Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2582-2845.8842

ISSN: 2582 – 2845

Ind. J. Pure App. Biosci. (2022) *10*(2), 29-37

Indian Journal of Pure & Applied Biosciences

Peer-Reviewed, Refereed, Open Access Journal

Research Article

A System Biology Approach to Construct a Gene Regulatory Network for C-Kit Mediated Proliferation in Hematopoietic Stem Cells

Anuradha Bhardwaj and Vikrant Nain*

Department of Biotechnology, Gautam Buddha University, Greater Noida, Uttar Pradesh, 201312, India

*Corresponding Author E-mail: vikrant.nain@gmail.com

Received: 14.12.2021 | Revised: 11.02.2022 | Accepted: 19.02.2022

ABSTRACT

Many human diseases are characterized by deviations in signaling pathway linked to cell proliferation and differentiation. The crucial interaction of the receptor tyrosine kinase, c-Kit, with its ligand steel factor regulates the homeostatic immune and hematopoietic systems, controlling their fascinating features of proliferation, differentiation, survival. The gene c-Kit has been reported to be associated with a spectrum of human diseases and most commonly observed in cancer. The use of molecular techniques like gene therapy to alter human hematopoietic stem and progenitor cells presents great possibilities for various genetic and oncologic illnesses. As a result, elucidating the molecular machinery that controls proliferation is critical. Understanding the mysterious mechanisms that underpin proliferation has long been a focus of basic and clinical research. A combination of computational biology tools and interaction discovery techniques is ideal for global molecular characterization of disease pathways. When primary events are ambiguous due to their enormous complexity, constructing and analyzing the gene regulatory network of proliferation can be the most effective technique to comprehend negative consequences. We identified a network of 356 nodes and 178 interactions as reported in the STRING database, a search engine for retrieving interacting gene/protein. The study pipeline then moved on to functional clustering of related partners utilizing molecular complex detection (MCODE). We then filtered ten hub genes from the network with strong associations and having a critical role in proliferation. Surprisingly, the associated protein we discovered through the network shared more functional similarities with known cancer-related genes. This network-based approach to our microarray data assists in the identification of novel genes/proteins and sheds light on their critical function in c-Kit-mediated hematopoietic stem cell proliferation.

Keywords: *Stem cells, Proliferation, Hematopoietic stem cells, Microarray, differentially expressed genes, Gene Regulatory Networks, Potential therapeutic targets.*

Cite this article: Bhardwaj, A., & Nain, V. (2022). A System Biology Approach to Construct a Gene Regulatory Network for C-Kit Mediated Proliferation in Hematopoietic Stem Cells, *Ind. J. Pure App. Biosci. 10*(2), 29-37. doi: http://dx.doi.org/10.18782/2582-2845.8842

INTRODUCTION

Stem cells are rarely present, accounting for about 0.01 % of bone marrow (Walasek et al., 2012). Proliferating or non-proliferating hematopoietic stem cells (HPSCs) can be found in the bone marrow. Hematopoiesis is the process by which proliferating HPSCs differentiate into blood cells (Shizuru et al., 2005). In the last few decades, knowledge, and techniques in the field of hematopoietic stem cells have expanded at an exponential rate (Dong et al., 2021; Hurwitz et al., 2020; & Liang et al., 2020). Many hematopoietic stem cell growth factors, including cytokines, have been identified. Human hematopoietic stem cells have been manipulated utilizing *in vitro* expansion and gene therapy in recent years. This study gives up new avenues for treating a variety of hereditary and oncologic illnesses. In general, "stem cells" refer to cells that are capable of long-term hematopoietic system restoration in recipient animals. Hematopoietic stem/progenitor cells (HSPCs) isolated from bone marrow in healthy people are a heterogeneous group of cells with the ability to self-renew and differentiate into all types of functional blood cells (Dahlberg et al., 2011) Despite significant *ex vivo* and *in vivo* investigations, limited success in clinical transplantation has been achieved, owing to a scarcity of HSPCs.

High-throughput array-based methods for detecting the expression of thousands of genes at once have been widely available in recent decades. Due to the development of such tools and technologies, as well as advancements in bioinformatics analysis, differentially expressed genes (DEGs) may now be easily identified and analyzed in terms of functional annotations and pathway enrichment. An improved approach of microarray technology was used in a study to unravel the unknown molecular targets driving replication of myelomonocytic U937 cells caused by SCF (Sharma et al., 2016). The genetic alterations at the DNA level have been extensively analyzed to identify differentially expressed genes (DEGs) and track their functional enrichment in relation to proliferation.

