# Identification of Potent Novel Biomarkers of Uterine Leiomyoma Through DEGs Screening and Networking.

**Sonal UPADHYAY**
Banaras Hindu University
**Anjali RANI**
Banaras Hindu University
**Pawan K. DUBEY**
Banaras Hindu University
**Ravi BHUSHAN** ( ravi.mhg.bhu13@gmail.com )
Banaras Hindu University

**Research Article**

**Keywords:** MED12, Uterine Leiomyoma, Differential gene expressions, Novel biomarkers.

**Posted Date:** May 12th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-484042/v1

# Abstract

**Background:** Uterine leiomyomas is a benign lesion arising in myometrium of the uterus. Various risk factors like stress, obesity, hormonal imbalance are involved in the progression of the uterine leiomyomas. Despite the significant research, the potential biomarkers related to uterine leiomyomas are yet to be discovered.

**Methods:** The present study deals with searching the common potent markers of Uterine Leiomyoma (ULM) that was responsible for their pathogenicity. The microarray dataset (GSEID: GSE30673) was fetched through Gene Expression Omnibus database. Comparing with normal myometrium samples, Principal Component Analysis (PCA) and heat map were constructed to obtain differential expressed genes (DEGs) for ULM. Common DEGs were obtained through Venny software v 2.1.0. Significant enriched pathways and ontological study for DEGs were also performed through online Database for Annotation, Visualization and Integrated Discovery tool. Based upon STRING v 10.5, protein- protein interaction network was constructed in order to predict functional interactions among proteins.

**Results:** 176 of total DEGs with 101 overexpressed and 75 under expressed genes were screened out with their official gene symbol. 2 DEGs were found as common genes in OMIM and Gene Cards.Only 9 DEGs were found to have combined score > 0.4 and hence included in interaction network. The present study revealed EGF, FYN, VCAN, TRIP13, FBXW7 as up-regulated genes and GATA2, JAG1, TLR3, APOL1 as down-regulated genes found to be expressed in samples with ULM disease. KEGG pathway enrichment analysis for DEGs revealed Focal adhesion, ECM-receptor interaction, long-term depression and Retinol metabolism are major pathways which have been enriched for these DEGs.

**Conclusion:** We conclude that TRIP13 and TLR3 might be the novel biomarkers related to ULM disease which were revealed through our findings. The present study provides us a new perspective to detect the potent biomarkers responsible for ULM and further in vitro and in vivo experiments needed to be performed to verify the results.


# Background

A benign lesion arising in outer smooth muscle layers (known as myometrium) of the uterus is known as Uterine leiomyomas (Uterine Fibroid). The symptoms of Uterine leiomyoma includes uneasiness in abdomen, frequent urination, pain and abnormal bleeding from uterus which were found to occur in child bearing age of women [1]. Various factors involve obesity, smoking, stress, race, age, hormonal (like estrogen) imbalance and even some genetic factors which were found to be more prone of leiomyoma occurrence. It was found that one of the main causes of 86% of uterine leiomyoma cases occurring is the mutation at specific sites of *mediator complex subunit 12* (*MED12*) [2]. Even overexpression of WNT4 with MED12 mutation was responsible for leiomyoma [3].

Moreover, to chromosomal aberration after exome sequencing of uterine fibroid tissues it was found somatic alteration in mediator complex subunit 12 (*MED12)*[4]. Makinen et al. investigated out of 225

uterine fibroid (71%) cases 159 were carrying mutation in exon 2 of MED12 [5].In the mediator of RNA polymerase II transcription subunit 12 (MED12) protein specific missense and inframe mutation was found that leads to 70% of the occurrence of this disease [6].On the basis of ethnicity of patients it was reported that frequency of occurrence of MED12 mutation differs from 50–80%[7].

Since, adjuvant therapy does not show favourable options in the treatment of uterine fibroid. Even these benign lesions were resistant to numerous chemotherapeutics[8]. So, in order to search more candidate genes and find their potent mechanism at molecular level, comprehensive in silico analysis was performed to identify the upregulated and downregulated genes through differential gene expression of MED12 mutation in comparison with MED12 wild type fibroid samples. Pathway enrichment and functional enrichment analysis, pathway related network and sub-network analysis was also performed. On the basis of combined score the construction of PPI (protein-protein interaction) networks were performed. The aim of present study is to screen key candidate genes of the dataset (ID: GSE30673) of uterine fibroid disease, mainly candidate genes associated with MED12 mutation fibroid samples and also mechanism behind this disease.

## Methods

# Microarray gene expression dataset retrieval.

One raw microarray gene expression dataset of GSE30673 was retrieved using NCBI Gene expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) [9] which includes the samples related to MED12 mutation and MED12 wild type leiomyoma samples. This dataset deposition was made on the GPL13916 platform of [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array [CDF: Brainarray ENSG Version 14.1]. Various bioinformatics tools were applied for the study of Differential gene expressions in wild type and mutational MED12 Leiomyoma samples.

# Dataset pre-processing and identification of DEGs.

After retrieval of raw gene expression profile preprocessing using Benjomini and Hochberg algorithm was performed. Moreover, log transformation by avoiding Limma precision weights and force normalization was implemented to the dataset. It involved the values of the dataset of probes associated with specific genes which were averaged and then selection of up-regulated and down regulated genes was performed using BiGGEsTs software. With the help of GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) tool the conversion of probe level symbols to gene level symbols was made. For the selection of DEGs the parameters include <0.05 adjusted p values and threshold logFC values >0.1 for up regulated and <-0.1 for down regulated genes. Screened DEGs were visualized through volcano plot which emphasizes the upregulated and downregulated genes. Even a mean difference (MD) graph with log2 fold variation versus average log2 expressed gene values were constructed through LIMMA package of GEO2R.[10].

# Heatmap plots construction and Principal component analysis

With the help of ClustVis an online tool[11], heatmap plot and Principal component analysis (PCA) plot was constructed for DEGs. Only upto maximum of 2MB of file size can be supported by this tool thus it was impossible to construct a PCA plot for total dataset profile.

# PPI network and co network construction

For finding among proteins their functional interactions, STRING v 10.5 [12] (https://www.string-db.org/) an online tool is used. This software yields combined scores between gene pairs for protein-protein interactions. For present work, the selected DEGs were submitted to this online database and included the parameter of combined score >0.4 for analysis. Then a bioinformatics software Cytoscape v 3.2.1(http://www.cytoscape.org) [13] , was used for various network and co networking creation. Degree and edge betweenness criteria were considered for constructing networks.

