Identification of Signature Predictive of Pathological Situation from Atherosclerosis Plaque Microarray Data

Amit Dubey*, Pramod W Ramteke

Abstract— Large-scale gene expression analysis provides significance to the differentially expressed genes regulated in disease processes such as Atherosclerosis. Many targeting the promise of medical informatics by developing individualized diagnostic and therapeutic techniques for Atherosclerosis plaque will provide a detailed understanding about the genes and its variants that played an imperative role to this disease vulnerable and progression. A computational approach that accounts for genetic network, incorporates non-transcriptionally regulated genes, and accommodate foreseeing offers many advantages. We performed a comprehensive gene level estimation of performed prediction analysis for early and advanced condition of atherosclerosis plaque among 29 Atherosclerosis patients. Predictive signature profiles could be developed from normal tissues in humans and this is ambiguous and interchangeably sometimes it predicts as Diagnostic signature and in this it, reveals from advanced atherosclerosis plaque condition compares with early one. And after implicative analysis of microarrays to reveal gene ontology analyses and relative gene expression profiles. Our studies revealed that immune responsive genes is more up regulated in advance atherosclerosis patients as compared to early atherosclerosis patients and these gene expression is the elementary expression signature of disease progression in atherosclerosis plaque formation. Furthermore, we provide vision to the mutual genes interaction network of many clusters of genes associated with Atherosclerosis, revealing an overflow of immune and inflammatory signaling genes. We present a novel approach to future Pathway circuit designing based on connectivity approach. In doing this, we identify 1151genes of total genes are differentially expressing rest of many genes shows no variation in expression in early and advanced stages of Atherosclerosis that are preprocessing candidates for therapeutic targeting to future relevance. This study also provide significant insights over for systems-based approach to analyzing complex disease and also to evaluate global gene expression patterns in the common iliac arteries of monkeys with a diverse expansion of atherosclerosis vulnerabilities which is the basis of new medical informatics approaches in terms of their further drug development.

Index Terms— Micoarray Data, Atherosclerosis, Compuatational Analysis, Gene expression, Signature Genes, Up-regulated Genes,Down Regulated Genes, Gene Ontology, Pathway analysis.

1 INTRODUCTION

ATHEROSCLEROSIS is a manifold pathological process that is affected by both environmental and genetic factors.

According to American Heart Association (2002), atherosclerosis is a disease of large and medium-sized arteries characterized by thickening and hardening of the vascular wall. It involves a substance called plaque in the inner lining of the arteries. Over time, this buildup grows large enough to narrow the artery and significantly decrease the
Blood flow through it. When atherosclerosis affects the arteries that supply blood to the heart, it ultimately restricts blood flow to the heart muscle, causing heart pain (angina), irregular heartbeat (arrhythmia) and other problems. The plaques may also become fragile and rupture. Rupturing plaques form blood clots (thrombus) that may block the blood flow through an artery or break off and travel to another part of the body (embolus).

*Amit Dubey**, *PH- 09198278069. E-mail: ameetbioinfo@gmail.com*
Pramod w. Rameke, PH-09415124985. E-mail: pwramteke@yahoo.com
Corresponding Author (*)

The most devastating consequences of atherosclerosis, such as heart attack and stroke, are caused by superimposed thrombosis [4].
A very large number of total deaths are caused by cardiovascular disease like atherosclerosis[1].Despite progression in medical and surgical treatments, our knowledge that lead the therapies that slow the formation of atherosclerotic plaques are not totally successful [2].Therefore, it is necessary to continue investigating the fundamental mechanisms that cause atherosclerosis to develop more effective forms of treatment e.g. [3].Lots of studies have been conducted

Over the past few years to investigate the factors involved in initiation and progression of atherosclerosis. Molecular pathway analysis is one of the several tools used for investigating the initiation and progression of atherosclerosis. The aim of this study was to evaluate, by microarray analysis, gene expression profiles tissue at early and advanced stage of atherosclerosis for matter will need to create these components, incorporating the applicable criteria that follow.