Microarray data was statistically examined in this work, and differentially expressed genes (DEGs) between wild type and mutant type were discovered. We also ran Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses on our DEGs using the DAVID bioinformatics tool. We then used the Cytohubba plugin of Cytoscape to build a protein-protein interaction (PPI) network of DEGs to find some core essential genes. Furthermore, utilizing Gene Expression Profiling Interactive Analysis, the expression of essential hub genes was compared between wild type and normal mutant type (GEPIA). CYCS, COX6B2, PRPF19, U2AF2, GSR, APOBEC3C, HNRNPA0, MAGOH, CSTF2T, and CNR1 were among the DEGs chosen for further investigation. The Kaplan Meier plotter was used to identify and examine the prognostic significance of these key genes. The Cytoscape software and its numerous plugins, including MCODE, GeneMANIA, and BiNGO, were also used to identify signaling pathways and genes linked with those pathways. Finally, this bioinformatic investigation discovered some intriguing biomarkers linked to hematopoietic stem cell proliferation.

MATERIALS AND METHODS

Microarray data

The high-throughput whole-genome microarray expression profiling used to uncover the critical targets mediating c-kit mutant D816V hematopoietic malignancy was previously reported and available in the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI) with GEO accession GSE76355 (Sharma and Gangenahalli, 2016). By comparing gene expression profiles of c-kit mutant D816V and wild-type c-kit, we hope to discover the important participants in c-kit-mediated proliferation in hematopoietic stem cells. Our study's microarray data was based on Agilent-027114 Genotypic Technology's

Custom Human Whole Genome 8x60k Microarray, which was designed by GPL19204 Platforms. There are six samples in all, three of each wild type (WT S) and three of each mutant type (Mut S).

Identification of DEGs

By comparing the two datasets (WT S and Mut S), the Differentially Expressed Genes (DEGs) across the experimental conditions were found. The log base 2 statistical parameters were used to compute the fold change value for each gene. The DEGs were then divided into two categories: up-regulated and down-regulated genes. We consider flag to be Detected in the treated sample and can be detected or compromised in the control sample, and |logFC| >= 0.6 in treated and control samples for filtering upregulation. We considered flags to be "Detected" in the Control sample and "Detected or Compromised" in the Treated sample for filtering Down Regulated, and |logFC|= - 0.6 in treated was regarded statistically significant.

Functional enrichment analysis

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics tool (http://david.ncifcrf.gov/,version 6.8) was used to perform Gene Ontology and Pathway Enrichment analysis of the up-regulated and down-regulated genes for functional characterization of DEGs (Dennis et al., 2003). DAVID is a web-based bioinformatics tool that uses clustering methods to assist researchers analyse large gene lists high in functional annotations. We classified our DEGs into four categories: biological process (BP), cellular component (CC), molecular function (MF), and KEGG (Kanehisa et al., 2009) pathway, with a P-Value of 0.05 considered statistically significant.

PPI network construction and module analysis

STRING-DB (http://string-db.org, version 10.0), an online Search Tool for the Retrieval of Interacting Genes (Szklarczyk & colleagues, 2014) By creating protein-protein interaction networks, was used to predict and understand the process of -kit mediated proliferation in hematopoietic stem cells. For subsequent analysis, the PPI network comprising differentially expressed genes and interactions with a total score > 0.4 were considered statistically significant. MCODE, Cytoscape's Molecular Complex Detection program, is a public bioinformatics software that is used to cluster the network's strongly connected sections (Bandettini et al., 2012) The PPI network was built using the Cytoscape programme (Smoot et al., 2011), with MCODE filter sets degree cutoff = 10, node score cutoff = 0.2, k-core = 2, max depth = 100 as selection criteria.

