Application of A Novel Triclustering Method in Analyzing Three Dimensional Transcriptomics Data

Dissertation

for the award of the degree

"Doctor of Philosophy" Ph.D. Division of Mathematics and Natural Sciences

of the Georg-August-University, Goettingen

within the doctoral Program for Environmental Informatics (PEI)

of the Georg-August University School of Science (GAUSS)

submitted by

Anirban Bhar

from Rishra, India

Goettingen, 2015

Thesis committee:

Prof. Dr. Edgar Wingender, Institute of Bioinformatics, University Medical Center, Georg-August University, Goettingen, Germany

Prof. Dr. Stephan Waack, Theory and Algorithmic Methods, Institute of Computer Science, Georg-August University, Goettingen, Germany

Members of the examination board:

Referee: Prof. Dr. Edgar Wingender, Institute of Bioinformatics, University Medical Center, Georg-August University, Goettingen, Germany

Co-referee: Prof. Dr. Stephan Waack, Theory and Algorithmic Methods, Institute of Computer Science, Georg-August University, Goettingen, Germany

Other members of the examination board:

Prof. Dr. Burkhard Morgenstern, Institute of Microbiology and Genetics, Department of Bioinformatics, Georg-August University, Goettingen, Germany

Prof. Dr. Anita Schoebel, Institute for Numerical and Applied Mathematics, Georg-August University, Goettingen, Germany

Prof. Dr. Tim Beissbarth, Department of Medical Statistics, University Medical Center, Georg-August University, Goettingen, Germany

Prof. Dr. Dieter Hogrefe, Institute of Computer Science, Georg-August University, Goettingen, Germany

Date of the oral examination : 24th March, 2015.

I hereby declare that I prepared the PhD thesis entitled "Application of A Novel Triclustering Method in Analyzing Three Dimensional Transcriptomics Data" on my own and with no other sources and aids than quoted.

Anirban Bhar

Dedicated to my family...

Contents

1 Introduction . . . 3
1.1 Central Dogma of Molecular Biology . . . 4
1.1.1 DNA . . . 4
1.1.2 RNA . . . 5
1.1.3 Protein . . . 6
1.1.4 Transcription . . . 6
1.1.5 Translation . . . 7
1.2 Microarray Technology . . . 8
1.2.1 cDNA Microarray . . . 8
1.2.2 Oligonucleotide Microarray . . . 9
1.3 Computational Analysis of Microarray Gene Expression Data . . . 10
1.3.1 Differential Expression Analysis . . . 10
1.3.2 Machine Learning in Mining Microarray Gene Expression Data . . . 11
1.3.3 Co-expression and Cluster Analysis . . . 13
1.4 Structure of the Thesis . . . 14
1.5 Bibliography . . . 15
2 A New Triclustering Approach for Unveiling Biological Processes of Disease Progression from Gene Expression Profiles . . . 19
2.1 Introduction . . . 20
2.2 Materials and Methods . . . 20
2.2.1 Materials . . . 20
2.2.2 Methods: δ-TRIMAX . . . 21
2.3 Results and Discussion . . . 28
2.3.1 Results on Simulated Dataset . . . 28
2.3.2 Results on Real-life Dataset . . . 29
2.3.3 Biological Significance . . . 31
2.4 Conclusion . . . 38
2.5 Bibliography . . . 38

3 Enhanced Multi-objective Triclustering Based on a Genetic Algorithm and Its Application in Revealing Biological Processes of Development 49
3.1 Introduction 50
3.2 Summary of δ-TRIMAX 50
3.2.1 Aim of δ-TRIMAX 50
3.2.2 Pitfalls of δ-TRIMAX 50
3.3 Materials and Methods 51
3.3.1 Materials 51
3.3.2 EMOA-δ-TRIMAX 54
3.3.3 Convergence of Solutions 57
3.4 Results and Discussions 58
3.4.1 Artificial Dataset 58
3.4.2 Real-life Datasets 60
3.4.3 Performance Comparison 61
3.4.4 Identifying Key Genes of Triclusters and Analyzing Their Roles During hiPSC Differentiation into Cardiomyocytes 72
3.5 Conclusion 74
3.6 Bibliography 75
3.7 Appendix 89
4 Speculating about the Role of *ZEB2* During Stem Cell Differentiation into Cardiomyocytes 99
4.1 Introduction 100
4.2 Materials and Methods 101
4.2.1 Dataset 101
4.2.2 Methods 101
4.3 Results and Discussion 102
4.3.1 SCCs of Module 1 104
4.3.2 SCCs of Module 2 104
4.3.3 SCCs of Module 3 105
4.3.4 SCCs of Module 4 105
4.3.5 Elucidating the Roles of *ZEB2* and *SMADs* Transcription Factors During Cardiac Development 110
4.4 Conclusion 114
4.5 Bibliography 115

5 Co-regulation Analysis of Time Series Transcriptomics Data Unveils the Roles of Three Node Feed-Forward Loops in Regulating Genes of Signaling Pathways of Breast Cancer Progression . . . 125
5.1 Introduction . . . 126
5.2 Materials and Method . . . 128
5.2.1 Dataset . . . 128
5.2.2 Methods . . . 129
5.3 Results and Discussion . . . 130
5.4 Conclusion . . . 142
5.5 Bibliography . . . 142
5.6 Appendix . . . 146
6 Unraveling Potential Signaling Pathways During the Exposure of Several Tissues to Different Toxicants in Different Species . . . 167
6.1 Introduction . . . 168
6.2 Materials and Methods . . . 168
6.2.1 Dataset 1 . . . 168
6.2.2 Dataset 2 . . . 168
6.2.3 Dataset 3 . . . 169
6.2.4 Dataset 4 . . . 169
6.2.5 Workflow . . . 169
6.3 Results and Discussion . . . 170
6.3.1 Results on Dataset 1 . . . 171
6.3.2 Results on Dataset 2 . . . 174
6.3.3 Results on Dataset 3 . . . 176
6.3.4 Results on Dataset 4 . . . 177
6.4 Conclusion . . . 181
6.5 Bibliography . . . 182
6.6 Appendix . . . 191
7 Elucidating the Importance of Clustering Replicates in Three Dimensional Microarray Gene Expression Data . . . 271
7.1 Introduction . . . 272
7.2 Materials and Methods . . . 272
7.2.1 Dataset 1 (Accession no.- GSE11324) . . . 272
7.2.2 Dataset 2 (Accession Number- GSE35671) . . . 272

7.2.3 Dataset 3 (Accession Number- GSE46280) 273
7.2.4 Dataset 4 (Accession Number- GSE17693) 273
7.2.5 Dataset 5 (Accession Number- GSE17933) 273
7.2.6 Dataset 6 (Accession Number- GSE18858) 273
7.2.7 Dataset 7 (Accession Number- GSE38513) 273
7.2.8 Workflow 274
7.3 Results and Discussion 275
7.4 Conclusion 279
7.5 Bibliography 279
8 Conclusion 281

List of Figures

1.1 The central dogma of molecular biology 4
1.2 The helical structure of DNA 5
1.3 Orientation of transcription 7
1.4 The workflow of cDNA microarray 9
1.5 The workflow of an Affymetrix microarray experiment 10
1.6 Co-regulation of co-expressed genes 13

2.1 Number of transactions in each month 23
2.2 Comparison in terms of Affirmation Scores 29
2.3 The Pearson correlation coefficient between tricluster eigen-gene vector and the tricluster key genes 33
2.4 Enriched KEGG pathway terms for the key genes 34
2.5 Enriched KEGG pathway terms for the key genes 35

3.1 Descriptions of the time points of Dataset 2 52
3.2 Descriptions of the time points of Dataset 3 52
3.3 Workflow used in this chapter 53
3.4 Steps of EMOA-δ-TRIMAX 54
3.5 Comparison in terms of affirmation score 59
3.6 Convergence of solutions towards the Pareto optimal front 60
3.7 Tricluster Diffusion scores for Dataset 1 62
3.8 Tricluster Diffusion scores for Dataset 2 63
3.9 Tricluster Diffusion scores for Dataset 3 64
3.10 Hit scores using KEGG pathway enrichment for Dataset 1 67
3.11 Hit scores using KEGG pathway enrichment for Dataset 2 68
3.12 Hit scores using KEGG pathway enrichment for Dataset 3 69
3.13 Hit scores using TFBS enrichment for Dataset 1 70
3.14 Hit scores using TFBS enrichment for Dataset 2 71

3.15 Hit scores using TFBS enrichment for Dataset 3 72
3.16 Average Pearson correlation coefficient between key genes and tricluster eigen-gene 73
3.17 Summarization of enriched GOBPs and metabolic pathways 74

4.1 Workflow applied in this chapter 102
4.2 Clustered replicates and time points of the identified modules 103
4.3 Median expression profiles of the genes belonging to the modules 103
4.4 The bow-tie structure of transcriptional networks inferred from the modules 104
4.5 SCC and OUT parts of the transcriptional regulatory networks for Module 2 106
4.6 SCC and OUT parts of the transcriptional regulatory networks for Module 4 107
4.7 SCC and OUT parts of the transcriptional regulatory networks for Module 4 108
4.8 SCC and OUT parts of the transcriptional regulatory networks for Module 5 109
4.9 SCC and OUT parts of the transcriptional regulatory networks for Module 5 110
4.10 Interactions between ZEB2 and SMADs TFs and their roles during the early stage 112
4.11 Interactions between ZEB2 and SMADs TFs and their roles during the late stage 113
4.12 Interactions between ZEB2 and SMADs TFs and their roles during the maturation stage 114

5.1 Schematic diagram of bow-tie structure 128
5.2 Workflow applied in this work 129
5.3 Overlap between the members belonging to each part of the bow-tie structure 131
5.4 Inter-modular switch of transcription factors 132
5.5 The strongly connected components 133
5.6 Feed-forward loop comprising three nodes 136

6.1 Workflow applied in this work 169
6.2 Convergence of solutions towards the Pareto optimal front 171
6.3 Venn diagram showing the overlap between the lists of pathways found to be perturbed 173
6.4 Venn diagram showing the overlap between the lists of pathways found to be perturbed 175

6.5 Venn diagram showing the overlap between the lists of pathways found to be perturbed . . . 178
6.6 Venn diagram showing the overlap between the lists of pathways found to be perturbed . . . 179
6.7 Venn diagram showing the overlap between the lists of pathways found to be perturbed . . . 180

7.1 Workflow used in this chapter . . . 274
7.2 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 1 . . . 275
7.3 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 2 . . . 276
7.4 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 3 . . . 276
7.5 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 4 . . . 277
7.6 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 5 . . . 277
7.7 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 6 . . . 278
7.8 Average Euclidean distances between the expression profiles of each pair of clustered and all replicates over the clustered genes and time points for Dataset 7 . . . 278

List of Tables

2.1 Number of transactions in each month 23
2.2 Comparison between δ-TRIMAX and TRICLUSTER algorithm 30
2.3 Comparison between δ-TRIMAX and TRICLUSTER algorithm in terms of p-values of GOBP and KEGG pathway term enrichment analysis 31
2.4 TRANSFAC Matrices for Triclusters 37

3.1 Values of input parameters of EMOA-δ-TRIMAX in case of the artificial dataset 58
3.2 Values of the input parameters in case of real-life datasets 61
3.3 Number of resultant triclusters, percentage of probe-ids, replicates and time points covered by the resultant triclusters for each of the real-life datasets . 61
3.4 Performance comparison between in terms of SDB score for Dataset 1 . . . 65
3.5 Performance comparison between in terms of SDB score for Dataset 2 . . . 65
3.6 Performance comparison between in terms of SDB score for Dataset 3 . . . 65
3.7 Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during cardiomyocytes differentiation 89
3.8 Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during cardiomyocytes differentiation 90
3.9 Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways since the first heart beat was observed 91
3.10 Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during all stages of hiPSC-derived cardiomyocyte differentiation 92

3.11 Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during maturation 93
3.12 Genes that might be associated with cardiac disorders, cardiovascular development 94
3.13 Genes that might be associated with cardiac disorders, cardiovascular development 95
3.14 Genes that might be associated with cardiac disorders, cardiovascular development 96
3.15 Genes that might be associated with cardiac disorders, cardiovascular development 97
3.16 Genes that might be associated with cardiac disorders, cardiovascular development 98

5.1 Signaling pathways regulated by three node feed-forward loop in the early module 136
5.2 Signaling pathways regulated by three node feed-forward loop in the late module 137
5.3 Signaling pathways regulated by three node feed-forward loop in the early module 146
5.4 Signaling pathways regulated by three node feed-forward loop in the late module 147

6.1 Values of input parameters of EMOA-δ-TRIMAX algorithm 170
6.2 Jaccard similarity coefficient (JSC) between the pathways found to be significant for every pair of datasets 181
6.3 Enriched pathways for over expressed genes for dataset 1 191
6.4 Enriched pathways for under expressed genes for dataset 1 197
6.5 Enriched pathways for over expressed genes for dataset 2 201
6.6 Enriched pathways for under expressed genes for dataset 2 207
6.7 Enriched pathways for over expressed genes for dataset 3 214
6.8 Enriched pathways for under expressed genes for dataset 3 228
6.9 Enriched pathways for over expressed genes for dataset 4 243
6.10 Enriched pathways for under expressed genes for dataset 4 244
6.11 Summarization of the enriched pathways for over expressed genes for all data sets 245

6.12 Summarization of the enriched pathways for under expressed genes for all data sets 258

7.1 Values of $divergence_i$ metric for each of the datasets used in this chapter . 279

Acknowledgement

Foremost, I would like to thank Prof. Dr. Edgar Wingender not only for the supervision of my PhD studies but also for his continuous support, enthusiasm and enormous knowledge without which I could not have finished my PhD thesis.

I would like to express my gratitude to Prof. Dr. Stephan Waack for his help, encouragement and insightful comments on my work carried out during my PhD studies.

Besides, I am very thankful to Martin Haubrock for sharing his work, giving me time for fruitful discussions and helping me from the very first day of joining this department. I thank Torsten Schoeps for providing computer resources which helped me a lot to carry out the experiment during my PhD studies. I would also like to convey my sincere thanks to Dr. Sebastian Zeidler for providing his very valuable suggestions in order to transform the results into biologically interesting and useful outcomes. I am also thankful to Carmen Modrok and Doris Waldmann for helping me in handling the bureaucratic procedures during my studies. Besides, I convey my sincere thanks to all of my present and former colleagues for providing me a convenient atmosphere to work on my PhD thesis.

I acknowledge the financial support from Erasmus Mundus External Cooperation Window, Lot 13 - India - EURINDIA project and ExiTox project funded by German Federal Ministry of Education and Research (BMBF) (Grant no.- 031 A269C).

Finally, I am grateful to my mother and each of my family members for their tremendous support which helped me a lot to reach this stage.

Abstract

Due to the advancement of microarray technology over the last decade, it is feasible to monitor the gene expression dynamics not only over a set of replicates but also either a set of time points or doses of chemical substances. In such three dimensional datasets, variations in the expression profiles can not only be observed across the time points or doses of the chemical substances but also across the replicates due to either abnormalities in the experimental protocol or the physiological variations. Thus, it is important to mine such three dimensional datasets in order to extract biologically meaningful information. In this work, I have proposed a novel triclustering algorithm δ-TRIMAX by introducing a mean squared residue (MSR) score as a coherence measure of the resultant triclusters. The application of this algorithm has been shown in the context of breast cancer progression in order to reveal potential biological processes driving breast cancer invasion. Moreover, I have proposed an improved version of δ-TRIMAX, the EMOA-δ-TRIMAX algorithm which effectively deals with the pitfalls of the former one. One artificial dataset and three real-life datasets have been used to compare the performance of the proposed algorithms with that of other existing algorithms. Besides, the improved version has been applied to one dataset monitoring expression profiles of genes during breast cancer progression for unveiling regulatory mechanisms. Furthermore, the application of the EMOA-δ-TRIMAX algorithm has been demonstrated in investigating the potential biological processes and transcriptional regulatory mechanisms involved in the adolescence of cardiomyocytes. Additionally, I have applied EMOA-δ-TRIMAX algorithm to four real-life datasets in order to provide hints on the pathways perturbed by different toxicants in different tissues. Overall, I could demonstrate that the results of the proposed algorithms for each of the real-life datasets and the artificial ones are promising and provide new insights into the context of breast cancer progression, cardiomyocytes generation and explaining inhalation toxicity.

1 Introduction

1.1 Central Dogma of Molecular Biology

1.1.1 DNA

Figure 1.1 shows the central dogma of molecular biology.

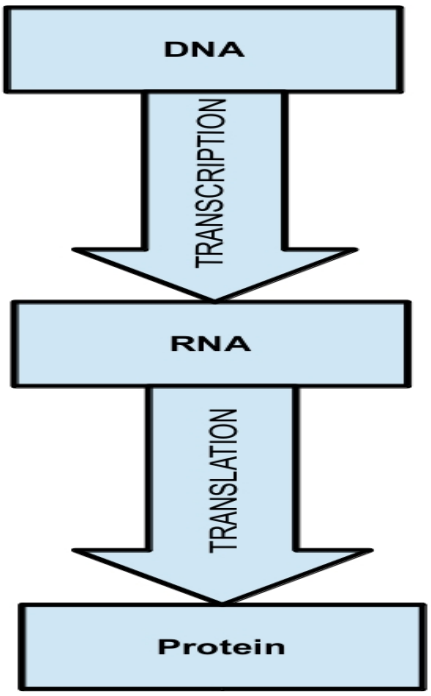

Figure 1.1: The central dogma of molecular biology. Transcription is the processes of making RNA from DNA, whereas RNA is translated into Protein.