# Functional and pathway enrichment analysis of DEGs.

Integration of large set of functionally annotated genes record , DAVID (Database for Annotation, Visualisation And Integrated Discovery) software [14](https://david.abcc.ncifcrf.gov) an in silico tool was used. For performing Gene Ontology (GO) enrichment analysis, DAVID v 6.8 and STRING v 10.5 tools helped in analysing molecular function (MF), cellular component (CC) and biological process (BP) of genes. DAVID includes hypergeometric distribution, by using a whole set of genes with similar or closely associated functions.

## Results

# Determination of DEGs for ULM.

Microarray datasets of ULM with MED12 wildtype and mutant were normalized (Fig. 1) through GEO2R. 176 of total DEGs with 101 overexpressed and 75 under expressed genes were screened out (Supplementary file 1). The p-value < 0.05 and | $\log_2$ FC > 0.1 as parameters for DEGs determination were used. Visualization of screened DEGs through constructed Volcano plot and MD plot reveals the up and down regulated genes (Fig. 2A and 2B)

# Principal component and Heat Map plot of DEGs

Principal Component plot for ULM disclose a scatter plot indicating total variance of 60.7% identical to the principal component 1 on x-axis while 6.5% identical to principal component 2 on y-axis respectively

(Fig. 3A). Heat-map plot was constructed for DEGs disclose a data matrix with coloring patterns provides an overview of numerical variation. (Fig. 3B).

# Visualization of known DEGs for ULM

Constructed Venn diagram revealed 2 DEGs as common genes found in OMIM and Gene Cards (Fig. 4). Moreover, 6 known disease gene were also found. DEGs were determined based upon average gene expression value.

# Protein-Protein Interaction network

Based on combined score generated by STRING, protein-protein interaction network was formed for all the DEGs with combined score > 0.9. This lead to construction of PPI network (Fig. 5A) consisting of 85 nodes and 157 edges. The red and blue color represents the over and under expression of DEGs. Degree and edge betweeness were employed to create hub nodes. Genes forming hub nodes were EGF (Epidermal Growth Factor), FYN (FYN Proto-Oncogene), VCAN (VERSICAN), TRIP13 (Thyroid Hormone Receptor Interactor 13), FBXW7 (F Box and WD Repeat Domain Containing 7) (up regulated genes) and GATA2 (GATA Binding Protein 2),JAG1 (Jagged Canonical Notch Ligand 1),TLR3 (Toll Like Receptor 3),APOL1 (Apolipoprotein L1) (under regulated genes). Since these hub nodes were directly associated with known disease gene and are novel hence can be considered potential target genes for ULM.

Few DEGs formed two separate network from main network and were treated as sub-network. Sub-network first contains six nodes and five edges while second network contain three nodes and three edges (Fig. 5B and C). Based on the combined score and degree, a total of ten DEGs were selected out as novel genes.

# GO Enrichment and KEGG pathway analysis

All DEGs in the network were enriched for their function, processes and component (Tables 1 and 2). Most significant ($p < 0.05$) processes being regulated by up-DEGs were cell adhesion, extracellular structure organisation, skeletal system morphogenesis while Notch signalling, JNK cascade, leukemia inhibitory factor signalling and NF-KB signalling are major significant ($p < 0.05$) processes which are down-regulated (Fig. 6A and B). Most of the DEGs are localised in either extracellular matrix or basement membrane.

Focal adhesion, ECM-receptor interaction, long-term depression and Retinol metabolism are major pathways which have been enriched for these DEGs as revealed through KEGG enrichment analysis (Fig. 6C).

Table 1
GO enrichment for up-regulated DEGs

| Category | Term | PValue |
|---|---|---|
| GOTERM_BP_FAT | GO:0030246 ~ carbohydrate binding | 0.023533 |
| GOTERM_BP_FAT | GO:0005539 ~ glycosaminoglycan binding | 0.026869 |
| GOTERM_BP_FAT | GO:0030145 ~ manganese ion binding | 0.034256 |
| GOTERM_BP_FAT | GO:0001871 ~ pattern binding | 0.034256 |
| GOTERM_BP_FAT | GO:0030247 ~ polysaccharide binding | 0.034256 |
| GOTERM_CC_FAT | GO:0031012 ~ extracellular matrix | 3.22E-07 |
| GOTERM_CC_FAT | GO:0005578 ~ proteinaceous extracellular matrix | 1.18E-06 |
| GOTERM_CC_FAT | GO:0005604 ~ basement membrane | 0.008926 |
| GOTERM_MF_FAT | GO:0007155 ~ cell adhesion | 7.71E-04 |
| GOTERM_MF_FAT | GO:0060021 ~ palate development | 0.011522 |
| GOTERM_MF_FAT | GO:0048705 ~ skeletal system morphogenesis | 0.019493 |
| GOTERM_MF_FAT | GO:0048562 ~ embryonic organ morphogenesis | 0.030409 |
| GOTERM_MF_FAT | GO:0048738 ~ cardiac muscle tissue development | 0.034176 |
| GOTERM_MF_FAT | GO:0001764 ~ neuron migration | 0.042203 |
| GOTERM_MF_FAT | GO:0009582 ~ detection of abiotic stimulus | 0.043402 |
| GOTERM_MF_FAT | GO:0043062 ~ extracellular structure organization | 0.050522 |
| GOTERM_MF_FAT | GO:0009581 ~ detection of external stimulus | 0.054748 |

