Microarray Profiling in Fibrogenesis Imperfecta Ossium

Vandana Dhiman
Post Graduate Institute of Medical Education and Research
Samarth Saini
University of Minnesota
Sanjay Bhadada (bhadadask@rediffmail.com)
Post Graduate Institute of Medical Education and Research
Manoj Bhasin
Emory School of Medicine Atlanta
Sudhaker Rao
Henry Ford Health System

Research Article

Keywords: Fibrogenesis Imperfecta Ossium, Microarray, Osteomalacia

Posted Date: January 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-138783/v1

Microarray Profiling in Fibrogenesis Imperfecta Ossium

Vandana Dhiman1, Samarth Saini 2, Sanjay Kumar Bhadada1, Manoj K. Bhasin 3, Sudhaker D. Rao4,

1Department of Endocrinology, Postgraduate Institute of Medical Education and Research, Chandigarh, India; 2 Department of Neuroscience ,University of Minnesota, Minnesota,USA; 3 Emory School of Medicine, Atlanta, USA, ;4Bone & Mineral Research Laboratory, Henry Ford Health System, Detroit, Michigan, USA

Abbreviated title: Microarray of FIO

Keywords: Fibrogenesis Imperfecta Ossium, Microarray, Osteomalacia

Type of Article: Original Article

Word Counts: Abstract: 110; Manuscript Text: 3552

Corresponding Address for correspondence:

1 **Dr. Sanjay Kumar Bhadada**
Professor of Endocrinology
PGIMER, Chandigarh, India
E-mail: bhadadask@rediffmail.com
Phone: +919876602448

2 **Manoj K. Bhasin, Ph.D.**
Associate Professor Of Biomedical Informatics,
Emory School of Medicine, Atlanta, GA 30322
Phone: 404-712-9849
E-mail: manoj.bhasin@emory.edu

Abstract

Fibrogenesis imperfecta ossium (FIO) is a rare, metabolic bone disease clinically characterized by generalized bone pain and fragility fractures in both the axial and appendicular skeleton. The etiology of the disease is unknown and the pathogenesis is poorly understood. In this disorder, collagen arrangement in the bone is disorganized , and causes abnormalities in the organic bone matrix. Genetic cues causing this defect in collagen arrangement are largely undetermined. Microarray analysis showed remarkable changes in gene expression , when we compared gene expression of FIO with the control. Differential gene expression analysis revealed 50 gene signatures, which can be used to organize novel ways to diagnose and treat FIO patients.

Introduction

Fibrogenesis Imperfecta Ossium is a progressively crippling skeletal disorder that is adult-onset and usually fatal. Only 29 total cases of FIO have been reported since it was first described in 1950 by Baker and Turnbull[1]. It's etiology is unknown but presumed to be genetic, as there have been multiple reported cases of FIO presenting itself in siblings as well as parents and their children. FIO manifests as generalized bone pain and fragility fractures, which can eventually cause skeletal deformities. These bone fractures happen under minimal stress or very minor traumas , and usually occur in weight-bearing and tendon insertion sites. It should be noted that the features of FIO are present everywhere except the skull and the teeth. Ultimately, those affected by FIO become incapable of doing daily activity of life, unable to stand, and are bedridden.

The pathogenesis of FIO is mostly understood, and can be pinpointed to the arrangement of collagen fibers. The primary abnormality in FIO is impaired bone mineralization due to defects in collagen arrangement. Bone is a dynamic connective tissue, this dynamicity is maintained by

osteoblast and osteoclast cells**.** Osteoblasts are in charge of laying down the collagen fibers in the bone matrix, but it is osteocalcin (OCN), osteonectin (ON) and osteopontin (OPN) that determine how the fibers are arranged. Currently, it cannot be said if this defect is caused by impaired osteoblasts or abnormalities in OCN, ON or OPN. Many diseases are similarly affected by impaired function of these cell types, such as osteogenesis imperfecta, osteopetrosis, osteoporosis, Paget's disease, fibrous dysplasia, osteomalacia and many more. The key difference between FIO and other osteomalacic conditions is the lack of the birefringence characteristic of normal, oriented collagen fibers. Transmission electron microscopy shows that bones of FIO patients are made up of abnormal collagen fibrils that are often curved and of variable size[41]. It is likely this function that enables fractures to occur in even minor trauma and weight bearing conditions.

The paucity of literature about the disease and low index of suspicion can also make difficult obtaining an accurate diagnosis. FIO is often misdiagnosed as osteomalacia or Paget's disease because the symptoms of the diseases are very similar, and all three occur in conjunction with increased levels of alkaline phosphate[42]. Usually, diagnosis is confirmed after a bone biopsy shows the reduced birefringence in the collagen fibers. Treatment for FIO is largely ineffective and a complete cure has not been found, mainly due to the scarcity of patients available for randomized trials. Since bone mineralization is defective, treatment usually includes vitamin D, metabolites, calcitonin, bisphosphonates, and corticosteroids, usually to little degrees of success. Recently, growth hormone (GH) has been found to be more effective in treating two brothers with FIO[43-45]. They have shown improvements in osteoid thickness and normalization of collagen arrangement, but long term benefits have yet to be determined.

The primary objective of this research is to fill in the lack of information about FIO gene data in the current time. Microarray data from this study identifies differentially expressed genes between FIO patients and healthy humans. Finding and indexing these genes may open the door to development of more effective novel therapies for FIO. In addition, this study may also provide insight into the importance of gene regulation in collagen, which is the main organic component of bone matrix responsible for skeletal integrity.

Materials and Methods:

Patients. *Ethics statement*. The research design was approved by the Institutional Ethics committee, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh India. Informed consent was obtained from all study participants. The study was carried out in accordance with the approved guidelines.

Patients and Clinical sample collection FIO patient were enrolled in this study. The diagnosis of FIO was established with the help of clinical feature, supported by radiological (Looser's zone) and Tc ^{99} MDP bone scan (super scan with absence of uptake in skull also called 'beheaded scan') and osteomalacia was confirmed by bone histomorphometry (thickened osteoid seam) with loss of birefringence on polarized light microscopy. Total RNA was isolated from Trans-Iliac Bone biopsy (Control Individual (n=1) & Patient (n=1)) using TRIZOL method and global microarray was performed on Affymetrix Genechip HumanPrimview platform and sample descriptions are provided in Table 1(a). The array was performed on duplicate.

Datasets

This study uses data from three datasets: GSE43861, GSE58474, and GSE30159. The primary dataset, GSE43861, contains data from a study conducted by department of Endocrinology comparing two FIO patients and two control samples. GSE58474 contains data comparing

mandibular and iliac bone derived cells. Both GSE58474 and GSE30159 were taken from the Gene Expression Omnibus (GEO) database and are used for co-expression network building. The workflow for the bioinformatics analysis in this paper is illustrated in (Figure 1).

