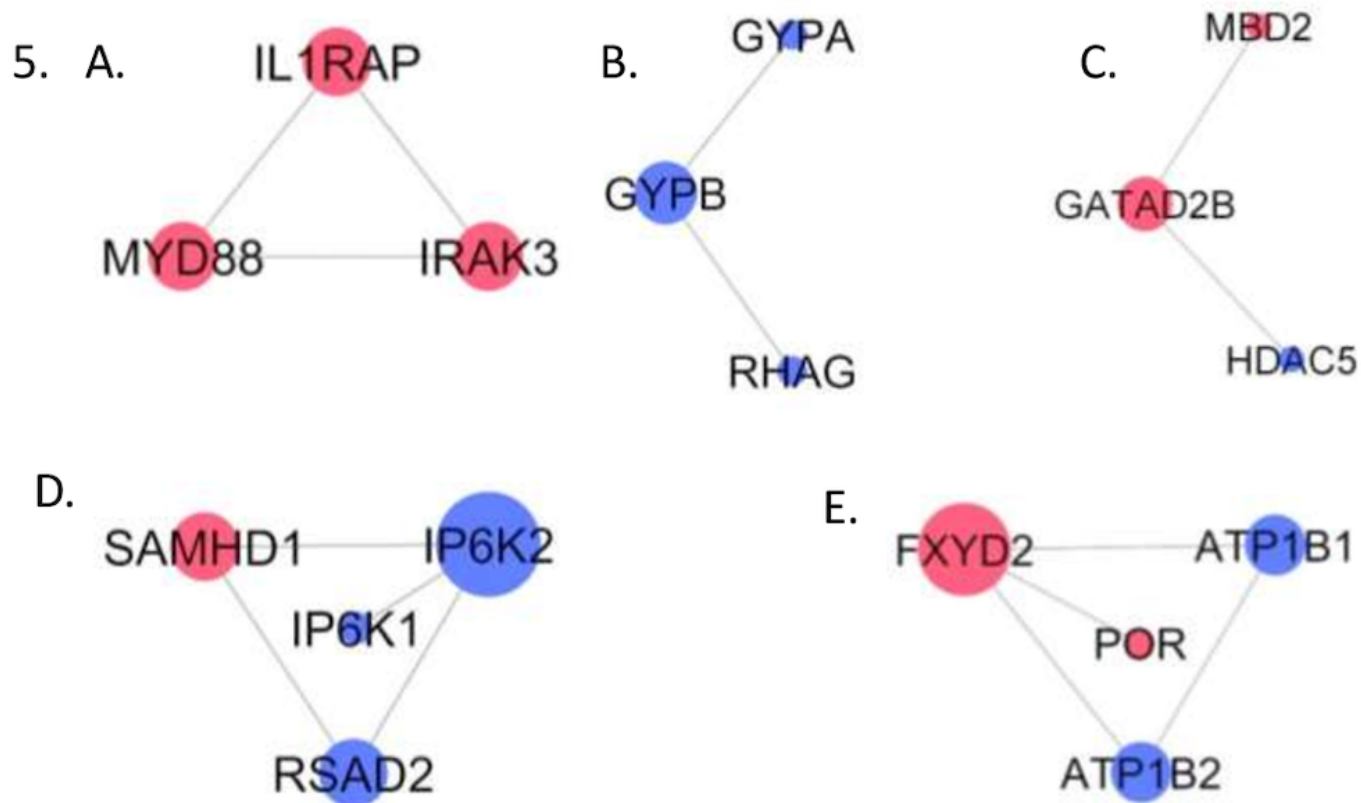See image above for figure legend.

Fig. 5. These five (A,B,C,D,E) sub-networks of differentially expressed genes (DEGs). Pink circle represents the up-regulated genes and blue circle represent the down-regulated genes and lines shows the correlation between these genes.

Figure 5

See image above for figure legend.

Figure 6

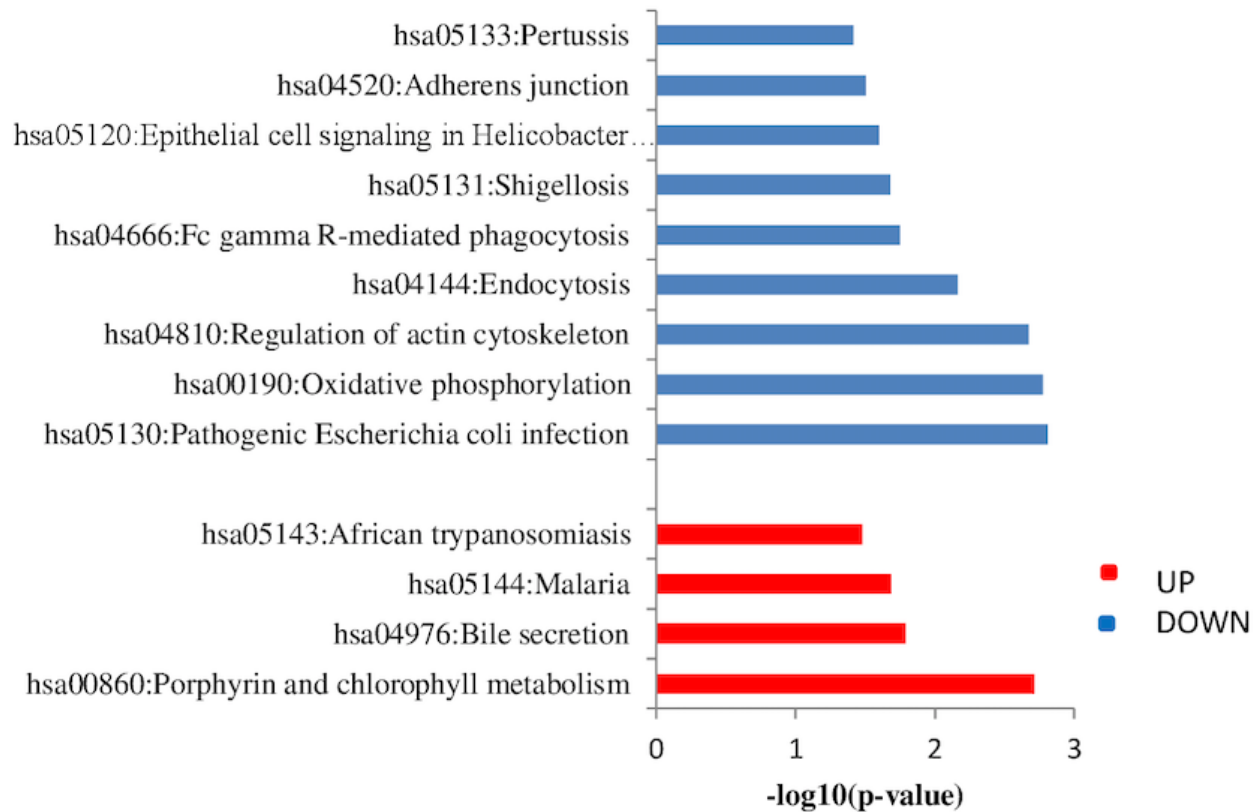See image above for figure legend.

7.

Fig. 7. KEGG pathway enrichment for DEGs

Figure 7

See image above for figure legend.

Figure 8

See image above for figure legend.

Figure 9

See image above for figure legend.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementryfile1.xlsx