Hub genes selection and analysis

cytoHubba (Chin et al., 2014) is a new software for Cytoscape that helps find hub objects and sub-networks from complicated interactomes by rating nodes in a network. We employed maximal clique centrality (MCC), maximum neighbourhood component (MNC), edge percolated component (EPC), and the degree to select our likely 20 important genes out of 11 topological analysis methods available in Cytoscape's cytohubba plug-in. GeneMANIA was used to perform Gene Ontology and Pathway analysis on the key genes found in the selected topological analysis results. Gene Expression Omnibus, BioGRID, and functional genomics datasets from multiple organisms are available in Cytoscape's GeneMania App (Warde-Farley et al., 2010). BiNGO, or Biological Networks Gene Tool, is a Cytoscape plugin that was used to visualise the biological processes of the 20 important genes. The GEPIA (http://gepia.cancer-pku.cn/index.html) (Tang et al., 2017) (an online server of interactions based on The Cancer Genome Atlas (TGCA) and Genome-Tissue Expression (GTEx) data) was used to validate our gene expression data. Plotting the Kaplan Meier-plot (Li et al., 2018) was used to examine the survival effect of the important genes.

RESULTS

Identification of DEGs in hematopoietic stem cells

Statistical analysis was used to identify differentially expressed genes in high-throughput microarray data of mutant vs wild

type groups. The log base 2 method was used to calculate the fold change values. We consider a fold >= 0.6 in the treatment sample and a fold >= 0.6 in the control sample when filtering upregulation. We consider fold = -0.6 in the Treated sample and fold = -0.6 in the Control sample for filtering Downregulated. In the whole-genome microarray data, 104 genes were discovered to be up-regulated and 357 genes were found to be down-regulated.

Enrichment analysis for DEGs

The biological activities of the differentially expressed genes, both up-regulated and down-regulated, were identified and extracted using DAVID Bioinformatics, a web-based program. The DEGs were subjected to functional annotation and pathway enrichment analysis. Upregulated genes in the biological process of gene enrichment were remarkably associated with autophagy, innate immune response, growth and inflammatory response regulation, and positive regulation of NF-kappaB transcription factor activity (Fig. 1a), whereas downregulated DEGs were associated with regulation of ion transmembrane transport, covalent chromatin modification, and detection of chemical stimuli involved in sensory perception (Fig. 1b) (Fig. 2a). Zinc ion binding, superoxide generating NADPH oxidase activity, superoxide generating NADPH oxidase activator activity, Toll like receptor 4 binding, phosphatidylinositol3 4bisphosphate binding, and arachidonic acid binding were among the molecular functions of gene enrichment (Fig 1b). G-protein coupled receptor activity, kinase activity, and hydrolase activity acting on carbon, nitrogen (but not peptide) bonds in cyclic amidines linked genes were discovered in the analysis of down-regulated genes in molecular function enrichment (Fig 2b). Upregulated DEGs were also shown to be linked with cytosol, nucleoplasm, and dendrite (Fig. 1c), while down-regulated genes were found to relate to anchoring membrane components, plasma membrane, and kinesin complex (Fig. 1d) (Fig. 2c). Furthermore, KEGG signaling pathway analysis revealed that up-regulated DEGs were associated with neuroactive ligand-receptor interaction, olfactory transduction, and osteoclast differentiation (Fig. 1d). Down-regulated genes were associated with neuroactive ligand-receptor interaction, olfactory transduction, and olfactory transduction (Fig. 2d).

PPI network construction and significant module identification

At the protein level, the database of interactions – STRING pre-computed database indicated probable relationships and interactions among differentially expressed genes with a combined probabilistic confidence score > 0.4 as the criterion. 356 nodes and 178 interactions make up the Protein-Protein Interaction Network created with Cytoscape software (Fig. 3a). The generated PPI network module consists of 6 nodes and 15 edges showing associations after trimming the most important network using MCODE (Fig. 3b). We used the Cytohubba plugin of the Cytoscape software to further filter the essential hub genes out of our top 20 nodes in the PPI network using four topological analysis methods: MCC, MNC, Degree, and EPC (Table 1). For additional investigation, we discovered 10 essential overlapping critical hub genes: CYCS, COX6B2, PRPF19, U2AF2, GSR, APOBEC3C, HNRNPA0, MAGOH, CSTF2T, and CNR1.