Data Quality Control Analysis

A PCA (Principal Component Analysis) was performed on GSE43861 to reduce dimensionality and bring out patterns in the dataset as shown in (Figure 2). PCA was also done on the publicly available datasets GSE72490, GSE30159, and GSE 58474 for dimensionality reduction and quality control of samples within the study. If the quality of the sample is not suitable for the study based on the distribution of PC (Principal Component), it was excluded from the subsequent analysis. Details of each microarray study, including sample descriptions are provided in (Table 2).

Batch Effect Adjustment

The primary goal of this study is to identify differentially expressed genes between normal control and FIO patients. However, data integration to increase the sample size by including the other bone related studies was hindered by batch effects. The pre-processed and normalized datasets were further subjected to ComBat in order to mitigate the batch effects and other undesired variation. To compare the sample clustering patterns, the results were visually examined using the principal component analysis (PCA). Multidimensional scaling of the datasets revealed that before application of the batch adjustment algorithm, each dataset clearly separated from all the others ("batch effect"), whereas when batch adjustment was applied, samples from each dataset were well intermixed.

Identification and Selection of Eligible Gene Expression Datasets for Meta-Analysis

The selection method of an eligible gene signature for FIO is also shown above in (Figure 2). After the microarray analysis of GSE43861 and two other publicly available datasets (GSE30159, GSE 58474), a co-expression network was built by utilizing a WGCNA package. Only interactions with a p-value < 0.05 were included into the expression network. Next, a clustering method was applied to the co expression network in order to identify the subnetwork with maximum connectivity. Cystoscope 3.0.1 was used as the platform for network analysis. The cluster was analyzed with the cystoscope package j Active Modules, which finds subnetworks based on expression activated subnetworks. jActiveModules also provides a ranking co-efficient for each subnetwork. The subnetwork (cluster) with the highest co-efficient was used for our biomarker selection. Subsequently, the CytoNCA cystoscope tool was employed to rank the genes within the selected subnetwork and the genes were ranked based on degree, from highest to lowest.

Microarray Meta-Analysis

To identify the Differentially Expressed Genes (DEGs), GEO2R was used to download the data and group the samples. Once samples are grouped and normalized, expression values were used for the identification of DEGs with the limma package of R based on the thresholds of fold change >2 or <0.5 and FDR adjusted p-value < 0.05. The heatmap of clustering genes were obtained using the heatmap 2 R function.

Functional Gene Set Enrichment Analysis of Shared Differentially Expressed Genes (DEGs)

For the exploration of processes involved in the development of FIO, a functional enrichment analysis was also conducted for DEGs after clustering through ShinyGO v0.51 with the

thresholds of p-value < 0.05. To obtain the heatmap enrichment terms and genes, GenCLiP2 was applied.

Results

Microarray data analysis

During the differential gene expression analysis, 6133 out of the 30,218 genes were shown to have statistically significant (adj. P-value cutoff< 0.05) differential expression, whereas 2220 genes out of the 35,307 genes show statistically significant (adj. P-value cutoff< 0.05) differential expression. To obtain a substantial gene list for analysis, the differentially expressed genes were further filtered by taking a fold change $\geq$ 1.5. Only 325 genes were expressed with a fold change $\geq$1.5 (311 down-regulated and 14 up-regulated genes; Table 1 (b and c). The gene list related to growth hormones was also extracted, as shown in Table 1(d). To facilitate the co-expression network, we also used other publicly available datasets related to bone disease GSE58474 and GSE30159.All the raw and normalized data were deposited at NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number: GSE43861 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?Acc =GSE43861.)

Co-expression Network analysis

To facilitate the list of genes, we built a co-expression network by using only the expression matrix from GSE43861. However, the co-expression network retrieved through this dataset yielded a disjoint network. Therefore, we pooled the expression matrix from GSE43861 with other publicly available datasets as shown in GSE58474 and GSE30159. These datasets contain information from studies with similar skeletal diseases. The gene list from Table 1 (b and c) which shows differentially expressed genes had 266 overlapping genes out of 325 from

GSE43861. Then, we built a co-expression network with p-value 0.05 for each interaction as shown in (Figure 3).

In order to rank the 248 genes based on their connectivity with co-expressed genes, we calculated centrality measures for the network (Figure 3). The outcome of the centrality measure is shown in (Table S1), where the genes are ranked based on degree of centrality. Furthermore, we applied the clustering method on the network as shown in (Figure 3A) in order to reduce the dimensionality of network. Then, we used jActiveModules, a clustering method which clusters the nodes based on the significance in expression levels. After clustering, we selected the cluster with maximum connectivity score as shown in (Figure3B &Table3) and (Figure3C). The maximum connectivity score in (Figure3C) is 7.159 as compared to another larger cluster Module_0_3 shown in (Figure S3C).

Thus, 50 genes were selected to be studied based on the clustering analysis. The ranked list of 50 genes can be seen in (Table 4). In this table, genes are ranked based on the degree of connectivity (Figure3C). Figure 4 shows the heatmap of the 50 genes in our study from GSE43861. The control samples are grouped as green and FIO patient samples are grouped as blue.

Gene Set Enrichment Analysis for Identification of Biological Pathways Gene Ontology Terms

To identify the overrepresented biological pathways and gene ontology terms associated with the 50 differentially expressed genes, we performed gene set enrichment analysis using the Shiny Go v0.50 tool by inputting the list of 50 DEGs. The outcome of this analysis is shown in (TableS2). Gene ontology terms and biological pathways were considered significantly overrepresented if they showed an adjusted p-value <0.05 for GO (BP, CC, MP), Reactome, KEGG, or MSigDB.