2 MATERIAL AND METHODS

2.1 Data Collection

The dataset, consisting of the microarray samples of Atheroma plaques at two development stages. The two sample groups are of limited size (n=13(Early atherosclerosis patients and n=16(Advanced atherosclerosis patients) with associated clinical data was obtained from public one from the Gene Expression Omnibus (GEO database). Its accession number is GSE28829. The downloaded data were already normalized by RMA method. And to correctly annotate the probe sets, it's necessary to download Affymetrix annotation files for the Affymetrix Human Genome U133 plus 2.0 Array.

2.2 Data Analysis

Following are the steps for data analysis.

Step1: Data download from GEO site: File downloaded =GSE28829_family.soft.gz

Step2: load data in R software: Package used =GEOquery
We have used R software for microarray data analysis. The software R consists of large number of packages according to analysis. We have installed biocondoctor that consists of packages for microarray data analysis. Data loading requires GEOquery package

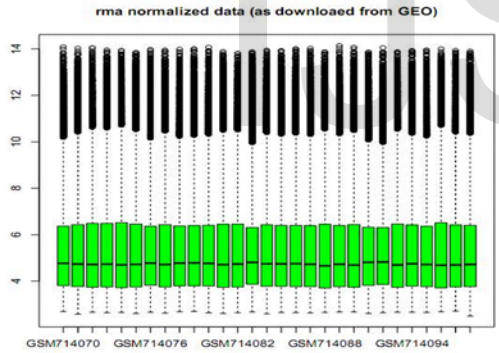

Fig. 1. Its represents the box plot of RMA normalized data as downloaded from GEO and it can be a significant source to-interpret information about a sample of size study. A box plot can provide information about a sample's range, median, normality of the distribution, and skew of the distribution. It can also identify and plot utmost cases within the sample.

Step3. Gene filtering: Package used= limma
Gene filtering: was done in four steps:
i) All those probes that do not have Entrez ID were removed from the analysis.
ii) All those probes without GO term annotation were also removed
iii) Uninformative probes were filter out (probes having (Interquartile range) IQR <= 0.2)
iv) If a gene is represented by more than two probes than only probe with highest IQR was selected for further analysis.

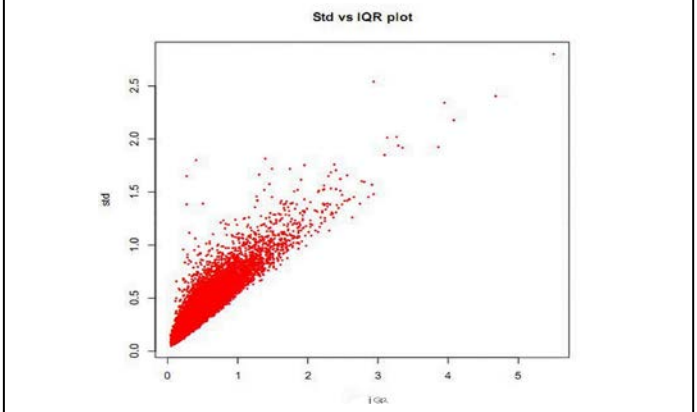

Fig. 2. It Represent the comparison between the standard and IQR plot and both were providing the similar results as we can see it from the plot that a diagonal can be possible to drawn which determine that it will provide same

Step4: Unsupervised clustering: Hierarchical Clustering and Visualization
This step is done to identify groups of co-expressed genes recognizes coherent expression patterns. However, the interpretation of co-expressed genes and coherent patterns strongly depends on the domain knowledge, which makes it difficult to fully automate and estimating the number of classes (groups or clusters) and assigning an object to these classes. This analysis is ideal for the discovery of novel classes among 29 Ath-