Deoxyribonucleic acid (DNA) is a molecule which comprises two biopolymer strands coiled around each other to shape a double helix and encodes the genetic instructions required for the development of all living organisms. The information in DNA is stored as sequence of four bases i.e. adenine(A), cytosine(C), guanine(G) and thymine(T). The double helix of DNA basically comprises two strands which form the helical structure as shown in Figure 1.2. Each of these two strands has two ends which are called 5' and 3' where the 5' and 3' carbon atoms of the deoxyribose are exhibited, respectively. In a DNA double helix, each type of nucleobase located on one strand bonds with just another type of nucleobase on another strand. For instance, adenine bonds with thymine in two hydrogen

bonds, whereas cytosine bonds with guanine in three hydrogen bonds and the arrangement of two nucleotides binding together is called a base pair.

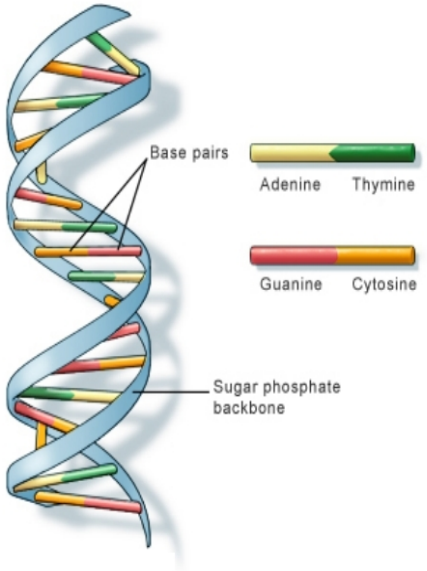

Figure 1.2: The helical structure of DNA. The figure has been taken from [2].

1.1.2 RNA

Ribonucleic acid (RNA) is a large biological constituent which is found in all living organisms and involved in regulation and expression of genes. RNA differs from DNA in terms of base pairing, stability and propagation. For instance, instead of the bonds between adenine and thymine in DNA, adenine bonds with uracil (U) in RNA. Because of the C-H bonds, deoxyribose sugar in DNA is less reactive, whereas ribose sugar in RNA is more reactive because of the hydroxyl group in the 2 position. Besides, DNA is known to be self-replicating, whereas RNA is synthesized from DNA. Messenger RNA (mRNA) plays instrumental roles to transfer genetic information from DNA to ribosome where the synthesis of polypeptides encoded by mRNA occurs, whereas transfer RNA (tRNA) delivers amino acids to the ribosomes and ribosomal RNA (rRNA) forms proteins by linking the

amino acids together. Another type of RNA is a small non-coding RNA also called micro RNA (miRNA). miRNA plays crucial roles in RNA silencing and the post transcriptional regulation of gene expression.

1.1.3 Protein

Proteins are also polymeric molecules consisting of one or more amino acid chains. Proteins are polymers built of 20 different amino acids. The linear sequence of amino acids are referred as primary structure of the proteins. Alpha helix and beta pleated-sheets are the most common types of secondary structure of proteins. The patterns of hydrogen bonds located between the main-chain peptide groups delineate these secondary structures. Moreover, the three dimensional folding of a protein is referred as the tertiary protein structure, whereas the association of two or more identical or different polypeptides forms the quaternary protein structure. Proteins play essential roles in any biological process including proper development of tissues and organs in the body, maintaining the metabolism of cells, mediating the communication between cells, etc.

1.1.4 Transcription

Transcription is the process of producing RNA from DNA. Transcription is usually accomplished by the steps described as follows. First a number of proteins called transcription factors (TFs) bind to specific DNA sequences. This facilitates the assembly of general transcription factors around TATA box and transcription start site (TSS) and the recruitment of RNA polymerase. Subsequently, RNA is produced from the transcribed region of a gene. Other proteins such as coactivators, corepressors also assist TFs in the regulation of a gene product either by enhancing or diminishing the rate of transcription, respectively. The region around the TSS, the TATA box and comprising all the nearby TF binding sites (TFBSs) constitute the promoter of a gene. Up to several mega bases away, either upstream or downstream, are additional control regions such as enhancers. They also consist of arrays of TFBSs and influence the efficiency of the promoter.

During transcription of the synthesis of the RNA occurs in the direction of 5' to the 3' as shown in Figure 1.3 [1] using the template strand of DNA for arranging the proper sequence of ribonucleotides according to the principles of base pairing. The coding strand of DNA is the complementary copy of the template strand and used to represent the sequence of a gene.

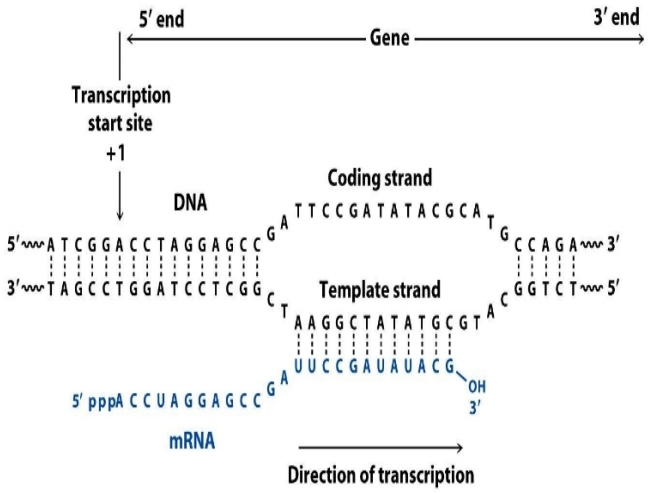

Figure 1.3: Orientation of transcription. The figure has been taken from [1].

1.1.5 Translation

In case of protein-coding genes, the transcript is an mRNA molecule.

RNA Splicing

The process of modifying the pre-messenger RNA (pre-mRNA) transcript through the deletion of introns and the joining of exons is referred as RNA splicing. This process is essential for the mRNA prior to its translation into a protein.

Attaching to Ribosomes

First, RNA from the nucleus enters the ribosome and attaches to the ribosomes in order to produce protein.

Polypeptide Synthesis

The process of producing long peptides or proteins is called protein biosynthesis.

1.2 Microarray Technology

Microarray technology is used to monitor the changes in the expression profiles of mRNA across a variety of experimental conditions. A microarray is a 2D array which consists of a solid substrate and the probes are attached in fixed positions. Based on the ways how probes are placed on the solid substrate, microarray can be of two different types.

1.2.1 cDNA Microarray

In case of a cDNA microarray (Figure 1.4), complementary DNA copies of mRNA transcripts are used and labeled with fluorescent dyes of two different colors, namely red and green. The red colored fluorescent dye is used to label the treated libraries, whereas the green colored dye is used to label the control or untreated ones. Once the hybridization of these two libraries is finished, the fluorescent dye is excited by a dual channel laser. The fluorescent intensity is used to represent the degree of hybridization occurred. Then the ratio of two fluorescent wavelengths is computed to measure the gene expression.

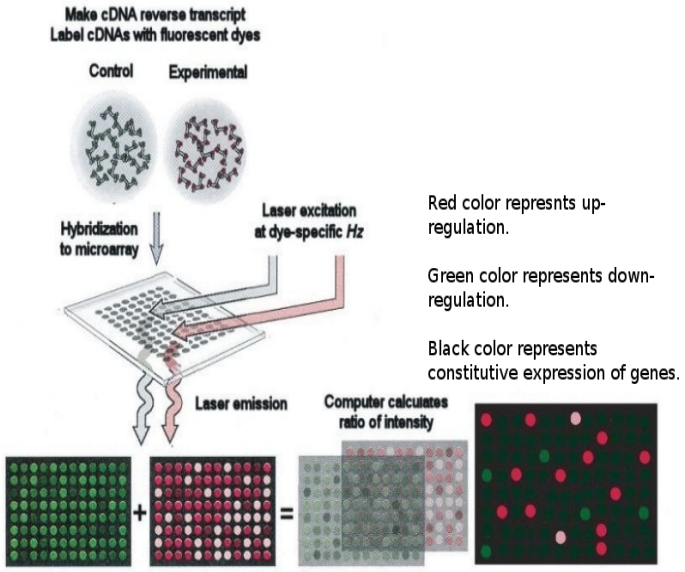

Figure 1.4: The workflow of cDNA microarray to measure gene expression. The figure has been taken from [3].

1.2.2 Oligonucleotide Microarray

These chips are manufactured by companies like Affymetrix, Illumina, Agilent etc. Figure 1.5 shows the workflow of an Affymetrix microarray experiment to monitor gene expression profiles. Oligonucleotide refers to the process of creating short sequences that are complementary to parts of a gene. Thus a gene can be represented by one or several probes. Here the oligonucleotides are synthesized by a photolithographic process. Once the sample has been hybridized to a chip, an Affymetrix scanner is used to quantify the fluorescence intensity at each spot. Evaluation of the readings is done by Affymetrix software which usually produces several files of different extensions such as EXP, DAT, CEL, CDF, CHP. An ".EXP" file contains information about the experiment. ".DAT" file stores the raw image of the scanned array and pixel values of the image are read to compute the intensities which are stored in the ".CEL" file. The content of the ".CHP" file is the gene expression levels, estimated by the Affymetrix software. ".CDF" file contains information

needed for mapping between features, probe-sets, genes.

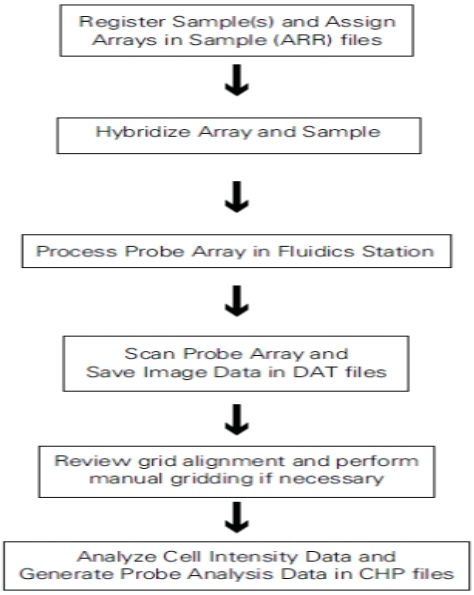

Figure 1.5: The workflow of an Affymetrix microarray experiment. The figure has been taken from [4].

1.3 Computational Analysis of Microarray Gene Expression Data

1.3.1 Differential Expression Analysis

Differential expression analysis is performed in order to identify genes which are up- or down-regulated between two experimental conditions such as untreated vs treated. Statistical methods such as ordinary t-test, Bayesian methods, rank-sum statistics, moderate t-test (used in linear models for microarray (limma) package available in R) are used to extract the groups of differentially expressed genes [5]. The ordinary t-test is able to consider only one gene or probe to test the differential expression, whereas limma can take all genes/ probes into consideration at each condition at a time. In case of each gene expression profile, limma fits a linear model, coefficients of which represent the differences between the sources of RNA hybridized to the arrays.

1.3.2 Machine Learning in Mining Microarray Gene Expression Data

In machine learning, the main task of supervised learning technique is to train a classifier using a set of training examples categories of which is known for inferring a function which can later be used to classify new examples properly. To evaluate the accuracy of the classifier K-fold cross validation is usually performed where the examples are divided into K subsets and during each iteration one subset is used as test, whereas (K-1) subsets of examples are used as training set [6]. One of the application of supervised learning in the field of bioinformatics is to assign genes into proper classes for instance, different tumor stages.

In contrast with the supervised learning technique, unsupervised learning refers to finding the proper categories of examples without having any prior knowledge about the categories of any one of the examples [6]. One of the applications of unsupervised learning in mining the microarray gene expression data is to find the associations between a set of genes which in turns may provide insights into the biological processes in which the groups of associated genes play major roles. In other words, unsupervised learning may facilitate to identify groups of co-expressed genes.

Role of Co-expression Analysis in the Field of Biology

Developmental Biology

Elucidating the biological processes that play crucial roles in the development of organs or tissues for instance, differentiation of a stem cell into cardiomyocytes, is the purpose of studying developmental biology. Disruption of these biological processes may lead to the developmental disorders which in turn cause disease like cardiac disorders etc. Either a single gene or a set of genes interacting with each other can act as driver genes which participate in the developmental processes. Moreover, revealing the genes that drive phenotypic changes during the development may also provide hints on potential therapeutic targets. Co-expression analysis thus may play major roles in unraveling the interactions between such driver genes in the context of a particular experimental study of interest [7–9].

Disease Biology

Co-expression analysis may also be beneficial in unraveling the interaction between key genes which mediate the potential biological processes that trigger the progression of diseases such as breast cancer etc. [10–12]. Any disease may be considered as a specific cellular state, alternative to the normal state of a cell (or a whole organ). Thus, having the full

gene activity profiles of a normal and a disease state available may enable us to characterize these states, identify the transition between them and to give hints on how to interfere with the transition and to prevent the development of a disease. These characterizations can be done by simply comparing gene activity lists or, more sophisticated, by constructing and analyzing the underlying regulatory networks.

Inference of Gene Regulatory Network

A gene regulatory network can be viewed as a mixed graph G=$\{V, U, D\}$, where V denotes a set of vertices each of which represents a gene, whereas U and D represent sets of undirected and directed edges, respectively. An undirected edge may refer to the association between genes in terms of their expression similarity. Each directed edge may refer to the causal relationship between two nodes which is a regulatory relationship between a transcription factor and a target gene [13]. Co-expression analysis facilitates in deducing the undirected edges, whereas the transcription factor binding site (TFBS) information may help us in inferring the directed edges in gene regulatory network. Thus integrating both TFBS and microarray gene expression data may assist in inferring the gene regulatory network where the co-expressed genes are co-regulated by a set of transcription factor as shown in Figure 1.6.

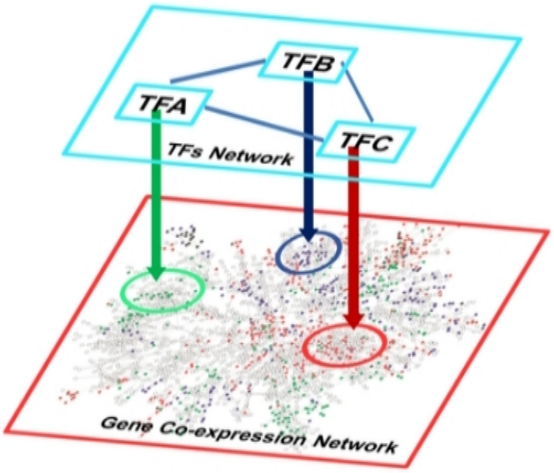

Figure 1.6: Co-regulation of co-expressed genes by a set of transcription factors. The figure has been taken from [14].

1.3.3 Co-expression and Cluster Analysis

To carry out this exploratory analysis, several clustering algorithms have been proposed to identify the groups of genes that have similar expression profiles over all experimental conditions or replicates based on some similarity or dissimilarity metrics. Bayesian Network (BN) and Dynamic Bayesian Network (DBN) are also used to perform this kind of exploratory analysis on microarray gene expression data [6]. By estimating the conditional dependencies among the genes, it can provide insightful knowledge about the gene interaction networks. In spite of having their own merit in mining 2D gene expression data, the clustering algorithms are not capable of grouping genes and experimental conditions/ replicates simultaneously. To cope with the problem, biclustering algorithms are used. Several biclustering algorithms have been propounded over the last decade. The advances in microarray technology over the past decade facilitate to monitor the expression profiles of thousands of genes not only over a set of replicates but also across a set of time

points or substances. These replicates may either be technical or biological. For instance, if the expression profiles of thousands of genes are measured from one source for a certain number of times, then the replicates are said to be technical, whereas if the expression profiles of thousands of genes are measured from different sources such as, patients etc. then the replicates are treated as biological ones. It is quite obvious that the variations in the expression profiles of genes over different biological replicates are present in such 3D gene expression data. However, the expression profiles of genes may also vary over the technical replicates in such datasets. Hence, it is required to mine all these three dimensions in order to extract potential biologically meaningful information from such datasets. To accomplish this goal, triclustering algorithms are used. A triclustering algorithm is capable of grouping genes, replicates and time points/ substances simultaneously. Zhao et al. proposed the TRICLUSTER algorithm to retrieve groups of genes that have similar expression profiles over a subset of replicates and across a subset of time points [15]. In a recent work, Tchagang et al. proposed a triclustering algorithm (OPTricluster) in order to mine short time series gene expression datasets. OPTricluster effectively mines time series gene expression data that have approximately 3-8 time points and 2-5 replicates. According to their definition of a tricluster, genes belonging to a tricluster must have constant, coherent or order preserving expression patterns over a subset of replicates during a subset of time points. In case of an order-preserving tricluster, there must be a permutation of the time points such that expression levels of genes form a monotonic function [16]. However, the OPTricluster algorithm is not computationally efficient to mine the long time series gene expression datasets. Furthermore, as the OPTricluster algorithm yields triclusters each of which always comprises all the time points, this algorithms is not applicable to the datasets which provide the expression profiles of genes measured during the exposure of a cell to the different doses of chemical substances. During my PhD studies, we have proposed a triclustering algorithm δ-TRIMAX and an improved version of δ-TRIMAX namely, EMOA-δ-TRIMAX (Evolutionary Multiobjective Optimization Algorithm for δ-TRIMAX) and both of these two algorithms use a novel coherence measure namely, mean squared residue (MSR) in order to mine three dimensional gene expression dataset.