The results from (Table S2) and (Figure 5) shows that the top enriched pathways are in BP (Biological Process): “Cellular catabolic process ","Catabolic process ","Negative regulation of metabolic process ","Macromolecule catabolic process ", and "Cellular macromolecule catabolic process" where enriched genes are more than 15 out of 50 genes. For the cellular component: “Non-membrane-bounded organelle ","Intracellular non-membrane-bounded organelle ","Vesicle ","Cytoskeleton ","Extracellular organelle ","Extracellular exosome ", and "Extracellular vesicle” where the minimum enriched genes are 12. For MP (molecular process): “Nucleic acid binding ","Metal ion binding ","RNAbinding ", and “Structural molecule activity” are the top enriched terms with minimum genes greater than 9. Pathways from Reactome and MSigDB (curated) are “Reactome: R-HSA-168256 Immune System","Reactome:R-HSA-162582 Signal Transduction","Reactome:R-HSA-392499 Metabolism of proteins","Reactome:R-HSA-168256 Immune System","Reactome:R-HSA-162582 Signal Transduction", and "Reactome:R-HSA-392499 Metabolism of proteins" for mapped genes more than 10 for each pathways out of the 50 genes. We also looked into Hallmark MSigDB and KEGG and obtained “MSigDB : Hallmark MYC Targets v1"," MSigDB: Hallmark Reactive Oxygen Species Pathway", " MSigDB: Hallmark epithelial mesenchymal transition", "Carbon metabolism", "Ribosome", "Autophagy", "RNA transport”, “Regulation of actin cytoskeleton”, “Huntington disease"," Bacterial invasion of epithelial cells", and "Fc gamma R-mediated phagocytosis" pathways where a minimum of 2 genes are mapped. The function of each gene is using Gene Clip 2.0**.** However, 5 genes—ATP5G2, MRPL23, SPOP, POLR2L, and SAP30BP—did not map to any functions.

Discussion

The current study centers around two patients of ages 35 and 40 who are currently living with FIO. We carried out a microarray study on the mRNA extracted from patients. Differential expression analysis and other network biology methods were employed to establish 50 differentially expressed gene signatures. Our results show a significant number of differentially expressed genes in the patient cells. The list of genes including rankings is available in (Table 4) along with (Table S2) and the most significant an Ingenuity Pathway Analysis (IPA) network diagram Figure 5a**.** The red circles in (Figure 5a) shows the down regulated genes whereas the blue shows up regulated genes. The retrieved genes, such as ATG4B, ATG3 [Table S2: KEGG], are associated with autophagy. Autophagy is closely associated with processes such as cellular catabolic and metabolic processes[3]. In our enrichment from (Table S2: GO BP), the enriched genes and biological processes are also associated with catabolic processes. There are about 15 genes (PGK1, FBXL5, SPOP, CAT, ATG3, PRDX2, ATG4B, DDX5 TCIRG1, CSDE1, VIM, EIF4G1, CLTA, RPS23, RPS9) that are associated with cellular catabolic processes. With this in mind, it has been studied in previous works that autophagy have a known role in multiple osteoblast functions [4]. In the current approach, due to the autophagy in macromolecules and organelles also reported in our results (Table S2: GO CC, Non-membrane-bounded organelle and Intracellular non-membrane-bounded organelle-23 genes: RPS23, RPS9, MRPL28, YY1,TPM4, POLR2L, MRPL23, AIF1, VIM, LIMCH1, DDX5, GNAI2, SAP30BP, CCDC85B,MYL12A, COTL1, CLTA, FGD3, TIAL1, WASF2, SF1, EIF4G1, LRRFIP1), these genes are degraded and have demonstrated their crucial role in bone and muscle cell homeostasis. Any kind of abnormality in autophagy leads to osteoporosis and sarcopenia[3]. Furthermore, the genes from (TableS2: GO MP Metal ion binding: SPARC, POLR2L, AIF1, FBXL5, EGR1, LIMCH1, RSBN1,YY1, MYL12A, PHC1, GNAI2, CAT, EIF2S2, FGD3, TPM4, SF1, SDHD) control

metal ions that play several major roles in proteins such as structural, regulatory, and enzymatic. For instance, the gene SPARC relates to osteonectin (ON), which is secreted from osteoblasts and bind calcium. More examples include AIF1, a calcium binding gene associated with synovial fluid (SF) derived from patients with either RA or osteoarthritis(OA) [5]. AIF1 is also reported to have a role in bone destruction as well as synovitis[5] In addition, the gene FBXL5 limits oxidative stress maintained in hematopoietic stem cells [6], the CAT gene variants are known to have a role in acatalasemia, and POLR2L, EGR1,YY1, and SF1 are all zinc binding genes. The results clearly show that the 50genes retrieved by our disease (FIO) have role in bone deformity and Ca+, Zn+, and metal ion enrichment. The overview in terms of gene functions of 45 of the 50 genes is shown in Figure 5(b). As per (Table 4), DDX5 is a top ranked gene on the list based on higher degree. DDX5 is upregulated in FIO patients and has a distinct association with Myotonic Dystrophy (C0027126: Myotonic Dystrophy, DisGeNET Curated, P-value: 8.02E-03). Myotonic Dystrophy is associated with increased risk of bone fragility and progressive muscle weakness on bone strength[7]. Similarly, the SPOP gene, known as a POZ domain gene is a regulator of skeletal development, and may lead to shorter digit bones and lower bone density[8]. SDHD has an association with abnormality of limb bone (HP:0040068 (HPO_PHENOTYPE_GENE)), SAP30BP [9], and EIF2S2, EIF4G1,and EIF4A2 all belong to a eukaryotic translation initiation factor gene family that are known to have role in apoptosis in osteosarcoma cells. Osteosarcoma is a type of cancer that produces immature and weak bones[10]. Similarly, the SRSF3 and FBXL5 genes are also associated with osteosarcoma, leading to under developed bones[20,21]. The CAT gene is related to catalase activities, and has a defined role in bone mineral density [11]. LRRFIP1 is a repressor gene and when it bounds with non-coding RNAs of TNF genes, it acts as stimulator in bone resorption.[12] COX2 is also known as Cyclo-Oxygenase 2, and is an essential

marker for bone fracture healing when upregulated, which is analogous to the case for our FIO patients[13]. WASF2 expression is related to bone formation in animals and humans with intermittent PTH administration via the WNT and phosphoinositide-dependent protein kinase signaling pathways[14]. YY1 is accurately known to have functions as growth factors and could generate bone morphogenetic protein (BMP) in synergy with SMAD1 and SMAD2 genes[15]. FAM134C is known to have associated with distal sensory loss in lower limbs[16]. In our FIO patients, TPM4 is down regulated and it has known effects on bone formation, for example, upregulation of TPM4 has already been reported to have higher expression of untreated osteoblasts[17]. Radial spoke proteins RSP23 and RSP9 are associated with controlling dynein (family of cytoskeletal motor proteins) activity and flagellar motility[18], which might disturb motor-specific functions in osteoclast formation and bone resorption[19]. The RSBN1 gene is downregulated in patients known to have marginal association in juvenile idiopathic arthritis (JIA)[22]. TMEM123 has been reported to have role in oncosis, a condition refers to have series of cellular reactions following injury that precedes cell death [23,24].TIAL1 and POLR2L are genes that are enriched with "Reactome:R-HSA-6803529 FGFR2 alternative splicing" as shown in (Table S2: REACTOME), which was previously reported to have role in bone formation[25]. SEC63 as shown in (Table S2: GO-MF) is enriched with metal ion binding. Metal ion binding is associated with bone health through various metal metabolisms. The ATG4B, ATG3, and CSDE1 genes have not been reported for any bone related diseases, however, as shown in (Table S2: GO-KEGG),these genes are enriched with autophagy and multiple studies have proved the role of autophagy in bone deformity as discussed earlier in this section.CSDE1, ATP5MC2 (ATP5G2) and PRDX2 are all enriched in "negative regulation of lipid localization" which is associated with energy metabolism[26] as shown in Figure 5(b). SF1, SRSF3, POLR2L, DDX5 as