Table 2
GO enrichment for down-regulated DEGs

| Category | Term | PValue |
|---|---|---|
| GOTERM_BP_FAT | GO:0002675 ~ positive regulation of acute inflammatory response | 9.69E-04 |
| GOTERM_BP_FAT | GO:0009967 ~ positive regulation of signal transduction | 9.92E-04 |
| GOTERM_BP_FAT | GO:0010647 ~ positive regulation of cell communication | 0.001734 |
| GOTERM_BP_FAT | GO:0010627 ~ regulation of protein kinase cascade | 0.002835 |
| GOTERM_BP_FAT | GO:0031349 ~ positive regulation of defense response | 0.002911 |
| GOTERM_BP_FAT | GO:0042157 ~ lipoprotein metabolic process | 0.003772 |
| GOTERM_BP_FAT | GO:0001709 ~ cell fate determination | 0.007357 |
| GOTERM_BP_FAT | GO:0048584 ~ positive regulation of response to stimulus | 0.01356 |
| GOTERM_BP_FAT | GO:0048861 ~ leukemia inhibitory factor signaling pathway | 0.015581 |
| GOTERM_BP_FAT | GO:0009611 ~ response to wounding | 0.016965 |
| GOTERM_BP_FAT | GO:0050778 ~ positive regulation of immune response | 0.019123 |
| GOTERM_BP_FAT | GO:0006869 ~ lipid transport | 0.019123 |
| GOTERM_BP_FAT | GO:0070120 ~ ciliary neurotrophic factor-mediated signaling pathway | 0.019439 |
| GOTERM_BP_FAT | GO:0010565 ~ regulation of cellular ketone metabolic process | 0.02095 |
| GOTERM_BP_FAT | GO:0046328 ~ regulation of JNK cascade | 0.026768 |
| GOTERM_BP_FAT | GO:0070302 ~ regulation of stress-activated protein kinase signaling pathway | 0.029894 |
| GOTERM_BP_FAT | GO:0042127 ~ regulation of cell proliferation | 0.032891 |
| GOTERM_BP_FAT | GO:0008593 ~ regulation of Notch signaling pathway | 0.042277 |
| GOTERM_BP_FAT | GO:0007242 ~ intracellular signaling cascade | 0.053462 |
| GOTERM_BP_FAT | GO:0060627 ~ regulation of vesicle-mediated transport | 0.054367 |
| GOTERM_BP_FAT | GO:0043123 ~ positive regulation of I-kappaB kinase/NF-kappaB cascade | 0.055376 |
| GOTERM_MF_FAT | GO:0004923 ~ leukemia inhibitory factor receptor activity | 0.007688 |
| GOTERM_MF_FAT | GO:0005127 ~ ciliary neurotrophic factor receptor binding | 0.01151 |
| GOTERM_MF_FAT | GO:0004924 ~ oncostatin-M receptor activity | 0.01151 |
| GOTERM_MF_FAT | GO:0008134 ~ transcription factor binding | 0.013453 |

| Category | Term | PValue |
|---|---|---|
| GOTERM_MF_FAT | GO:0019863 ~ IgE binding | 0.019111 |
| GOTERM_MF_FAT | GO:0005496 ~ steroid binding | 0.025177 |
| GOTERM_MF_FAT | GO:0008289 ~ lipid binding | 0.02885 |
| GOTERM_MF_FAT | GO:0004029 ~ aldehyde dehydrogenase (NAD) activity | 0.030406 |
| GOTERM_MF_FAT | GO:0004620 ~ phospholipase activity | 0.040586 |
| GOTERM_MF_FAT | GO:0003712 ~ transcription cofactor activity | 0.050591 |
| GOTERM_MF_FAT | GO:0019865 ~ immunoglobulin binding | 0.056266 |
| GOTERM_MF_FAT | GO:0016298 ~ lipase activity | 0.056666 |
| GOTERM_MF_FAT | GO:0030528 ~ transcription regulator activity | 0.059696 |

# Discussion

ULM, a benign lesion present frequently in the muscular part of the uterine wall. The prevalence of their existence is increasing and still no proper therapeutics was found [15]. Recently, various bioinformatics studies were applied to search the molecular mechanism of ULM disease. In the current in silico analysis, EGF, FYN, VCAN, TRIP13, FBXW7, GATA2, JAG1, TLR3, APOL1 were detected as DEGs expressed in samples with ULM disease.

EGF gene is found to be associated with the growth, development and differentiation of various types of cells. Takeshi Maruo et al. studies suggested EGF important role in regulating fibroid development[16]. According to Jiayin Wang et al. observed the overexpression of EGF in cultured fibroid cells [17]. Likewise, Johnston et al. revealed the dysfunctional role of EGF related signaling pathways in cancer[18]. John Mendelsohn et al. also found EGF as target for cancer therapeutics [19]. Thus, we infer that EGF gene may be a potent candidate of ULM by participating in various growth and development related signaling.

FYN gene belongs to tyrosine kinase family encodes protein that participates in cell growth, adhesion and immune response. Additionally, overexpression of this gene may pose negative impact in ULM cases [20].Khush R Mittal et al. also found that FYN gene upregulation leads to uterine fibroid [21]. Moreover, Giulio Poli et al. showed FYN as target for many cancers [22].Ye-Gong Xie et al.showed that FYN gene upregulation promotes invading of cells and cell migration in breast carcinoma [23]. Therefore, we speculate that FYN gene may be related with cell growth and adhesion and would prove to be one of the potential key of uterine fibroid for treatment.

VCAN being an important component of extracellular matrix (ECM) and involves in cell growth and motility. Reported work of Md Soriful Islam et al. revealed VCAN involvement in fibroid pathogenicity [24]. John M. Norian MD et al. studies indicated VCAN variants in ECM rich uterine fibroid [25]. Moreover,

through our findings VCAN as involved in ECM interaction pathway may prove to be promising target of ULM.

TRIP13 gene encodes protein which shows thyroxine receptor interaction. Anna P. Ponnampalam et al. suggested the transcription factor TRIP13 dysfunction lead to endometriosis [26]. Kenji Kurita et al. studies indicate TRIP13 important role in promoting colorectal carcinoma[27]. According to S.Lu et al. overexpression of TRIP13 leads to human carcinoma [28]. Though till now no studies could find overexpressed TRIP13 role in uterine leiomyoma hence, may prove to be novel target for treatment.

FBXW7 gene, a F box protein plays an important role in ubiquitin ligase mechanism.Cuevas et al.found FBXW7 gene involvement with uterine sarcomas at molecular level [29]. Roland P. Kuiper et al. investigated that FBXW7 disruption lead to renal cell carcinoma [30]. Since no studies till now indicate the role of FBXW7 in uterine leiomyoma. However, FBXW7 might be an attractive key for ULM by participating in ubiquitin mechanism.

GATA2 encodes protein involved in differentiation and proliferation in cells. Alan A. Arslan et al. indicates GATA2 underexpression in uterine fibroid [31]. Cory A. Rubel et al. studies provided GATA2 transcript dysfunctional role in endometriosis [32]. Feng-Ming Tien et al. suggested that GATA2 mutation may lead to acute myeloid lymphoma [33]. Since, GATA2 was found through studies that its mutational effect leads to uterine leiomyoma so, may prove to be key target for fibrosis treatment.

JAG1 gene found to be as calcium binder and as growth factor. JAG1loss of function was studied in leiomyoma by Xiaoping Luo et al [34]. According to Sandro Santagata et al. JAG1 expression was found to be related to prostrate carcinoma [35]. Also, Jungwhoi Lee et al. through their work found JAG1 gene association with pancreatic carcinoma [36]. Although, some studies revealed the association of JAG1 with uterine fibroid thus it may prove to be one of the candidate gene of uterine leiomyoma.