1.4 Structure of the Thesis

The thesis has been structured in the following way.
Chapter 2 provides the mathematical background, definition of the mean squared residue

(MSR) and the steps of δ-TRIMAX triclustering algorithm. Moreover, it provides an application of the δ-TRIMAX algorithm in elucidating the biological processes involved in the progression of breast cancer.
In Chapter 3, we describe the pitfalls of δ-TRIMAX algorithm and provided the details of its improved version, the EMOA-δ-TRIMAX algorithm. Furthermore, this chapter shows an application of EMOA-δ-TRIMAX algorithm in investigating the biological processes and the corresponding key genes involved in the phenotypic divergence during stem cell differentiation into cardiomyocytes.
Being motivated by the biological processes of development obtained in the previous chapter, in Chapter 4, we have applied the EMOA-δ-TRIMAX algorithm to the same dataset in order to investigate the transcriptional regulatory mechanisms during the adolescence of cardiomyocytes.
Being inspired by the promising results in the context of revealing transcriptional regulatory mechanisms, in Chapter 5, we have used EMOA-δ-TRIMAX in order to unveil the transcriptional regulatory mechanisms governing the progression of a breast cancer cell exposed to estrogen.
To show the capability of EMOA-δ-TRIMAX algorithm in mining other types gene expression dataset, Chapter 6 emphasizes an application of EMOA-δ-TRIMAX in the context of toxicology i.e. we have applied the algorithm to four different datasets which comprise expression profiles of genes measured during the exposure of different tissues to different doses of chemical substances in order to reveal the lists of perturbed signaling pathways.
Chapter 7 emphasizes the importance of clustering replicates in three dimensional gene expression datasets.
Finally, Chapter 8 summarizes the conclusions of the individual applications.

1.5 Bibliography

[1] Moran, L.: Gene and Transcription Orientation. `http://sandwalk.blogspot.de/2007/03/gene-and-transcription-orientation.html`

[2] What is DNA? `http://ghr.nlm.nih.gov/handbook/basics/dna`. Published: September 8, 2014

[3] Principle of cDNA microarray assay of gene expression. `https://www.mun.ca/biology/scarr/cDNA_microarray_Assay_of_Gene_Expression.html`. Figure after Gibson Muse 2002

[4] *Affymetrix*: *Affymetrix GeneChip* Command *Console* 3.0 User Manual. (2009). *Affymetrix*

[5] Suarez, E., Burguete, A., Mclachlan, G.J.: Microarray data analysis for differential expression: a tutorial. Puerto Rico Health Sciences Journal **28**(2), 89–104 (2009)

[6] Molla, M., Waddell, M., Page, D., Shavlik, J.: Using Machine Learning to Design and Interpret Gene-expression Microarrays. A I Magazine **25**(1), 23–44 (2004)

[7] Saris, C.G., Horvath, S., van Vught, P.W., van Es, M.A., Blauw, H.M., Fuller, T.F., Langfelder, P., DeYoung, J., Wokke, J.H., Veldink, J.H., van den Berg, L.H., Ophoff, R.A.: Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. BMC Genomics **10**(405) (2009). doi:10.1186/1471-2164-10-405

[8] Min, J.L., Nicholson, G., Halgrimsdottir, I., Almstrup, K., Petri, A., Barrett, A., Travers, M., Rayner, N.W., Maegi, R., Pettersson, F.H., Broxholme, J., Neville, M.J., Wills, Q.F., Cheeseman J; GIANT Consortium; MolPAGE, C., Allen, M., Holmes, C.C., Spector, T.D., Fleckner, J., McCarthy, M.I., Karpe, F., Lindgren, C.M., Zondervan, K.T.: Coexpression network analysis in abdominal and gluteal adipose tissue reveals regulatory genetic loci for metabolic syndrome and related phenotypes. PLoS Genetics **8**(2), e1002505 (2012). doi:10.1371/journal.pgen.1002505

[9] de Jong, S., Boks, M.P., Fuller, T.F., Strengman, E., Janson, E., de Kovel, C.G., Ori, A.P., Vi, N., Mulder, F., Blom, J.D., Glenthoj, B., Schubart, C.D., Cahn, W., Kahn, R.S., Horvath, S., Ophoff, R.A.: A gene co-expression network in whole blood of schizophrenia patients is independent of antipsychotic-use and enriched for brain-expressed genes. PLoS One **7**(6), e39498 (2012). doi:10.1371/journal.pone.0039498

[10] Reedijk, M., Odorcic, S., Chang, L., Zhang, H., Miller, N., McCready, D.R., Lockwood, G., Egan, S.E.: High-level coexpression of JAG1 and NOTCH1 is observed in human breast cancer and is associated with poor overall survival. Cancer Research **65**(18), 8530–8537 (2005)

[11] Wolf, D.M., Lenburg, M.E., Yau, C., Boudreau, A., van 't Veer, L.J.: Gene co-expression modules as clinically relevant hallmarks of breast cancer diversity. PLoS One **9**(2), e88309 (2014). doi:10.1371/journal.pone.0088309

[12] Penrod, N.M., Moore, J.H.: Influence networks based on coexpression improve drug target discovery for the development of novel cancer therapeutics. BMC Systems Biology **12**, 12 (2014). doi:10.1186/1752-0509-8-12

[13] de la Fuente, A.: What are Gene Regulatory Networks? In: Das, S., Caragea, D., Welch, S., Hsu, W. (eds.) Handbook of Research on Computational Methodologies in Gene Regulatory Networks, pp. 1–27. IGI Global, The United States of America (2010). doi:10.4018/978-1-60566-685-3.ch001

[14] Ma, S., Shah, S., Bohnert, H.J., Snyder, M., Dinesh-Kumar, S.P.: Incorporating motif analysis into gene co-expression networks reveals novel modular expression pattern and new signaling pathways. PLoS Genetics **9**(10), e1003840 (2013). doi:10.1371/journal.pgen.1003840

[15] Zhao, L., Zaki, M.J.: triCluster: An Effective Algorithm for Mining Coherent Clusters in 3D Microarray Data. In: In Proc. of the 2005 ACM SIGMOD International Conference on Management of Data, pp. 694–705 (2005). ACM Press

[16] Tchagang, A.B., Phan, S., Famili, F., Shearer, H., Fobert, P., Huang, Y., Zou, J., Huang, D., Cutler, A., Liu, Z., Pan, Y.: Mining biological information from 3D short time-series gene expression data: the OPTricluster algorithm. BMC Bioinformatics **13**, 54 (2012). doi:10.1186/1471-2105-13-54

2 A New Triclustering Approach for Unveiling Biological Processes of Disease Progression from Gene Expression Profiles

2.1 Introduction

Estrogen a chemical messenger, plays an instrumental role in the development of many tissues such as breast, uterus etc. through binding to the estrogen receptor. In the context of breast cancer originates in the breast tissue, the abnormal cells grow and divide i.e. cell proliferation is observed. Estrogen enhances the risk of breast cancer by triggering cellular proliferation. Moreover, cellular proliferation is also controlled by nutrients. Thus oncogenes which play crucial roles in promoting cellular proliferation may also be involved in metabolic alteration. Hence, understanding the metabolic and signaling pathways nay facilitate the development of novel therapeutic strategies for the breast cancer treatment. In this chapter, we have applied our proposed triclustering algorithm to a dataset in which the expression profiles of genes were measured during different stages of estrogen's effects on a MCF-7 breast cancer cell. Our results not only provide novel insights into the metabolic pathways most of which are already known to be associated with breast cancer but also elucidate the regulatory roles of several transcription factors in estrogen induced breast cancer cell.

2.2 Materials and Methods

2.2.1 Materials

Artificial dataset: We have used one simulated dataset of size $2000 \times 30 \times 30$. At first we have inserted three perfect shifting triclusters of size $100 \times 6 \times 6$, $80 \times 6 \times 6$ and $60 \times 5 \times 5$ into the simulated dataset and then implanted three noisy shifting triclusters having the same size and different levels of noise i.e. standard deviations (σ = 0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7) into the dataset [1, 2].

Real-life dataset: This publicly available gene expression data (dataset id- GSE 11324) contains 54675 affymetrix probe-set ids, 3 biological replicates and 4 time points. In this experiment, MCF7 cells were stimulated with 100 nm estrogen during 0, 3, 6 and 12 hours and the expression values are measured in triplicate. The experiment aims at discovering the cis-regulatory sites in previously uninvestigated regions and cooperating transcription factors underlying estrogen signaling in breast cancer [3].

2.2.2 Methods: δ-TRIMAX

Definitions

Definition 1 (Time Series Microarray Gene Expression Dataset) *A time series gene expression data can be considered as a $G \times C \times T$ matrix, each element of which (d_{ijk}) represents the expression value of the gene i over the jth sample/experimental condition across the time point k and $i \in (g_1,g_2,...,g_G)$, $j \in (c_1,c_2,...,c_C)$, $k \in (t_1,t_2,...,t_T)$ [1, 2].*

Definition 2 (Tricluster) *A tricluster can be delineated as a sub-matrix M(I,J,K) = [m_{ijk}], where $i \in I$, $j \in J$ and $k \in K$. The sub-matrix M corresponds to a subset of genes (I) that are co-expressed over a subset of conditions (J) across a subset of time points (K) [1, 2].*

Definition 3 (Perfect Shifting Tricluster) *A Tricluster M(I,J,K) = [m_{ijk}], where $i \in I$, $j \in J$ and $k \in K$, is a perfect shifting tricluster if each element of the sub-matrix M can be formulated as: $m_{ijk} = \Gamma + \alpha_i + \beta_j + \eta_k$, where Γ corresponds to a constant value for the tricluster, α_i, β_j and η_k represent shifting factors of the ith gene, the jth samples/experimental condition and the kth time point, respectively [1, 2].*

But in case of a noisy microarray gene expression dataset, the deviation from actual value and expected value of each element is usually observed. This deviation may prevent retrieving a perfect shifting tricluster from a gene expression dataset.

In this work, we present a novel definition of Mean Squared Residue (MSR) score for 3D microarray gene expression datasets. The MSR score we defined here, is a generalization of the one proposed by Cheng and Church [4, 5]. The MSR of a perfect shifting tricluster becomes zero, where each element, $m_{ijk} = \Gamma + \alpha_i + \beta_j + \eta_k$. To delineate the MSR score for 3D gene expression datasets, at first we need to define the residue score as follows:

Let the mean of ith gene (m_{iJK}): $m_{iJK} = \frac{1}{|J||K|} \sum_{j \in J, k \in K} m_{ijk}$, the mean of jth sample/experimental condition (m_{IjK}): $m_{IjK} = \frac{1}{|I||K|} \sum_{i \in I, k \in K}$
m_{ijk}, the mean of kth time point (m_{IJk}): $m_{IJk} = \frac{1}{|I||J|} \sum_{i \in I, j \in J} m_{ijk}$, and the mean of tricluster (m_{IJK}): $m_{IJK} = \frac{1}{|I||J||K|} \sum_{i \in I, j \in J, k \in K} m_{ijk}$. Now the mean of the tricluster can be considered as the value of constant i.e. $\Gamma = m_{IJK}$. We can define the shifting factor for the ith gene (α_i) as the difference between m_{iJK} and m_{IJK} i.e. $\alpha_i = m_{iJK} - m_{IJK}$ [1, 2]. Similarly, we can define shifting factor for the jth condition (β_j) as $\beta_j = m_{IjK} - m_{IJK}$ and shifting factor for the kth time point (η_k) can be defined as $\eta_k = m_{IJk} - m_{IJK}$. Hence,

we can define each element of a perfect shifting tricluster as $m_{ijk} = \Gamma + \alpha_i + \beta_j + \eta_k = m_{IJK}+(m_{iJK}-m_{IJK})+(m_{IjK}-m_{IJK})+(m_{IJk}-m_{IJK}) = (m_{iJK}+m_{IjK}+m_{IJk}-2m_{IJK})$.

As noise is discernible in microarray gene expression dataset, we need to compute the difference between the actual value of an element (m_{ijk}) and its expected value, obtained from the above equation. This deviation can be termed as *"residue"* [5]. Thus, the residue of a tricluster (r_{ijk}) can be formulated as follows: $r_{ijk} = m_{ijk} - (m_{iJK} + m_{IjK} + m_{IJk} - 2m_{IJK}) = (m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})$.

Definition 4 (Mean Squared Residue) *We delineate the Mean Squared Residue MSR(I,J,K) of a tricluster M(I,J,K) to measure the quality of a tricluster i.e. the level of coherence among the elements of a tricluster as follows [1, 2]:*

$$\begin{aligned} MSR &= \frac{1}{|I||J||K|} \sum_{i\in I, j\in J, k\in K} r_{ijk}^2 \\ &= \frac{1}{|I||J||K|} \sum_{i\in I, j\in J, k\in K} (m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 \end{aligned} \quad (2.1)$$

Lower residue score represents larger coherence and better quality of a tricluster.

Theoretical Background of the Proposed Mean Squared Residue Score

Table 2.1 shows the numbers of transactions in each month. Now, no straight line will pass exactly through the four blue colored points shown in Figure 2.1. However, there are many straight lines that will pass close to the four blue colored points. For instance one of them is represented by the red colored line and the values are P = {62, 70, 78, 86}. In a linear regression model, we could consider the set P as predicted values. So the better choice of line would be the one which is closer to the observed values (minimize the value of r).
As real-life gene expression datasets are noisy, it is hard to find the perfect additive tricluster. So we defined a novel MSR score that aims to minimize the deviation between observed value (noisy) and the computed value (perfect additive tricluster). Note that MSR does not use any distance measure directly. It only computes the deviation of a detected tricluster from a ideal tricluster and we try to minimize this deviation.

Table 2.1: Number of transactions in each month

Month	**Number of transactions in this month (Observed) (O)**	**Number of transactions (Predicted) (P)**	**Residue** $(r = O - P)$
January	68	62	6
February	72	70	2
March	80	78	2
April	83	86	-3

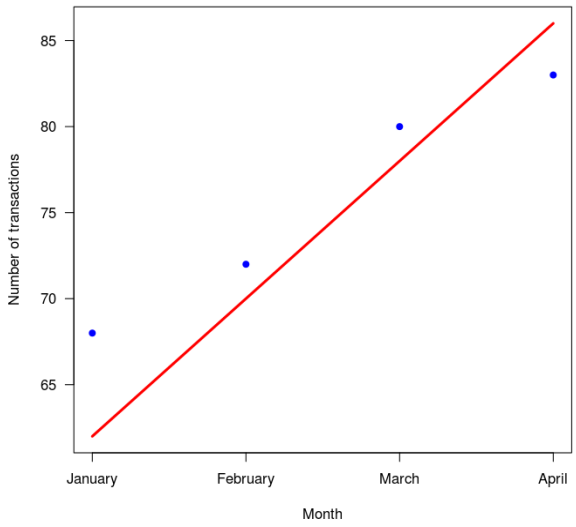

Figure 2.1: Number of transactions in each month. Blue colored points and red colored line represent the observed and predicted numbers of transactions, respectively.

Proposed Method

δ-TRIMAX aims at finding the largest and maximal triclusters in a 3D microarray gene expression dataset [1, 2]. It is a generalization of the biclustering algorithm proposed by Cheng and Church [4] that copes with 2-D microarray datasets. In contrast, our algorithm is capable to mine 3D gene expression dataset. There is always a sub-matrix in an expression dataset that has a perfect MSR(I,J,K) = 0 and this sub-matrix is each element of

the dataset [1,2]. But as mentioned above, our algorithm finds maximal triclusters having MSR score under a threshold δ, hence we have used a greedy heuristic approach to find triclusters. Our algorithm therefore starts with the entire dataset containing all genes, all samples/experimental conditions and all time points.

Algorithm I (δ-TRIMAX):

Input. D, a matrix that represents 3D microarray gene expression dataset, $\lambda > 1$, an input parameter for multiple node deletion algorithm, $\delta \geq 0$, maximum allowable MSR score.

Output. All possible δ-triclusters.

Initialization. Missing elements in D $\leftarrow$ random numbers, D' $\leftarrow$ D

Repeat a. D'$_1$ $\leftarrow$ Results of Algorithm II on D' using *delta* and λ. If the no. of genes (conditions/samples and/or no. of time points) is 50 (This value can be choosen experimentally. Large value increases the execution time of the algorithm as it then executes more number of iterations.), then do not apply Algorithm II on genes (conditions/samples and/or time points).

b. D'$_2$ $\leftarrow$ Results of Algorithm III on D'$_1$ using δ.

c. D'$_3$ $\leftarrow$ Results of Algorithm IV on D'$_2$.

d. Return D'$_3$ and replace the elements that exist in D' and D'$_3$ with random numbers.