Re-analysis of the ten selected genes

The GEPIA online tool was used to validate the expression of the 10 essential genes that were shortlisted. This study confirmed that the expression of the hub genes differed significantly between normal and malignant cells (Fig. 4a). We discovered that three of the 10 essential genes were associated with poor survival using the Via Kaplan Meier plotter online tool and the training set LAML (Fig. 5). The GeneMANIA program of Cytoscape was used to create a protein/gene regulatory network of the 10 selected genes APOBEC3C and their relationships (Fig. 6a) The ranks were assigned in the following manner: BCAS2, DNAJC7, RBM5, HNRNPK, DDX42, PRCC, PJA1, HNRNPH2, and OCM2 based on the score order from low to

high. Furthermore, the BiNGO plugin of Cytoscape's biological process analysis of the hub genes revealed that these essential genes are involved in mRNA processing, mRNA metabolic process, RNA splicing, and positive control of astrocyte differentiation (Fig. 6b).

DISCUSSION

Hematopoiesis is a delicate and intricate process. Because of the selectivity and low toxicity of the mutant D816V c-Kit receptor, molecular targeted therapy is becoming increasingly popular (Babaei et al., 2016). More research is needed to fully comprehend the function and mechanism of D816V c-regulator Kit's gene sets, which could reveal a method for regulating HSCs with targeted medications to achieve maximum efficacy with the least dose. As a result, this gene expression data was thoroughly analyzed utilizing extremely powerful bioinformatics and system biology tools to comprehend the regulatory mechanism of c-Kit-mediated proliferation. We identified and extensively studied the associations between the functional components of a hematopoietic cell like genes, transcription factors and proteins. We used computational tools and statistical analysis to try to identify the complex cytokine SCF signaling to develop new strategies for improved hematopoietic stem cell proliferation and to characterize key players in c-Kit-mediated proliferation in HSCs, allowing us to find new potential drug targets. To get mechanistic insights into these interactions, it is necessary to characterize the interaction interface.

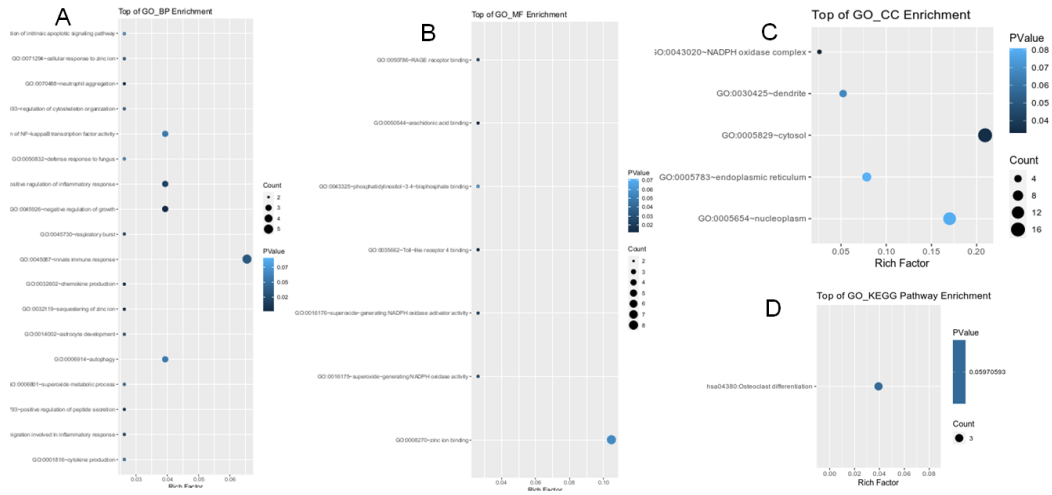

Fig. 1 GO and KEGG analysis of the up-regulated DEGs in hematopoietic stem cells. a Biological process. b Molecular function. c Cellular component d KEGG pathway. All of the enrichment pathways were generated using the ggplot2 package in R language

