Original Article

Bioinformatical enrichment analysis of genes involved in the pathway of endometriosis disease

Kusum Kusum[1]*, Ashish Ashish[2*], Ravi Bhushan[3], Radha Chaube[1], Sangeeta Rai[4], Royana Singh [2]*

***The First Two Authors Have Equal Contribution**

[1] Department of Zoology, Institute of Science, Banaras Hindu University, Varanasi-221005 Uttar Pradesh, India.

[2] MRU Lab, Department of Anatomy, Institute of Medical Science, Banaras Hindu University, Varanasi, 221005, Uttar Pradesh, India.

[3] Department of Chromosomal Genetics Disorder Institute of Science, Banaras Hindu University, Varanasi-221005 Pradesh, India.

[4] obstetrics and gynecology, Institute of Medical Science, Banaras Hindu University Varanasi-221005 Pradesh, India.

Corresponding Author
royanasingh@bhu.ac.in

Keywords: Endometriosis, mRNA seq analysis, reproductive women, estrogen, aromatase

Abstract

Endometriosis, is a gynecological disease, where uterine (eutopic) endometrial glands and tissues are present outside the intra-uterine locations, ectopic regions (pelvic peritoneum, fallopian tubes or ovaries). About 5-10% reproductive and 20-50% infertile women have endometriosis. Several factors like hormonal, environmental, genetic and immune system are involved in the pathogenesis of endometriosis both directly or indirectly altering the estrogen level and thus affecting the reproductive health of women. Current study was done with an aim to identify novel and potential biomarker for endometriosis using mRNA seq analysis. From raw gene expression profiles differentially expressed genes (DEGs) were identified and further their functional analysis was conducted. A total of 552 (312 up and 240 downregulated) DEG's were identified in samples from endometriosis suffering women when compared with control subjects. Major DEGs forming hubnodes like C3, PSAP, APP, GNG12 were found to be involved in various functions such as, epithelia cell differentiation and development, proteolysis, gland development, muscle fiber development, response to hormone stimulus. The identified DEGs can be directly or indirectly involved in the pathway of pathogenesis of endometriosis and can act as a potential biomarker for ectopic endometrium. Current study will provide a preliminary insight into the mechanism of endometriosis disease; however, it will require further detailed studies for its complete path of action.

Introduction

Endometriosis is a pronouncing gynecological disease presently concerning about the women health. It's a disease with presence of endometrial glands and stromal tissues outside the uterine endometrium (eutopic region) to ectopic regions (pelvic peritonium, on fallopian tubes or on ovaries)(1–3). It is an estrogen dependent disorder, with which about 5-10% reproductively active and 20-50% women diagnosed with infertility suffer worldwide (4). The etiology and pathogenesis of endometriosis is still unclear, and it's directly or indirectly burdens economic as well as reproductive health of the women(5). A number of theories have been postulated to explain its cause of occurrence, but none has been accepted widely. Sampson's theory of retrograde mensuration is based on the endometrial cells being transported from eutopic regions (uterine location) to ectopic regions (outside uterine locations), causing the endometriosis disease(6). However extensive studies have been needed to differentiate between expression patterns in eutopic endometrium and normal uterine endometrium (ectopic) (7). A non-invasive diagnostic marker using serum, urine or endometrial tissue samples is now needed for early diagnosis of the disease(8). The precise mechanism affecting the overall mechanism is still unknown. RNA-Seq analysis is recently developed new approach to study transcriptomes using deep-sequencing technologies(9). Transcriptome is the complete set of transcripts in a cell and their total quantity for a particular physiological condition, and it help us to understand the functional elements working at that time(10) . The molecular analysis of endometrium of diseased women in comparison to normal endometrium as control will be helpful in understanding the disease pathophysiology with specific biomarker for its diagnosis.

Material and Methods

Microarray data and samples

From the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www. ncbi.nlm.nih.gov/geo/), the raw gene expression profile datasets (ID: GSE7305) were obtained for this study (11). For this data set samples were from tissues of normal eutopic endometrium, and ectopic together with eutopic endometrium from endometriosis patients. In the study, genes involved in eutopic and ectopic endometrium of control and endometriosis patient samples were undertaken and analyzed using several bioinformatic tools. The microarray study was duly approved by ethical committee of the Institution, Banaras Hindu University **((No. Dean/2018/EC/936)).**

Data processing and DEG's screening

From the datasets of raw gene expression, the expression values of probes corresponding to specific gene were averaged to obtain the gene expression values. Then by using BiGGEsTs software analysis tool, up and down regulated genes were selected. Further with GEO2R (https://www. ncbi.nlm.nih.gov/geo/geo2r/) the probe-level symbols were converted into gene-level symbols, where all the selected DEG's were having p-values<.05, with threshold logFC values>0.1 for up and <-0.1 for down regulated genes(12).

Principal component analysis and heat map generation

This is a provisional file, not the final typeset article

Using the online tool ClustVis (13), Principal component analysis (PCA) was performed and heat map was generated. PCA was plotted only for DEG's, and it was not possible to prepare PCA plot for the total gene expression as because ClustVis supports only file size up to 2MB.

Identification of novel biomarkers related to endometriosis

To filter out the novel bio-markers related to endometriosis, list of differentially expressed genes were compared to the reported gene list obtained from OMIM (https://www.omim.org/) and Gene Cards (https://www.genecards.org/) (14). For comparison and venn diagram construction Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) (15)was used.

Construction of protein-protein interaction (PPI) network and sub-network mining

DEG's were identified and uploaded on STRING v 10.5 (http://www. string-db.org/) (16). STRING is an online database that predicts functional interactions between proteins, giving a combined score for PPI among gene pairs. Combined score was set as >0.4, as the baseline criteria and further network and sub-network were created using Cytoscape v 3.2.1. Cytoscape (http://www.cytoscape.org/) (17) is a software for visualization and analysis of biological networks with high throughput data, where clustering coefficient and edge betweenness were taken as criteria for network constructions.

Functional and pathway enrichment analysis of DEG's

Gene ontology (GO) enrichment analysis includes biological processes (BP), molecular functions (MF) and cellular components (CC), and it was performed using DAVID v 6.8. DAVID (database for annotation, visualization and intergrated discovey, http://david.abcc.ncifcrf.gov/)(18), a programme integrates comprehensive set of functional annotation of a large gene list. On the basis of hypergenometric distribution, DAVID takes the genes with similar or related functions as a whole set.

Results

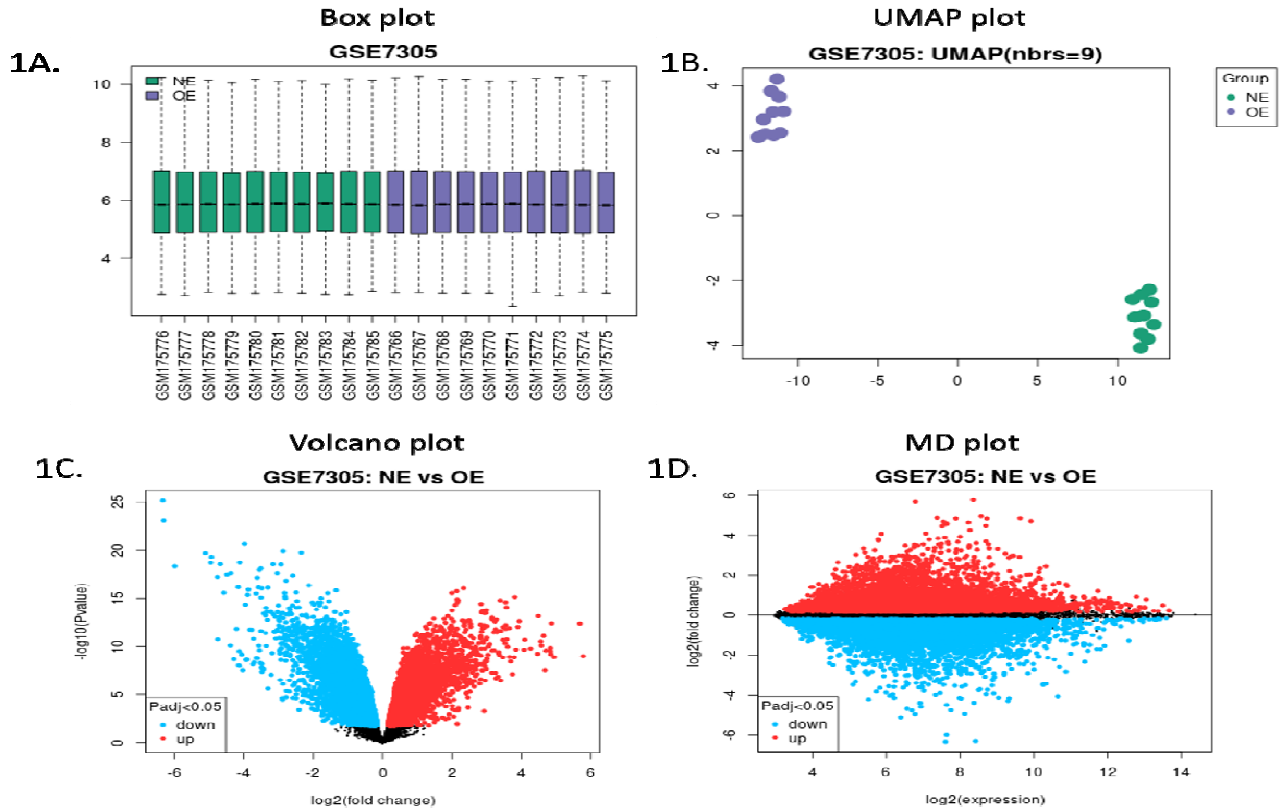


Figure 1: Data normalization and DEG's visualization via Box plot (A), UMAP plot (B), Volcano plot (C) and MD plot (D).