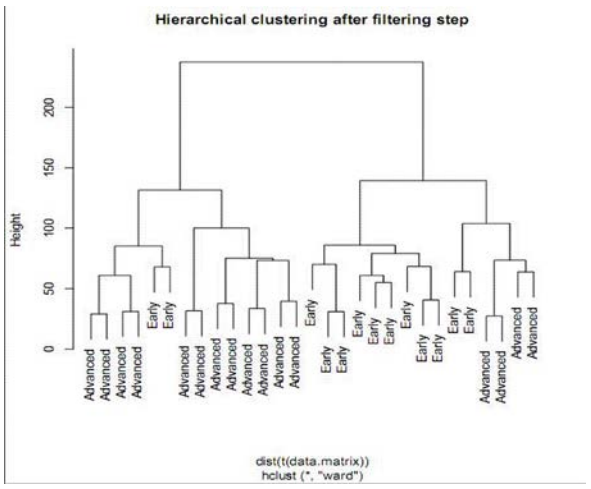

Fig. 3. It Represents the Hierarchical clustering that shows sample relationships using microarray data as evidenced by clustering in which two large dendogram shows that the atherosclerosis patients' data is categorized into two sample conditions->one is for advanced atherosclerosis stage and another is for early atherosclerosis stage.

Step 5: t-test: to differentially expressed genes: Package Used= limma

This approach in addition to ease of interpretation. Since the number of gene cluster verified is usually much less than the number of genes represented on the Microarray, the consequence of the multiple analogy problem is diminished. Also, expression patterns of genes in a gene cluster can emphasize Each other and do not have to be individually significant at a very inflexible level as required for the post annotation methods. Genes differentially expressed between 'Advanced' and 'Early' stage samples were identified by modified t-test as provided in R-package 'limma'. Using the criterion of adjusted P-value < 0.01 for gene selection

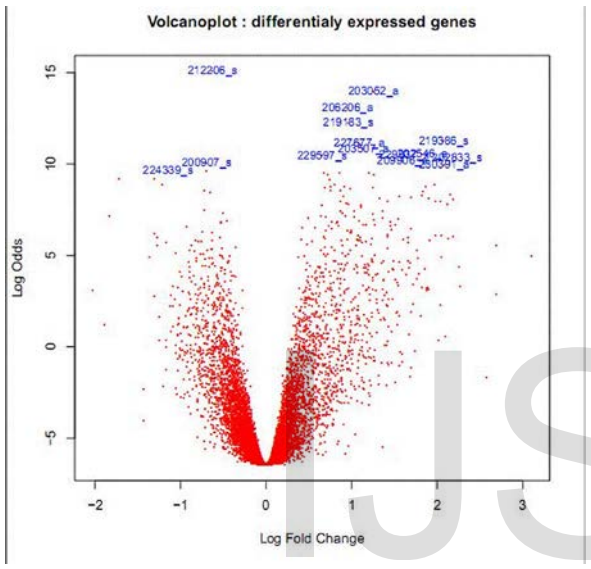

Fig4: It Represent volcano plot which displays fold changes versus a measure of statistical importance of the change representing the expressed genes. It also combines a statistical test with the magnitude of the change enabling quick visual identification of those data-points (expressed genes) that display large-magnitude changes that are also statistically significant in identification to their rate of expression.

Step 6: Gene Ontology analyses: Package used =GOstats
Each list of differentially expressed genes was analyzed in the background of gene ontology (GO) to identify groups of genes with similar functions or processes.Genes, differentially expressed between 'Advanced' and 'Early' stage samples were identified by modified t-test as provided in R-package 'limma' .The differentially expressed genes were analyzed using gene ontology (GO) analysis using GOstat package.

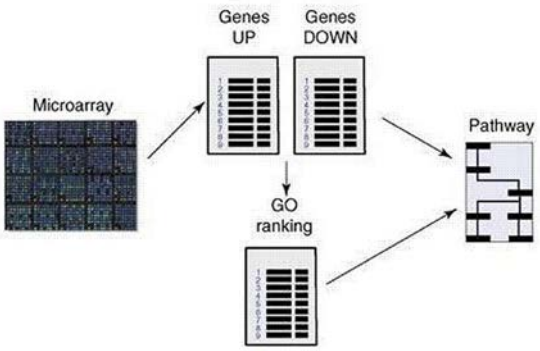

Fig5: It Represent microarray to biological interpretation of results. Genes up Regulated, Genes down Regulated, and GO ranking symbolize lists, which are essentially one- dimensional annotation of the data, though GO ranking already contains groups of genes on the basis of their P-value significance. The arrows indicate that there are many ways to develop the more complex structures from the initial data.