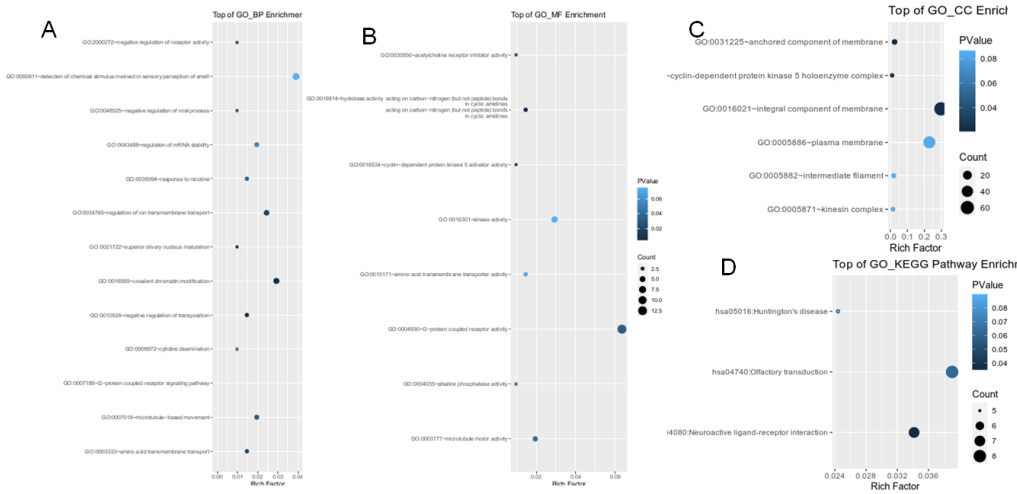

Fig. 2 GO and KEGG analysis of the Down-regulated DEGs in hematopoietic stem cells. a Biological process. b Molecular function. c Cellular component d KEGG pathway. All of the enrichment pathways were generated using the ggplot2 package in R language

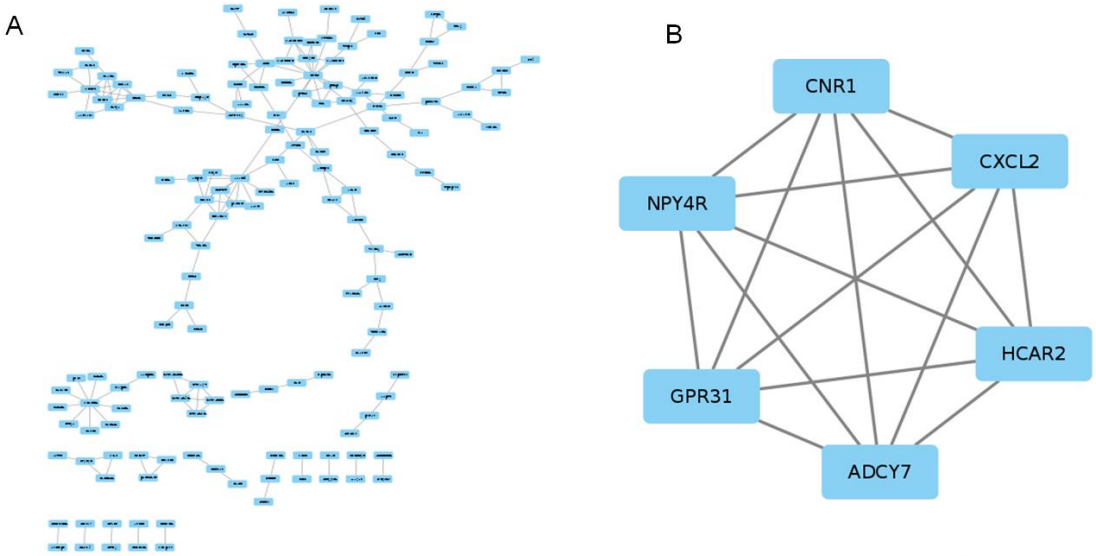

Fig. 3 Common DEGs PPI network construction and module analysis. a. A total of 104 genes were found to be up regulated and 357 genes were found to be down regulated were visualized in the DEGs PPI network complex: the nodes represent proteins, the edges represent the interaction of the proteins. b Module analysis using MCODE: degree cutoff =10, node score cutoff = 0.2, k-core = 2, max depth = 100