TLR3 gene activate innate immune response by producing cytokines. Svenja Allhorn et al. revealed through studies that TLR3 decreased expression in endometriosis [37]. Bruno Salaun et al. investigated TLR3 activate apoptosis in human carcinoma [38]. Francesca Bianchi et al. detected TLR3 as biomarker of lung carcinoma [39]. However, till date no studies revealed TLR3 gene association with uterine fibroid thus would help in finding a novel candidate for Leiomyoma cases.

APOL1 plays role in lipid transport. Chien-An A. Hu et al. found that APOL1 overexpression leads to cell death in various carcinomas and kidney problems [40]. Though no findings till date could associate APOL1 gene with Uterine Leiomyoma. Since, the present work suggests the APOL1 downregulation is closely related to Uterine Leiomyoma thus, may prove to be key target for Uterine Fibroid treatment.

# Conclusion

We conclude that TRIP13 and TLR3 might be the novel biomarkers related to ULM disease which were revealed through our findings. The present study provides us a new perspective to detect the potent

biomarkers responsible for ULM and further in vitro and in vivo experiments needed to be performed to verify the results.

# Abbreviations

ULM: Uterine Leiomyoma; PCA: Principal Component Analysis; DEGs: Differentially Expressed Genes; GO: Gene Ontology; GEO: Gene expression Omnibus; PPI: Protein-Protein Interaction; NM: Normal Myometrium; ECM :Extra Cellular Matrix; DAVID: Database for Annotation, Visualization And Integrated Discovery; MF: molecular function; CC: Cellular Component; BP: Biological Process; EGF :Epidermal Growth Factor;FYN: FYN Proto-Oncogene; +VCAN:VERSICAN; TRIP13:Thyroid Hormone Receptor Interactor 13;FBXW7 :F Box and WD Repeat Domain Containing 7; GATA2:GATA Binding Protein 2;JAG1:Jagged Canonical Notch Ligand 1;TLR3:Toll Like Receptor 3;APOL1 :Apolipoprotein L1

# Declarations

### Ethics approval and consent to participate

The present work is totally computational and hence does not require any ethical approval and consent of participation.

### Consent for publication

Not applicable

### Availability of data and material

The datasets generated and/or analysed during the current study are available in the [GEO Datasets, NCBI] repository, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30673]". The dataset analysed have the accession number GSE30673.

### Competing interests

The authors declare that they have no any conflict of interest.


### Funding

Not applicable


### Authors' contributions

SU: Methodology, Investigation, Writing - review & editing RB: Conceptualization, Data curation, Writing-review & editing. AR: Visualization, Investigation, Validation. PKD: Review & editing. All Authors have read and approved the manuscript.

### Acknowledgements

The authors are highly thankful to all study participants and acknowledge their valuable help for this study.

### Author Details

[1]Centre for Genetic Disorders, Institute of Science, Banaras Hindu University Varanasi 221005, Uttar Pradesh, India.

# References

1. Spencer TE, Hayashi K, Hu J, Carpenter KD. Comparative Developmental Biology of the Mammalian Uterus.:38.
2. Commandeur AE, Styer AK, Teixeira JM. Epidemiological and genetic clues for molecular mechanisms involved in uterine leiomyoma development and growth. Hum Reprod Update. 2015;21:593–615.
3. Heinonen H-R, Pasanen A, Heikinheimo O, Tanskanen T, Palin K, Tolvanen J, et al. Multiple clinical characteristics separate MED12-mutation-positive and -negative uterine leiomyomas. Sci Rep. 2017;7:1015.
4. Lee M, Cheon K, Chae B, Hwang H, Kim H-K, Chung Y-J, et al. Analysis of *MED12* Mutation in Multiple Uterine Leiomyomas in South Korean patients. Int J Med Sci. 2018;15:124–8.
5. Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, et al. MED12, the Mediator Complex Subunit 12 Gene, Is Mutated at High Frequency in Uterine Leiomyomas. Science. 2011;334:252–5.
6. Ciavattini A, Di Giuseppe J, Stortoni P, Montik N, Giannubilo SR, Litta P, et al. Uterine Fibroids: Pathogenesis and Interactions with Endometrium and Endomyometrial Junction. Obstet Gynecol Int. 2013;2013:1–11.
7. Mäkinen N, Vahteristo P, Kämpjärvi K, Arola J, Bützow R, Aaltonen LA. MED12 exon 2 mutations in histopathological uterine leiomyoma variants. Eur J Hum Genet. 2013;21:1300–3.
8. Hensley ML, Maki R, Venkatraman E, Geller G, Lovegren M, Aghajanian C, et al. Gemcitabine and Docetaxel in Patients With Unresectable Leiomyosarcoma: Results of a Phase II Trial. J Clin Oncol. 2002;20:2824–31.
9. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles–database and tools update. Nucleic Acids Res. 2007;35 Database:D760–5.
10. Bhushan R, Rani A, Ali A, Singh VK, Dubey PK. Bioinformatics enrichment analysis of genes and pathways related to maternal type 1 diabetes associated with adverse fetal outcomes. J Diabetes Complications. 2020;34:107556.

11. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 2015;43:W566–70.
12. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2012;41:D808–15.
13. Kohl M, Wiese S, Warscheid B. Cytoscape: Software for Visualization and Analysis of Biological Networks. In: Hamacher M, Eisenacher M, Stephan C, editors. Data Mining in Proteomics. Totowa, NJ: Humana Press; 2011. p. 291–303. doi:10.1007/978-1-60761-987-1_18.
14. Jr GD, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;:11.
15. Leibsohn S, d'Ablaing G, Mishell DR, Schlaerth JB. Leiomyosarcoma in a series of hysterectomies performed for presumed uterine leiomyomas. Am J Obstet Gynecol. 1990;162:968–76.
16. Maruo T, Matsuo H, Shimomura Y, Kurachi O, Gao Z, Nakago S, et al. Effects of progesterone on growth factor expression in human uterine leiomyoma. Steroids. 2003;68:817–24.
17. Wang J, Ohara N, Wang Z, Chen W, Morikawa A, Sasaki H, et al. A novel selective progesterone receptor modulator asoprisnil (J867) down-regulates the expression of EGF, IGF-I, TGFβ3 and their receptors in cultured uterine leiomyoma cells. Hum Reprod. 2006;21:1869–77.
18. Johnston J, Navaratnam S, Pitz M, Maniate J, Wiechec E, Baust H, et al. Targeting the EGFR Pathway for Cancer Therapy. Curr Med Chem. 2006;13:3483–92.
19. Mendelsohn J, Baselga J. The EGF receptor family as targets for cancer therapy. Oncogene. 2000;19:6550–65.
20. Chen H-W, Liu JCC, Chen JJW, Lee Y-M, Hwang J-L, Tzeng C-R. Combined differential gene expression profile and pathway enrichment analyses to elucidate the molecular mechanisms of uterine leiomyoma after gonadotropin-releasing hormone treatment. Fertil Steril. 2008;90:1219–25.
21. Mittal KR, Chen F, Wei JJ, Rijhvani K, Kurvathi R, Streck D, et al. Molecular and immunohistochemical evidence for the origin of uterine leiomyosarcomas from associated leiomyoma and symplastic leiomyoma-like areas. Mod Pathol. 2009;22:1303–11.
22. Poli G, Tuccinardi T, Rizzolio F, Caligiuri I, Botta L, Granchi C, et al. Identification of New Fyn Kinase Inhibitors Using a FLAP-Based Approach. J Chem Inf Model. 2013;53:2538–47.
23. Xie Y-G, Yu Y, Hou L-K, Wang X, Zhang B, Cao X-C. FYN promotes breast cancer progression through epithelial-mesenchymal transition. Oncol Rep. 2016;36:1000–6.
24. Islam MS, Ciavattini A, Petraglia F, Castellucci M, Ciarmela P. Extracellular matrix in uterine leiomyoma pathogenesis: a potential target for future therapeutics. Hum Reprod Update. 2018;24:59–85.
25. Norian JM, Malik M, Parker CY, Joseph D, Leppert PC, Segars JH, et al. Transforming Growth Factor β3 Regulates the Versican Variants in the Extracellular Matrix-Rich Uterine Leiomyomas. Reprod Sci. 2009;16:1153–64.

26. Ponnampalam AP, Weston GC, Trajstman AC, Susil B, Rogers PAW. Molecular classification of human endometrial cycle stages by transcriptional profiling. MHR Basic Sci Reprod Med. 2004;10:879–93.
27. Kurita K, Maeda M, Mansour MA, Kokuryo T, Uehara K, Yokoyama Y, et al. TRIP13 is expressed in colorectal cancer and promotes cancer cell invasion. Oncol Lett. 2016;12:5240–6.
28. Lu S, Qian J, Guo M, Gu C, Yang Y. Insights into a Crucial Role of TRIP13 in Human Cancer. Comput Struct Biotechnol J. 2019;:8.
29. Cuevas IC, Sahoo SS, Kumar A, Zhang H, Westcott J, Aguilar M, et al. Fbxw7 is a driver of uterine carcinosarcoma by promoting epithelial-mesenchymal transition. Proc Natl Acad Sci. 2019;116:25880–90.
30. Kuiper RP, Vreede L, Venkatachalam R, Ricketts C, Kamping E, Verwiel E, et al. The tumor suppressor gene FBXW7 is disrupted by a constitutional t(3;4)(q21;q31) in a patient with renal cell cancer. Cancer Genet Cytogenet. 2009;195:105–11.
31. Arslan AA, Gold LI, Mittal K, Suen T-C, Belitskaya-Levy I, Tang M-S, et al. Gene expression studies provide clues to the pathogenesis of uterine leiomyoma: new evidence and a systematic review. Hum Reprod. 2005;20:852–63.
32. Rubel CA, Wu S-P, Lin L, Wang T, Lanz RB, Li X, et al. A Gata2-Dependent Transcription Network Regulates Uterine Progesterone Responsiveness and Endometrial Function. Cell Rep. 2016;17:1414–25.
33. Tien F-M, Hou H-A, Tsai C-H, Tang J-L, Chiu Y-C, Chen C-Y, et al. GATA2 zinc finger 1 mutations are associated with distinct clinico-biological features and outcomes different from GATA2 zinc finger 2 mutations in adult acute myeloid leukemia. Blood Cancer J. 2018;8:87.
34. Luo X, Ding L, Xu J, Chegini N. Gene Expression Profiling of Leiomyoma and Myometrial Smooth Muscle Cells in Response to Transforming Growth Factor-. :22.
35. Santagata S, Demichelis F, Riva A, Varambally S, Hofer MD, Kutok JL, et al. JAGGED1 Expression Is Associated with Prostate Cancer Metastasis and Recurrence. Cancer Res. 2004;64:6854–7.
36. Lee J, Lee J, Kim JH. Association of Jagged1 expression with malignancy and prognosis in human pancreatic cancer. Cell Oncol. 2020;43:821–34.
37. Allhorn S, Böing C, Koch AA, Kimmig R, Gashaw I. TLR3 and TLR4 expression in healthy and diseased human endometrium. Reprod Biol Endocrinol. 2008;6:40.
38. Salaun B, Coste I, Rissoan M-C, Lebecque SJ, Renno T. TLR3 Can Directly Trigger Apoptosis in Human Cancer Cells. J Immunol. 2006;176:4894–901.
39. Bianchi F, Milione M, Casalini P, Centonze G, Le Noci VM, Storti C, et al. Toll-like receptor 3 as a new marker to detect high risk early stage Non-Small-Cell Lung Cancer patients. Sci Rep. 2019;9:14288.
40. Hu C-AA, Klopfer EI, Ray PE. Human apolipoprotein L1 (ApoL1) in cancer and chronic kidney disease. FEBS Lett. 2012;586:947–55.