shown in Figure 5(b) enriched through “Formation of the Spliceosomal E complex”.The spliceosome is an essential step in gene expression where it combines snRNAs and it removes introns from transcribed pre-mRNA.While that may be true, the role of these spliceosomal genes are poorly studied. However, the presence of a few spliceosomes indicates its position in autosomal recessive developmental defects[27]. The CHID1, CLTA, TCIRG1, VIM, FCGR1A, COTL1, and PGK1 genes have been linked with “Vesicle, Cytoskeleton, Extracellular organelle, Extracellular exosome, and extracellular vesicles” as shown in (Table S2: GO-CC). The extracellular vesicles and associated pathways have a recognized role in osteoblast defects, osteoclast development and other bone related diseases[28]. LIMCH1 -- an actin stress fiber-associated protein -- and PCH1are downregulated in samples in this study in both the patients. LIMCH1 and PHC1 both are acknowledged to have to role in skeletal muscle[29], and PHC1 also has a function in osteoporosis, which is characterized by low bone mass and an increased risk of fracture.[30] MRPL23 and MRPL28 are mitochondrial ribosomal proteins that are reported to play a role in motor dysfunction and bone disorders[31,32]. MYL12A has a significant function in muscle accretion[33], and FGD3 is associated with bone-enlargement. TRPT1 is a downregulated gene in this study and is enriched mostly with tRNA processing, splicing, via endonucleolytic cleavage and ligation (GO: 0006388, 0008033). Although TRPT1 is not associated with bone disease, it is highly enriched with tRNA metabolic pathways which have been reported to be present in related diseases such as Paget's disease [34]. It was also found that INTS1 is upregulated in FIO patients, and upregulation in INTS1 known to have a role in the bone disposition process in cell line studies[35]. LST1 doesn’t have any direct relation with bone diseases, however, it has been correlated with autoimmune diseases such as rheumatoid arthritis[36]. GNAI2 is another downregulated gene in both the patients and it has been reported earlier to have role in

development of synovitis in osteoarthritis[37]. EGR1 is a growth response gene that has been revealed to associate with early-life loss of the bone and mechanical deformities in the bone[38]. EGR1 could be one of the genes responsible for such rapid decay and deformity. Lastly, the CCDC85B gene is the part of slow-skeletal troponin I (Troponin Subunit) that is known to be a protein complex that calcium binds to, thereby triggering the production of muscular force and required skeleton growth[39,40].

Our study has certain limitations like small sample size, although disease is extremely rare and the samples were selected from one platform. This may result in a high rate of false positive results.

In summary, our study provides a comprehensive analysis of DEGs, biological processes terms, clustering genes, and pathways, which are associated with FIO. Most of the differentially expressed genes were found to be related to deficiency in bone cells and mineralization, and can be used to explain the symptoms of FIO patients. These results could facilitate an improvement in our comprehension of the underlying molecular mechanisms of FIO, and could stimulate novel treatments for FIO. The genetic and functional analyses suggested that FIO is associated with defects in osteoblast maturation, collagen fibril arrangement, matrix organization and bone mineralization. The 50 gene biomarker, and their related biological process terms and pathways, such as autophagy, and other signaling pathways, may represent potential targets for FIO treatment and diagnosis. Additional experimental and genetic studies with a larger sample size are required to confirm our results.

Reference

1. Oliver, W. C. & Pharr, G. M. An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mater. Res.***7**, 1564–1583 (1992).

2. Das, R. *et al.* Biomechanical characterization of spider webs. *J. Mech. Behav. Biomed. Mater.***67**, 101–109 (2017).

3. Valenti, M. T., Dalle Carbonare, L. & Mottes, M. Role of autophagy in bone and muscle biology. *World J. Stem Cells***8**, 396–398 (2016).

4. Morello, R. & Esposito, P. W. Osteogenesis Imperfecta. *Osteogenes. Imperfecta* 30

5. Kimura, M. *et al.* A Critical Role for Allograft Inflammatory Factor-1 in the Pathogenesis of Rheumatoid Arthritis. *J. Immunol.***178**, 3316–3322 (2007).

6. Muto, Y., Nishiyama, M., Nita, A., Moroishi, T. & Nakayama, K. I. Essential role of FBXL5-mediated cellular iron homeostasis in maintenance of hematopoietic stem cells. *Nat. Commun.***8**, 16114 (2017).

7. Ward, L. M., Hadjiyannakis, S., McMillan, H. J., Noritz, G. & Weber, D. R. Bone Health and Osteoporosis Management of the Patient With Duchenne Muscular Dystrophy. *Pediatrics***142**, S34–S42 (2018).

8. Cai, H. & Liu, A. Spop promotes skeletal development and homeostasis by positively regulating Ihh signaling. *Proc. Natl. Acad. Sci. U. S. A.***113**, 14751–14756 (2016).

9. Tsuchie, H. *et al.* Prognosis of Primary Osteosarcoma in Elderly Patients: A Comparison between Young and Elderly Patients. *Med. Princ. Pract. Int. J. Kuwait Univ. Health Sci. Cent.* (2019). doi:10.1159/000500404

10. Choi, Y. J., Lee, Y. S., Lee, H. W., Shim, D. M. & Seo, S. W. Silencing of translation initiation factor eIF3b promotes apoptosis in osteosarcoma cells. *Bone Jt. Res.***6**, 186–193 (2017).

11. Oh, B. *et al.* Associations of catalase gene polymorphisms with bone mineral density and bone turnover markers in postmenopausal women. *J. Med. Genet.***44**, e62 (2007).

12. Shi, L. *et al.* Non-coding RNAs and LRRFIP1 Regulate TNF Expression. *J. Immunol. Baltim. Md 1950***192**, 3057–3067 (2014).

13. Simon, A. M., Manigrasso, M. B. & O'Connor, J. P. Cyclo-oxygenase 2 function is essential for bone fracture healing. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.***17**, 963–976 (2002).

14. Uyama, M., Kawanami, M. & Tamura, M. Wasf2: A novel target of intermittent parathyroid hormone administration. *Int. J. Mol. Med.***31**, 1243–1247 (2013).