Until(No gene is found for δ-tricluster)

Initially, our algorithm removes genes or conditions or time points from the dataset to effectuate largest diminishing of the MSR score; this step is described in the following section in which a node represents a gene or an experimental condition or a time point in the 3D microarray gene expression dataset.

Algorithm II (Multiple node deletion):

Input. D, a matrix of real numbers that represents 3D microarray gene expression dataset; $\delta \geq 0$, maximum allowable MSR threshold, $\lambda > 1$, threshold for multiple node deletion. The value of λ has been set experimentally to optimize the speed and performance (to avoid falling into local optimum) of the algorithm.

Output. M_{IJK}, a δ-tricluster, consisting of a subset(I) of genes, a subset(J) of samples/ experimental conditions and a subset of time points, having MSR score less than or equal to δ.

Initialization. I $\leftarrow$ {set of all genes}, J $\leftarrow$ {set of all experimental conditions/ samples} and K $\leftarrow$ {set of all time points} and to M(I,J,K) $\leftarrow$ D

Repeat Calculate m_{iJK}, $\forall$ i $\in$ I; m_{IjK}, $\forall$ j $\in$ J; m_{IJk}, $\forall$ k $\in$ K; m_{IJK} and MSR.

If MSR $\leq \delta$ return M(I,J,K)
Else Delete genes i $\in$ I that satisfy the following inequality

$$\frac{1}{|J||K|}\Sigma_{j\in J,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 > \lambda MSR$$

Recalculate m_{iJK}, $\forall$ i $\in$ I; m_{IjK}, $\forall$ j $\in$ J; m_{IJk}, $\forall$ k $\in$ K; m_{IJK} and MSR Delete samples/experimental conditions j $\in$ J that satisfy the following inequality

$$\frac{1}{|I||K|}\Sigma_{i\in I,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 > \lambda MSR$$

Recalculate m_{iJK}, $\forall$ i $\in$ I; m_{IjK}, $\forall$ j $\in$ J; m_{IJk}, $\forall$ k $\in$ K; m_{IJK} and MSR Delete time points k $\in$ K that satisfy the following inequality

$$\frac{1}{|I||J|}\Sigma_{i\in I,j\in J}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 > \lambda MSR$$

End if
Until(There is no change in I, J and/or K)

In the second step, we delete one node at each iteration from the resultant sub-matrix, produced by Algorithm II, until the score MSR of the resultant sub-matrix is less than or equal to δ. This step yields a δ-tricluster.

Algorithm III (Single node deletion):
Input. D, a matrix of real numbers that represents 3D microarray gene expression dataset; $\delta \geq 0$, maximum allowable MSR threshold.
Output. M_{IJK}, a δ-tricluster, consisting of a subset(I) of genes, a subset(J) of samples/experimental conditions and a subset of time points, having MSR score less than or equal to δ.
Initialization. I $\leftarrow$ {set of all genes in D}, J $\leftarrow$ {set of experimental conditions/samples in D} and K $\leftarrow$ {set of time points in D} and to M(I,J,K) $\leftarrow$ D

Calculate m_{iJK}, $\forall$ i $\in$ I; m_{IjK}, $\forall$ j $\in$ J; m_{IJk}, $\forall$ k $\in$ K; m_{IJK} and MSR.
While MSR $> \delta$ Detect gene i $\in$ I that has the highest score

$$\mu(i) = \frac{1}{|J||K|}\Sigma_{j\in J,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2$$

Detect sample/experimental condition j ∈ J that has the highest score

$$\mu(j) = \frac{1}{|I||K|}\Sigma_{i\in I,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2$$

Detect time point k ∈ K that has the highest score

$$\mu(k) = \frac{1}{|I||J|}\Sigma_{i\in I,j\in J}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2$$

Delete gene or sample/experimental condition or time point that has highest μ score and modify I or J or K. Recalculate m_{iJK}, ∀ i ∈ I; m_{IjK}, ∀ j ∈ J; m_{IJk}, ∀ k ∈ K; m_{IJK} and MSR.

End while

Return M(I,J,K)

As the goal of our algorithm is to find the maximal triclusters, having MSR score below the threshold δ, the resultant tricluster M(I,J,K) may not be the largest one. That means some genes and/or experimental conditions/samples and/or time points may be included in the resultant tricluster T produced by the node deletion algorithm, so that the MSR score of new tricluster T' produced after node addition does not exceed the MSR score of T.

Now the third step of our algorithm is described below.

Algorithm IV (Node addition):

Input. D, a matrix of real numbers that represents δ-tricluster, having a subset of genes (I), a subset of experimental conditions/samples (J) and a subset of time points (K).

Output. $M_{I'J'K'}$, a δ-tricluster, consisting of a subset of genes (I') , a subset of samples/experimental conditions (J') and a subset of time points (K'), such that I ⊂ I', J ⊂ J', K ⊂ K' and MSR(I',J',K') ≤ MSR of D.

Initialization. M(I,J,K) ← D

Repeat Calculate m_{iJK}, ∀ i; m_{IjK}, ∀ j; m_{IJk}, ∀ k; m_{IJK} and MSR. Add genes i ∉ I that satisfy the following inequality

$$\frac{1}{|J||K|}\Sigma_{j\in J,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 \leq MSR$$

Recalculate m_{IjK}, ∀ j; m_{IJk}, ∀; m_{IJK} and MSR Add samples/experimental conditions j ∉

J that satisfy the following inequality

$$\frac{1}{|I||K|}\Sigma_{i\in I,k\in K}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 \leq MSR$$

Recalculate m_{iJK}, $\forall$ i; m_{IJk}, $\forall$ k; m_{IJK} and MSR Add time points k $\notin$ K that satisfy the following inequality

$$\frac{1}{|I||J|}\Sigma_{i\in I,j\in J}(m_{ijk} - m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 \leq MSR$$

Add genes i $\notin$ I that satisfy the following inequality

$$\frac{1}{|J||K|}\Sigma_{j\in J,k\in K}(-m_{ijk} + m_{iJK} - m_{IjK} - m_{IJk} + 2m_{IJK})^2 \leq MSR$$

Until(There is no change in I, J and/or K)
I' $\leftarrow$ I, J'$\leftarrow$ J, K' $\leftarrow$ K
Return I', J', K'

Complexity of the Proposed Algorithm

The complexity of Algorithm II is O(max(m,n,p)), where m, n and p correspond to the number of genes, samples and time points in the 3D microarray dataset. The complexity of first and second steps of Algorithm III is O(mnp) as those will iterate (m+n+p) times. The complexity of selection of best genes, samples and time points is O(log m + log n + log p). So it is recommended to use the Algorithm II before Algorithm III. The complexity of Algorithm IV is O(mnp) as each step iterates (m+n+p) times [1, 2].

Tricluster Eigen-gene

We applied singular value decomposition method (SVD) on the expression data of each tricluster to find the tricluster eigen-gene [6]. For instance, $X^i_{g\times(c*t)}$ corresponds to the expression matrix of ith tricluster, where g, c and t represent the number of genes, samples and time points of ith tricluster. Now we apply SVD on the data matrix (normalized to mean=0 and variance=1) and the SVD of ith tricluster can be computed as,

$$X^i = UDV^T, \tag{2.2}$$

where U and V are the orthogonal matrices. U^i is a g $*$ (c $*$ t) matrix with orthonormal columns, V^i is a (c $*$ t) $\times$ (c $*$ t) orthogonal matrix and D^i is (c $*$ t) $\times$ (c $*$ t) diagonal matrix of singular values.

Assuming that singular values of matrix D^i are arranged in non-decreasing order, we can represent eigen-gene of the ith tricluster by the first column of matrix V^i, i.e.

$$E^i = V_1^i, \tag{2.3}$$

2.3 Results and Discussion

2.3.1 Results on Simulated Dataset

To measure the degree of similarity between the implanted and produced triclusters, we delineate the *affirmation score* in the same way as Prelic et. al. defined for two sets of biclusters [5, 7]. Suppose, we have two sets of triclusters T_{im} and T_{res} where T_{im} is the set of implanted triclusters and T_{res} corresponds to the set of triclusters retrieved by any triclustering algorithm. So, overall average affirmation score of T_{im} with respect to T_{res} is as follows [1, 2], where ($\mathrm{SM^*}_G(\mathrm{T}_{im}, \mathrm{T}_{res})$) is the average gene affirmation score, ($\mathrm{SM^*}_C(\mathrm{T}_{im}, \mathrm{T}_{res})$) is the average sample affirmation score and ($\mathrm{SM^*}_K(\mathrm{T}_{im}, \mathrm{T}_{res})$) is the average time point affirmation score of T_{im} with respect to T_{res}:

$$SM^*(T_{im}, T_{res}) = \sqrt{(SM^*_G(T_{im}, T_{res}) \times SM^*_C(T_{im}, T_{res}) \times SM^*_T(T_{im}, T_{res}))} \tag{2.4}$$

Hence, $\mathrm{SM}^*(T_{im}, T_{res})$ denotes how well the triclustering algorithm finds the implanted triclusters from the dataset. This score ranges from 0 to 1 (if $\mathrm{T}_{im} = \mathrm{T}_{res}$). To compute the value of δ, we have first clustered the genes over all time points and then the time points over the subset of genes for each gene cluster in each sample plane using the K-means algorithm. Then we have measured the MSR of the sub-matrix, considering a randomly selected sample plane, gene and time-point cluster for 100 times. Then we have taken the lowest value as the value of δ [1, 2]. For the dataset containing perfect shifting triclusters, the parameters δ and λ are set to 0.35 and 1.0005, respectively whereas in case of the noisy datasets, we have assigned 3.75 and 1.004 to the parameters δ and λ, respectively [1, 2]. Figure 2.2 shows the comparison between the performance of our algorithm with that of the *TRICLUSTER* algorithm [8] in terms of affirmation score using the artificial dataset.

Our δ-TRIMAX algorithm outperforms *TRICLUSTER* algorithm for the noisy dataset used in this chapter.

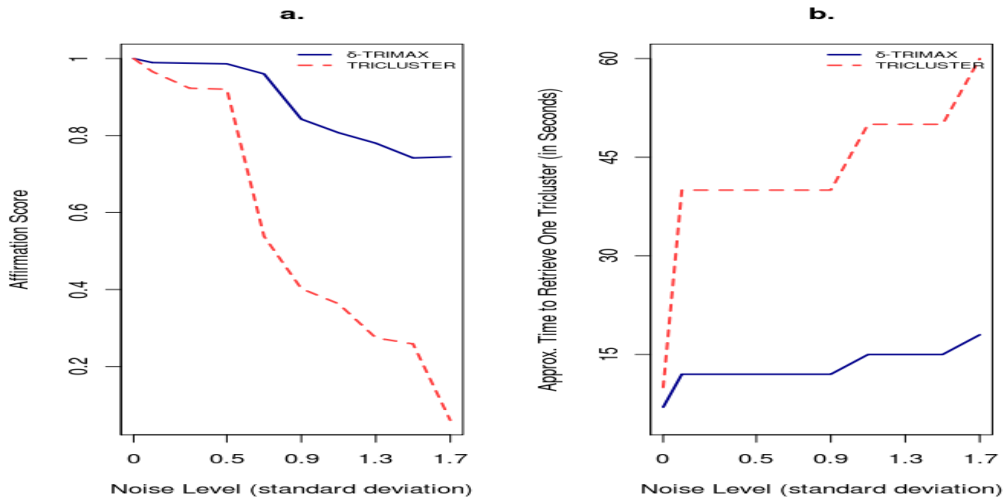

Figure 2.2: Comparison in terms of Affirmation Scores. **a.** Comparison of Affirmation scores produced by δ-TRIMAX and TRICLUSTER algorithm. **b.** Comparison of running time of δ-TRIMAX and TRICLUSTER algorithm on the synthetic dataset. The figure has been taken from [1, 2].

2.3.2 Results on Real-life Dataset

We assigned 0.012382 and 1.2 to the parameters δ and λ, respectively [1, 2] and our algorithm yields 115 triclusters. To collate the performance of our proposed algorithm with TRICLUSER algorithm on the real-life dataset, we have used three validation indexes.

Coverage

Coverage for any triclustering algorithm can be delineated as

$$Coverage = (\frac{g_{alg} \times c_{alg} \times t_{alg}}{G \times C \times T}) \times 100, \tag{2.5}$$

where g_{alg}, c_{alg} and t_{alg} represent the total number of genes, experimental samples and time points retrieved by the triclustering algorithm. G, C and T denote the number of all genes, experimental samples and time points in the dataset.

Triclustering Diffusion (TD) Score

We can elucidate Triclusering Diffusion score of a tricluster by equation 6 [1, 2, 9].

$$TD_i = \frac{MSR_i}{Volume_i}, \tag{2.6}$$

where MSR_i and $Volume_i$ correspond to the mean-squared residue and volume of the ith tricluster. Lower TD score represents better quality of tricluster.

Statistical Difference from Background (SDB)

Here we have introduced another quality measurement, called as Statistical Differences from Background (SDB) as [9]

$$SDB = \frac{1}{n}\sum_{i=1}^{n} \frac{\frac{1}{r}\sum_{j=1}^{r} RMSR_j - MSRi}{MSR_i}, \tag{2.7}$$

where n is the total number of triclusters extracted by the algorithm. MSR_i represents mean squared residue of ith tricluster retrieved by the algorithm and $RMSR_j$ represents mean squared residue of jth random tricluster having the same number of genes, experimental samples and time points as that of ith resultant tricluster. Here a lower value of the denominator denotes better quality of the resultant tricluster. Hence, a higher SDB score signifies better performance of the algorithm. Table 2.2 shows the comparison between proposed δ-TRIMAX algorithm and TRICLUSTER algorithm in terms of coverage, SDB and TQI score.

Table 2.2: Comparison between δ-TRIMAX and TRICLUSTER algorithm using coverage, Statistical Difference of from Background (SDB) and Triclustering Quality Index (TQI). The contents of this table have been taken from [2].

Algorithm	**Coverage**	**SDB**	**Average TD**
δ-TRIMAX	93.7412	2.140935	3.082684e-05
TRICLUSTER	72.34019	2.094091	3.348486e-05

2.3.3 Biological Significance

To establish the biological significance of genes belonging to each resultant tricluster, we have carried out the following: (a) Gene Ontology (GO) and KEGG pathway enrichment analysis, (b) identifying key genes of each tricluster and the corresponding enriched KEGG pathway terms and (c) Transcription Factor Binding Site (TFBS) enrichment analysis.

GO and KEGG Pathway Enrichment Analysis

GOStats package [10] in R has been used to perform GO and KEGG pathway enrichment analysis for establishing biological significance of genes belonging to each tricluster. Adjustment of the p-values has been done using FDR method proposed by Benjamini-Hochberg [11]. The terms having a p-value below a threshold of 0.05 are considered as significant ones. We have found statistically enriched GO terms for genes belonging to each tricluster. Moreover, we have collated the performance of our proposed δ-TRIMAX algorithm with that of TRICLUSTER algorithm on the real-life dataset used in this chapter. To compare the performances, we have considered GO Biological Processes (GOBP) and KEGG pathway terms that have already been reported to play an important role in estrogen induced breast cancer cell. Table 2.3 shows the comparison between δ-TRIMAX and TRICLUSTER algorithm in terms corrected p-values of GOBP and KEGG pathway terms *cell adhesion* and *Wnt signaling pathway* that have already been inferred to play a crucial role in estrogen induced breast cancer [12, 13], respectively.

Table 2.3: Comparison between δ-TRIMAX and TRICLUSTER algorithm in terms of p-values of GOBP and KEGG pathway term enrichment analysis. The contents of this table have been taken from [1, 2].

Algorithm	**GOBP term**	**KEGG pathway terms**
δ-TRIMAX	GO:0007155: cell adhesion (4.31e-08)	KEGG:04310: Wnt signaling pathway (0.011)
TRICLUSTER	GO:0007155: cell adhesion (0.00022)	KEGG:04310: Wnt signaling pathway (0.03)

Identifying Key Genes and Their Corresponding Pathways During Early, Middle and Late Stages of Estrogen Treatment

The Pearson correlation coefficients between the expression profiles of genes over the clustered time points, samples and the tricluster eigen gene vector have been computed to

identify the key genes of each of the resultant triclusters. The genes having the Pearson correlation coefficient closed to 1 or -1 can be considered as tricluster key genes. We have considered the gene symbols of the 10 topmost probe-ids as key genes of each tricluster. From Figure 2.3, we can see that the identified tricluster key genes are highly correlated with the corresponding eigene-gene vector. Moreover, we have performed KEGG pathway enrichment analysis on the sets of identified key genes to reveal the potential pathways during different stages of estrogen treatment. Through our analysis we have found several pathways such as Huntington's disease, lysosome, oxidative phosphorylation, phagosome, MAPK signaling pathway, endometrial cancer, non-small cell lung cancer, ether lipid metabolism, adherens junction, natural killer cell mediated cytotoxicity, chemokine signaling pathway, adipocytokine signaling pathway, steroid hormone biosynthesis, Hepatitis C, chronic myeloid leukemia, toll-like receptor signaling pathway, toxoplasmosis, Jak-STAT signaling pathway, rheumatoid arthritis, leukocyte transendothelial migration, cell adhesion molecules (CAMs), glioma, arginine and proline metabolism, N-Glycan biosynthesis, mTOR signaling pathway, tight junction, osteoclast differentiation, leishmaniasis, ErbB signaling pathway, GnRH signaling pathway, circadian rhythm, viral myocarditis, arrhythmogenic right ventricular cardiomyopathy (ARVC), hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy, mucin type O-Glycan biosynthesis, D-glutamine and D-glutamate metabolism, etc. that are already known to be used for breast cancer therapeutics [14–68]. Figures 2.4-2.5 show all the pathways found to be enriched for the triclusters key genes. Though it is astonishing to us to find the cardiac disorders related pathways such as dilated cardiomyopathy, hypertrophic cardiomyopathy etc. as the enriched ones for tricluster 4 key genes (Figure 2.4) in the context of breast cancer, the association between cardiovascular diseases and cancer therapy was inferred by a previous study [63].