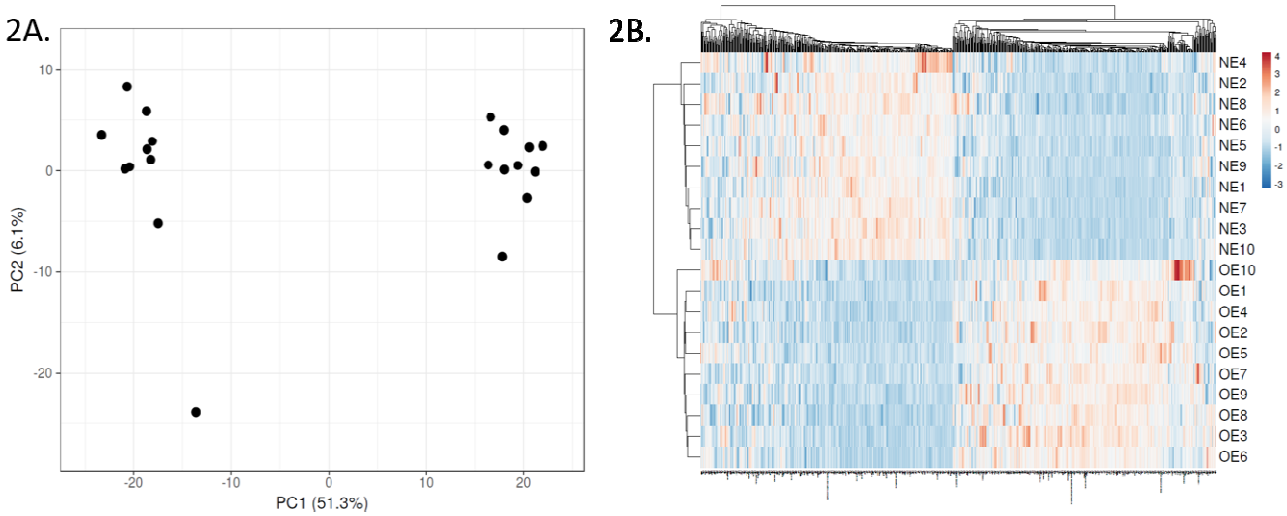
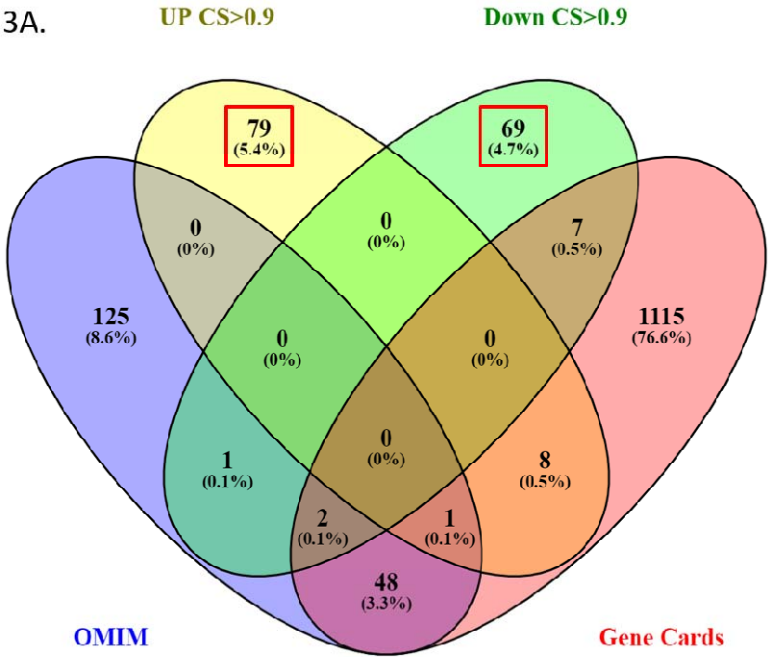


Figure 2: Heat map of DEG's. Blue to orange gradation is for small to large changes in gene expression values.

This is a provisional file, not the final typeset article

3A.



DEGs	Number
Total DEGs	552
UP DEGs	312
Down DEGs	240
UP DEGs with cs>0.9	88
Down DEGs with cs>0.9	79
Reported DEGs	19
Total Novel DEGs	148
Novel UP DEGs	79
Novel Down DEGs	69

Figure 3: Venn diagram of the differentially expressed genes (DEGs).

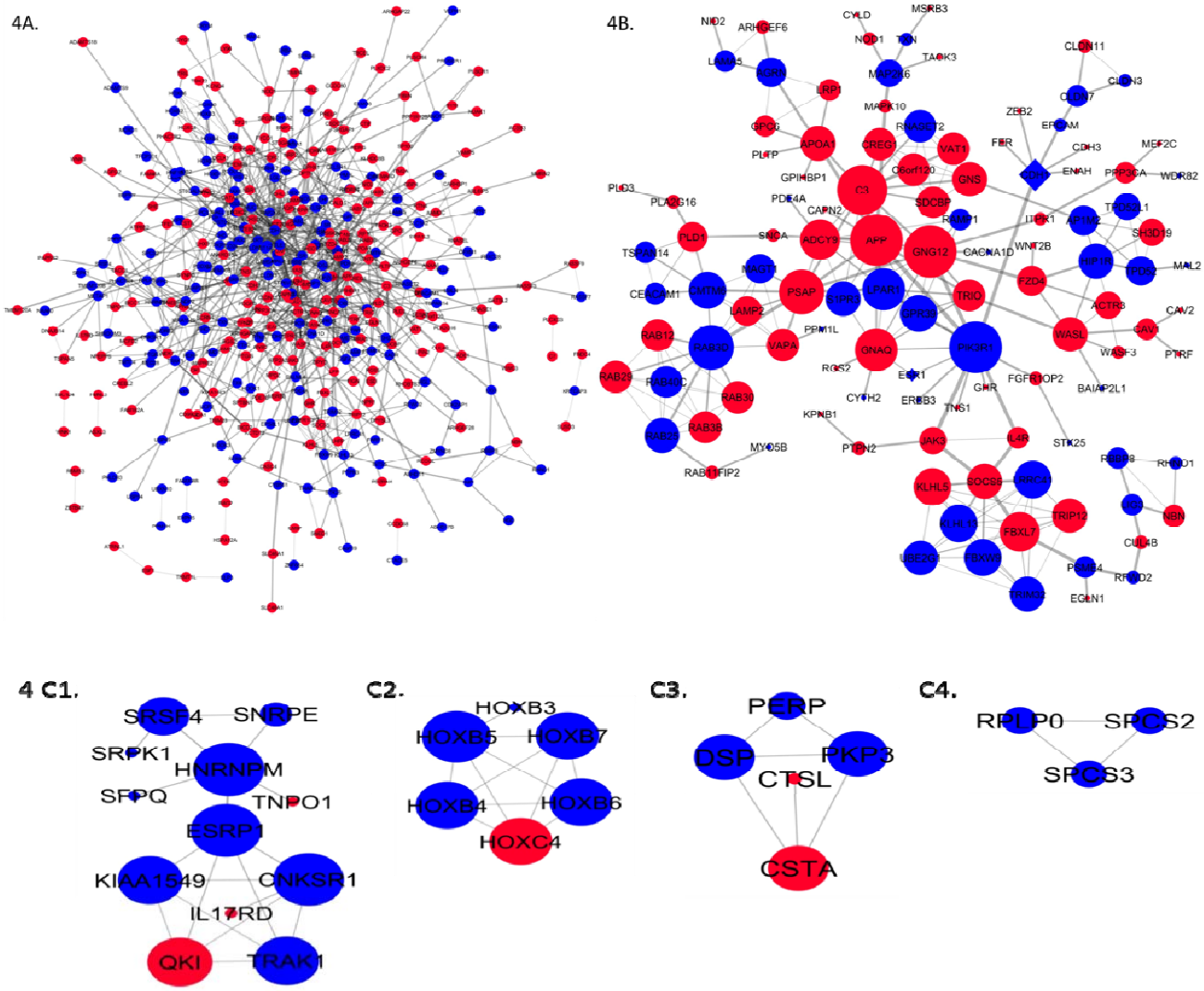


Figure 4: (A and B) Protein-protein interaction (PPI) of DEG's. C1, C2, C3 and C4 are four sub-networks of DEG's: C1; sub-network 1, C2; sub-network 2, C3; sub-network 3, C4; sub-network 4. Red-circle for up and blue-circle for down-regulated genes respectively. Blue diamond for similar related genes and lines shows the correlation between genes, where thickness of lines (edges), is proportional to the combined scores.

This is a provisional file, not the final typeset article

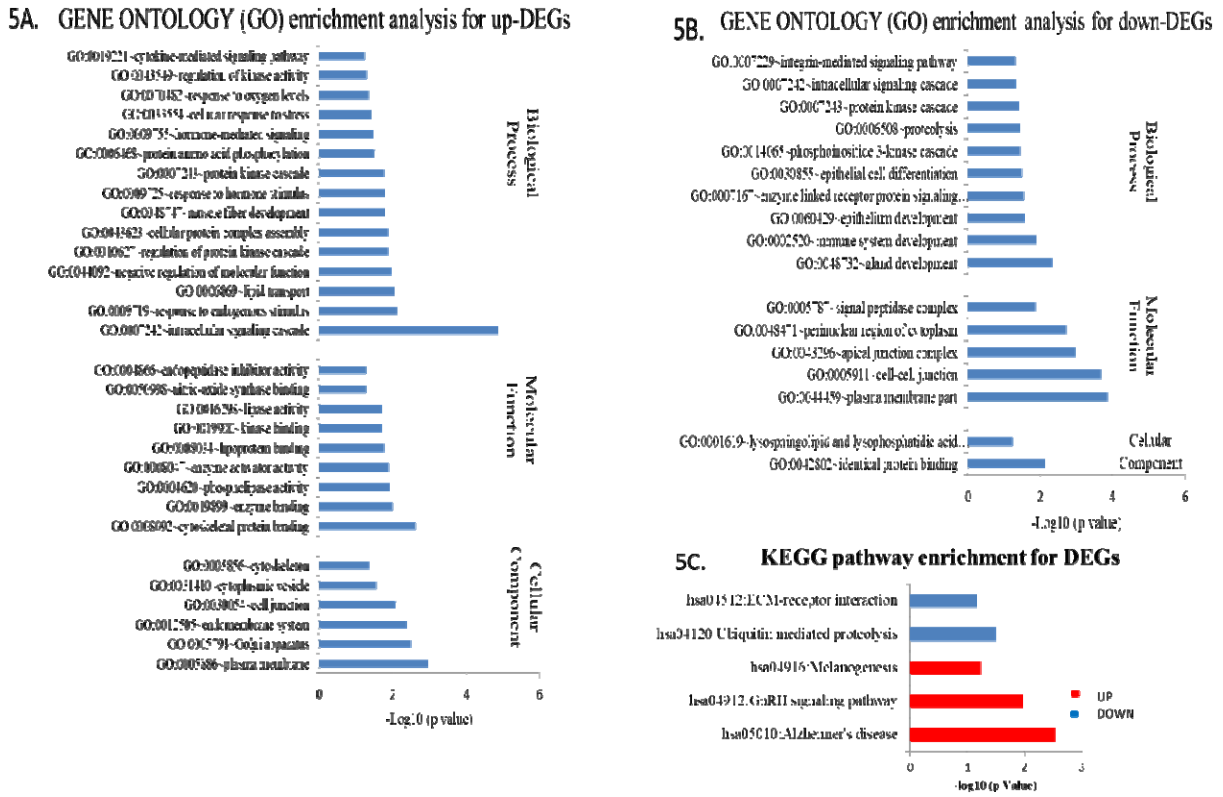


Figure 5: Gene Ontology (GO) enrichment analysis for up-regulated (5A) and down-regulated DEG's (5B), and KEGG pathway enrichment for DEG's (5C).