3 RESULTS

3.1 Patient Population

Microarray Gene Expression data from early and advanced atherosclerotic plaque from humans carotid samples of Atherosclerosis patients at two development stages. The two sample groups are of limited size and Histological grading according to GEO record GSE28829 Affymetrix U133 plus 2 arrays for histo pathological grading of atherosclerosis was obtained for 29 atherosclerosis' patients', samples size is (n=16) Advanced atherosclerosis patients samples were restricted as advanced lesions and (n=13) Early atherosclerosis patients as with Early lesions.

3.2 Differential Gene Expression According to Atherosclerosis Disease

Place our first set of analyses focused on identifying differential gene expression among various disease harshness classes. In this prediction, the two samples classified as disease early sample, and advanced samples were excluded because the Up regulated and down regulated genes access did not match up with our quality parameters for prediction. The lack of altogether normal arteries, such as might be obtained from young individuals, plays as a constraint to our discovery of early stage candidate genes.
We first analyzed gene expression with reference to histological grading Early versus advanced, t-test analysis revealed 1151 differentially regulated genes criterion of adjusted P-value < 0.01 Six hundred sixty Seven genes were up regulated in advanced lesions. Specifically, fifty four genes in up regulated and eight genes in down regulated genes a to be the most significant biological processes over immune system showed the P-value score (0.00) and eight genes showed the P-value scores (0.00) in down regulated genes, as the most com-

pelling molecular function terms regulated by these genes and three other immune responsive genes such as mast cell activation involved in immune response, supervision of T-helper 1 type immune response, and regulation of type 2 immune response genes showed the P-value scores (0.01). We also found cell cycle regulatory genes such as cellular component organization at cellular level, actin cytoskeleton organization, adhesion the most significantly down regulated gene in advanced grade lesions P-value scores (0.01)

In contrast, an analysis of early sample versus advanced samples generated significant differences in gene expression, with maximum genes with high False Detection Rate (FDR). Our findings announce that the microscopic progression of disease severity in atherosclerosis is nearly related with the expression profile; however, in advanced sample lesions are more in compared with early sample lesions. This result is persistent with the fact that the presence of lipid Pools (rather than cell type changes) is the differentiating component of advanced atherosclerosis disease.

TABLE 1
GO-DISCOVERED EXTRACTED REPORT FOR DISEASE HARSHNESS GRADE UP REGULATION OF GENES SHOWING SIGNIFICANCE OVER IMMUNE RESPONSELEAD TO PLAQUE FORMATION