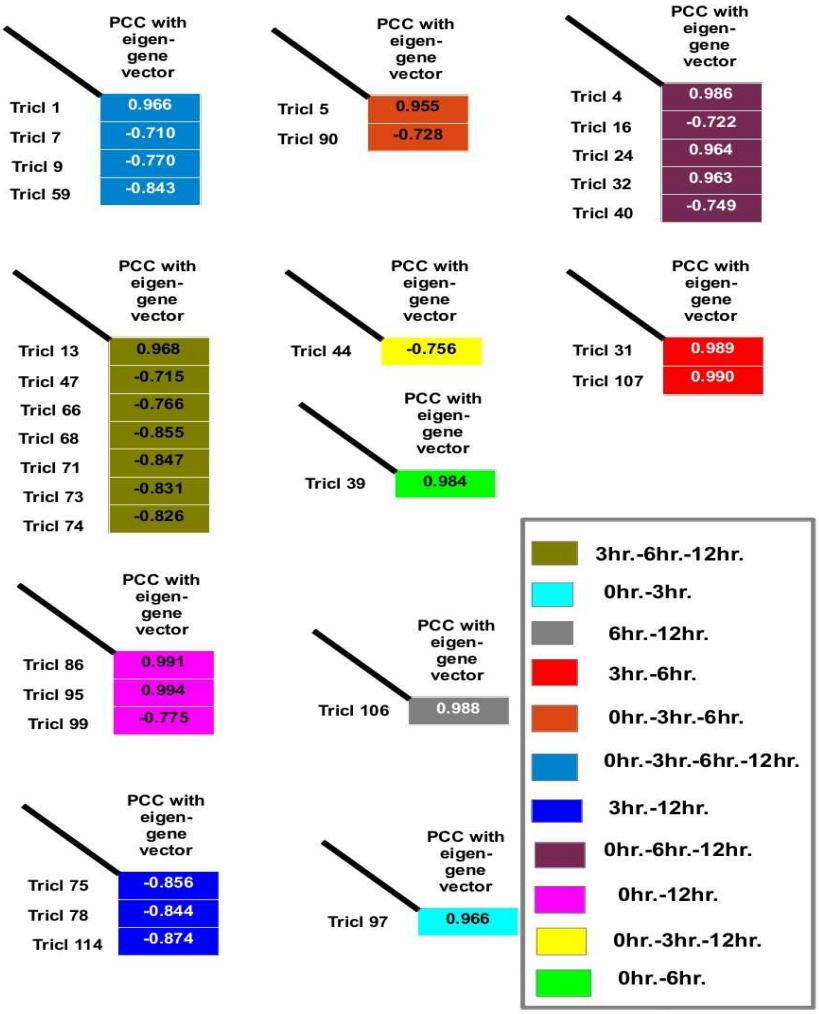

Figure 2.3: The Pearson correlation coefficient between tricluster eigen-gene vector and the tricluster key genes

Figure 2.4: Enriched KEGG pathway terms for the key genes of different triclusters during cellular response to estrogen at different time points

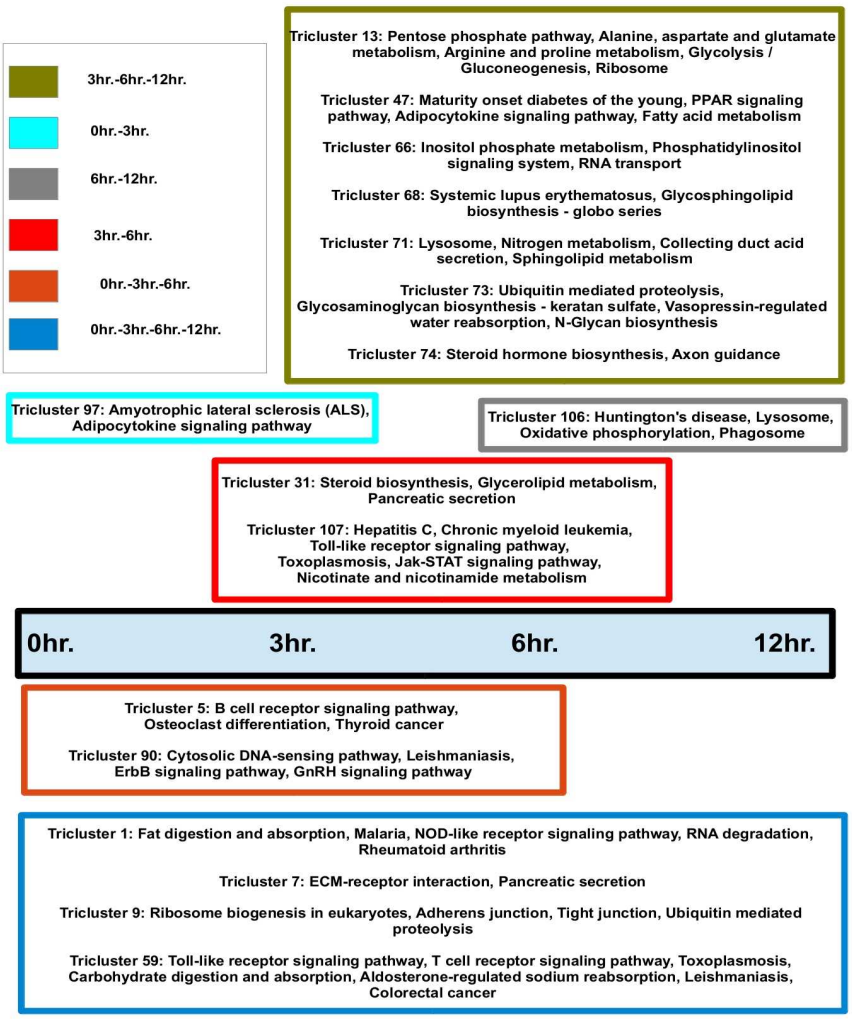

Figure 2.5: Enriched KEGG pathway terms for the key genes of different triclusters during cellular response to estrogen at different time points

TFBS Enrichment Analysis

The transcription factor binding site (TFBS) enrichment analysis using the TRANSFAC library (version 2009.4) has been performed to establish the potential co-regulation of the genes belonging to each of the resultant triclusters [69]. Here we used 42,544,964 TFBS predictions that have high affinity scores and are conserved between human, mouse, dog

and cow [70]. Out of these 42 million conserved TFBSs, we have chosen the best 1% for each TRANSFAC matrix individually to identify the most specific transcription factor - target interactions. We have applied the hyper-geometric test [71] and Benjamini Yekutieli-FDR method [72] for p-value correction to find over-represented binding sites (p-value ≤ 0.05) in the upstream regions of genes belonging to each tricluster. Table 2.4 shows the list of triclusters where we have found statistically enriched TFBSs. From Table 2.4, we can observe that the genes in tricluster 26 are enriched with helix-turn-helix, zinc-coordinating DNA-binding and basic domain transcription factors. The helix-turn-helix domain transcription factor E2F1, to which TRANSFAC matrix V$E2F_Q2 is associated acts as a regulator of cell proliferation in estrogen-induced breast cancer cell [73]. The zinc finger transcription factors Sp1 and Sp4, associated with matrix V$SP1_Q6_01 have already been reported to play an instrumental role in estrogen-induced MCF-7 breast cancer cell line [74, 75]. In tricluster 17, the basic domain transcription factor CREB (matrix V$CREB_01) is important for malignancy in breast cancer cell. ATF1, ATF2, ATF3, ATF4, ATF5 (matrix V$CREBATF_Q6) likewise play a crucial role in breast cancer cell [76]. Moreover,we have observed the enrichment for matrix V$NFAT1_Q6 and the corresponding transcription factor (NFATC1) has been inferred to be associated with clinical characteristics in breast cancer cell [77]. In tricluster 4 POU2F1, the TF associated with matrix V$OCT1_03 is a helix-turn-helix domain transcription factor (Oct-1) and has been reported to be estrogen-responsive in a previous study [78].

Table 2.4: TRANSFAC Matrices for Triclusters, having statistically enriched TFBS for real-life dataset. The contents of this table have been taken from [1, 2].

Tricluster (no. of genes)	**20 most significant TRANSFAC matrices (in ascending order of p-values)**	**FDR-BY corrected p-value of top-most matrix**
Tricluster 3 (875)	VNCX_02, VMSX1_02, V$PAX4_02, V$POU3F2_01, VTBP_01, VBRN3C_01, V$BARX2_01, V$HB24_02, V$HOXD10_01, V$BARX1_01, V$DBX1_01, V$HMBOX1_01, VHDX_01, VBSX_01, V$NKX52_01, V$HMX3_02, V$LBX2_01, V$HOXD13_01, V$NFAT1_Q6, V$HOXD8_01	4.29e-08
Tricluster 1 (4477)	VNCX_02, VHDX_01, V$BCL6_01, V$ZNF333_01, V$DLX2_01, V$DLX7_01, V$DLX5_01, V$SRY_02, V$BARX1_01, V$SOX4_01, V$NKX24_01, V$HOXD3_01, V$LBX2_01, V$LHX61_02, VSRY_01, VTST1_01, V$DLX3_01, V$XVENT1_01, V$EVX1_01, V$BARX2_01	1.27e-05
Tricluster 26 (3177)	V$E2F_Q2, V$ZF5_01,V$USF2_Q6, V$SP1_Q6_01, V$KID3_01,V$CHCH_01	2.99e-05
Tricluster 4 (3482)	V$BCL6_01, V$HOXA10_01, VSRY_01, VNKX23_01, V$WT1_Q6, V$HOXB9_01, V$ISL2_01, V$HOXD10_01, V$HOXD8_01, V$NCX_02, V$MSX1_02, V$PAX4_04, V$BARHL2_01, V$DLX1_01, VSRY_02, VOCT1_03, V$DLX5_01, V$LHX9_01, V$DBX2_01, V$HMGIY_Q6	9.51e-05
Tricluster 2 (2186)	V$CHCH_01, V$MOVOB_01, VMAZ_Q6, VPAX4_03, V$CACD_01, V$GEN_INI3B_B, VGEN_INI_B, VCKROX_Q2	0.0001

Table 2.4 continued....

Tricluster (no. of genes)	**20 most significant TRANSFAC matrices (in ascending order of p-values)**	**FDR-BY corrected p-value of top-most matrix**
Tricluster 12 (476)	VSRY_02, VNCX_02, V$BCL6_01, V$HB24_01, V$HOXA10_01, V$NKX25_02, VSRY_01, VPBX1_02, V$HOXD10_01	0.002
Tricluster 17 (999)	V$CREB_01,V$CREBATF_Q6, V$SP1_Q6_01,V$ATF3_Q6, V$CREBP1CJUN_01	0.004
Tricluster 50 (182)	V$ETF_Q6	0.006
Tricluster 18 (260)	V$STAT1STAT1_Q3	0.042
Tricluster 31 (2465)	V$SP1_Q6_01	0.046

2.4 Conclusion

In this chapter, we have applied our proposed triclustering algorithm δ-TRIMAX to a time-series gene expression data which contains expression values of genes during adaptation of a MCF-7 breast cancer cell to possession of estrogen receptor alpha. The proposed algorithm not only outperforms an existing triclustering algorithm in case of both artificial and real-life datasets but also provides propitious results in terms of both co-expression and co-regulation. Further analysis of the groups of co-expressed genes yielded by δ-TRIMAX algorithm provides insights into several metabolic, biosynthetic processes which can be used as therapeutic targets for the treatment of breast cancer.

2.5 Bibliography

[1] Bhar, A., Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: δ-TRIMAX: Extracting triclusters and analysing coregulation in time series gene expression data. In: Raphael, B., Tang, J. (eds.) Algorithms in Bioinformat-

ics, 12th International Workshop, WABI 2012, Ljubljana, Slovenia, September 10-12, 2012, pp. 165–177 (2012). LNBI 7534

[2] Bhar, A., Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: Coexpression and coregulation analysis of time-series gene expression data in estrogen-induced breast cancer cell. Algorithms for molecular biology **8**(1), 9 (2013)

[3] Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., , Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., FOX, E.A., Silver, P.A., Gingeras, T.R., Liu, X.S., Brown, M.: Genome-wide analysis of estrogen receptor binding sites. Nature Genetics **38**(11), 1289–1297 (2006)

[4] Cheng, Y., Church, G.M.: Biclustering of expression data. Proceedings of International Conference of Intelligent Systems for Molecular Biology **8**, 93–103 (2000)

[5] Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S.: A novel coherence measure for discovering scaling biclusters from gene expression data. Journal of Bioinformatics and Computational Biology **7**(5), 853–868 (2009)

[6] Langfelder, P., Horvath, S.: Eigengene networks for studying the relationships between co-expression modules. BMC Systems Biology **1**, 54 (2007)

[7] Prelic, A., Bleuler, S., Zimmermann, P., Wille, A., Buehlmann, P., Gruissem, W., Hennig, L., Thiele, L., Zitzler, E.: A systematic comparison and evaluation of biclustering methods for gene expression data. Bioinformatics **22**(9), 1122–1129 (2006)

[8] Zhao, L., Zaki, M.J.: triCluster: An Effective Algorithm for Mining Coherent Clusters in 3D Microarray Data. ACM Press (2005)

[9] Bhar, A., Haubrock, M., Mukhopadhyay, A., Wingender, E.: Multiobjective Triclustering Of Time-Series Transcriptome Data Reveals Key Genes Of Biological Processes. BMC Bioinformatics, (in press) (2015)

[10] Falcon, S., Gentleman, R.: Using GOSTATS to test gene lists for GO term association. Bioinformatics **23**(2), 257–258 (2007)

[11] Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: a practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society **57**(1), 289–300 (1995). doi:10.2307/2346101

[12] Schlange, T., Matsuda, Y., Lienhard, S., Huber, A., Hynes, N.E.: Autocrine WNT signaling contributes to breast cancer cell proliferation via the canonical WNT pathway and EGFR transactivation. Breast Cancer Research **9**(5), 63 (2007)

[13] Maynadier, M., Nird, P., Ramirez, J.M., Cathiard, A.M., Platet, N., Chambon, M., Garcia, M.: Role of estrogens and their receptors in adhesion and invasiveness of breast cancer cells. Advances in Experimental Medicine and Biology **617**, 485–491 (2008). doi:

[14] Moreira Sousa, C., McGuire, J.R., Thion, M.S., Gentien, D., de la Grange, P., Tezenas du Montcel, S., Vincent-Salomon, A., Durr, A., Humbert, S.: The Huntington disease protein accelerates breast tumour development and metastasis through ErbB2/HER2 signalling. EMBO Molecular Medicine **5**(2), 309–325 (2013). doi:10.1002/emmm.201201546

[15] Fehrenbacher, N., Jaeaettelae, M.: Lysosomes as targets for cancer therapy. Cancer Research **65**(8), 2993–2995 (2005)

[16] Janjic, B.M.: Analysis of the Level of Oxidative Phosphorylatin in Breast Cancer Cell Types (2012). `http://d-scholarship.pitt.edu/11545/`

[17] Montcourrier, P., Mangeat, P.H., Valembois, C., Salazar, G., Sahuquet, A., Duperray, C., Rochefort, H.: Characterization of very acidic phagosomes in breast cancer cells and their association with invasion. Journal of cell science **107**(Pt 9), 2381–2391 (1994)

[18] Dubey, S., Siegfried, J.M., Traynor, A.M.: Non-small-cell lung cancer and breast carcinoma: chemotherapy and beyond. The Lancet Oncology **7**(5), 416–424 (2006)

[19] Persson, I., Weiderpass, E., Bergkvist, L., Bergstroem, R., Schairer, C.: Risks of breast and endometrial cancer after estrogen and estrogen-progestin replacement. Cancer Causes and Control: CCC **10**(4), 253–260 (1999)

[20] Katz-Brull, R., Seger, D., Rivenson-Segal, D., Rushkin, E., Degani, H.: Metabolic markers of breast cancer: enhanced choline metabolism and reduced choline-ether-phospholipid synthesis. Cancer Research **62**(7), 1966–1970 (2002)

[21] Haidari, M., Zhang, W., Caivano, A., Chen, Z., Ganjehei, L., Mortazavi, A., Stroud, C., Woodside, D.G., Willerson, J.T., Dixon, R.A.: Integrin alpha2-beta1 Mediates Ty-

rosine Phosphorylation of Vascular Endothelial Cadherin Induced by Invasive Breast Cancer Cells. The Journal of Biological Chemistry **287**(39), 32981–32992 (2012)