# Figures

# Figure. 1

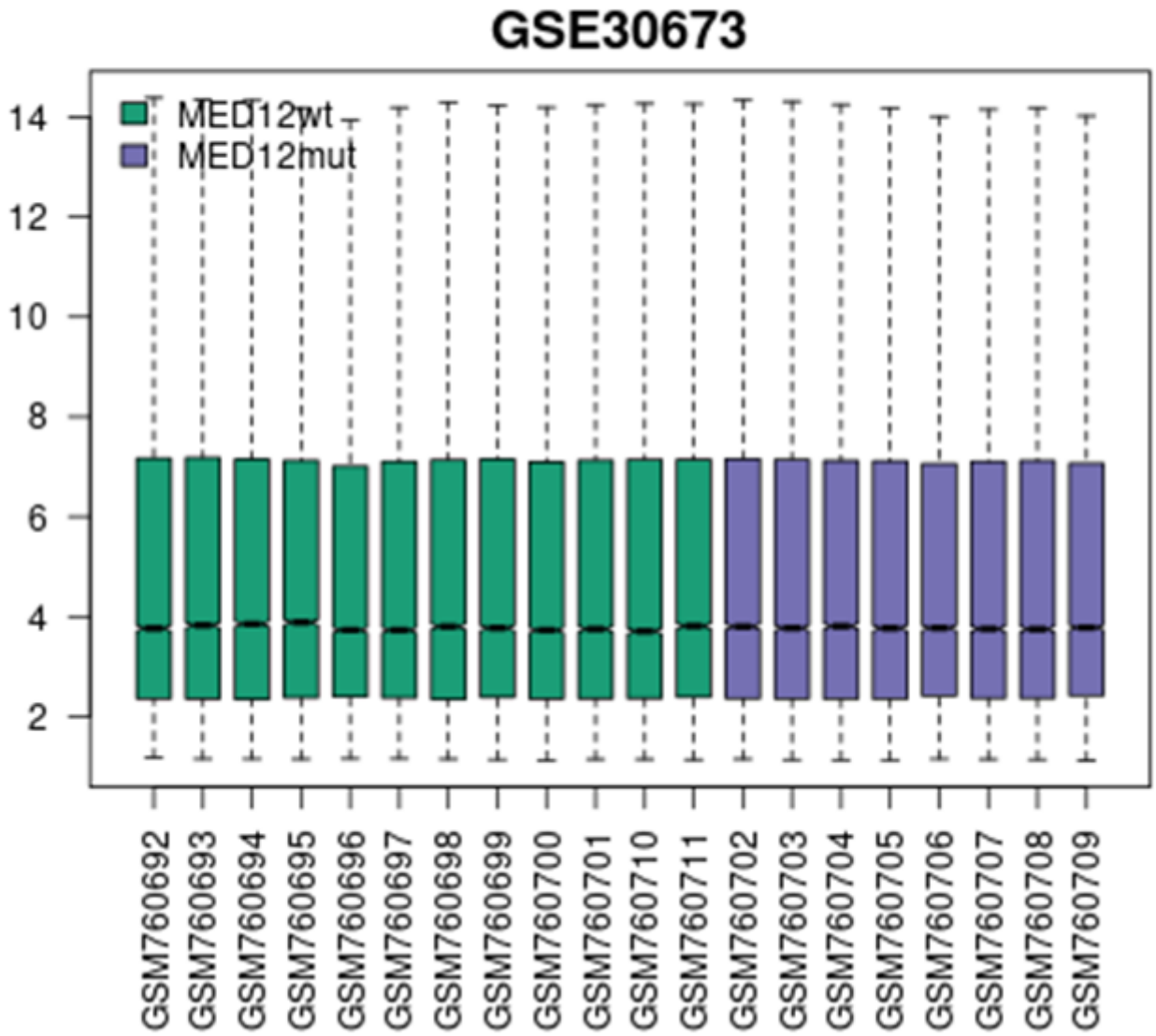
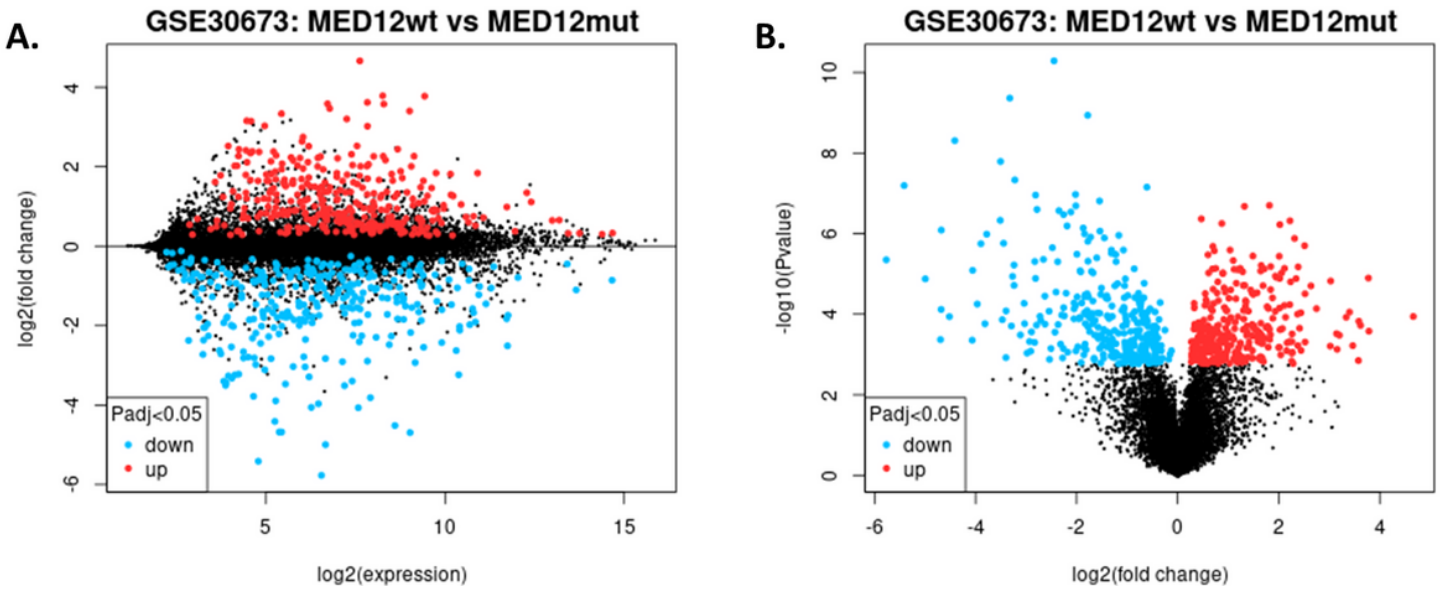



**Figure 1**

Microarray datasets of ULM with MED12 wildtype and mutant were normalized (Figure 1) through GEO2R. 176 of total DEGs with 101 overexpressed and 75 under expressed genes were screened out (Supplementary file 1). The p-value < 0.05 and | log2 FC > 0.1 as parameters for DEGs determination were used