GOBPID	Pvalue	Odds Ratio	ExpCount	Count	Size	Term
GO:0002376	0.000	5.661	58	204	1200	immune system process
GO:0006955	0.000	6.795	37	160	765	immune response
GO:0050776	0.000	7.373	18	93	381	regulation of immune response
GO:0002682	0.000	5.580	29	114	593	regulation of immune system process
GO:0002684	0.000	6.125	19	84	391	positive regulation of immune system process
GO:0050778	0.000	7.098	13	64	260	positive regulation of immune response
GO:0002253	0.000	7.415	10	55	214	activation of immune response
GO:0045087	0.000	5.155	17	67	349	innate immune response
GO:0002757	0.000	7.714	9	47	176	immune response-activating signal transduction
GO:0002764	0.000	7.480	9	47	180	immune response-regulating signaling pathway
GO:0002429	0.000	9.074	5	32	105	immune response-activating cell surface receptor signaling pathway
GO:0002768	0.000	8.599	5	32	109	immune response-regulating cell surface receptor signaling pathway
GO:0002460	0.000	7.366	5	29	110	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
GO:0002520	0.000	3.581	17	50	343	immune system development
GO:0002263	0.000	8.460	3	20	68	cell activation involved in immune response
GO:0002366	0.000	8.460	3	20	68	leukocyte activation involved in immune response
GO:0006959	0.000	7.232	4	21	80	humoral immune response
GO:0045088	0.000	4.913	7	29	150	regulation of innate immune response
GO:0002697	0.000	5.065	7	27	136	regulation of immune effector process
GO:0016064	0.000	8.279	3	18	62	immunoglobulin mediated immune response
GO:0045089	0.000	4.962	5	22	112	positive regulation of innate immune response
GO:0002683	0.000	4.978	4	18	91	negative regulation of immune system process
GO:0002758	0.000	4.978	4	18	91	innate immune response-activating signal transduction
GO:0002218	0.000	4.844	4	18	93	activation of innate immune response
GO:0002699	0.000	5.918	3	15	66	positive regulation of immune effector process
GO:0002822	0.000	6.867	2	13	51	regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
GO:0002455	0.000	8.811	2	11	36	humoral immune response mediated by circulating immunoglobulin
GO:0002819	0.000	6.690	3	13	52	regulation of adaptive immune response
GO:0002703	0.000	6.066	3	13	56	regulation of leukocyte mediated immunity
GO:0002275	0.000	10.572	1	9	26	myeloid cell activation involved in immune response
GO:0002285	0.000	6.879	2	11	43	lymphocyte activation involved in immune response
GO:0002440	0.000	5.005	3	12	60	production of molecular mediator of immune response
GO:0002448	0.000	11.926	1	6	16	mast cell mediated immunity
GO:0002286	0.000	8.707	1	7	23	T cell activation involved in immune response
GO:0002700	0.000	5.984	2	9	39	regulation of production of molecular mediator of immune response
GO:0002824	0.000	6.930	1	8	31	positive regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
GO:0002702	0.000	10.841	1	6	17	positive regulation of production of molecular mediator of immune response
GO:0002821	0.000	6.641	2	8	32	positive regulation of adaptive immune response
GO:0002367	0.000	7.738	1	7	25	cytokine production involved in immune response
GO:0042092	0.000	14.177	1	5	12	type 2 immune response
GO:0042088	0.000	9.172	1	6	19	T-helper 1 type immune response
GO:0002720	0.000	12.404	1	5	13	positive regulation of cytokine production involved in immune response
GO:0002706	0.000	4.849	2	9	46	regulation of lymphocyte mediated immunity
GO:0002704	0.000	19.819	0	4	8	negative regulation of leukocyte mediated immunity
GO:0002707	0.000	19.819	0	4	8	negative regulation of lymphocyte mediated immunity
GO:0002282	0.000	59.370	0	3	4	microglial cell activation involved in immune response
GO:0002313	0.000	59.370	0	3	4	mature B cell differentiation involved in immune response
GO:0052552	0.000	59.370	0	3	4	modulation by organism of immune response of other organism involved in symbiotic interac
GO:0052555	0.000	59.370	0	3	4	positive regulation by organism of immune response of other organism involved in symbiot interaction
GO:0052556	0.000	59.370	0	3	4	positive regulation by symbiont of host immune response
GO:0002279	0.001	9.921	1	5	15	mast cell activation involved in immune response
GO:0002825	0.001	13.210	0	4	10	regulation of T-helper 1 type immune response
GO:0002828	0.001	13.210	0	4	10	regulation of type 2 immune response