[22] Schramm, G., Surmann, E.M., Wiesberg, S., Oswald, M., Reinelt, G., Eils, R., Koenig, R.: Analyzing the regulation of metabolic pathways in human breast cancer. BMC Medical Genomics **3**, 39 (2010). doi:10.1186/1755-8794-3-39

[23] Ames, E., Hallett, W.H., Murphy, W.J.: Sensitization of human breast cancer cells to natural killer cell-mediated cytotoxicity by proteasome inhibition. Clinical and Experimental Immunology **155**(3), 504–513 (2009)

[24] Fang, W.B., Jokar, I., Zou, A., Lambert, D., Dendukuri, P., Cheng, N.: CCL2/CCR2 chemokine signaling coordinates survival and motility of breast cancer cells through Smad3 protein- and p42/44 mitogen-activated protein kinase (MAPK)-dependent mechanisms. The Journal of Biological Chemistry **287**(43), 36593–36608 (2012)

[25] Kaur, T., Zhang, Z.F.: Obesity, breast cancer and the role of adipocytokines. Asian Pacific Journal of Cancer Prevention **6**(4), 547–552 (2005)

[26] Kijima, Y., Yoshinaka, H., Higuchi, I., Owaki, T., Aikou, T.: A case of amyotrophic lateral sclerosis and breast cancer. Breast Cancer **12**(1), 57–59 (2005)

[27] Abdelhadi, M.S., Bukharie, H.A.: Breast infections in non-lactating women. Journal of Family and Community Medicine **12**(3), 133–137 (2005)

[28] Zhang, F., Chen, J.Y.: Discovery of pathway biomarkers from coupled proteomics and systems biology methods. BMC Genomics **11 Suppl 2**, 12 (2010). doi:10.1186/1471-2164-11-S2-S12

[29] Cooney, C.A., Jousheghany, F., Yao-Borengasser, A., Phanavanh, B., Gomes, T., Kieber-Emmons, A.M., Siegel, E.R., Suva, L.J., Ferrone, S., Kieber-Emmons, T., Monzavi-Karbassi, B.: Chondroitin sulfates play a major role in breast cancer metastasis: a role for CSPG4 and CHST11 gene expression in forming surface P-selectin ligands in aggressive breast cancer cells. Breast Cancer Research **13**(3), 58 (2011). doi:10.1186/bcr2895

[30] Foster, P.A.: Steroid metabolism in breast cancer. Minerva Endocrinologica **33**(1), 27–37 (2008)

[31] Adachi, T., Yasuda, K., Mori, C., Yoshinaga, M., Aoki, N., Tsujimoto, G., Tsuda, K.: Promoting insulin secretion in pancreatic islets by means of bisphenol A and nonylphenol via intracellular estrogen receptors. Food and Chemical Toxicology **43**(5), 713–719 (2005)

[32] Morrow, P.K., Tarrand, J.J., Taylor, S.H., Kau, S.W., Theriault, R.L., Hortobagyi, G.N., Broglio, K.R., Hahn, K.M.: Effects of chronic hepatitis C infection on the treatment of breast cancer patients. Annals of Oncology **21**(6), 1233–1236 (2010). doi:10.1093/annonc/mdp458

[33] Sansone, P., Bromberg, J.: Targeting the interleukin-6/Jak/stat pathway in human malignancies. Journal of Clinical Oncology **30**(9), 1005–1014 (2012). doi:10.1200/JCO.2010.31.8907

[34] Ma, S., Kosorok, M.R.: Detection of gene pathways with predictive power for breast cancer prognosis. BMC Bioinformatics **11**, 1 (2010). doi:10.1186/1471-2105-11-1

[35] Peliknov, G., Pelikn, A., Bolgc, A., Sitr, A.: Toxoplasmosis as a cause of pseudotumor of the breast in women. Ceskoslovenska Gynekologie **49**(10), 737–740 (1984)

[36] Khan, W.A., Khan, M.W.A.: Cancer Morbidity in Rheumatoid Arthritis: Role of Estrogen Metabolites. BioMed Research International **2013**, 748178 (2013). doi:10.1155/2013/748178

[37] Nathan, L., Pervin, S., Singh, R., Rosenfeld, M., Chaudhuri, G.: Estradiol inhibits leukocyte adhesion and transendothelial migration in rabbits in vivo : possible mechanisms for gender differences in atherosclerosis. Circulation Research **85**(4), 377–385 (1999)

[38] Saadatmand, S., de Kruijf, E.M., Sajet, A., Dekker-Ensink, N.G., van Nes, J.G., Putter, H., Smit, V.T., van de Velde, C.J., Liefers, G.J., Kuppen, P.J.: Expression of cell adhesion molecules and prognosis in breast cancer. The British Journal of Surgery **100**(2), 252–260 (2013). doi:10.1002/bjs.8980

[39] Engin, H., Akoez, A.G.: Breast Cancer and Chronic Myeloid Leukemia: A Short Review. International Journal of HEMATOLOGY **86**, 468–469 (2007). doi:10.1532/IJH97.A20712

[40] Kidd, L.C.R., Rogers, E.N., Yeyeodu, S.T., Jones, D.Z., Kimbro, K.S.: Contribution of toll-like receptor signaling pathways to breast tumorigenesis and treatment. Breast Cancer: Targets and Therapy **5**, 43–51 (2013). doi:10.2147/BCTT.S29172

[41] Vucenik, I., Shamsuddin, A.M.: Cancer inhibition by inositol hexaphosphate (IP6) and inositol: from laboratory to clinic. The Journal of Nutrition **133**(11 Suppl 1), 3778–3784 (2003)

[42] Baselga, J.: Targeting the phosphoinositide-3 (PI3) kinase pathway in breast cancer. The Oncologist **16 Suppl 1**, 12–19 (2011). doi:10.1634/theoncologist.2011-S1-12

[43] Piccirilli, M., Salvati, M., Bistazzoni, S., Frati, A., Brogna, C., Giangaspero, F., Frati, R., Santoro, A.: Glioblastoma multiforme and breast cancer: report on 11 cases and clinico-pathological remarks. Tumori **91**(3), 256–260 (2005)

[44] Chen, Y.F., Chen, Y.T., Chiu, W.T., Shen, M.R.: Remodeling of calcium signaling in tumor progression. Journal of Biomedical Science **20**, 23 (2013). doi:10.1186/1423-0127-20-23

[45] Beltran, A.S., Graves, L.M., Blancafort, P.: Novel role of Engrailed 1 as a prosurvival transcription factor in basal-like breast cancer and engineering of interference peptides block its oncogenic functio (2013). doi:10.1038/onc.2013.422

[46] Michels, K.B., Solomon, C.G., Hu, F.B., Rosner, B.A., Hankinson, S.E., Colditz, G.A., Manson, J.E.: Type 2 Diabetes and Subsequent Incidence of Breast Cancer in the Nurses' Health Study. Diabetes Care **26**(6), 1752–1758 (2003)

[47] Drabovich, A.P., Pavlou, M.P., Dimitromanolakis, A., Diamandis, E.P.: Quantitative analysis of energy metabolic pathways in MCF-7 breast cancer cells by selected reaction monitoring assay. Molecular and Cellular Proteomics **11**(8), 422–434 (2012). doi:10.1074/mcp.M111.015214

[48] Phang, J.M., Liu, W., Hancock, C., Christian, K.J.: The proline regulatory axis and cancer. Frontiers in Oncology **2**, 60 (2012). doi:10.3389/fonc.2012.00060

[49] Harburg, G.C., Hinck, L.: Navigating breast cancer: axon guidance molecules as breast cancer tumor suppressors and oncogenes. Journal of Mammary Gland Biology and Neoplasia **16**(3), 257–270 (2011). doi:10.1007/s10911-011-9225-1

[50] Chen, Y.Z., Xue, J.Y., Chen, C.M., Yang, B.L., Xu, Q.H., Wu, F., Liu, F., Ye, X., Meng, X., Liu, G.Y., Shen, Z.Z., Shao, Z.M., Wu, J.: PPAR signaling pathway may be an important predictor of breast cancer response to neoadjuvant chemotherapy. Cancer Chemotherapy and Pharmacology **70**(5), 637–644 (2012). doi:10.1007/s00280-012-1949-0

[51] Mitra, D., Bernhardt, S., Soons, Z., Poschet, G., Hell, R., Koenig, R., Korf, U., Wiemann, S.: Metabolic transformations in breast cancer subtypes. Cancer and Metabolism **2**(Suppl 1), 48 (2014). doi:10.1186/2049-3002-2-S1-P48

[52] Liu, X., Nie, H., Zhang, Y., Yao, Y., Maitikabili, A., Qu, Y., Shi, S., Chen, C., Li, Y.: Cell Surface-Specific N-Glycan Profiling in Breast Cancer. PLoS One **8**(8), e72704 (2013). doi:10.1371/journal.pone.0072704

[53] Alphen, R.J.V., Wiemer, E.A.C., Burger, H., Eskens, F.A.L.M.: The spliceosome as target for anticancer treatment. British Journal of Cancer **100**, 228–232 (2009). doi:10.1038/sj.bjc.6604801

[54] Azab, S.S.: Targeting the mTOR Signaling Pathways in Breast Cancer: More Than the Rapalogs. Journal of Biochemical and Pharmacological Research **1**(2), 75–83 (2013)

[55] Lu, P., Weaver, V.M., Werb, Z.: The extracellular matrix: a dynamic niche in cancer progression. The Journal of Cell Biology **196**(4), 395–406 (2012). doi:10.1083/jcb.201102147

[56] Brennan, K., Offiah, G., McSherry, E.A., Hopkins, A.M.: Tight junctions: a barrier to the initiation and progression of breast cancer? Journal of Biomedicine and Biotechnology **2010**, 460607 (2010). doi:10.1155/2010/460607

[57] Gu-Trantien, C., Loi, S., Garaud, S., Equeter, C., Libin, M., de Wind, A., Ravoet, M., Le Buanec, H., Sibille, C., Manfouo-Foutsop, G., Veys, I., Haibe-Kains, B., Singhal, S., Michiels, S., Roth, F., Salgado, R., Duvillier, H., Ignatiadis, M., Desmedt, C., Bron, D., Larsimont, D., Piccart, M., Sotiriou, C., Willard-Gallo, K.: $CD4_+$ follicular helper T cell infiltration predicts breast cancer survival. The Journal of Clinical Investigation **123**(7), 2873–2892 (2013). doi:10.1172/JCI67428

[58] Ouellet, V., Tiedemann, K., Mourskaia, A., Fong, J.E., Tran-Thanh, D., Amir, E., Clemons, M., Perbal, B., Komarova, S., Siegel, P.M.: CCN3 impairs osteoblast and stimulates osteoclast differentiation to favor breast cancer metastasis to bone. The American Journal of Pathology **178**(5), 2377–2388 (2011). doi:10.1016/j.ajpath.2011.01.033

[59] Hynes, N.E., MacDonald, G.: ErbB receptors and signaling pathways in cancer. Current Opinion in Cell Biology **21**(2), 177–184 (2009). doi:10.1016/j.ceb.2008.12.010

[60] Everest, H.M., Hislop, J.N., Harding, T., Uney, J.B., Flynn, A., Millar, R.P., McArdle, C.A.: Signaling and antiproliferative effects mediated by GnRH receptors after expression in breast cancer cells using recombinant adenovirus. Endocrinology **142**(11), 4663–4672 (2001)

[61] Wang, X., Qian, H., Zhang, S.: Discovery of significant pathways in breast cancer metastasis via module extraction and comparison. IET Systems Biology **8**(2), 47–55 (2014). doi:10.1049/iet-syb.2013.0041

[62] Ramaswamy, B., Lu, Y., Teng, K.Y., Nuovo, G., Li, X., Shapiro, C.L., Majumder, S.: Hedgehog signaling is a novel therapeutic target in tamoxifen-resistant breast cancer aberrantly activated by PI3K/AKT pathway. Cancer Research **72**(19), 5048–5059 (2012). doi:10.1158/0008-5472.CAN-12-1248

[63] Yeh, E.T., Tong, A.T., Lenihan, D.J., Yusuf, S.W., Swafford, J., Champion, C., Durand, J.B., Gibbs, H., Zafarmand, A.A., Ewer, M.S.: Cardiovascular complications of cancer therapy: diagnosis, pathogenesis, and management. Circulation **109**(25), 3122–3131 (2004)

[64] Brockhausen, I.: Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. EMBO Reports **7**(6), 599–604 (2006)

[65] Shajahan-Haq, A., Cook, K., Schwartz-Roberts, J., Eltayeb, A., Demas, D., Warri, A., Hilakivi-Clarke, L., Clarke, R.: Glutamine metabolism and the unfolded protein response in MYC-driven breast cancer. Cancer and Metabolism **2**(Suppl 1), 66 (2014). doi:10.1186/2049-3002-2-S1-P66

[66] Zhao, Y., Butler, E.B., Tan, M. Cell Death and Disease **4**, 532 (2013). doi:10.1038/cddis.2013.60

[67] Li, S., Ao, X., Wu, H.: The role of circadian rhythm in breast cancer. Chinese Journal of Cancer Research **25**(4), 442–450 (2013). doi:10.3978/j.issn.1000-9604.2013.08.19

[68] Gonzalez-Angulo, A.M., Iwamoto, T., Liu, S., Chen, H., Do, K.A., Hortobagyi, G.N., Mills, G.B., Meric-Bernstam, F., Symmans, W.F., Pusztai, L.: Gene expression, molecular class changes, and pathway analysis after neoadjuvant systemic therapy for breast cancer. Clinical Cancer Research **18**(4), 1109–1119 (2012). doi:10.1158/1078-0432.CCR-11-2762

[69] Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhaeuser, R., Pruess, M., Schacherer, F., Thiele, S., Urbach, S.: The TRANSFAC system on gene expression regulation. Nucleic Acids Research **29**(1), 281–283 (2001)

[70] Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., Kellis, M.: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. NATURE **434**(7031), 338–345 (2005)

[71] Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., Sherlock, G.: GO::TermFinder-open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics **20**(18), 3710–3715 (2004)

[72] Benjamini, Y., Yekutieli, D.: The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics **29**(4), 1165–1188 (2001). doi:10.1214/aos/1013699998

[73] Stender, J.D., Frasor, J., Komm, B., Chang, K.C., Kraus, W.L., Katzenellenbogen, B.S.: Estrogen-regulated gene networks in human breast cancer cells: involvement of E2F1 in the regulation of cell proliferation. Molecular Endocrinology **21**(9), 2112–2123 (2007)

[74] Khan, S., Wu, F., Liu, S., Wu, Q., Safe, S.: Role of specificity protein transcription factors in estrogen-induced gene expression in MCF-7 breast cancer cells. Journal of Molecular Endocrinology **39**(4), 289–304 (2007)

[75] Kim, K., Barhoumi, R., Burghardt, R., Safe, S.: Analysis of estrogen receptor α-Sp1 interactions in breast cancer cells by fluorescence resonance energy transfer. Molecular Endocrinology **19**(4), 843–854 (2005)

[76] Haakenson, J.K., Kester, M., Liu, D.X.: The ATF/CREB family of transcription factors in breast cancer. InTech (2012)

[77] Mancini, M., Toker, A.: NFAT proteins: emerging roles in cancer progression. Nature Reviews. Cancer **9**(11), 810–820 (2009). doi:10.1038/nrc2735

[78] Wang, C., Yu, J., Kallen, C.B.: Two estrogen response element sequences near the PCNA gene are not responsible for its estrogen-enhanced expression in MCF7 cells. PLoS One **3**(10), e3523 (2008). doi:10.1371/journal.pone.0003523

3 Enhanced Multi-objective Triclustering Based on a Genetic Algorithm and Its Application in Revealing Biological Processes of Development

3.1 Introduction

In this chapter, we have leveraged the proven benefits of Genetic Algorithm (GA) to conquer the disadvantages of the proposed triclustering algorithm δ-TRIMAX and described an improved version of it, namely Evolutionary Multi-objective Optimization Algorithm for δ-TRIMAX (EMOA-δ-TRIMAX). Here we have used one artificial dataset by implanting different levels of noise and three real-life datasets to compare the performance of EMOA-δ-TRIMAX with that of the existing triclustering algorithms. Capability of a stem cell to differentiate into a specific tissue has been emerged by several studies such as cardiac myocytes, neural progenitors etc. Development of cardiomyocytes from a stem cell is attained by several steps such as epithelial to mesenchymal transition, mesoderm differentiation, cardiac specification, differentiation and maturation [1]. Several metabolic processes, signaling pathways are found to be instrumental in driving the phenotypic changes of the adolescent cell [2]. Moreover, unveiling the signaling pathways involved in such differentiation may provide insightful knowledge into the therapeutic targets for the treatment of cardiac disorders [3, 4]. In this chapter, we have applied the improved version of δ-TRIMAX algorithm to a time-series gene expression dataset obtained during the adolescence of cardiomyocytes from human induced pluripotent stem cells (hiPSCs) in order to reveal potential key genes, biological processes and signaling pathways which may play instrumental roles during different stages of cardiomyocytes development.

3.2 Summary of δ-TRIMAX

3.2.1 Aim of δ-TRIMAX

The proposed algorithm δ-TRIMAX aims at finding maximal triclusters having mean squared residue (MSR) score delineated in the previous chapter below a threshold δ [5, 6]. Thus δ-TRIMAX deals with two conflicting objectives i.e. minimization of MSR score of the resultant triclusters and enhancement of the size of the triclusters.