Category	Term	P-Value
GOTERM_BP_FAT	GO:0007242~intracellular signaling cascade	1.34E-05
GOTERM_BP_FAT	GO:0014706~striated muscle tissue development	5.44E-04
GOTERM_BP_FAT	GO:0042692~muscle cell differentiation	5.87E-04
GOTERM_BP_FAT	GO:0060537~muscle tissue development	6.80E-04
GOTERM_BP_FAT	GO:0042391~regulation of membrane potential	9.31E-04
GOTERM_BP_FAT	GO:0007517~muscle organ development	0.001183
GOTERM_BP_FAT	GO:0051146~striated muscle cell differentiation	0.001501
GOTERM_BP_FAT	GO:0010033~response to organic substance	0.002202
GOTERM_BP_FAT	GO:0048878~chemical homeostasis	0.00222
GOTERM_BP_FAT	GO:0010647~positive regulation of cell communication	0.002453
GOTERM_BP_FAT	GO:0006873~cellular ion homeostasis	0.004966
GOTERM_BP_FAT	GO:0006928~cell motion	0.005073
GOTERM_BP_FAT	GO:0055082~cellular chemical homeostasis	0.00541
GOTERM_BP_FAT	GO:0016044~membrane organization	0.005487
GOTERM_BP_FAT	GO:0007519~skeletal muscle tissue development	0.00605
GOTERM_BP_FAT	GO:0060538~skeletal muscle organ development	0.00605
GOTERM_BP_FAT	GO:0009967~positive regulation of signal transduction	0.006229
GOTERM_BP_FAT	GO:0009719~response to endogenous stimulus	0.007589
GOTERM_BP_FAT	GO:0050801~ion homeostasis	0.007991
GOTERM_BP_FAT	GO:0032273~positive regulation of protein polymerization	0.008595
GOTERM_BP_FAT	GO:0006869~lipid transport	0.008949
GOTERM_BP_FAT	GO:0042592~homeostatic process	0.009037

GOTERM_BP_FAT	GO:0048741~skeletal muscle fiber development	0.010712
GOTERM_BP_FAT	GO:0044092~negative regulation of molecular function	0.011126
GOTERM_BP_FAT	GO:0001765~membrane raft formation	0.011205
GOTERM_BP_FAT	GO:0045807~positive regulation of endocytosis	0.011464
GOTERM_BP_FAT	GO:0010876~lipid localization	0.011736
GOTERM_BP_FAT	GO:0052547~regulation of peptidase activity	0.012493
GOTERM_BP_FAT	GO:0010627~regulation of protein kinase cascade	0.012994
GOTERM_MF_FAT	GO:0008092~cytoskeletal protein binding	0.002399
GOTERM_MF_FAT	GO:0016504~peptidase activator activity	0.007702
GOTERM_MF_FAT	GO:0019899~enzyme binding	0.010389
GOTERM_MF_FAT	GO:0003779~actin binding	0.01129
GOTERM_MF_FAT	GO:0004620~phospholipase activity	0.012226
GOTERM_MF_FAT	GO:0008047~enzyme activator activity	0.012783
GOTERM_MF_FAT	GO:0008289~lipid binding	0.015001
GOTERM_MF_FAT	GO:0008034~lipoprotein binding	0.017323
GOTERM_MF_FAT	GO:0019900~kinase binding	0.019893
GOTERM_MF_FAT	GO:0016298~lipase activity	0.020073
GOTERM_MF_FAT	GO:0070290~NAPE-specific phospholipase D activity	0.028556
GOTERM_MF_FAT	GO:0042802~identical protein binding	0.030878
GOTERM_MF_FAT	GO:0070064~proline-rich region binding	0.03417
GOTERM_MF_FAT	GO:0008035~high-density lipoprotein binding	0.039752
GOTERM_MF_FAT	GO:0004630~phospholipase D activity	0.039752
GOTERM_MF_FAT	GO:0008139~nuclear localization sequence binding	0.039752
GOTERM_MF_FAT	GO:0019904~protein domain specific binding	0.042369
GOTERM_MF_FAT	GO:0050998~nitric-oxide synthase binding	0.050821
GOTERM_MF_FAT	GO:0004866~endopeptidase inhibitor activity	0.05148
GOTERM_MF_FAT	GO:0030414~peptidase inhibitor activity	0.058629
GOTERM_CC_FAT	GO:0030055~cell-substrate junction	4.14E-04
GOTERM_CC_FAT	GO:0005886~plasma membrane	0.001132
GOTERM_CC_FAT	GO:0005912~adherens junction	0.001789
GOTERM_CC_FAT	GO:0005626~insoluble fraction	0.002358
GOTERM_CC_FAT	GO:0005925~focal adhesion	0.002595
GOTERM_CC_FAT	GO:0070161~anchoring junction	0.002815
GOTERM_CC_FAT	GO:0005924~cell-substrate adherens junction	0.002983
GOTERM_CC_FAT	GO:0005794~Golgi apparatus	0.00324
GOTERM_CC_FAT	GO:0009986~cell surface	0.003371
GOTERM_CC_FAT	GO:0012505~endomembrane system	0.004123
GOTERM_CC_FAT	GO:0031982~vesicle	0.004141
GOTERM_CC_FAT	GO:0005624~membrane fraction	0.005317
GOTERM_CC_FAT	GO:0048471~perinuclear region of cytoplasm	0.00559
GOTERM_CC_FAT	GO:0016323~basolateral plasma membrane	0.00569
GOTERM_CC_FAT	GO:0044431~Golgi apparatus part	0.006169
GOTERM_CC_FAT	GO:0030027~lamellipodium	0.007163
GOTERM_CC_FAT	GO:0044459~plasma membrane part	0.008355
GOTERM_CC_FAT	GO:0030054~cell junction	0.008464

This is a provisional file, not the final typeset article

GOTERM_CC_FAT	GO:0009925~basal plasma membrane	0.009319
GOTERM_CC_FAT	GO:0045178~basal part of cell	0.012293
GOTERM_CC_FAT	GO:0031088~platelet dense granule membrane	0.016805
GOTERM_CC_FAT	GO:0000267~cell fraction	0.017154
GOTERM_CC_FAT	GO:0044421~extracellular region part	0.017907
GOTERM_CC_FAT	GO:0000139~Golgi membrane	0.020736
GOTERM_CC_FAT	GO:0031410~cytoplasmic vesicle	0.027418
GOTERM_CC_FAT	GO:0005955~calcineurin complex	0.027853
GOTERM_CC_FAT	GO:0042827~platelet dense granule	0.033332
GOTERM_CC_FAT	GO:0005901~caveola	0.03462
GOTERM_CC_FAT	GO:0016023~cytoplasmic membrane-bounded vesicle	0.034768
GOTERM_CC_FAT	GO:0005615~extracellular space	0.038156
GOTERM_CC_FAT	GO:0031988~membrane-bounded vesicle	0.040294
GOTERM_CC_FAT	GO:0042995~cell projection	0.0416
GOTERM_CC_FAT	GO:0031252~cell leading edge	0.042819
GOTERM_CC_FAT	GO:0005856~cytoskeleton	0.043217
GOTERM_CC_FAT	GO:0005576~extracellular region	0.051766
GOTERM_CC_FAT	GO:0019898~extrinsic to membrane	0.059319

Table 1: GENE ONTOLOGY (GO) analysis for down regulated DEGs.

Category	Term	P-Value
GOTERM_BP_FAT	GO:0048704~embryonic skeletal system morphogenesis	1.89E-04
GOTERM_BP_FAT	GO:0048706~embryonic skeletal system development	6.00E-04
GOTERM_BP_FAT	GO:0048705~skeletal system morphogenesis	0.002413
GOTERM_BP_FAT	GO:0048562~embryonic organ morphogenesis	0.004473
GOTERM_BP_FAT	GO:0048732~gland development	0.004717
GOTERM_BP_FAT	GO:0009952~anterior/posterior pattern formation	0.005363
GOTERM_BP_FAT	GO:0043009~chordate embryonic development	0.006282
GOTERM_BP_FAT	GO:0009792~embryonic development ending in birth or egg hatching	0.006557
GOTERM_BP_FAT	GO:0048568~embryonic organ development	0.010931
GOTERM_BP_FAT	GO:0002520~immune system development	0.012496
GOTERM_BP_FAT	GO:0003002~regionalization	0.01722
GOTERM_BP_FAT	GO:0048598~embryonic morphogenesis	0.018973
GOTERM_BP_FAT	GO:0001501~skeletal system development	0.021983
GOTERM_BP_FAT	GO:0060429~epithelium development	0.027277
GOTERM_BP_FAT	GO:0007167~enzyme linked receptor protein signaling pathway	0.02859
GOTERM_BP_FAT	GO:0030097~hemopoiesis	0.030857
GOTERM_BP_FAT	GO:0030855~epithelial cell differentiation	0.031589
GOTERM_BP_FAT	GO:0014065~phosphoinositide 3-kinase cascade	0.034668
GOTERM_BP_FAT	GO:0060216~definitive hemopoiesis	0.034668
GOTERM_BP_FAT	GO:0015031~protein transport	0.036558
GOTERM_BP_FAT	GO:0006508~proteolysis	0.036929
GOTERM_BP_FAT	GO:0007243~protein kinase cascade	0.038188
GOTERM_BP_FAT	GO:0045184~establishment of protein localization	0.038276
GOTERM_BP_FAT	GO:0006465~signal peptide processing	0.039523
GOTERM_BP_FAT	GO:0009057~macromolecule catabolic process	0.041345
GOTERM_BP_FAT	GO:0048534~hemopoietic or lymphoid organ development	0.04171