15. SMAD-mediated modulation of YY1 activity regulates the BMP response and cardiac-specific expression of a GATA4/5/6-dependent chick Nkx2.5 enhancer. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/15329343. (Accessed: 18th April 2019)

16. Kurth, I. *et al.* Mutations in FAM134B, encoding a newly identified Golgi protein, cause severe sensory and autonomic neuropathy. *Nat. Genet.***41**, 1179–1181 (2009).

17. Carpi, D. *et al.* Dioxin-Sensitive Proteins in Differentiating Osteoblasts: Effects on Bone Formation In Vitro. *Toxicol. Sci.***108**, 330–343 (2009).

18. Wirschell, M. *et al.* Building a radial spoke: flagellar radial spoke protein 3 (RSP3) is a dimer. *Cell Motil. Cytoskeleton***65**, 238–248 (2008).

19. Ng, P. Y. *et al.* Disruption of the dynein-dynactin complex unveils motor-specific functions in osteoclast formation and bone resorption. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.***28**, 119–134 (2013).

20. Ajiro, M., Jia, R., Yang, Y., Zhu, J. & Zheng, Z.-M. A genome landscape of SRSF3-regulated splicing events and gene expression in human osteosarcoma U2OS cells. *Nucleic Acids Res.***44**, 1854–1870 (2016).

21. Wang, Z., Chen, X., Zhao, Y., Jin, Y. & Zheng, J. G-protein-coupled estrogen receptor suppresses the migration of osteosarcoma cells via post-translational regulation of Snail. *J. Cancer Res. Clin. Oncol.***145**, 87–96 (2019).

22. Prahalad, S. *et al.* Variants in TNFAIP3, STAT4 and c12orf30 loci associated with multiple auto-immune diseases are also associated with Juvenile Idiopathic Arthritis. *Arthritis Rheum.***60**, 2124–2130 (2009).

23. Ma, F., Zhang, C., Prasad, K. V., Freeman, G. J. & Schlossman, S. F. Molecular cloning of Porimin, a novel cell surface receptor mediating oncotic cell death. *Proc. Natl. Acad. Sci. U. S. A.***98**, 9778–9783 (2001).

24. Zhang, C., Xu, Y., Gu, J. & Schlossman, S. F. A cell surface receptor defined by a mAb mediates a unique type of cell death similar to oncosis. *Proc. Natl. Acad. Sci. U. S. A.***95**, 6290–6295 (1998).

25. Eswarakumar, V. P. *et al.* The IIIc alternative of Fgfr2 is a positive regulator of bone formation. *Dev. Camb. Engl.***129**, 3783–3793 (2002).

26. Bredella, M. A. *et al.* Ectopic and Serum Lipid Levels Are Positively Associated with Bone Marrow Fat in Obesity. *Radiology***269**, 534–541 (2013).

27. Jafarifar, F., Dietrich, R. C., Hiznay, J. M. & Padgett, R. A. Biochemical defects in minor spliceosome function in the developmental disorder MOPD I. *RNA***20**, 1078–1089 (2014).

28. Liu, M., Sun, Y. & Zhang, Q. Emerging Role of Extracellular Vesicles in Bone Remodeling. *J. Dent. Res.***97**, 859–868 (2018).

29. Kassianidou, E. & Kumar, S. A biomechanical perspective on stress fiber structure and function. *Biochim. Biophys. Acta***1853**, 3065–3074 (2015).

30. Karasik, D., Rivadeneira, F. & Johnson, M. L. The genetics of bone mass and susceptibility to bone diseases. *Nat. Rev. Rheumatol.***12**, 323–334 (2016).

31. Wang, W. *et al.* Ribosomal Proteins and Human Diseases: Pathogenesis, Molecular Mechanisms, and Therapeutic Implications. *Med. Res. Rev.***35**, 225–285 (2015).

32. Serre, V. *et al.* Mutations in mitochondrial ribosomal protein MRPL12 leads to growth retardation, neurological deterioration and mitochondrial translation deficiency. *Biochim. Biophys. Acta***1832**, 1304–1312 (2013).

33. Pérez O'Brien, A. M. *et al.* Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet. Sel. Evol. GSE***46**, 19 (2014).

34. Piccirilli, J. A. & Koldobskaya, Y. Crystal structure of an RNA polymerase ribozyme in complex with an antibody fragment. *Philos. Trans. R. Soc. B Biol. Sci.***366**, 2918–2928 (2011).

35. Caputo, M. *et al.* Effect of low frequency (LF) electric fields on gene expression of a bone human cell line. *Electromagn. Biol. Med.***33**, 289–295 (2014).

36. Liu, G. *et al.* Cis-eQTLs regulate reduced LST1 gene and NCR3 gene expression and contribute to increased autoimmune disease risk. *Proc. Natl. Acad. Sci.***113**, E6321–E6322 (2016).

37. Huang, H. *et al.* Identification of pathways and genes associated with synovitis in osteoarthritis using bioinformatics analyses. *Sci. Rep.***8**, (2018).

38. Reumann, M. K. *et al.* Early growth response gene 1 regulates bone properties in mice. *Calcif. Tissue Int.***89**, 1–9 (2011).

39. Johnston, J. R., Chase, P. B. & Pinto, J. R. Troponin through the looking-glass: emerging roles beyond regulation of striated muscle contraction. *Oncotarget***9**, 1461–1482 (2017).

40. Ju, Y. *et al.* Troponin T3 expression in skeletal and smooth muscle is required for growth and postnatal survival: Characterization of Tnnt3tm2a(KOMP)Wtsi mice. *Genes. N. Y. N 2000***51**, 667–675 (2013).

41. Ralphs JR , Stamp TC , Dopping-Hepenstal PJ , Ali SY . Ultrastructural features of the osteoid of patients with fibrogenesis imperfecta ossium. Bone . 1989;10(4):243–249.

42. Carr AJ , Smith R , Athanasou N , Woods CG . Fibrogenesis imperfecta ossium. J Bone Joint Surg Br . 1995;77(5):820–829.

43. Bhadada, Sanjay Kumar et al. "Fibrogenesis Imperfecta Ossium and Response to Human Growth Hormone: A Potential Therapy." *The Journal of clinical endocrinology and metabolism* 102 5 (2017): 1750-1756 .

44. Bhadada SK, Dhaliwal R, Dhiman V, Rao SD. Fibrogenesis Imperfecta Ossium. Calcif Tissue Int. 2019;104:561-569.

45. Dhaliwal R, Dhiman V, Rao DS, Bhadada SK. Fibrogenesis Imperfecta Ossium: Clinical Approach to Diagnosis and Management of a Rare Skeletal Disorder. The Journal of Clinical Endocrinology & Metabolism.2019;104:4005-40015.