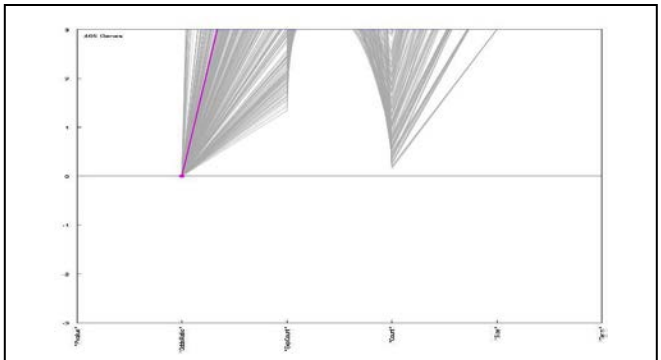

Fig. 6 Connectivity odd rankings were evaluated for all significantly genes differentially up regulated across disease harshness classes. Genes showed are rated according to the average p-value score. The aggregate score represents the sum of all the significance values within a given genetic network, and the average score shows the aggregate score divided by the number of connections

3.3 Pathway recognition by Profusion

Looking for profusion of differentially regulated genes in created pathways can be beneficial over genomic evidence. We carried out profusion analysis according to the hyper geometric circulation for disease harshness and Atherosclerosis groups of genes for pathways within KEGG Database. This database providing insights to understand the metabolic pathways. Within the disease severity group (differentially regulated genes between early and advanced sample of Atherosclerosis among 29 patients groups), the cell cycle pathway was the most significantly over interpreted, persistent with an strongly uncertain cell phenotype and providing further support for the concept that mast cell and smooth muscle dedifferentiation is a basic process in disease progression.

4 CONCLUSION

We occupied a computational-based approach toward the study of atherosclerosis Microarray Data by applying statistical tools to a large, manifold data set derived from human carotid artery tissue (with atherosclerosis disease) the major goal of expression data analysis is to provide the biologists biologically consequential information about the cluster of genes and related things related with the vulnerability of atherosclerosis plaque formation. Through this intelligence, biologists are able to discover unknowns and disclose previous Knowledge. Microarray analysis has the potential to reveal atherosclerosis vulnerability and its prognosis, well beyond the currently used clinical parameters to predict Disease results. Diagnostic assays developed on gene expression profiling Studies will therefore assistance to areas of medicine.
Compelling genes were studied in the broader background of ontology and by mapping onto known and innovative pathways. Output revealed insights into biology not obtained and possible through basic gene lists alone; instead, highly associated cluster of genes were identified and placed in the unified context of the disease whole.

TABLE 2
GO-DISCOVERED EXTRACTED REPORT FOR DISEASE HARSHNESS GRADE DOWN REGULATION OF GENES SHOWING SIGNIFICANCE OVER IMMUNE RESPONSELEAD TO PLAQUE FORMATION

GOBPID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0007010	0.000	2.283	17	36	507	cytoskeleton organization
GO:0006936	0.000	3.099	6	18	188	muscle contraction
GO:0003012	0.000	2.875	7	18	201	muscle system process
GO:0030029	0.000	2.430	10	23	301	actin filament-based process
GO:0007160	0.000	3.665	4	12	107	cell-matrix adhesion
GO:0006996	0.000	1.610	50	74	1475	organelle organization
GO:0007155	0.000	1.866	24	42	713	cell adhesion
GO:0022610	0.000	1.866	24	42	713	biological adhesion
GO:0071842	0.001	1.472	80	107	2359	cellular component organization at cellu
GO:0030036	0.001	2.356	9	20	268	actin cytoskeleton organization
GO:0031589	0.001	2.983	5	13	139	cell-substrate adhesion