GOTERM_BP_FAT	GO:0007389~pattern specification process	0.045236
GOTERM_BP_FAT	GO:0007242~intracellular signaling cascade	0.047661
GOTERM_BP_FAT	GO:0007229~integrin-mediated signaling pathway	0.048307
GOTERM_BP_FAT	GO:0045022~early endosome to late endosome transport	0.04916
GOTERM_BP_FAT	GO:0008380~RNA splicing	0.054478
GOTERM_CC_FAT	GO:0044459~plasma membrane part	1.39E-04
GOTERM_CC_FAT	GO:0005911~cell-cell junction	2.12E-04
GOTERM_CC_FAT	GO:0043296~apical junction complex	0.001043
GOTERM_CC_FAT	GO:0016327~apicolateral plasma membrane	0.001165
GOTERM_CC_FAT	GO:0048471~perinuclear region of cytoplasm	0.001881
GOTERM_CC_FAT	GO:0030057~desmosome	0.003648
GOTERM_CC_FAT	GO:0005787~signal peptidase complex	0.013552
GOTERM_CC_FAT	GO:0070695~FHF complex	0.022487
GOTERM_CC_FAT	GO:0030054~cell junction	0.02936
GOTERM_CC_FAT	GO:0005886~plasma membrane	0.036058
GOTERM_CC_FAT	GO:0070161~anchoring junction	0.043114
GOTERM_CC_FAT	GO:0045177~apical part of cell	0.047573
GOTERM_CC_FAT	GO:0030897~HOPS complex	0.053135
GOTERM_MF_FAT	GO:0042802~identical protein binding	0.007236
GOTERM_MF_FAT	GO:0001619~lysosphingolipid and lysophosphatidic acid receptor activity	0.05657

Table 2: GENE ONTOLOGY (GO) analysis for down regulated DEGs.

Differentially expressed genes (DEGs)

A total of 552 DEG's were identified with p-value<0.05, out of which 312 were found to be up-regulated while 240 were down-regulated. Out of total 552 DEGs, 148 DEG's were found to be novel of which 79 were up and 69 were down-regulated DEG's, and these were selected on the basis of their average gene expression values.

Principal component and hierarchical clustering analysis of DEGs

Uniform manifold approximation and projection plot (umap) shows clustering of data with neighborhood scoring 9. Heat-map was constructed for DEG's, which shows a data matrix where coloring gives an overview of the numeric differences for gene expressions. The Up regulated DEG's are shown in orange and while down regulated are shown in blue color. The blue to orange gradation shows the values changing from small to large for the gene expressions.

The protein-protein interaction network

Based on the combined score calculated by STRING, a total of 297 gene pairs (combined score>0.9) were found to interact with each other, forming a prime network having 114 nodes and 237 edges respectively (Figure 4) while 4 sub-networks were extracted separately. Upregulated DEGs like C3, APP, GNG12, PSAP, GNAQ were forming hub nodes while down-regulated DEG's forming hub nodes were LPAR1, RAB3D, PIK3R1, TRIM32, CMTM6. The clustering coefficient and edge betweenness were taken as a basic criterion for the selection of hub nodes.

Sub-network extraction

Four sub-networks C1, C2, C3 and C4 were extracted from the main network with the use of Cytoscape (Fig. 4). All genes in sub-network C4 were down regulated while in sub-networks C1, C2, C3, 3, 1, 2, genes were up-regulated and 9, 5, 3 genes were down-regulated respectively.

This is a provisional file, not the final typeset article

Gene Ontology and pathway enrichment analysis

Functional enrichment analysis was performed and significantly enriched major molecular functions, biological processes and cellular components of DEGs (FDR<0.05) were listed in Tables 1 and 2. The results of GO enrichment analyses identifies hormone mediated-signaling, muscle fiber development, intracellular signaling cascade and cellular protein complex assembly as major significant biological processes for upregulated DEGs (5A). While processes like, proteolysis, immune system development, epithelial cell differentiation and development are major significant biological processes related to downregulated DEGs (5B). Furthermore, KEGG pathway enrichment analysis (Fig.5C) revealed ECM-receptor interaction and upiquitin mediated proteolysis as downregulated pathways while melanogenesis, GnRH signaling pathway and alzimers disease pathway are upregulated pathways.

Discussion

Endometriosis is a gynecological disease that affects the reproductive health of a women. The present study was performed to understand the molecular mechanisms involved in the complexity of the endometriosis disease progression in the reproductive women. In this study gene expression datasets were analyzed thoroughly by using different bioinformatics tools, which has revealed total of 552 DEGs, in which 312 were up and 240 down-regulated respectively. The up-regulated DEG's are involved in Biological processes like cytokine mediated signaling pathway, cellular response to stress, hormone-mediated signaling, lipid transport, response to endogenous stimulus etc, Molecular Functions like lipase, kinase, lipoprotein, enzyme and cytoskeleton protein binding, enzyme activator and phospholipase activity. These up-regulated DEG's are part of cellular components that includes cytoskeleton, cytoplasmic vesicle, cell junctions, endomembrane system, golgi apparatus and plasma membrane. The down-regulated DEG's are involved in biological processes like proteolysis, glands and immune system development, integrin-mediated, intracellular and protein kinase cascade etc. These DEG's in molecular functions are a part of signal peptidase, apical junction, cell-cell junction complex, perinuclear regions and plasma membrane. While these as cellular components are involved in lysosphingolipid and lysophosphatidic acid and in identical protein binding. Major DEGs forming hub nodes are C3, APP, GNG12, PSAP, GNAQ, LPAR1, RAB3D, PIK3R1, TRIM32, CMTM6. In these C3, APP, GNG12, PSAP, GNAQ, are up-regulated and LPAR1, RAB3D, PIK3R1, TRIM32, CMTM6 are down-regulated DEG's.

C3 (Complement Component 3) is one of the important complement proteins out of 30 recognized till date. C3 has intensive and pivotal role in complement activation in both classical and alternative pathways. The alternative pathway is independent of antigen-antibody complexes and can directly be induced by components of cell wall of bacteria or present on the surface of damaged host cell via C3 unlike classical and lectin pathway (19). Altered immune system is among the various risk factors which are involved in pathophysiology of endometriosis and hence deregulated C3 (which is an important player of immune system) might be involved in the progression of endometriosis and can be considered as a potent biomarker.

APP (Amyloid precursor protein), plays an important role in synaptic activity and neuronal plasticity, but upto this time it's not completely revealed (20). APP gene has been found to be associated with down's syndrome (21). Mutation in APP gene were also found to be associated with dementia and Alzimers disease including amyloid deposition, neurofibrillary tangle formation and

cerebral amyloid angiopathy (CAA)(22). APP is involved in the proliferation, migration and adhesion of endothelial cells. It mediates stability to focal adhesion and cell-cell junctions, while its necessary for VEGF-A growth factor responses (23).Thus, it can be said that APP may be a factor to be involved in the adhesion and cellular junction formation of endometrial tissues during endometriosis.

GNG12 is known as the c12 subunit of G protein, and G proteins are made by three different subunits a, b and c making heterotrimeric. The different subunits are interchangeable, making possible combinations and wide array of effects. GNG12 is highly conserved, as and its homologs are present in human, rat, cow, frog, chicken and zebrafish. C12 is expressed differentially in mammalian brain, where its localized in glial cells and expressed in reactive astrocytes (24).

PSAP (Prosaposin, also known as SGP-1) is an intriguing multifunctional protein that plays roles intracellularly in regulation of lysosomal enzyme functions, and extracellularly, as a secretary factor having both neuroprotective and glioprotective effects. PSAP gene encodes a 524 amino-acid precursor protein prosaposin (pSAP), that gives 4 small glycoproteins-saposins (SapA, B, C and D) (25). Saposins, a sphingolipid activator protein that's required for the function of lysosomal hydrolases. Mutation in PSAP gene (two null alleles) in an individual have pSap deficiency and suffer with fatal infantile lysosomal storage disease(26). Defects of Sap A and SapC leads to atypical Krabbe and Gaucher disease (27). Any defect in SapB results in MLD (metachromatic leukodystrophy) due to impaired degradation and accumulation of cerebroside-3 sulfate (sulfatide) (28), while SapD deficiency causes Farber disease in mice (29).

GNAQ (Gnaq) is a member of guanine nucleotide binding protein (G protein) subunits and is found abundantly in brain. Gnaq gene knocking shows serious nervous dysfunction and endocrine system in mice (30) .Gnaq gene mutation has been reported mostly in uveal melanoma (31). Gnaq, Gq protein a-subunit, encoded by GNAQ gene, is a member of Gq/11subfamily of heterotrimeric G proteins and its ubiquitously expressed in mammalian cells (30,32). GNAQ has role in cardiovascular system (31), cancer and autoimmune diseases (33). It is found to be coupled with GPCRs in viral infections(34), but no any expression or mutation studies has been yet reported in endometriosis, adenomyosis or ovarian carcinoma.

PIK3R1

PI3K enzymes are lipid kinases, having a conserved sequences and phosphorylating the inositol 30-OH groups of membrane phosphoinositides (PI).