Figure Legends

Figure 1. Chart showing the workflow used to do bioinformatics analysis

Figure 2. Shown PCA (Principal Component Analysis) was performed on GSE43861 to reduce dimensionality and bring out patterns in the dataset.

Figure 3 Co-expression network interaction. Figure 3a Showing clustering method and Figure 3b Showing the analysis of iActiveModules based on expression levels. Figure 3c. Clustering showing the cluster with maximum connectivity

Figure 4 Heat map showing the expression of the 50 genes that were selected based on the clustering analysis and compares the control samples with the two FIO patient samples.

Figure 5 Pathways/Area of effect associated with FIO Figure 5a show an Ingenuity Pathway Analysis (IPA) network diagram. Figure 5b shows a heatmap of the 50 selected genes with their associated pathways/areas of effect.

Figures

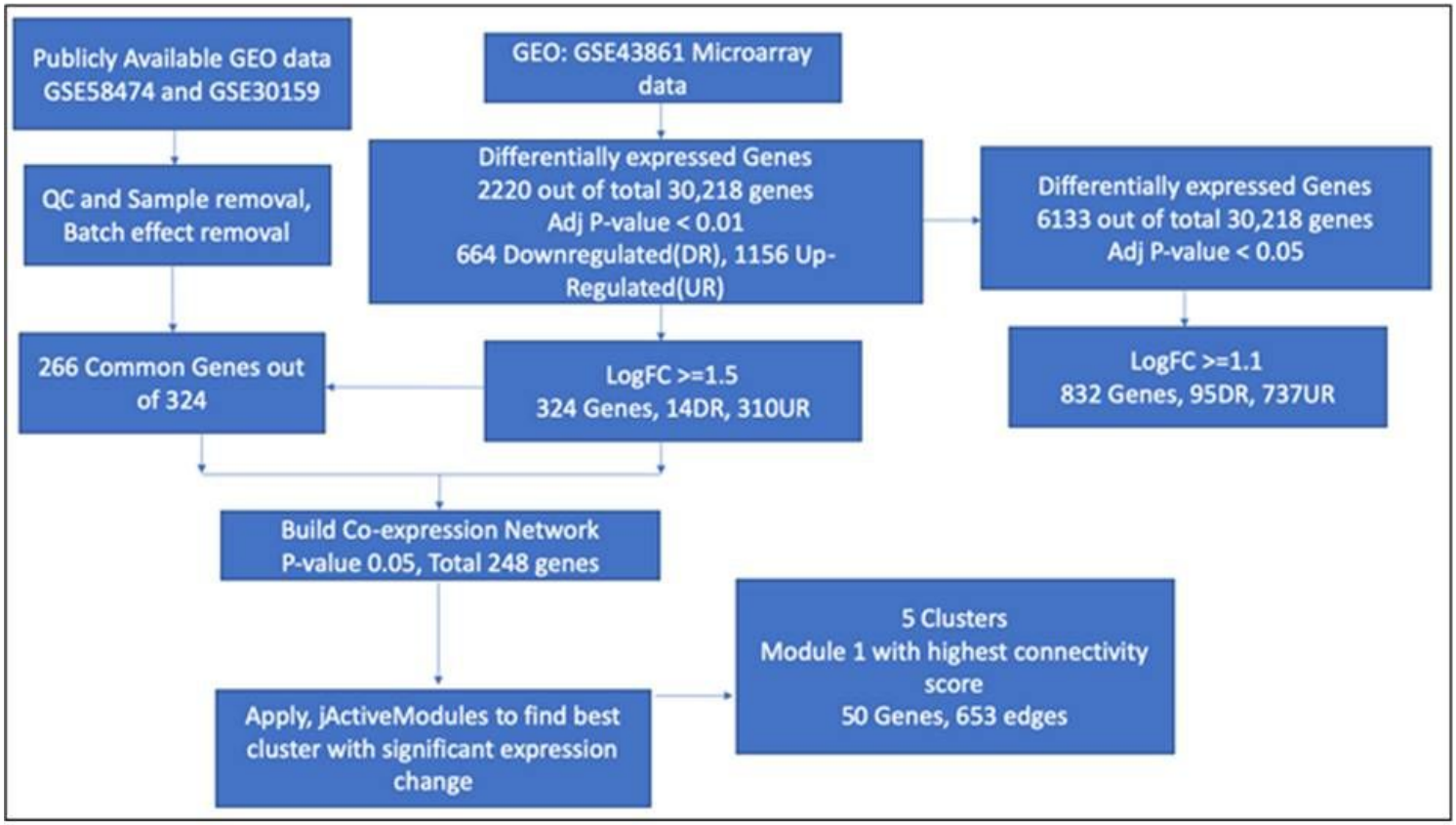

Figure 1

Chart showing the workflow used to do bioinformatics analysis

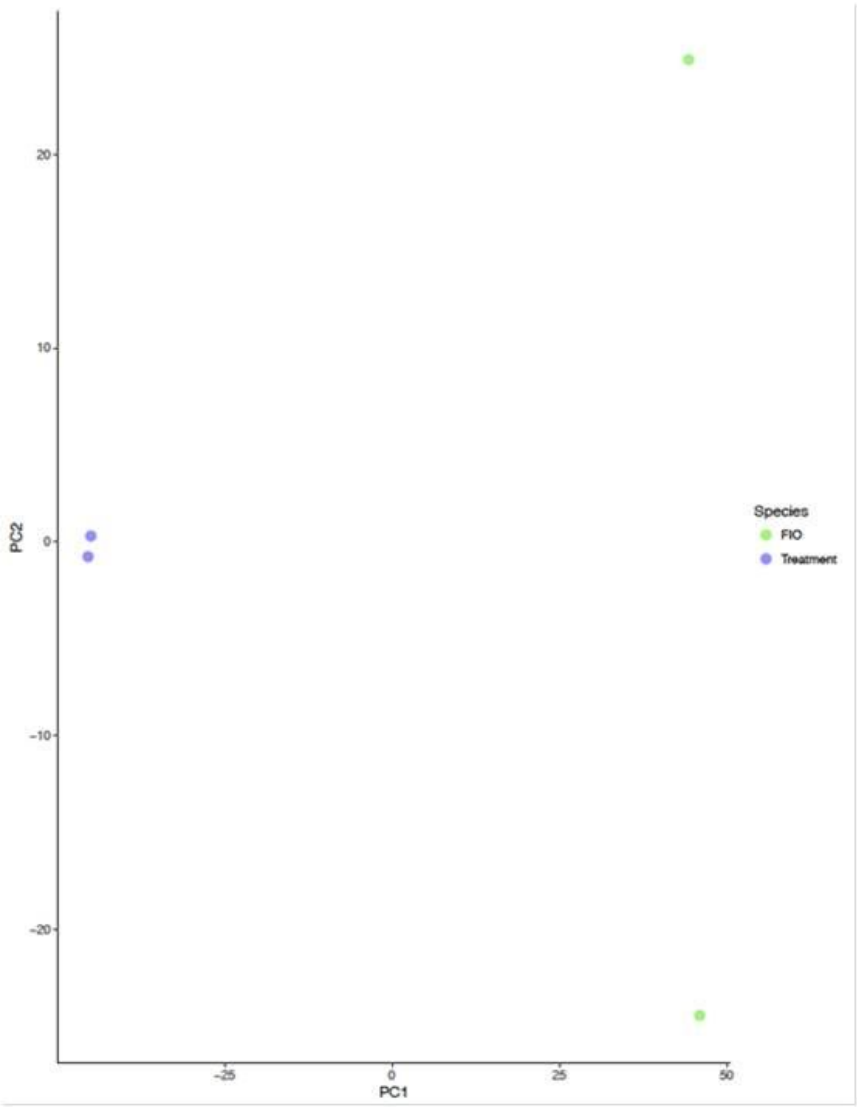

Figure 2

Shown PCA (Principal Component Analysis) was performed on GSE43861 to reduce dimensionality and bring out patterns in the dataset.

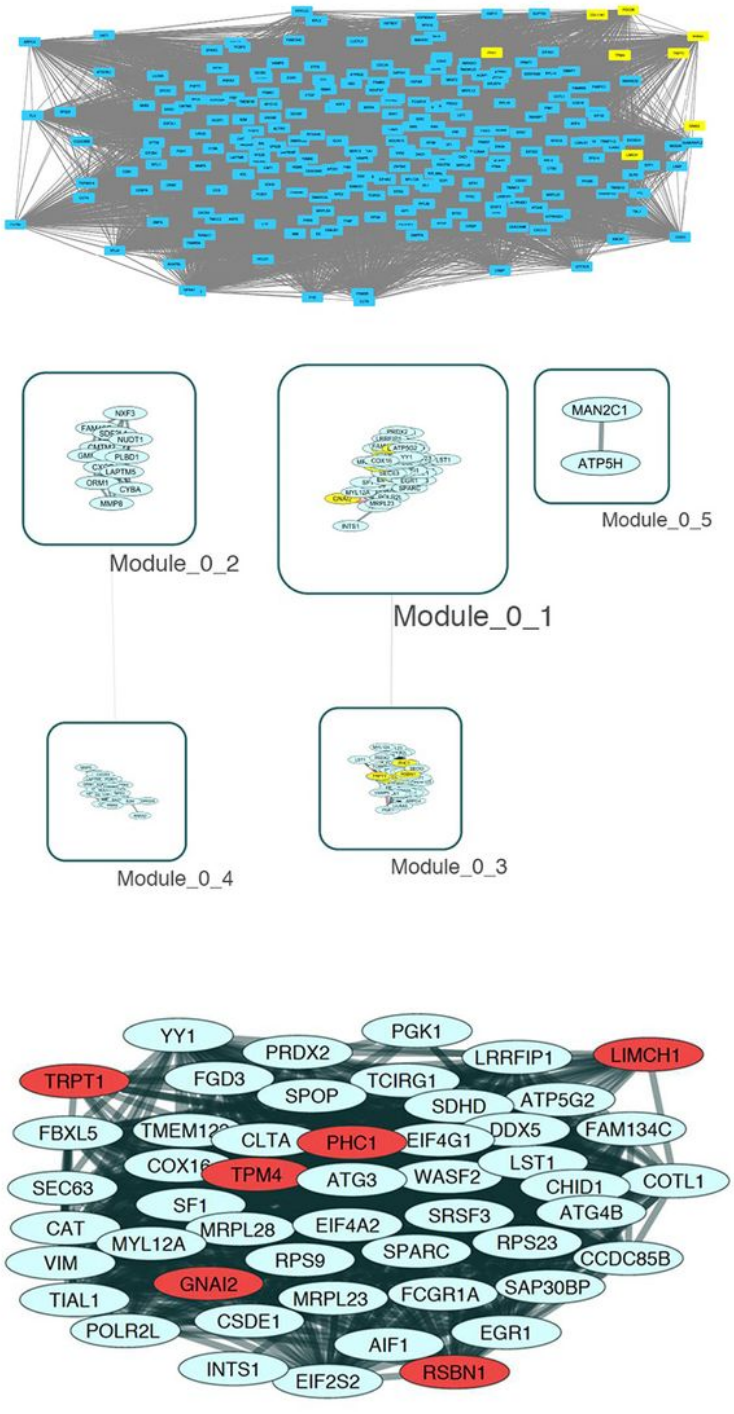

Figure 3

Co-expression network interaction. Figure 3a Showing clustering method and Figure 3b Showing the analysis of iActiveModules based on expression levels. Figure 3c. Clustering showing the cluster with maximum connectivity