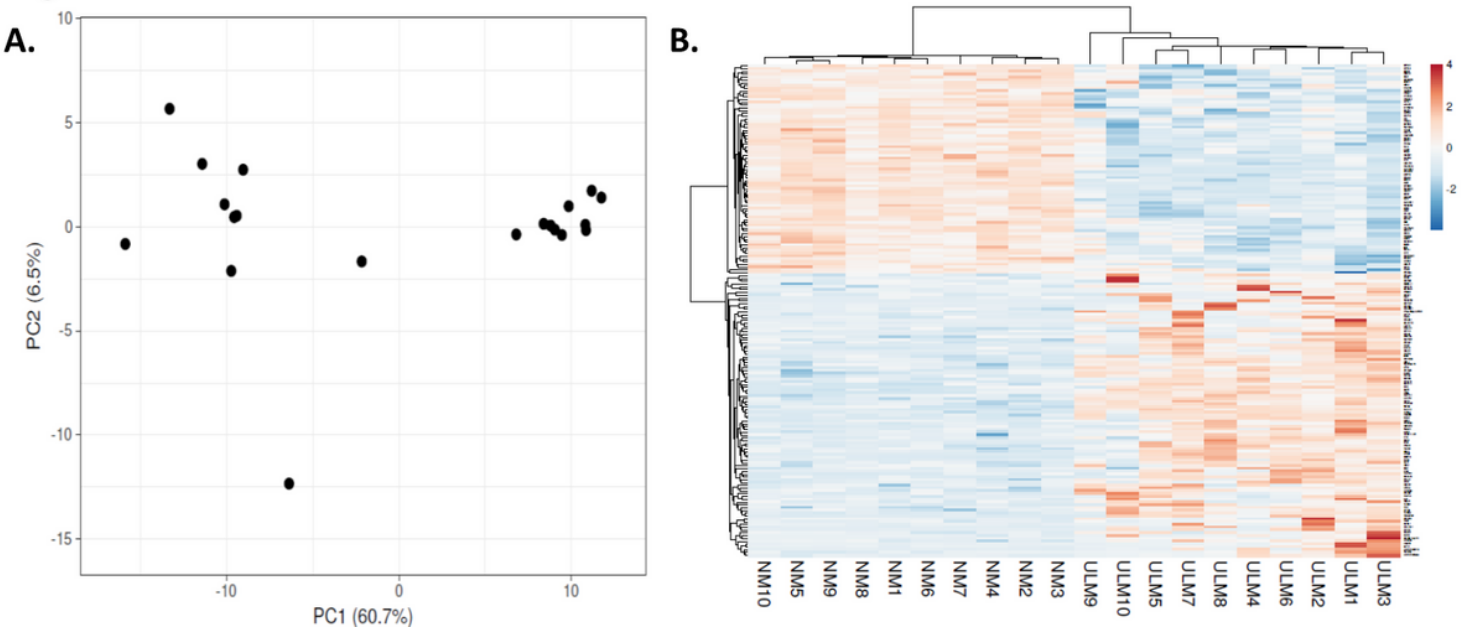
Figure. 2




## Figure 2

Visualization of screened DEGs through constructed Volcano plot and MD plot reveals the up and down regulated genes (Figure 2A and 2B)

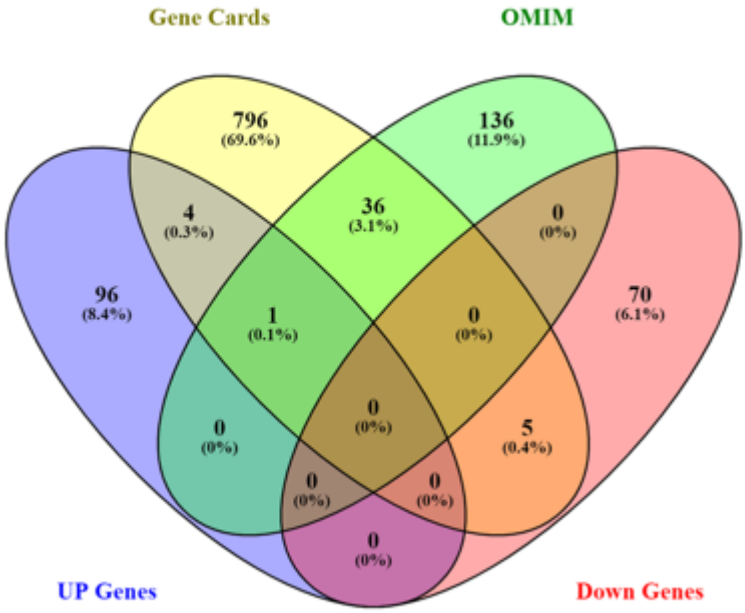
Figure. 3




## Figure 3

Principal Component plot for ULM disclose a scatter plot indicating total variance of 60.7% identical to the principal component 1 on x-axis while 6.5% identical to principal component 2 on y-axis respectively (Figure 3A). Heat-map plot was constructed for DEGs disclose a data matrix with coloring patterns provides an overview of numerical variation. (Figure 3B).