Class I PI3K convert phosphoinositol bisphosphate (PIP2) (4,5) into phosphoinositol trisphosphate (PIP3) (3,4,5) a second messenger (35). Class IA PI3K is composed of a heterodimer having a p110 catalytic subunit and a p85 regulatory subunit. Out of four PI3K catalytic subunit isoforms (PI3Ka, PI3Kb, PI3Kg, and PI3Kd), only PI3Ka and PI3Kb are ubiquitously expressed in the body and are frequently altered in cancer disease (36). Three different genes PIK3R1, PIK3R2, and PIK3R3, encode p85-type subunits; p85a, p85b and p55g, respectively. The two of PIK3R1 and PIK3R2 are widely expressed in the body, whereas the third one PIK3R3 is expressed only in testis and brain of adults (37). PIK3R1/p85a is isoform found abundantly, but in cancer patients its expression is reduced (38). It's a Tumor-Suppressor Gene and the most striking difference between p85a and p85b is that PIK3R1/p85a acts as a tumor suppressor while, PIK3R2/p85b is a cancer driver. The recent finding say that p85a subunit restrains the catalytic activity of PI3K (39) encourage testing

This is a provisional file, not the final typeset article

the consequences of reducing p85a levels. Other study tells deletion of Pik3r1 led to a gradual change in hepatocyte morphology (liver) and with over time, mice develops hepatocellular carcinoma (40). These two observations confirm the deletion of Pik3r1 gene is a cause of tumor development, as similarly observed for tumor suppressors. Pik3r1 loss in mouse accelerates HER2/neu-induced mammary cancer development, in cultured human epithelial cells PIK3R1 knockdown cause transformational changes, while hemi-zygous deletion is a frequent event in breast cancer samples (41). Mutation in PIK3R1 has been reported in breast, pancreatic, colon cancers (42)and in 8.5% cases of ovarian carcinoma (43) The PIK3 pathway is downstream regulated from RTK (receptor tyrosine kinase), that's active in cancer lineages, including endometrial cancer (EC) (44). PIK3 mutation has been reported in ovarian carcinoma but not in cases where patients suffer with endometriosis only(45) . While PI3K pathway has been found to be strongly implicated during the development of endometriosis (46).

LPAR1

Lysophosphatidic Acid, is a small phospholipid present in many mammalian cells and tissues (47). It is involved in cell migration, survival, proliferation, cellular interactions and cytoskeleton changes (48). A study on mare, suffering with endometriosis has indicated that concentration of LPA, its receptors, PGE2/PGF2 ration and CTGF secretion is altered during endometriosis (49). LPA induces IL-8 (Interleukin-8) expression via LPA-1 receptor, Gi protein, MAPK/p38 and NF-kB signalling pathway. Where IL-8 protein stimulates endometrial cell migration, permeability, capillary tube formation and proliferation leading to angiogenesis during pregnancy (50). Overexpression of all LPARs and enzymes occur responsible for LPA synthesis, during endometrial cancers, showing a positive correlation with myoinvasion and FIGO (International Federation of Gynecology and Obstetrics) stage (51). Recent study has demonstrated that LPA, acts as a mitogen and pro-invasive stimulus for endometrial and endometriotic cells acts via LPAR1 and LPAR3 (Lysophosphatidic Acid Receptors). It has demonstrated that LPA-dependent stimulus causes secretion of cathepsin B, a protease, which acts as a factor for endometriotic invasion (52).

RAB3D, is a GTPase of Rab family, plays a role as a central regulator of vesicular transport (53).A study reveals that the Rab3D is implicated in the subcellular localization and maturation of Immature Secretory granules (ISGs) (54). Rabs are involved in membrane trafficking, cellular signalling, growth and development. Rabs and their effectors has been found to be overexpressed or undergone for loss of function in many disease including cancers and causes the disease progression(55) .Endometriosis is also an inflammatory disease and similarly Rab can play its role in endometriosis disease progression and adhesion.

GPR39, are G-protein coupled receptors present in the plasma membrane, which is distinct target for binding of extracellular Zn^{+} ions (56–58). G-protein coupled receptors are a large family of seven-transmembrane proteins, which are all involved in a diverse array of extracellular stimuli (59). Zinc is an important component of enzymes and proteins. It is also required for intracellular message transmission, protein synthesis, cell membrane maintenance and transport, regulation of neuronal, endocrinal system (60). In brain GPR39 zinc sensing receptor has a significant role in Alzheimer and Epilepsy (61). Further investigations showed that GPR39 contributes its part in skin wound healing, thus having a positive role in the area of dermatology and stem cell therapy (62). A

study has revealed that GPR39 might be a direct promising target for therapy of zincergic dyshomeostasis observed in Alzheimer's disease, but more studies are still needed over it (63).

HOXC4, is actively transcribed during the development and differentiation of Lymphoid, myeloid and erythroid cells. It helps to maintain proliferation in hematopoietic cells(64).

Cystatin A (**CSTA**), is a type 1 cystatin super-family member, and is expressed mainly in epithelial and lymphoid tissues. It prevents the proteolysis of cytoplasmic and cytoskeletal proteins in cells. It acts a tumor suppressor in esophageal and lung cancer (65). The risk of disease recurrence and death are found to be higher in patients suffering with squamous cell carcinoma of head and neck, having low CSTA(66) .Here expression analysis of DEGs in endometriosis, CSTA was found to be downregulated but detailed studies for the role of CSTA in endometriosis disease progression is still has to be done.

Epithelial splicing regulatory protein 1 (**ESRP1**), plays crucial role during organogenesis of craniofacial and epidermal development, branching morphogenesis in the lungs and salivary glands. It also plays role during cancer progression, and its expression found to be low in normal epithelium but upregulated during carcinoma (67).

Conclusion

From this study, the conclusion can be derived that these genes are playing diversified roles, but these may be involved directly or indirectly in the pathophysiology of the endometriosis disease. There may be that many genes together playing their parts for the disease progression. This study will help us to understand the disease physiology in a better way but further a detailed study about it is still needed.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Institutional Ethical committee before starting the study (No. Dean/2018/EC/936).

Author contributions

RS, RCand SR conceived and designed the project. KK and AA performed all operations. RB analyzed the data and drew the figures. KK and RB wrote the manuscript. AA, KK, RB and RS revised the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgement

We want to extend our sincere gratitude to Multi-Disciplinary Research Units (MRUs) Laboratory, a grant by ICMR-Department of Health Research.

Conflict of interest

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

References

This is a provisional file, not the final typeset article

1. Parasar P, Ozcan P, Terry KL. Endometriosis: Epidemiology, Diagnosis and Clinical Management. Curr Obstet Gynecol Rep. 2017;6(1):34–41.
2. Bulun SE. Endometriosis. N Engl J Med. 2009 Jan;360(3):268–79.
3. Nnoaham KE, Hummelshoj L, Webster P, D'Hooghe T, De Cicco Nardone F, De Cicco Nardone C, et al. Impact of endometriosis on quality of life and work productivity: A multicenter study across ten countries. Fertil Steril. 2011;96(2).
4. Meuleman C, Vandenabeele B, Fieuws S, Spiessens C, Timmerman D, D'Hooghe T. High prevalence of endometriosis in infertile women with normal ovulation and normospermic partners. Fertil Steril. 2009;92(1):68–74.
5. Dai Y, Li X, Shi J, Leng J. A review of the risk factors, genetics and treatment of endometriosis in Chinese women: A comparative update. Vol. 15, Reproductive Health. 2018. p. 82.
6. Mathew AG. A Case Exemplifying Sampson's Theory of the Aetiology of Endometriosis. Aust New Zeal J Obstet Gynaecol. 1963;3(4):159–61.
7. Liu H, Lang JH. Is abnormal eutopic endometrium the cause of endometriosis? The role of eutopic endometrium in pathogenesis of endometriosis. Med Sci Monit. 2011;17(4):92–9.
8. Gupta D, Hull ML, Fraser I, Miller L, Bossuyt PMM, Johnson N, et al. Endometrial biomarkers for the non-invasive diagnosis of endometriosis. Vol. 2016, Cochrane Database of Systematic Reviews. 2016.
9. Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. Vol. 8, Wiley Interdisciplinary Reviews: RNA. 2017.
10. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Vol. 10, Nature Reviews Genetics. 2009. p. 57–63.
11. 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(SUPPL. 1):D760–5.
12. GEO2R - GEO - NCBI [Internet]. [cited 2021 Mar 13]. Available from: https://www.ncbi.nlm.nih.gov/geo/geo2r/
13. 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(W1):W566–70.
14. GeneCards - Human Genes | Gene Database | Gene Search [Internet]. [cited 2021 Mar 13]. Available from: https://www.genecards.org/
15. Venny 2.1.0 [Internet]. [cited 2021 Mar 13]. Available from:

https://bioinfogp.cnb.csic.es/tools/venny/

16. 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. 2013;41(D1).

17. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. Methods Mol Biol. 2011;696:291–303.

18. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.

19. Fishelson Z, Müller-Eberhard HJ. Regulation of the alternative pathway of human complement by C1q. Mol Immunol. 1987;24(9):987–93.

20. Müller UC, Deller T, Korte M. Not just amyloid: Physiological functions of the amyloid precursor protein family. Vol. 18, Nature Reviews Neuroscience. 2017. p. 281–98.

21. Hof PR, Bouras C, Perl DP, Sparks DL, Mehta N, Morrison JH. Age-Related Distribution of Neuropathologic Changes in the Cerebral Cortex of Patients with Down's Syndrome: Quantitative Regional Analysis and Comparison with Alzheimer's Disease. Arch Neurol. 1995;52(4):379–91.

22. Abbatemarco JR, Jones SE, Larvie M, Bekris LM, Khrestian ME, Krishnan K, et al. Amyloid Precursor Protein Variant, E665D, Associated With Unique Clinical and Biomarker Phenotype. Am J Alzheimers Dis Other Demen. 2021;36.

23. Ristori E, Cicaloni V, Salvini L, Tinti L, Tinti C, Simons M, et al. Amyloid-β Precursor Protein APP Down-Regulation Alters Actin Cytoskeleton-Interacting Proteins in Endothelial Cells. Cells. 2020;9(11).

24. Morishita R, Saga S, Kawamura N, Hashizume Y, Inagaki T, Kato K, et al. Differential localization of the γ3 and γ12 subunits of G proteins in the mammalian brain. J Neurochem. 1997;68(2):820–7.

25. Hiraiwa M, O'brien JS, Kishimoto Y, Galdzicka M, Fluharty AL, Ginns EI, et al. Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposins (sphingolipid activator proteins). Arch Biochem Biophys. 1993;304(1):110–6.

26. Motta M, Tatti M, Furlan F, Celato A, Di Fruscio G, Polo G, et al. Clinical, biochemical and molecular characterization of prosaposin deficiency. Clin Genet. 2016;90(3):220–9.

27. Spiegel R, Bach G, Sury V, Mengistu G, Meidan B, Shalev S, et al. A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: First report of saposin A deficiency in humans. Mol Genet Metab. 2005;84(2):160–6.