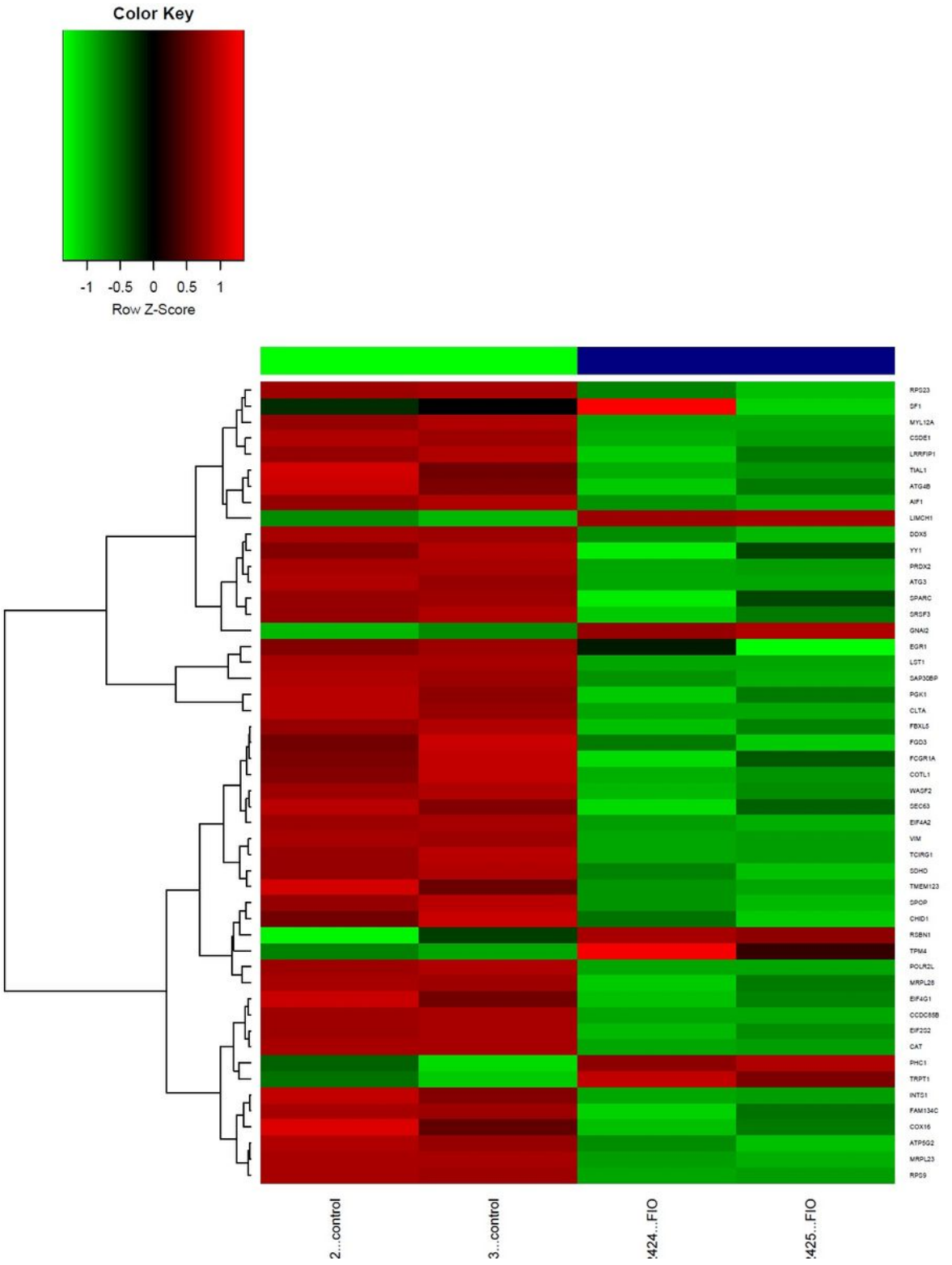

Figure 4

Heat map showing the expression of the 50 genes that were selected based on the clustering analysis and compares the control samples with the two FIO patient samples.

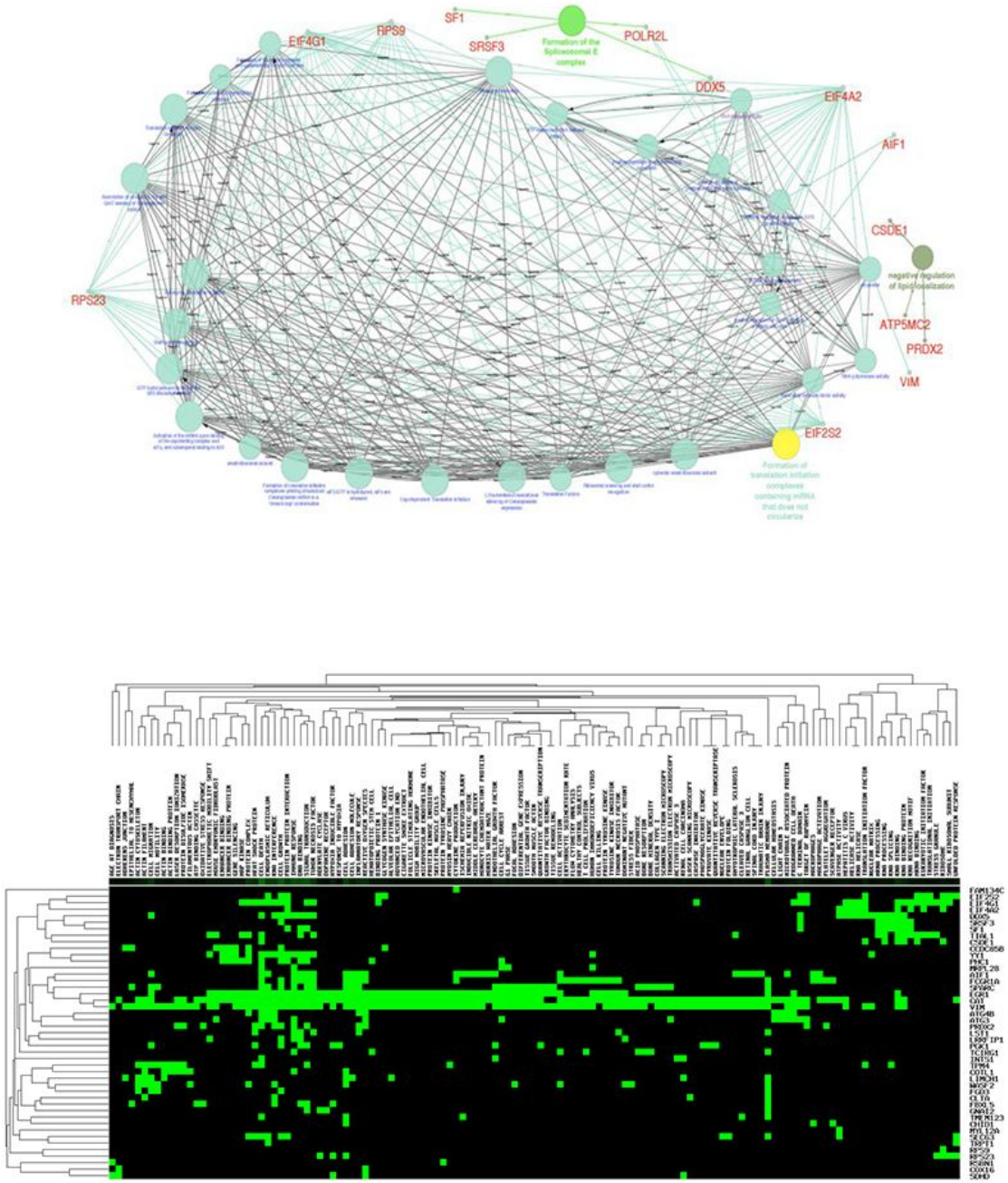

Figure 5

Pathways/Area of effect associated with FIO Figure 5a show an Ingenuity Pathway Analysis (IPA) network diagram. Figure 5b shows a heatmap of the 50 selected genes with their associated pathways/areas of effect.

Supplementary Files

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

- TableS1.txt
- TableS2.xlsx
- FigureS3C.jpg
- FigureS3C.jpg
- Table4.jpg