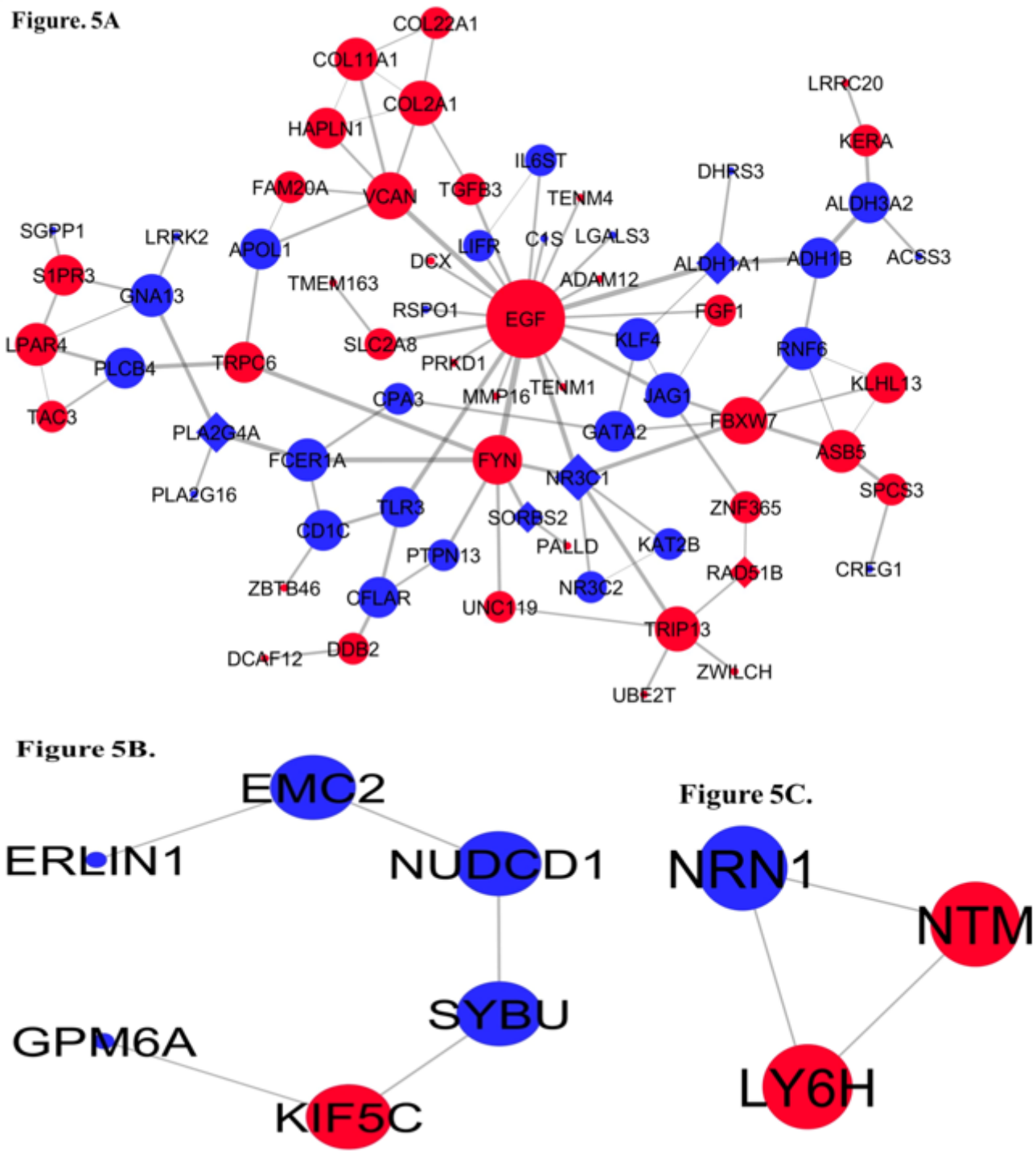
Figure. 4




| | |
|---|---|
| Total DEGs | 176 |
| UP Genes | 101 |
| Down Genes | 75 |
| Known Disease genes | 6 |

## Figure 4

Constructed Venn diagram revealed 2 DEGs as common genes found in OMIM and Gene Cards (Figure 4). Moreover, 6 known disease gene were also found. DEGs were determined based upon average gene expression value.




## Figure 5

Based on combined score generated by STRING, protein-protein interaction network was formed for all the DEGs with combined score >0.9. This lead to construction of PPI network (Figure 5A) consisting of 85 nodes and 157 edges. The red and blue color represents the over and under expression of DEGs. Degree and edge betweeness were employed to create hub nodes. Genes forming hub nodes were EGF (Epidermal Growth Factor), FYN (FYN Proto-Oncogene), VCAN (VERSICAN), TRIP13 (Thyroid Hormone Receptor Interactor 13), FBXW7 (F Box and WD Repeat Domain Containing 7) (up regulated genes) and

GATA2 (GATA Binding Protein 2),JAG1 (Jagged Canonical Notch Ligand 1),TLR3 (Toll Like Receptor 3),APOL1 (Apolipoprotein L1) (under regulated genes). Since these hub nodes were directly associated with known disease gene and are novel hence can be considered potential target genes for ULM. Few DEGs formed two separate network from main network and were treated as sub-network. Sub-network first contains six nodes and five edges while second network contain three nodes and three edges (Figure 5B and C). Based on the combined score and degree, a total of ten DEGs were selected out as novel genes.

Figure. 6A Gene Ontology enrichment analysis for up regulated genes

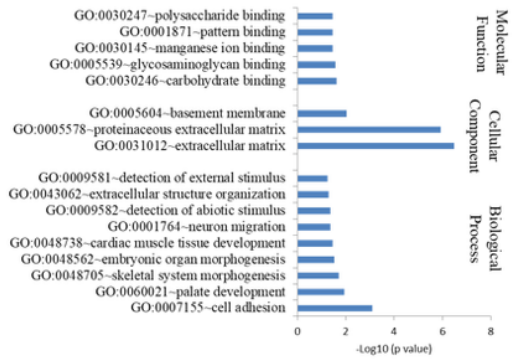



Figure. 6B Gene Ontology enrichment analysis for down regulated genes

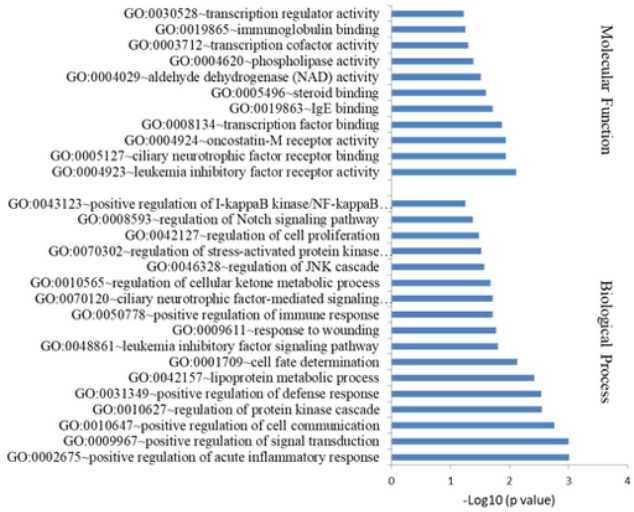



Figure. 6C KEGG pathway enrichment analysis

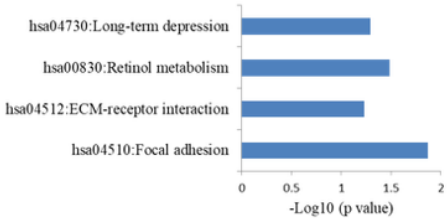



**Figure 6**

Most significant (p<0.05) processes being regulated by up-DEGs were cell adhesion, extracellular structure organisation, skeletal system morphogenesis while Notch signalling, JNK cascade, leukemia inhibitory factor signalling and NF-KB signalling are major significant (p<0.05) processes which are down-regulated (Figure 6A and B). Most of the DEGs are localised in either extracellular matrix or basement membrane. Focal adhesion, ECM-receptor interaction, long-term depression and Retinol metabolism are major pathways which have been enriched for these DEGs as revealed through KEGG enrichment analysis (Figure 6C).

# Supplementary Files

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

- Supplementryfile.xlsx