This is a provisional file, not the final typeset article

28. Shaimardanova AA, Chulpanova DS, Solovyeva V V., Mullagulova AI, Kitaeva K V., Allegrucci C, et al. Metachromatic Leukodystrophy: Diagnosis, Modeling, and Treatment Approaches. Vol. 7, Frontiers in Medicine. 2020. p. 576221.
29. Matsuda T, Cepko CL. Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc Natl Acad Sci U S A. 2004;101(1):16–22.
30. Wilkie TM, Scherle PA, Strathmann MP, Slepak VZ, Simon MI. Chararacterization of G-protein α subunits in the Gq class: Expression in murine tissues and in stromal and hematopoietic cell lines. Proc Natl Acad Sci U S A. 1991;88(22):10049–53.
31. D'Angelo DD, Sakata Y, Lorenz JN, Boivin GP, Walsh RA, Liggett SB, et al. Transgenic Gαq overexpression induces cardiac contractile failure in mice. Proc Natl Acad Sci U S A. 1997;94(15):8121–6.
32. Wang Y, Li Y, He Y, Sun Y, Sun W, Xie Q, et al. Expression of G protein αq Subunit is Decreased in Lymphocytes from Patients with Rheumatoid Arthritis and is Correlated with Disease Activity. Scand J Immunol. 2012;75(2):203–9.
33. Wang Y, Xiao H, Wu H, Yao C, He H, Wang C, et al. G protein subunit α q regulates gastric cancer growth via the p53/p21 and MEK/ERK pathways. Oncol Rep. 2017;37(4):1998–2006.
34. Lai W, Cai Y, Zhou J, Chen S, Qin C, Yang C, et al. Deficiency of the G protein Gαq ameliorates experimental autoimmune encephalomyelitis with impaired DC-derived IL-6 production and Th17 differentiation. Cell Mol Immunol. 2017;14(6):557–67.
35. Hirsch E, Ciraolo E, Franco I, Ghigo A, Martini M. PI3K in cancer-stroma interactions: Bad in seed and ugly in soil. Vol. 33, Oncogene. Nature Publishing Group; 2014. p. 3083–90.
36. Vallejo-Díaz J, Chagoyen M, Olazabal-Morán M, González-García A, Carrera AC. The Opposing Roles of PIK3R1/p85α and PIK3R2/p85β in Cancer. Vol. 5, Trends in Cancer. 2019. p. 233–44.
37. Vanhaesebroeck B, Vogt PK, Rommel C. PI3K: From the Bench to the Clinic and Back. In: Current topics in microbiology and immunology. 2010. p. 1–19.
38. Ueki K, Algenstaedt P, Mauvais-Jarvis F, Kahn CR. Positive and Negative Regulation of Phosphoinositide 3-Kinase-Dependent Signaling Pathways by Three Different Gene Products of the p85α Regulatory Subunit. Mol Cell Biol. 2000;20(21):8035–46.
39. Fruman DA. Regulatory Subunits of Class IA PI3K. In: Current topics in microbiology and immunology. Curr Top Microbiol Immunol; 2010. p. 225–44.
40. Ueki K, Fruman DA, Brachmann SM, Tseng Y-H, Cantley LC, Kahn CR. Molecular Balance between the Regulatory and Catalytic Subunits of Phosphoinositide 3-Kinase Regulates Cell

Signaling and Survival. Mol Cell Biol. 2002;22(3):965–77.

41. Alcázar I, Cortés I, Zaballos A, Hernandez C, Fruman DA, Barber DF, et al. P85β phosphoinositide 3-kinase regulates CD28 coreceptor function. Blood. 2009;113(14):3198–208.

42. Jaiswal BS, Janakiraman V, Kljavin NM, Chaudhuri S, Stern HM, Wang W, et al. Somatic Mutations in p85α Promote Tumorigenesis through Class IA PI3K Activation. Cancer Cell. 2009;16(6):463–74.

43. Sweeney SM, Cerami E, Baras A, Pugh TJ, Schultz N, Stricker T, et al. AACR project genie: Powering precision medicine through an international consortium. Cancer Discov. 2017;7(8):818–31.

44. Cheung LWT, Hennessy BT, Li J, Yu S, Myers AP, Djordjevic B, et al. High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discov. 2011;1(2):170–85.

45. Yamamoto S, Tsuda H, Takano M, Iwaya K, Tamai S, Matsubara O. PIK3CA mutation is an early event in the development of endometriosis-associated ovarian clear cell adenocarcinoma. J Pathol. 2011;225(2):189–94.

46. Matsumoto T, Yamazaki M, Takahashi H, Kajita S, Suzuki E, Tsuruta T, et al. Distinct β-catenin and PIK3CA mutation profiles in endometriosis-associated ovarian endometrioid and clear cell carcinomas. Am J Clin Pathol. 2015;144(3):452–63.

47. Lin ME, Herr DR, Chun J. Lysophosphatidic acid (LPA) receptors: Signaling properties and disease relevance. Vol. 91, Prostaglandins and Other Lipid Mediators. 2010. p. 130–8.

48. Moolenaar WH, Van Meeteren LA, Giepmans BNG. The ins and outs of lysophosphatidic acid signaling. Vol. 26, BioEssays. 2004. p. 870–81.

49. Szóstek-Mioduchowska A, Leciejewska N, Zelmańska B, Staszkiewicz-Chodor J, Ferreira-Dias G, Skarzynski D. Lysophosphatidic acid as a regulator of endometrial connective tissue growth factor and prostaglandin secretion during estrous cycle and endometrosis in the mare. BMC Vet Res. 2020;16(1).

50. Chen SU, Lee H, Chang DY, Chou CH, Chang CY, Chao KH, et al. Lysophosphatidic acid mediates interleukin-8 expression in human endometrial stromal cells through its receptor and nuclear factor-κB- dependent pathway: A possible role in angiogenesis of endometrium and placenta. Endocrinology. 2008;149(11):5888–96.

51. Wasniewski T, Woclawek-PotocKa I, Boruszewska D, Kowalczyk-Zieba I, Sinderewicz E, Grycmacher K. The significance of the altered expression of lysophosphatidic acid receptors,

This is a provisional file, not the final typeset article

autotaxin and phospholipase A2 as the potential biomarkers in type 1 endometrial cancer biology. Oncol Rep. 2015;34(5):2760–7.

52. Dietze R, Starzinski-Powitz A, Scheiner-Bobis G, Tinneberg HR, Meinhold-Heerlein I, Konrad L. Lysophosphatidic acid triggers cathepsin B-mediated invasiveness of human endometriotic cells. Biochim Biophys Acta - Mol Cell Biol Lipids. 2018;1863(11):1369–77.

53. Li G, Marlin MC. Rab family of GTpases. Methods Mol Biol. 2015;1298.

54. Hodneland RR, Copier E, Regazzi J. Rab3D Is Critical for Secretory Granule Maturation in PC12 Cells. PLoS One. 2013;8(3):57321.

55. Qin X, Wang J, Wang X, Liu F, Jiang B, Zhang Y. Targeting Rabs as a novel therapeutic strategy for cancer therapy. Vol. 22, Drug Discovery Today. 2017. p. 1139–47.

56. Hershfinkel M, Moran A, Grossman N, Sekler I. A zinc-sensing receptor triggers the release of intracellular Ca2+ and regulates ion transport. Proc Natl Acad Sci U S A. 2001;98(20):11749–54.

57. Sunuwar L, Gilad D, Hershfinkel M. The zinc sensing receptor, ZnR/GPR39, in health and disease. Vol. 22, Frontiers in Bioscience - Landmark. Frontiers in Bioscience; 2017. p. 1469–92.

58. Hershfinkel M. The zinc sensing receptor, ZnR/GPR39, in health and disease [Internet]. Vol. 19, International Journal of Molecular Sciences. 2018. p. 439. Available from: www.mdpi.com/journal/ijms

59. Hilger D, Masureel M, Kobilka BK. Structure and dynamics of GPCR signaling complexes. Nat Struct Mol Biol. 2018;25(1):4–12.

60. Takeda A. Movement of zinc and its functional significance in the brain. Vol. 34, Brain Research Reviews. 2000. p. 137–48.

61. Khan MZ. A possible significant role of zinc and GPR39 zinc sensing receptor in Alzheimer disease and epilepsy. Vol. 79, Biomedicine and Pharmacotherapy. 2016. p. 263–72.

62. Zhao H, Qiao J, Zhang S, Zhang H, Lei X, Wang X, et al. GPR39 marks specific cells within the sebaceous gland and contributes to skin wound healing. Sci Rep. 2015;5.

63. Rychlik M, Mlyniec K. Zinc-mediated Neurotransmission in Alzheimer’s Disease: A Potential Role of the GPR39 in Dementia. Curr Neuropharmacol. 2019;18(1):2–13.

64. Zeng J, Sun W, Chang J, Yi D, Zhu L, Zhang Y, et al. HOXC4 up-regulates NF-κB signaling and promotes the cell proliferation to drive development of human hematopoiesis, especially CD43+ cells. Blood Sci. 2020;2(4):117–28.

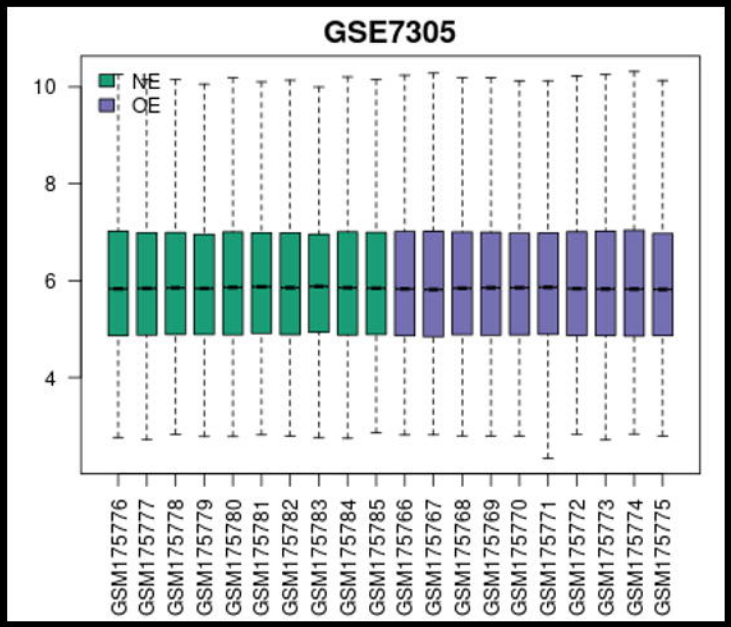
65. Ma Y, Chen Y, Li Y, Grün K, Berndt A, Zhou Z, et al. Cystatin A suppresses tumor cell