Rank Method in Cytohubba			
MCC	*EPC*	*MNC*	*Degree*
ADCY7	**CYCS**	**U2AF2**	**CYCS**
CNR1	**COX6B2**	**PRPF19**	**PRPF19**
NPY4R	**PRPF19**	ADCY7	ADRBK1
GPR31	**U2AF2**	NPY4R	ADCY7
CXCL2	**GSR**	**CYCS**	**U2AF2**
HCAR2	**APOBEC3C**	**CNR1**	**CNR1**
PRPF19	**HNRNPA0**	CXCL2	GPR31
U2AF2	APOBEC3B	HCAR2	NPY4R
HNRNPA0	MT1M	GPR31	**HNRNPA0**
CYCS	**MAGOH**	**COX6B2**	CXCL2
MAGOH	**CSTF2T**	**HNRNPA0**	HCAR2
CSTF2T	MT1E	**MAGOH**	**COX6B2**
COX6B2	SCD	**CSTF2T**	BRCA2
ADRBK1	NCF1	**GSR**	SCD
GSR	CYP2B6	KRTAP6-3	**MAGOH**
KRTAP6-3	NDUFC2	KRTAP19-1	**CSTF2T**
KRTAP19-1	APOBEC3F	GFRA3	**APOBEC3C**
APOBEC3C	BRCA2	PSPN	ZFYVE26
KRTAP10-1	ZFYVE26	KRTAP10-1	PDLIM7
APOBEC3B	**CNR1**	**APOBEC3C**	**GSR**

Table 1 Hub genes for highly differentiated expressed genes ranked in Cytohubba plugin of Cytoscape

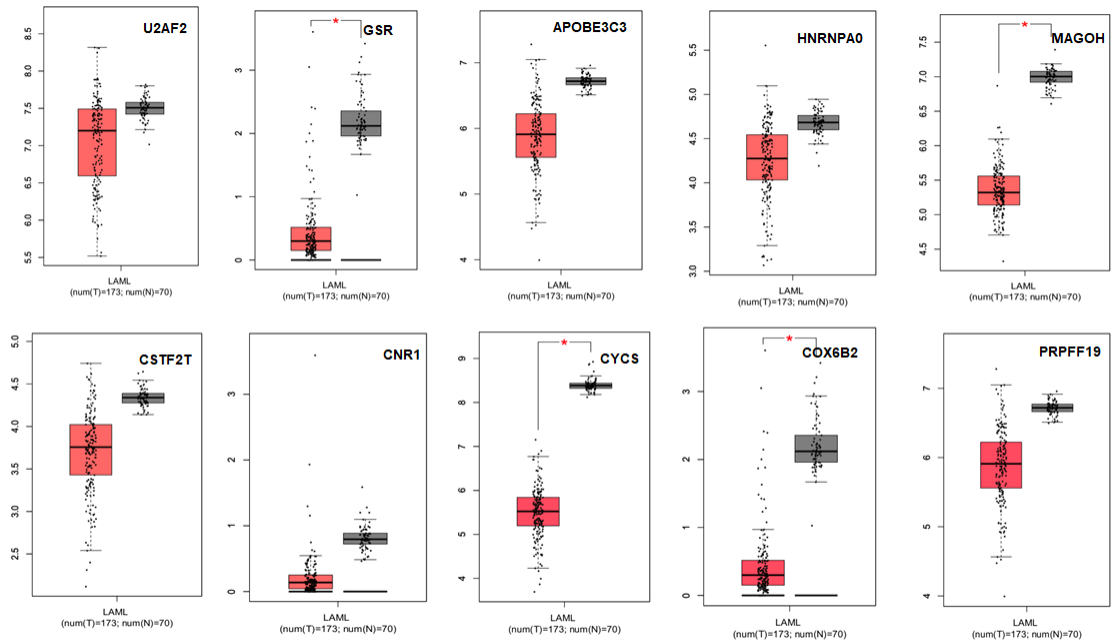

Fig. 4 The expression level of hub genes between cancer and normal tissue according to GEPIA database.