3.2.2 Pitfalls of δ-TRIMAX

Though the proposed triclustering algorithm δ-TRIMAX algorithm described in the previous chapter has some advantages, it has few drawbacks such as (a) because of using a greedy search heuristic, it often plunges into local optima and (b) it can not retrieve overlapping triclusters which is an important aspect from biological point of view. For

instance, in case of a time-series gene expression dataset same set of genes may be involved in two different biological processes with some other genes over two overlapping subsets of time points and replicates.

3.3 Materials and Methods

3.3.1 Materials

Artificial Dataset

We have used an artificial dataset having 1000 genes, 5 replicates and 4 time points. First we have implanted 3 perfect shifting triclusters (standard deviation (σ) = 0) of size 100 $\times$ 4 $\times$ 4, 80 $\times$ 4 $\times$ 4 and 60 $\times$ 4 $\times$ 4 into the dataset. In the next step, we have added different levels of noise to the implanted triclusters (σ = 0.1, 0.3, 0.5, 0.7, 0.9) [7].

Real-life Datasets

Dataset 1: The description of this dataset was given in the previous chapter (GSE11324) [8].
Dataset 2: The aim of the experiment was to provide new insights into the regulation of gene expression during the differentiation of a pluripotent stem cell into cardiomyocytes. This dataset contains 48803 Illumina HumanWG-6 v3.0 probe ids, 3 technical replicates and 12 time points (days 0, 3, 7, 10, 14, 20, 28, 35, 45, 60, 90 and 120) (GSE35671) [9]. A detailed description of the time points is shown in Figure 3.1.

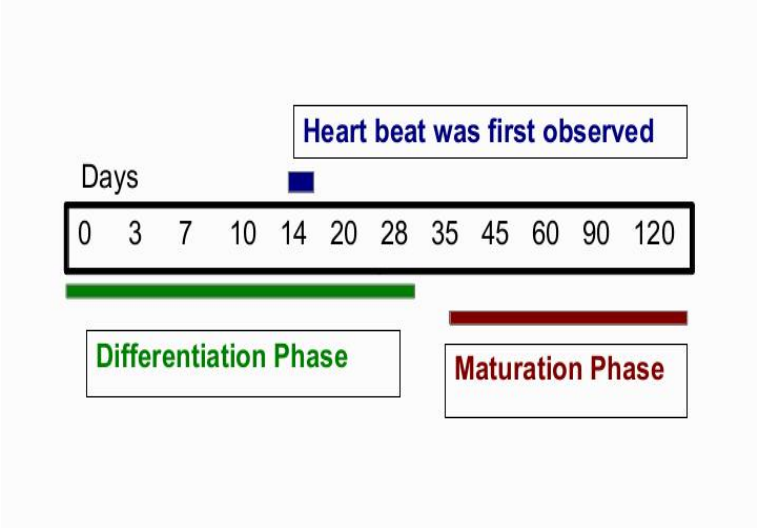

Figure 3.1: Descriptions of the time points of Dataset 2

Dataset 3: To monitor the dynamics of expression profiles of 54675 Affymetrix human genome U133 plus 2.0 probe ids in response to IFN-beta-1b treatment across four time points over 6 patients, the experiment was carried out (GSE46280) [10]. Figure 3.2 provides details of the different time points used in this experiment.

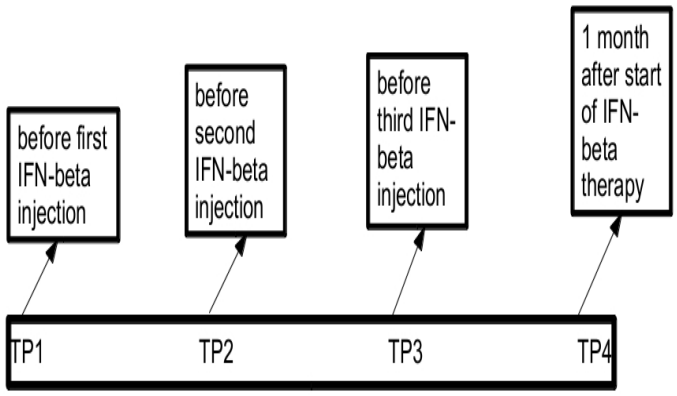

Figure 3.2: Descriptions of the time points of Dataset 3

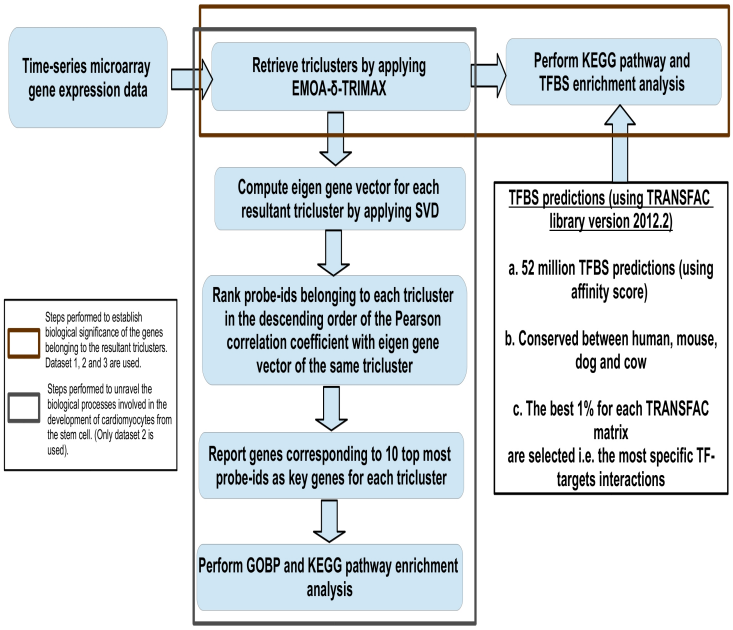

Figure 3.3: Workflow used in this chapter. The figure has been taken from [7].

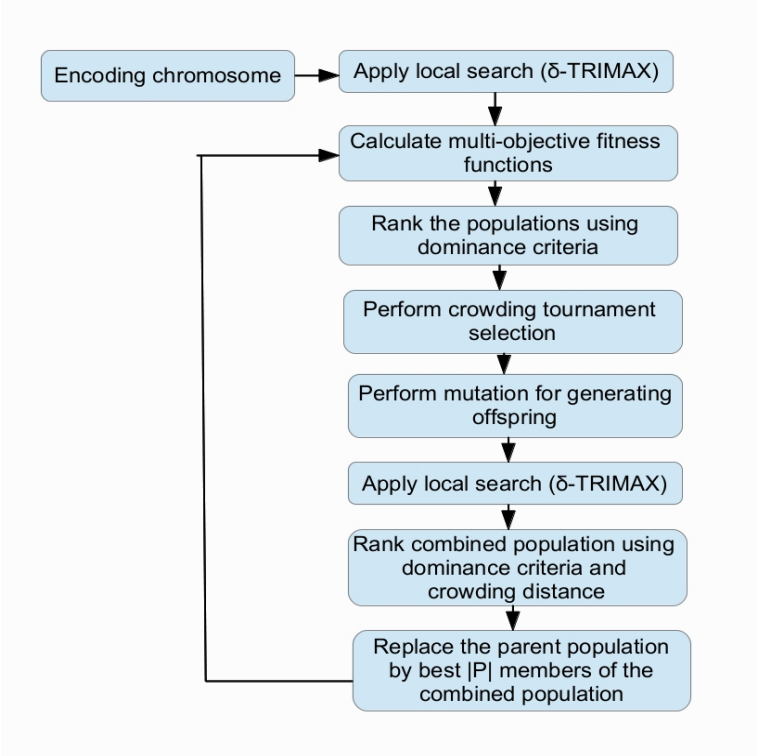

Figure 3.4: Steps of EMOA-δ-TRIMAX. The figure has been taken from [7].

3.3.2 EMOA-δ-TRIMAX

Figure 3.3 represents the workflow applied in this chapter, whereas Figure 3.4 shows the steps of the EMOA-δ-TRIMAX algorithm [7].

Multi-objective Optimization Problem

It refers to the problem of optimizing multiple conflicting objectives where a set of alternative solutions having identical calibre exists instead of one single optimal solution. Hence, the multi-objective optimization problem is to find the vector $\bar{x}^* = [x_1^*, x_2^*, ..., x_n^*]^T$ of decision variables that meets a number of quality and inequality constraints by optimizing the vector function $\bar{f}(\bar{x}) = [f_1(\bar{x}), f_2(\bar{x}), ..., f_r(\bar{x})]^T$ subject to some constraints. Here constraints denote the possible region F which holds all acceptable solutions and $\bar{x}^*$ represents an optimal solution. In case of a minimization problem, a decision vector $\bar{x}^*$ is called Pareto optimal if and only if there is no $\bar{x}$ such that $\forall i \in \{1, 2, .., r\}$, $f_i(\bar{x}) \leq f_i(\bar{x}^*)$ and $\exists i \in \{1, 2, ..., r\}$, $f_i(\bar{x}) < f_i(\bar{x}^*)$. In words, $\bar{x}^*$ is referred to as Pareto optimal if there

exists no feasible vector $\bar{x}$ that causes the curtailment of some criterion without a simultaneous increase of one other criterion [11, 12]. In this chapter, we have used non-dominated sorting genetic algorithm II (NSGA-II) [11] as a multi-objective optimization method to develop an improved version of δ-TRIMAX algorithm, named evolutionary multi-objective optimization for δ-TRIMAX (EMOA-δ-TRIMAX) which copes with the disadvantages of δ-TRIMAX algorithm.

Genetic Algorithm

Genetic algorithm is a search heuristic which follows the process of Darwin evolution [11,12]. Here the parameters of the search space are encoded by a set of randomly initialized chromosomes which composes the population. The goodness of each chromosome in the population is measured by a fitness function also called an objective function. Genetic operators such as selection, crossover and mutation are used to evolve the succeeding generations. If certain criterion is satisfied or the maximum number of generation is reached, the algorithm stops its execution.

Encoding Chromosome

In this work, we used a binary string to represent each chromosome in the population and each of these chromosomes denotes one possible tricluster. Suppose, a time series gene expression data contains G, C and T number of genes, replicates and time points, respectively. Hence, each string consists of (G + C + T) number of bits and each bit has either a value 1 or 0. A value 1 represents the presence of the corresponding gene/ sample/ time point in the triclusters. For example, in case of a time series gene expression data having 15 genes, 3 replicates and 10 time points, the binary string 100101001100010011010010011 0 denotes that genes $\{g_1, g_4, g_6, g_9, g_{10}, g_{14}\}$, replicates $\{c_2, c_3\}$ and time points $\{t_2, t_5, t_8, t_9\}$ are the members of the corresponding tricluster. As the initial population is generated randomly, some genes and/ or replicates and/or time points may be implanted to the initial population inspite of being far away from the feature space. To overcome this problem, we have used the proposed triclustering algorithm δ-TRIMAX as a local search heuristic.

Fitness Functions

In the next step, the goodness of each individual of the population is measured by computing the following three objective functions [7].

$$obj_1 = \frac{MSR}{\delta}, \tag{3.1}$$

where MSR is the mean squared residue score of the resultant tricluster. Hence, obj_1 needs to be minimized.

The second objective function is

$$obj_2 = \frac{|I| * |J| * |K|}{|G| * |C| * |T|}, \tag{3.2}$$

where $(|I| * |J| * |K|)$ is the volume of the resultant tricluster and $(|G| * |C| * |T|)$ is the volume of the dataset. Here we aim at maximizing the value of obj_2.

Finally, the third objective function is

$$obj_3 = \left| 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \right|, \tag{3.3}$$

where d_i is the difference between the ranks of average expression values (sorted either in ascending or descending order) over a subset of samples at ith time point of each pair of genes in one tricluster and n is the number of time points in that tricluster. Here the goal is to increase the non-parametric Spearman correlation coefficient (obj_3) of the resultant triclusters [13].

Motivation of Fitness Functions

Optimizing the first two objective functions obj_1 and obj_2 (equation (3.1)-(3.2)), accomplishes the goals of the δ-TRIMAX algorithm i.e. maximization of size of the tricluster and minimization of the MSR score of the resultant tricluster. The third objective function obj_3 (equation (3.3)) aims at assuring the monotonic changes in the expression profiles of the co-expressed genes over a subset of clustered replicates and time points. Moreover, the absolute value of the correlation coefficient has been considered due to the fact that co-regulated genes can be both up- and down-regulated by the transcription factors across a subset of time points.

Genetic Operators

Selection Operator: In this work, we have applied crowded binary tournament selection in which the binary tournament selection operator is performed based on the crowding distance [11]. Suppose, two solutions sol_1 and sol_2 are being compared and the solution sol_1 wins the tournament if and only if either of the following two is satisfied:
(a) The rank of sol_1 is better (less) than that of sol_2.
and (b) If the solutions sol_1 and sol_2 belong to the same non-dominated front and sol_1 has the higher crowding distance than that of sol_2. In other words, the solution located in the less crowded region will be selected.

Crossover: A previous study reports crossover as a generalization of several mutations performed at once. Being motivated by this fact, we did not use crossover operator [14].

Mutation: The bit string mutation has been used with a higher mutation probability. Suppose, we are performing mutation operator on a bit string {1000011100} using a probability 0.8. For each bit of the string we first generates a random number ranges from 0 to 1. Suppose, the generated random numbers are {0.85, 0.5, 0.7, 0.9. 0.6, 0.95, 0.55, 0.84, 0.3, 0.65}. Now in case of a particular bit of the string if the random number is less than or equal to the mutation probability, then the bit will flip from 1 to 0 or 0 to 1. For this particular example, the offspring will be 1110110111. After performing the mutation operator on each individual of the population, some genes / samples / time points may be inserted into the population that lies far away from the feature space. To deal with that problem δ-TRIMAX algorithm has been applied as a local search heuristic.

Elitism

Elitism has been performed to keep track of the non-dominated Pareto optimal solutions after each generation and carry over to the next generation [11].

3.3.3 Convergence of Solutions

In order to check whether solutions converge towards the Pareto optimal front around its center region, we have calculated $minSum$ values in each generation as follows (equation (3.4)) [7, 15]

$$minSum(\Psi) = \min_{x \in \Psi}(f_1(x) + (1 - f_2(x)) + (1 - f_3(x))), \tag{3.4}$$

where Ψ corresponds to the current population and k denotes the number of given objectives. We have observed the convergence of solutions towards the Pareto optimal front (Figure 3.6).

3.4 Results and Discussions

3.4.1 Artificial Dataset

To estimate the performance of the proposed algorithm EMOA-δ-TRIMAX for the artificial dataset used in this work, we have used the affirmation score defined in the previous chapter [5, 6]. Moreover, we used the same metric to compare the performance of the proposed algorithm (EMOA-δ-TRIMAX) with that of the other existing algorithms. The value of δ has been computed in the same way as described in the previous chapter, whereas the value of λ was set experimentally to optimize the speed and performance (to avoid falling into a local optima) of the proposed algorithm. Table 3.1 shows the values of two input parameters λ and δ for different levels of noise. From Figure 3.5 we can observe that the proposed algorithm performs better than the other triclustering algorithms and one biclustering algorithm proposed by Cheng and Church [16].

Table 3.1: Values of input parameters of EMOA-δ-TRIMAX namely, λ and δ for different levels of noise in case of the artificial dataset. The contents of the table have been taken from [7].

Noise Levels (σ)	Values of λ	Values of δ
0	1.2	0.00002
0.1	1.2	0.021
0.3	1.2	0.1075
0.5	1.2	0.2295
0.7	1.2	0.4
0.9	1.2	0.7

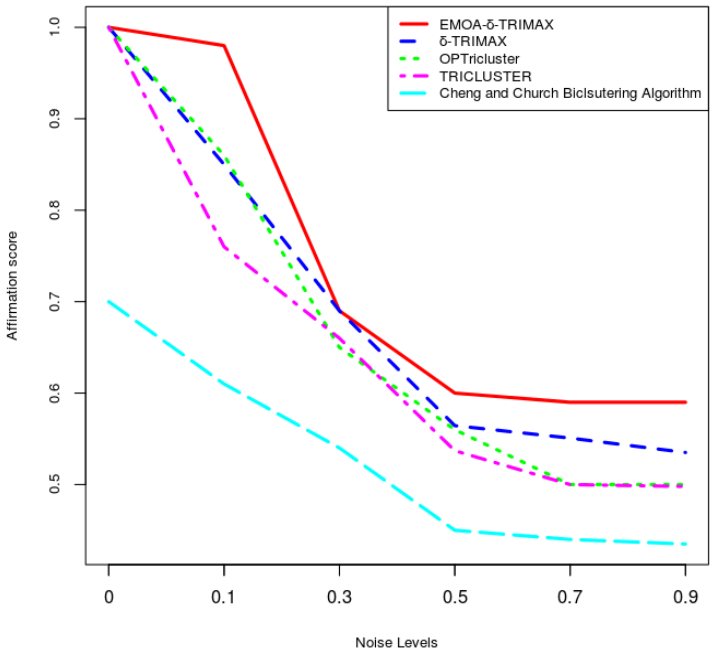

Figure 3.5: Comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER, OPTricluster and the biclustering algorithm proposed by Cheng and Church in terms of affirmation score for artificial dataset. The figure has been taken from [7].