Disease specificity

Analysis of histopathology graded samples revealed that many genes expressed at a relatively high level in early sample were important marker genes of smooth muscle cell differentiation, regulation, or activation (15). Moreover, many of these genes were classify under ontology terms such as muscle development, and actin filament etc. Over account of cell cycle pathway components also advised active cellular differentiation processes were present, and our connectedness ranking obtained cell cycle and immune networks most apparently adequate. Immuno-histochemistry (IHC) of selected proteins showed that the cell of origin of apparently down regulated genes such as actin was certainly the smooth muscle cell. There are still many processes are established to be involved in the development of atherosclerosis. Our prediction reveals, in which key smooth muscle genes and ontologies are outstanding over and above those of other cells or accepted immune signals, suggests that the key process in the progression of atherosclerosis relates to smooth muscle cell dedifferentiation. Expression analysis showed that 1151 genes of the total genes showed differential expression. Out of that 667 genes were up-regulated and 484 genes were down-regulated in advanced atherosclerotic Patients to early one. After performing prediction analysis for early and advanced condition also reveal that immune response processes is more active in advance stage its is responsive upon the mutual genes interaction network.
In arbitrary, insight into the true condition of a disease can only be achieved by connecting multiple approaches to investigation. Conventional techniques to the molecular genetics of atherosclerosis have focused on employ single genes with transgenic technology. The main cause of atherosclerosis is yet to be revealed, but is hypothesized that it is proposed by incendiary processes in the vessel wall in response to maintained low-density lipoprotein (LDL) molecules. LDL molecules become affected inside the wall and oxidized by free radicals, and become lethal to the cells. The destruction caused by the oxidized LDL molecules triggers a cascade of immune responses which over time can produce an atherosclerosis. To uncover this, we used transcription profiling, a technique that allows concurrent assay of more than fifty six thousands of genes, to inspect important pathways in atherosclerosis without the need for an a priori target. We implement widely recognized computational approaches to reveal false discovery (FD) and then identified outstanding signals in ontologies and known pathways, finally analyzing cluster of genes through the adoption of a connectivity analysis and visualization approach. Our access takes benefits of the unmatched high throughput of expression profiling while computing for its significant defect: many genes are regulated at the posttranslational level. Because our access connects genes whose gene expression is compelling changed, overall patterns of pathway activation become possible even if several representative of the network are not regulated at the transcriptional level.

ACKNOWLEDGMENT

We wish to thank my coleuges, friends who gave me their valuable support and guidelines. This work was supported in part by a self grant from me and my guide.

REFERENCES

[1] Braunwald E. Shattuck lecture, "cardiovascular medicine at the turn of the millennium: triumphs, concerns, and opportunities," *New Engl J Med*, 337, 1360–1369, 1997.

[2] CK, Witztum JL, "Atherosclerosis: the road ahead," *Cell*, 104,503–516, 2001.

[3] Jakulj L, Vissers MN, Tanck MWT, Hutten BA, Stellaard F, Kastelein JJP, Dallinga-Thie GM,"ABCG5/G8 polymorphisms and markers of cholesterol metabolism: systematic review and meta-analysis," *J Lipid Res*, 51, 3016–3023, 2010

[4] Gong-Qing Shen, Christine Moravec, Stephen Ellis, Eric J. Topol and Qing Wang Stephen R. Archacki, George Angheloiu, Xiao-Li Tian, Fen Lai Tan, Nick DiPaola,"Identification of new genes differentially expressed in coronary artery disease by expression profiling,"*Physiol. Genomics*, 15, 65-74, 2003.

[5]Suzanne Smit, Huub C.J. Hoefsloot, Age K. Smilde,"Statistical data processing in clinical proteomics,"*Journal of Chromatography B* , 866, 77-88, 2008.

[6] Miroslava Cuperlovic-Culf, Nabil Belacel and Rodney J. Ouellette," Determination of tumour marker genes from gene expression data,"*Drug Discov Today*., 10(6), 429-37, 2005.

[7] YingLu, Jiawei Han,"Cancer classification using gene expression data," *Information Systems*, 28, 243-268, 2003.

[8] Georgios Papachristoudis, Sotiris Diplaris, Pericles A. MitkasSoFoCles,"Feature filtering for microarray classification based on Gene Ontology,"*Journal of Biomedical Informatics*, 43, 1-14, 2010.

[9] Eleftherios Panteris, Stephen Swift, Annette Payne, Xiaohui Liu,"Mining pathway signatures from microarray data and relevant biological knowledge," *Journal of Biomedical Informatics*, 40, 698-706, 2007.