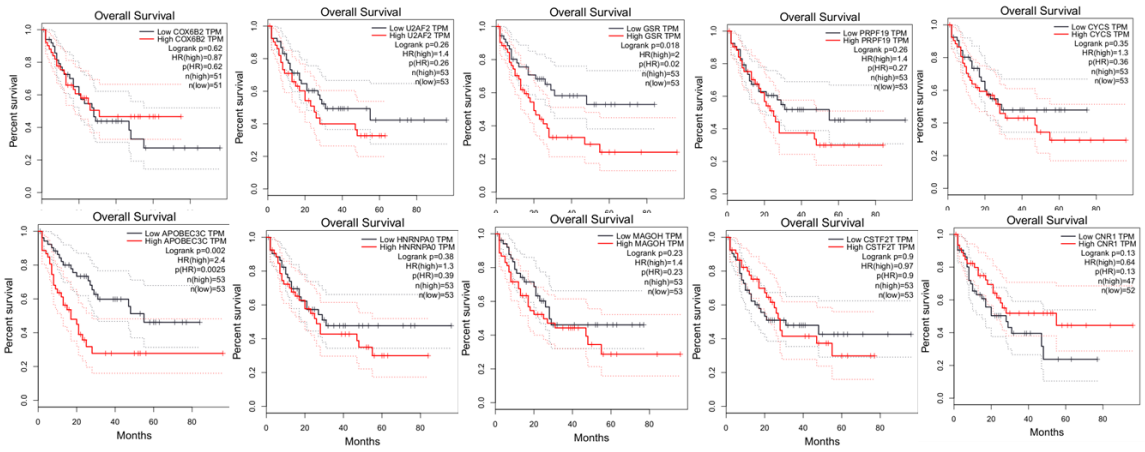

Fig. 5 The prognostic information of the 10 key genes. The online tool Kaplan Meier plotter was applied for identification of the prognostic value of key genes and 3 of 10 were correlated with worse survival. (P < 0.05)

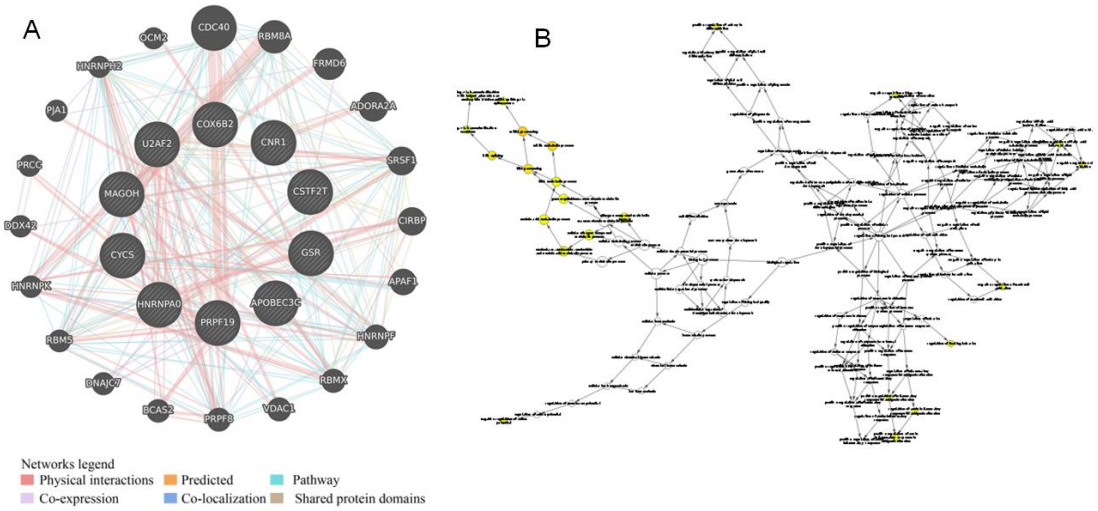

Fig. 6 a. The network of key genes and their related genes constructed by GeneMANIA . b. The biological process of key genes analyze by BiNGO (P < 0.01). The color depth of node represents the corrected P-value. The size of nodes represents the number of genes involved

CONCLUSION

These connections, on a systems level, produce a complex network that responds to both intracellular and extracellular disturbances. To untangle complicated processes and their exposure-related perturbations from gene expression data, we built proliferation active sub-networks models based on activated c-Kit in HSCs. These networks have been visualized as graphs, with each node representing a gene, protein, or transcription factor, and edges depicting interactions between them. Furthermore, the proteins/genes in these networks are not distributed randomly; rather, a protein linked with a specific function tends to cluster. In conclusion, we filtered out ten critical key players associated with HSC proliferation using extensive bioinformatics analysis of high-throughput gene expression profiles in wild type and mutant type hematopoietic stem cells, namely CYCS, COX6B2, PRPF19, U2AF2, GSR, APOBEC3C, HNRNPA0, MAGOH, CSTF2T, and CNR1. These important hub genes can also be thought of as new indicators for hematopoietic stem cell growth. However, more research is essential to understand the true biological mechanism of our shortlisted critical genes for c-kit mediated proliferation in hematopoietic stem cells, as well as to validate the results (*in vivo* and *in vitro* experiments).