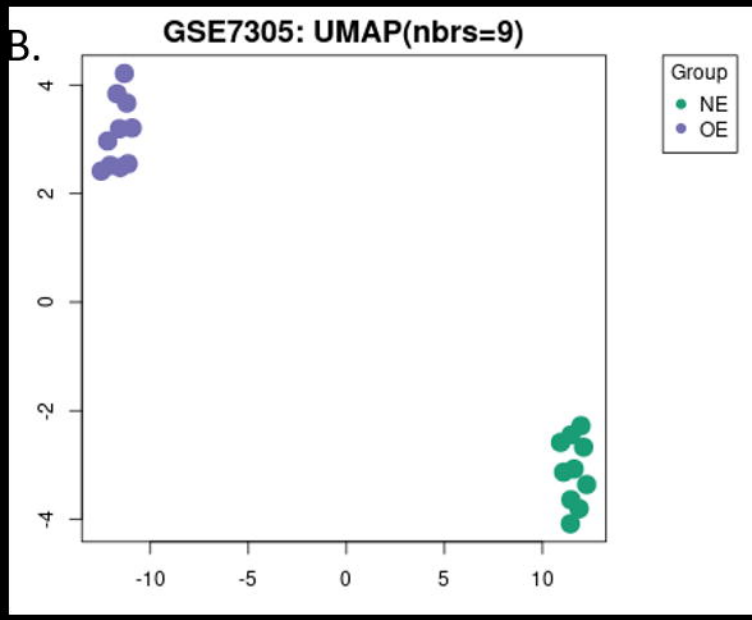
growth through inhibiting epithelial to mesenchymal transition in human lung cancer. Oncotarget. 2018;9(18):14084–98.

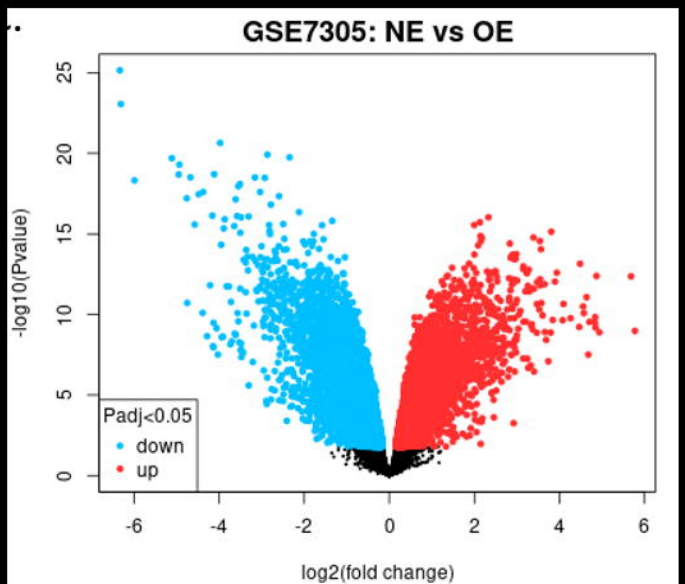
66. Budihna M, Strojan P, Èmid L, Èkrk J, Vrhovec I, Zupevc A. Prognostic value of cathepsins B, H, L, D and their endogenous inhibitors stefins A and B in head and neck carcinoma. Biol Chem Hoppe Seyler. 1996;377(6):385–90.

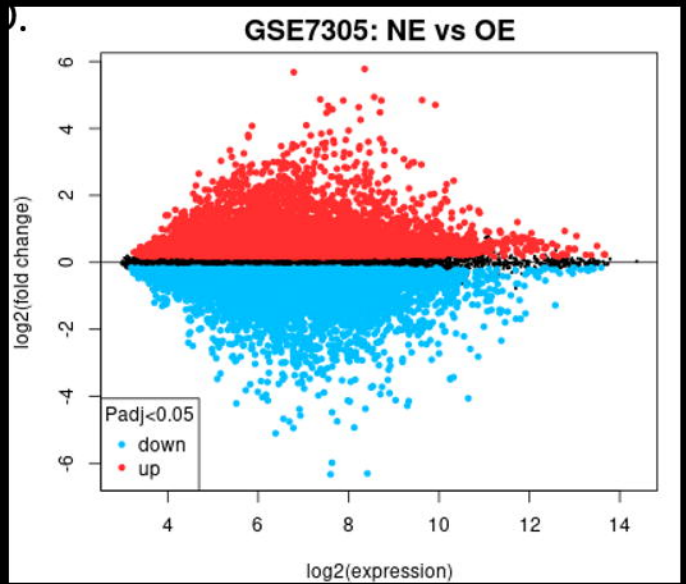
67. Hayakawa A, Saitoh M, Miyazawa K. Dual roles for epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) in cancer progression. In: Advances in Experimental Medicine and Biology [Internet]. Springer New York LLC; 2017 [cited 2021 Mar 13]. p. 33–40. Available from: https://pubmed.ncbi.nlm.nih.gov/27401076/

This is a provisional file, not the final typeset article

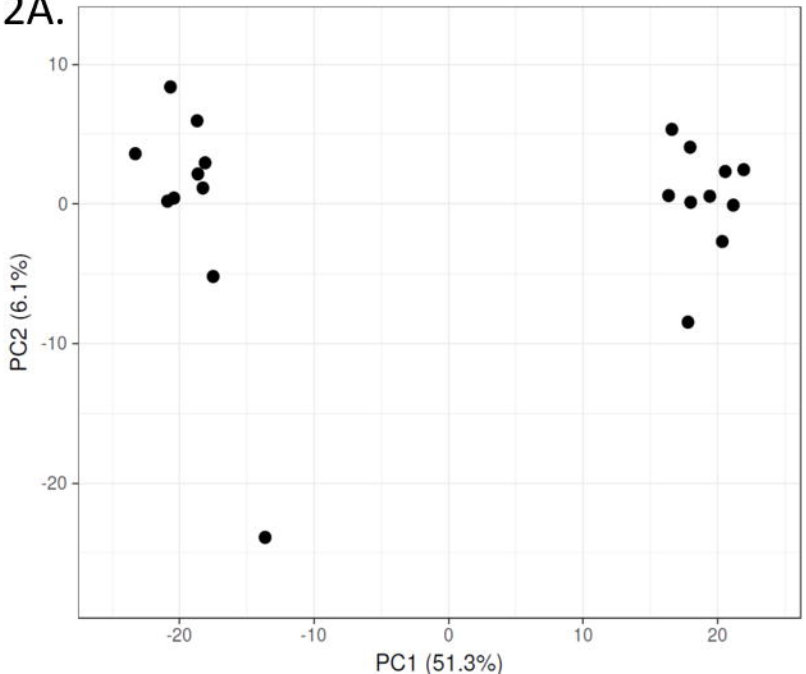

GSE7305
NE
OE
10
8
6
4
GSM175776
GSM175777
GSM175778
GSM175779
GSM175780
GSM175781
GSM175782
GSM175783
GSM175784
GSM175785
GSM175766
GSM175767
GSM175768
GSM175769
GSM175770
GSM175771
GSM175772
GSM175773
GSM175774
GSM175775

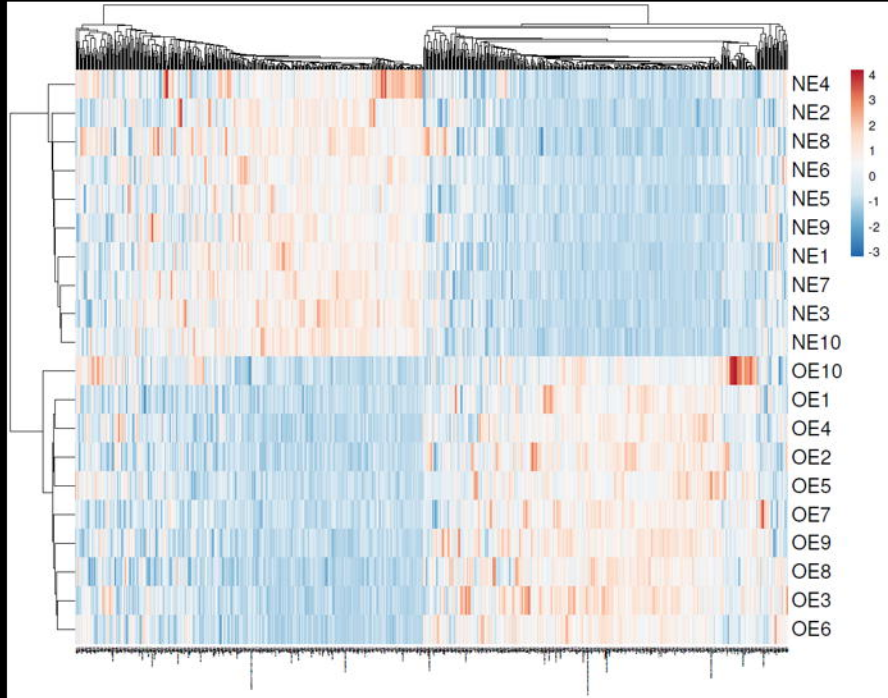

B.
GSE7305: UMAP(nbrs=9)
Group
NE
OE
4
2
0
-2
-4
-10
-5
0
5
10


GSE7305: NE vs OE
-log10(Pvalue)
25
20
15
10
5
0
Padj<0.05
down
up
-6
-4
-2
0
2
4
6
log2(fold change)