[10] Esther Lutgens, Robert-Jan van Suylen, Birgit C. Faber, Marion J. Gijbels, Petra M. Eurlings, Ann-Pascale Bijnens, Kitty B. Cleutjens, Sylvia Heeneman, Mat J.A.P. Daemen,"Atherosclerotic Plaque Rupture Local or Systemic Process?,"*Arterioscler Thromb Vasc Biol*., 23, 2123-2130, 2003.

[11] David Seo, Tao Wang, Holly Dressman, Edward E. Herderick, Edwin S. Iversen, Chunming Dong, Korkut Vata, Carmelo A. Milano, Fabio Rigat, Jennifer Pittman, Joseph R. Nevins, Mike West, Pascal J. Goldschmidt Clermont,"Gene Expression Phenotypes of Atherosclerosis.,"*Arterioscler. Thromb. Vasc. Biol*., 24, 1922-1927, 2004.

[12] Marianna Papaspyridonos, Alberto Smith, Kevin G. Burnand, Peter Taylor, Soundrie Padayachee, Keith E. Suckling, Christian H. James, David R. Greaves, Lisa PateL," Novel Candidate Genes in Unstable Areas of Human Atherosclerotic Plaques,"*Arteriosclerosis,Thrombosis, and Vascular Biology*, 26, 837-1844, 2006.

[13] Erling Falk, MD, PHD ,"Pathogenesis of Atherosclerosis," *Journal of the American College of Cardiology*, 47(8), C7–C12, 2006.

[14] Salim Yusuf Stephanie Ôunpuu Sonia Anand,"The Global Epidemic of Atherosclerotic Cardiovascular Disease: International Conference on Coronary Heart Disease in the New Millennium: Epidemiology: *Med Principles Pract*; 11(suppl 2), 3–8, 2002.

[15] Spin JM, Nallamshetty S, Tabibiazar R, Ashley EA, King JY, Chen M, Tsao PS, and Quertermous T," Transcriptional profiling of in vitro smooth muscle cell differentiation identifies specific patterns of gene and pathway activation," *Physiol Genomics* 19: 292–302, 2004.

[16] Andrie G. Panayiotou." Novel Biomarkers and Subclinical Atherosclerosis ", *Ultrasound and Carotid Bifurcation Atherosclerosis*, 2011.

[17] Juan I. Montoya-Burgos, "Patterns of Positive Selection and Neutral Evolut ion in the Protein-Coding Genes of Tetraodon and Takifugu", *PLoS ONE*, 09/ 13/ 2011

[18] Gaoc huan Zhang," Induced pluripotent stem (iPS) cell consensus genes: Implication for the risk of tumor genes is and cancers in iPS cell therapy", *Stem Cells and Development*, 12/ 20/ 2011

[19] Miroslava Cuperlovic-Culf, "Microarray Analysis of Alternative Splicing," *Omics A Journal of Integrative Biology*, 09/ 2006

[20] Kathleen M. Eyster, "Gene expression signatures differ with extent of atherosclerosis in monkey iliac artery:,", *Menopause*, 06/ 2011.

[21] Redel l, John B., Anthony N. Moore, Raymond J. Grill ,Daniel Johns on, Jing Zhao, Yin Liu, and Pramod K .Dash. "Analysis of Functional Pathways Altered Following Mild Traumatic Brain Injury ", *Journal of Neurotrauma*, 2012.

[22] Lu, Y,"Cancer classification using gene expression data", *Information Systems*, 200306

[23] J.Y. King, "Pathway analysis of coronary atherosclerosis", *Physiological Genomics*, 09/ 21/ 2005.

[24] Li, Xinrui, Patricia Schulte, David V Godin, and Kimberly M Cheng,"Differential mRNA expression of seven genes involved in cholesterol Metabolism and transport in the liver of atherosclerosis-susceptible and-Resistant Japanese quail strains ", *Genetics Selection*

[25] Werner, T. "Bioinformatics applications for pathway analysis of microarray data", *Current Opinion in Biotechnology*, 200802

[26] Richard Simon."Challenges of Microarray Data and the Evaluation of Gene Expression Profile Signatures", *Cancer Investigation*, 5/ 2008