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' contributions

AB conceived the idea, performed data collection, analysis, presentation and manuscript writing.

VN involved in overall supervision of research work and manuscript writing.

All authors read and approved the final manuscript

REFERENCES

Babaei, M. A., Kamalidehghan, B., Saleem, M., Huri, H. Z., & Ahmadipour, F. (2016). Receptor tyrosine kinase (c-Kit) inhibitors: a potential therapeutic target in cancer cells. Drug Design, *Development and Therapy 10*, 2443. https://doi.org/10.2147/DDDT.S89114.

Bandettini, W. P., Kellman, P., Mancini, C., Booker, O. J., Vasu, S., Leung, S. W., Wilson, J. R., Shanbhag, S. M., Chen, M. Y., & Arai, A. E. (2012). MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: A clinical validation study. *J. Cardiovasc. Magn. Reson. 14*. https://doi.org/10.1186/1532-429X-14-83.

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., & Lin, C. Y. (2014). cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol. 8*, S11. https://doi.org/10.1186/1752-0509-8-S4-S11.

Dahlberg, A., Delaney, C., & Bernstein, I. D. (2011). Ex vivo expansion of human hematopoietic stem and progenitor cells. Blood. https://doi.org/10.1182/blood-2011-01-283606.

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol. 4*, R60. https://doi.org/10.1186/gb-2003-4-9-r60.

Dong, S., Wang, Q., Kao, Y. R., Diaz, A., Tasset, I., Kaushik, S., Thiruthuvanathan, V., Zintiridou, A.,

Nieves, E., Dzieciatkowska, M., Reisz, J. A., Gavathiotis, E., D'Alessandro, A., Will, B., & Cuervo, A. M. (2021). Chaperone-mediated autophagy sustains haematopoietic stem-cell function. *Nature 591*, 117–123. https://doi.org/10.1038/S41586-020-03129-Z.

Hurwitz, S. N., Jung, S. K., & Kurre, P. (2020). Hematopoietic stem and progenitor cell signaling in the niche. *Leukemia 34*, 3136–3148. https://doi.org/10.1038/S41375-020-01062-8.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., & Hirakawa, M. (2009). KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res. 38*. https://doi.org/10.1093/nar/gkp896.

Liang, R., Arif, T., Kalmykova, S., Kasianov, A., Lin, M., Menon, V., Qiu, J., Bernitz, J. M., Moore, K., Lin, F., Benson, D. L., Tzavaras, N., Mahajan, M., Papatsenko, D., & Ghaffari, S. (2020). Restraining Lysosomal Activity Preserves Hematopoietic Stem Cell Quiescence and Potency. *Cell stem cell 26*, 359-376. e7. https://doi.org/10.1016/J.STEM.2020.01.013.

Li, T., Gao, X., Han, L., Yu, J., & Li, H. (2018). Identification of hub genes with prognostic values in gastric cancer by bioinformatics analysis. *World J. Surg. Oncol. 16*. https://doi.org/10.1186/s12957-018-1409-3.

Sharma, S., & Gangenahalli, G. (2016). Gene Expression Profiling of Human c-Kit Mutant D816V. *J. Cancer Ther. 07*, 439–454. https://doi.org/10.4236/jct.2016.76046.

Shizuru, J. A., Negrin, R. S., & Weissman, I. L. (2005). Hematopoietic Stem and Progenitor Cells: Clinical and Preclinical Regeneration of the Hematolymphoid System. *Annu. Rev. Med. 56*, 509–538. https://doi.org/10.1146/annurev.med.54.101601.152334.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics 27*, 431–432. https://doi.org/10.1093/bioinformatics/btq675.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., & Jensen, L. J., & Von Mering, C. (2014). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res. 43*, 447–452. https://doi.org/10.1093/nar/gku1003.

Tang, Z., Li, Chenwei, Kang, B., Gao, G., Li, Cheng, & Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res. 45*, W98–W102. https://doi.org/10.1093/nar/gkx247.

Walasek, M. A., van Os, R., & de Haan, G. (2012). Hematopoietic stem cell expansion: Challenges and opportunities. *Ann. N. Y. Acad. Sci. 1266*, 138–150. https://doi.org/10.1111/j.1749-6632.2012.06549.x.

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D., & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res. 38*, W214-20. https://doi.org/10.1093/nar/gkq537.