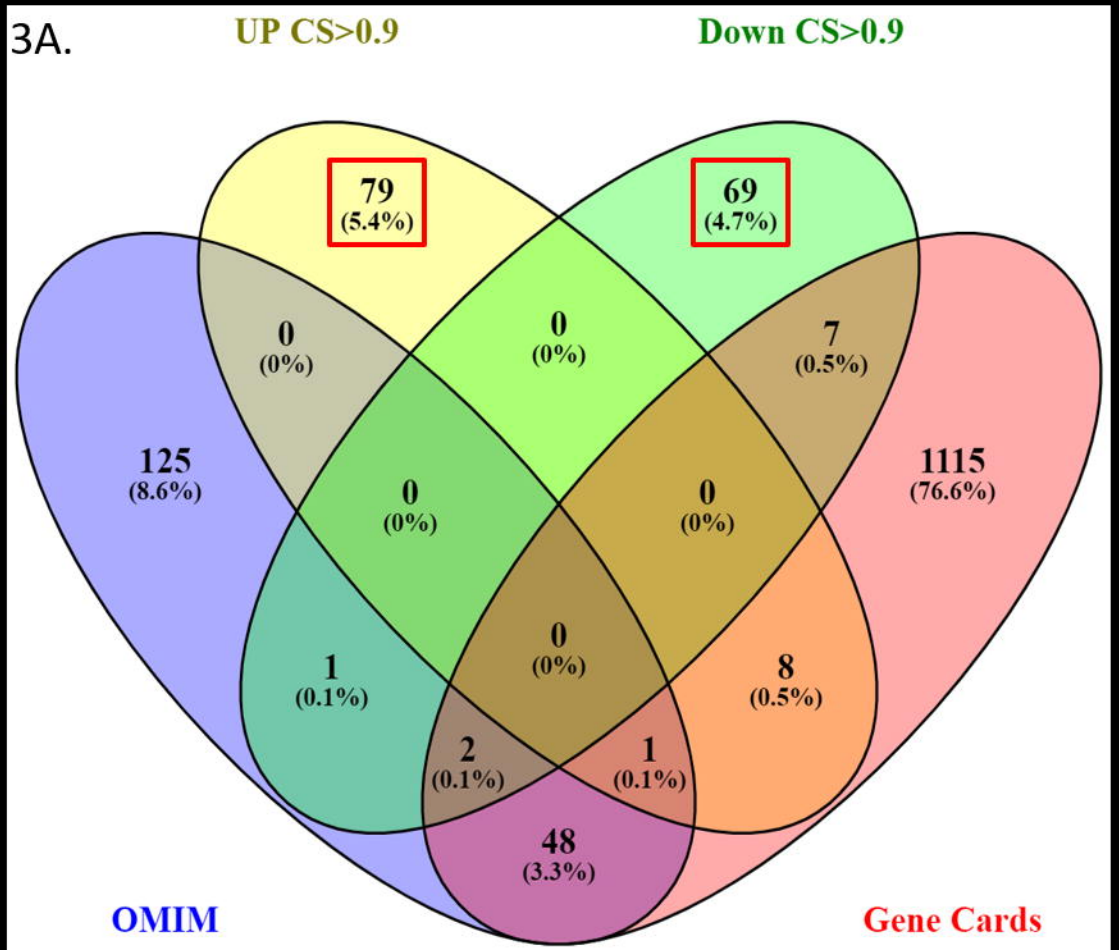

GSE7305: NE vs OE
log2(fold change)
6
4
2
0
-2
-4
-6
Padj<0.05
down
up
4
6
8
10
12
14
log2(expression)

2A.

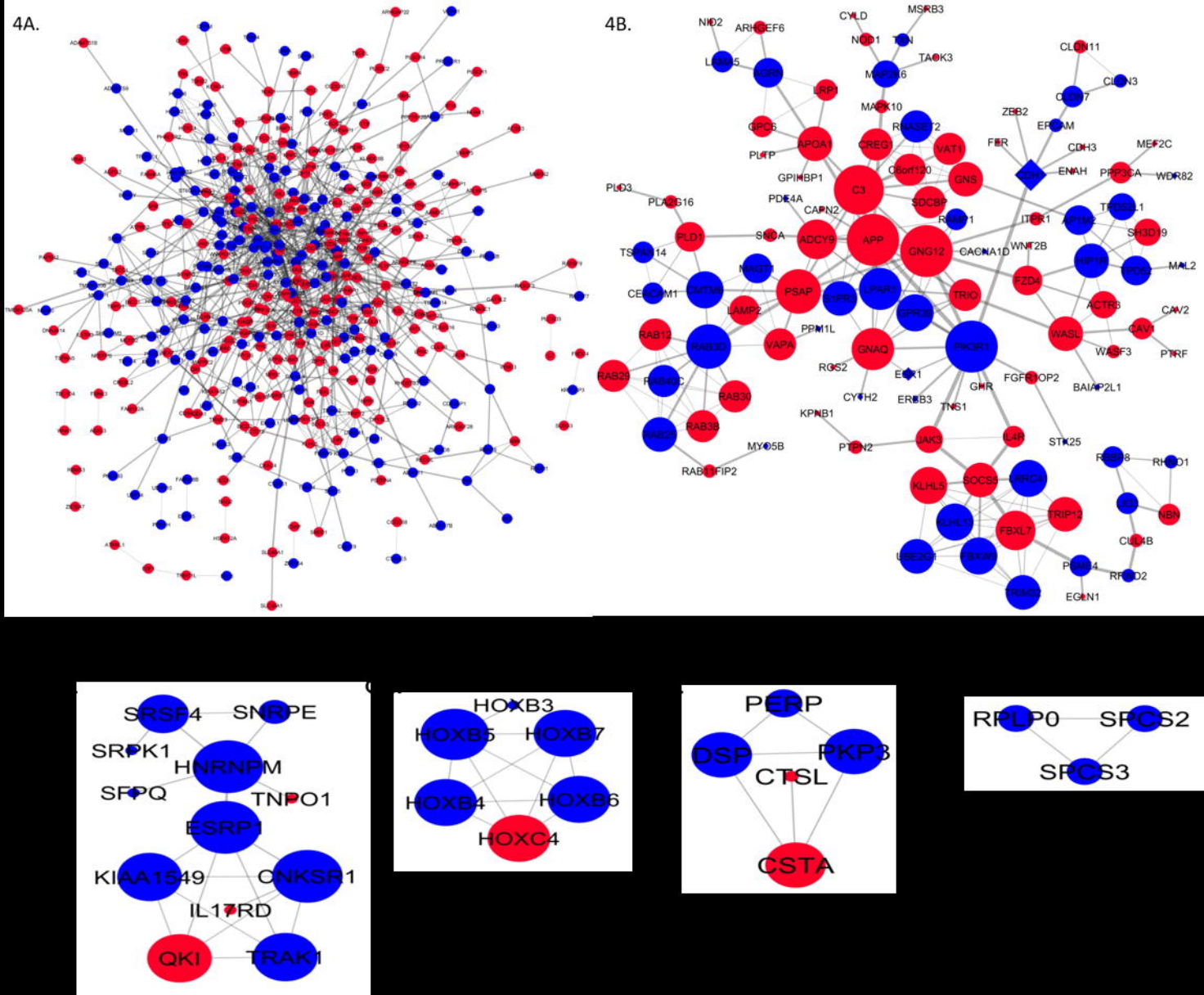
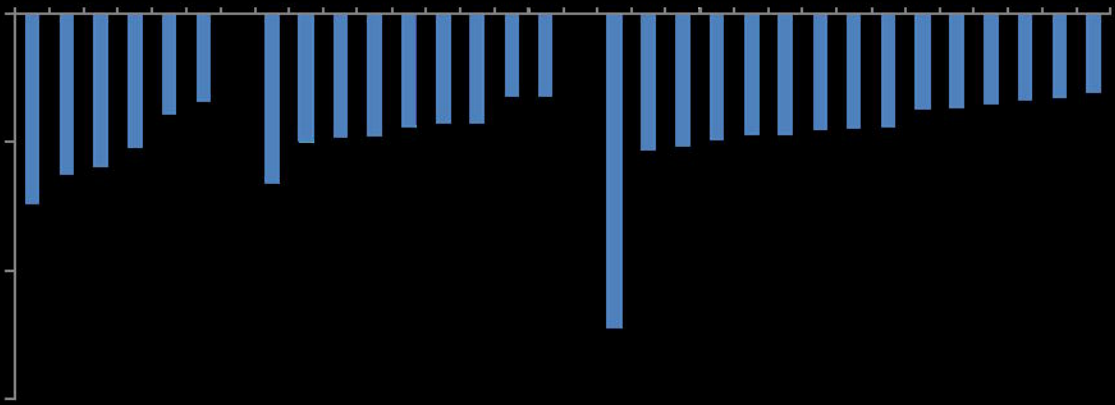
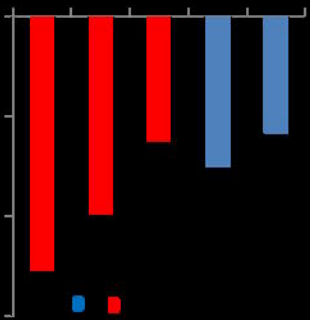
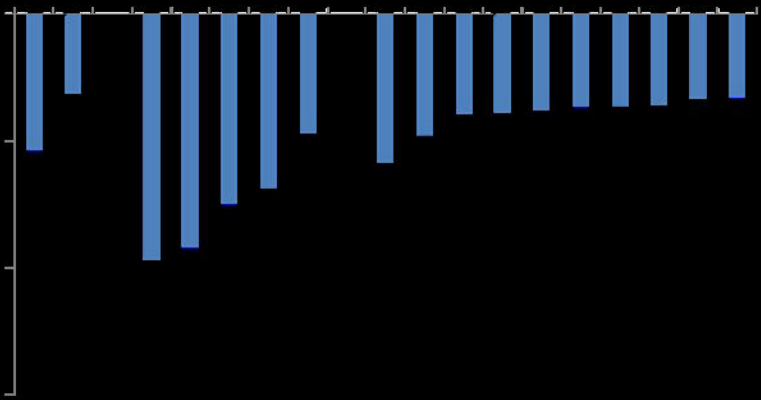

10
0
-10
-20
PC2 (6.1%)
-20
-10
0
10
20
PC1 (51.3%)


NE4
NE2
NE8
NE6
NE5
NE9
NE1
NE7
NE3
NE10
OE10
OE1
OE4
OE2
OE5
OE7
OE9
OE8
OE3
OE6
4
3
2
1
0
-1
-2
-3

3A.



DEGs	Number
Total DEGs	552
UP DEGs	312
Down DEGs	240
UP DEGs with cs>0.9	88
Down DEGs with cs>0.9	79
Reported DEGs	19
Total Novel DEGs	148
Novel UP DEGs	79
Novel Down DEGs	69


4A.
4B.
SRSF4
SNRPE
SRPK1
HNRNPM
SFPQ
TNPO1
ESRP1
KIAA1549
CNKSR1
IL17RD
QKI
TRAK1
HOXB3
HOXB5
HOXB7
HOXB4
HOXB6
HOXC4
PERP
DSP
PKP3
CTSL
CSTA
RPLP0
SPCS2
SPCS3