3.4.2 Real-life Datasets

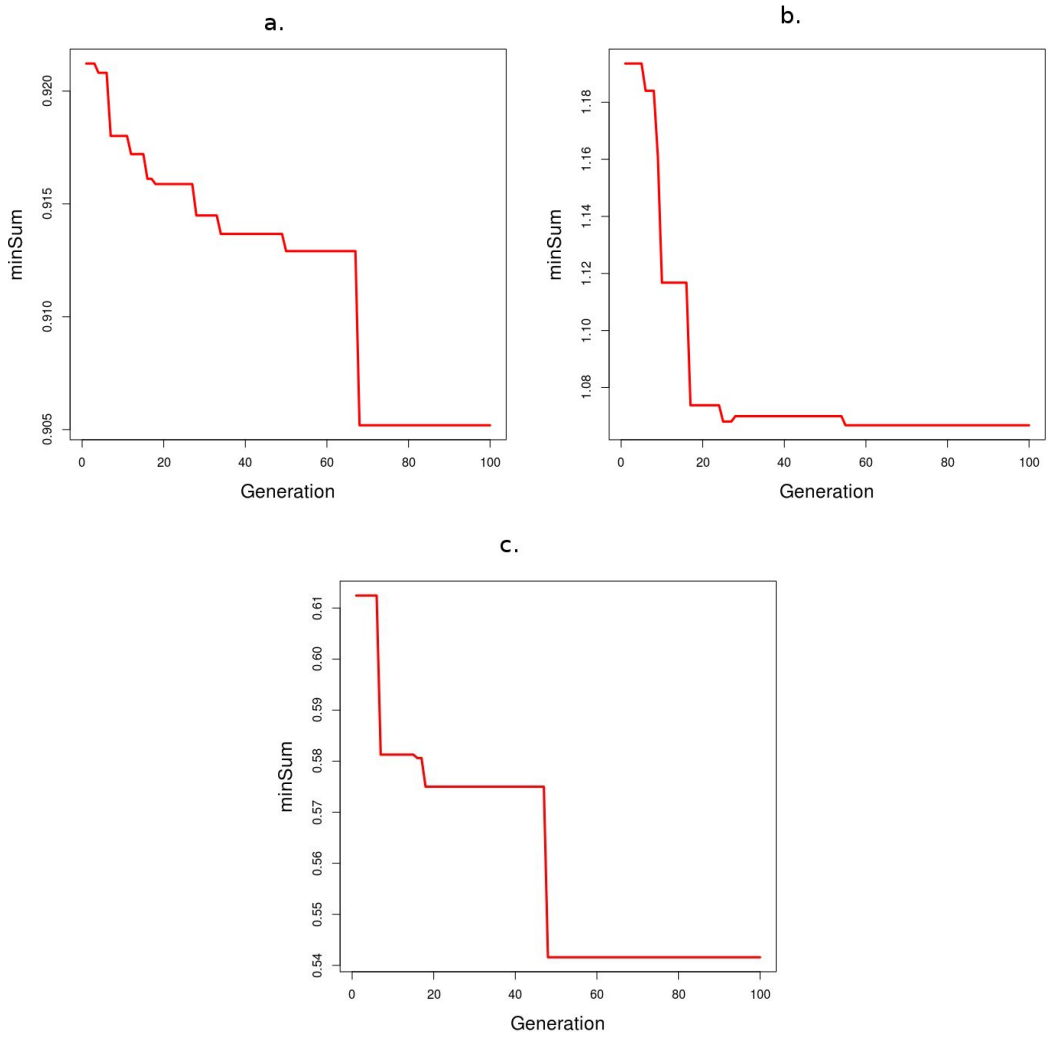

Figure 3.6: Convergence of solutions towards the Pareto optimal front. *minSum* values are plotted for dataset 1 (a), dataset 2 (b) and dataset 3 (c). The figure has been taken from [7].

Table 3.2 enlists the values of input parameters and Table 3.3 shows the number of resultant triclusters and the percentage of probe-ids, samples, time points for the aforementioned real-life datasets.

Table 3.2: Values of the input parameters in case of real-life datasets. The contents of the table have been taken from [7].

Dataset	Dataset 1	Dataset 2	Dataset 3
Values of λ	1.2	1.2	1.2
Values of δ	0.012382	0.008	0.008754
number of generations	100	100	100
population size	100	100	100
mutation probability	0.9	0.9	0.9

Table 3.3: Number of resultant triclusters, percentage of probe-ids, replicates and time points covered by the resultant triclusters for each of the real-life datasets. The contents of the table have been taken from [7].

Dataset	Dataset 1	Dataset 2	Dataset 3
Number of resultant triclusters	100	100	32
Probe-id coverage	99.02	88.14	93
Replicates Coverage	100	100	100
Time point coverage	100	100	100

3.4.3 Performance Comparison

To compare the performance of the proposed triclustering algorithm EMOA-δ-TRIMAX with that of the previously published ones for the real-life datasets we have used the following three metrics.

Tricluster Diffusion (TD) Score

The definition of this metric has been given in the previous chapter. From Figures 3.7-3.9, we can see that EMOA-δ-TRIMAX yields triclusters with a lower TD score compared to the other algorithms for each of the real-life datasets used here.

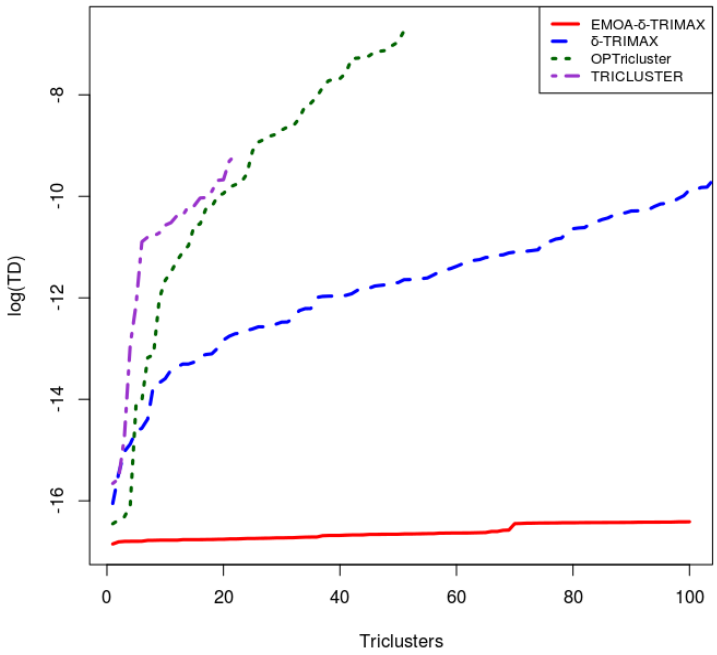

Figure 3.7: Tricluster Diffusion scores for Dataset 1. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of TD (in log scale) score for Dataset 1. The figure has been taken from [7].

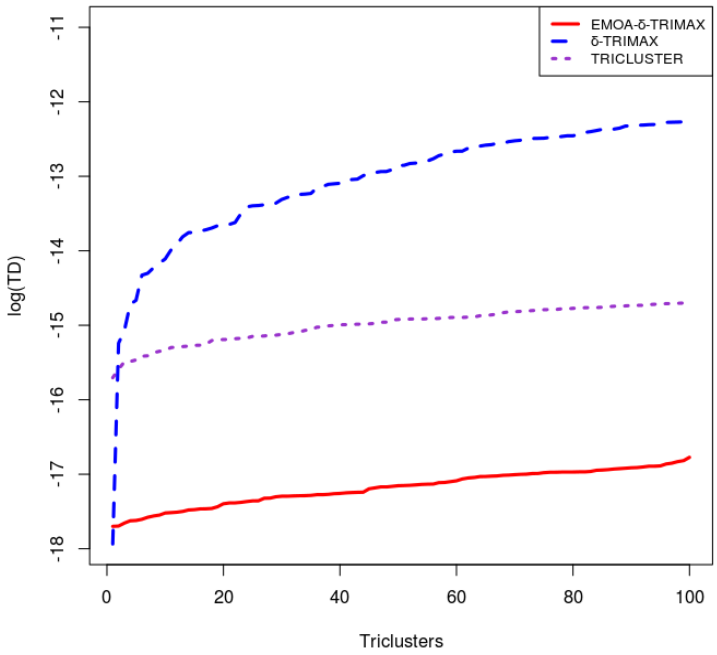

Figure 3.8: Tricluster Diffusion scores for Dataset 2. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX and TRICLUSTER in terms of TD score (in log scale) for Dataset 2. The figure has been taken from [7].

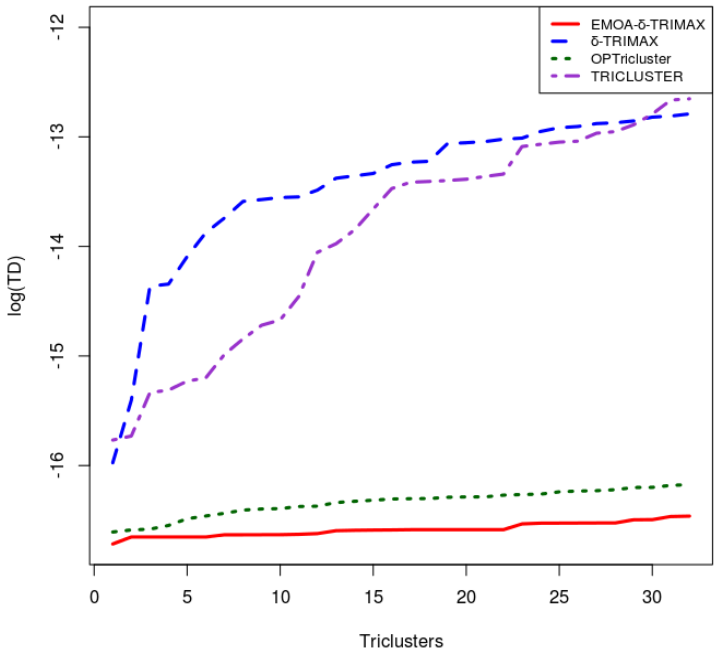

Figure 3.9: Tricluster Diffusion scores for Dataset 3. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of TD score (in log scale) for Dataset 3. The figure has been taken from [7].

Statistical Difference from Background (SDB)

This metric has been delineated in the previous chapter. From Tables 3.4-3.9, we can observe that EMOA-δ-TRIMAX results in triclusters with a higher SDB score compared to the other existing algorithms for each of the real-life datasets.

Table 3.4: Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of SDB score for Dataset 1. The contents of the table have been taken from [7].

Algorithms	SDB score
EMOA-δ-TRIMAX	2.49851
δ-TRIMAX	2.140935
TRICLUSTER	2.094091
OPTricluster	0.4956035

Table 3.5: Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX and TRICLUSTER in terms of SDB score for Dataset 2. The contents of the table have been taken from [7].

Algorithms	SDB score
EMOA-δ-TRIMAX	13.88559
δ-TRIMAX	12.10529
TRICLUSTER	7.520363

Table 3.6: Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of SDB score for Dataset 3. The contents of the table have been taken from [7].

Algorithms	SDB score
EMOA-δ-TRIMAX	9.454915
δ-TRIMAX	8.945816
TRICLUSTER	7.076184
OPTricluster	0.4383489

Hit Score Using KEGG Pathway and TFBS Enrichment Analysis

To establish the biological relevance of the groups of co-expressed genes, first we have done KEGG pathway enrichment analysis using GOStats package in R using a FDR-corrected p-value cutoff 0.05 [17, 18]. Using the results of KEGG pathway enrichment analysis, we

have computed hit score [19] (equation (3.5)) for each of the resultant triclusters [7].

$$Hit(K) = \frac{\max\{\left|N_T^1\right|, \left|N_T^2\right|, ..., |N_T^n|\}}{|T|}, \tag{3.5}$$

where N_T^i corresponds to the intersection gene set of tricluster T and its enriched KEGG pathway term i; $|T|$ is the total number of genes in tricluster T. A higher hit score means more genes in T participate in a canonical pathway.

In the next step, we have carried out TFBS enrichment analysis on the groups of co-expressed genes to establish their potential co-regulation using the TRANSFAC library (version 2012.2) [20]. A hypergeometric test has been used to detect the over-represented binding sites in the upstream regions of the groups of co-expressed genes [21, 22]. Here we have used 52 million TFBS predictions having high affinity scores and are conserved between four species, namely human, mouse, dog and cow [23]. Out of these 52 million conserved TFBSs, we have then selected the highest-scoring 1% for each TRANSFAC matrix to identify the most specific transcription factor and target interactions. We have then measured the hit score (equation (3.6)) in the same way that we did in case of the KEGG pathway enrichment analysis [7].

$$Hit(TF) = \frac{\max\{\left|P_T^1\right|, \left|P_T^2\right|, ..., |P_T^n|\}}{|T|}, \tag{3.6}$$

where P_T^i represents the intersection gene set of tricluster T and its enriched TRANSFAC matrix i; $|T|$ corresponds to the total number of genes in tricluster T. A higher hit score signifies that more genes in T are regulated by a common transcription factor.

At first we have computed the hit scores (Hit (K) and Hit(TF)) for each resultant tricluster using KEGG pathway and TFBS enrichment results. Then we generated 100 random gene lists having the same size of that of the each resultant tricluster (T). The Hit scores for each randomly generated gene list were computed using KEGG pathway and TFBS enrichment results. As a final step we have applied the non-parametric Mann-Whitney-Wilcoxon test [24] to measure the significance between these two sets of hit scores in terms of p-values. From Figures 3.10-3.15, we can deduce the fact that EMOA-δ-TRIMAX results in the groups of co-expressed genes which show better potential co-regulation than that of the others yielded by the other existing triclustering algorithms.

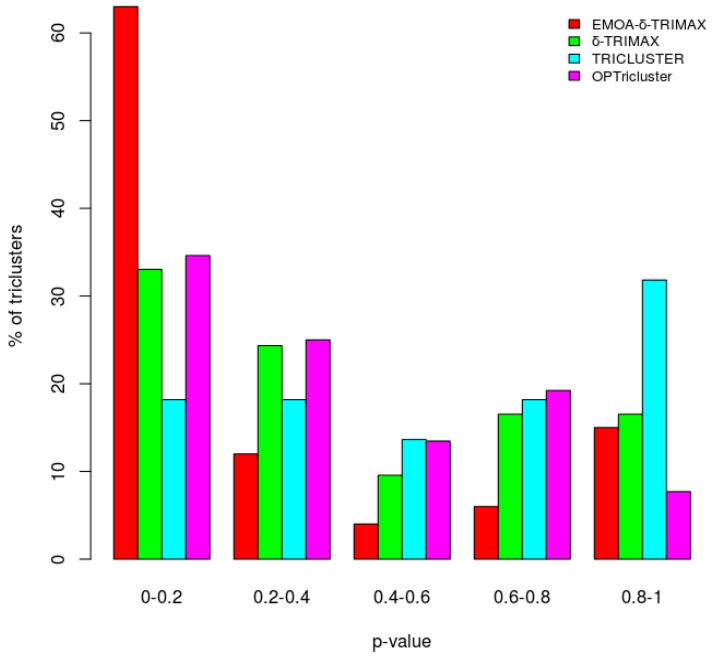

Figure 3.10: Hit scores using KEGG pathway enrichment for Dataset 1. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of Hit scores for Dataset 1. The figure has been taken from [7].

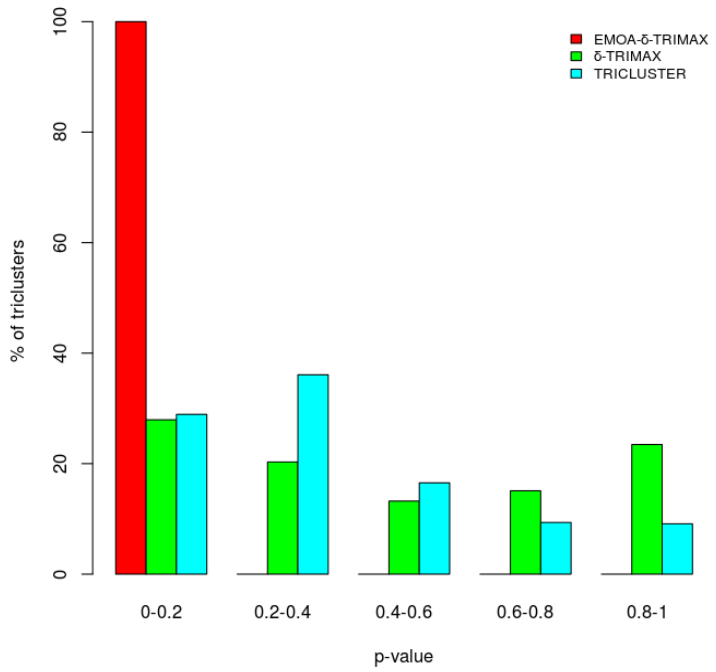

Figure 3.11: Hit scores using KEGG pathway enrichment for Dataset 2. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX and TRICLUSTER in terms of Hit scores for Dataset 2. The figure has been taken from [7].

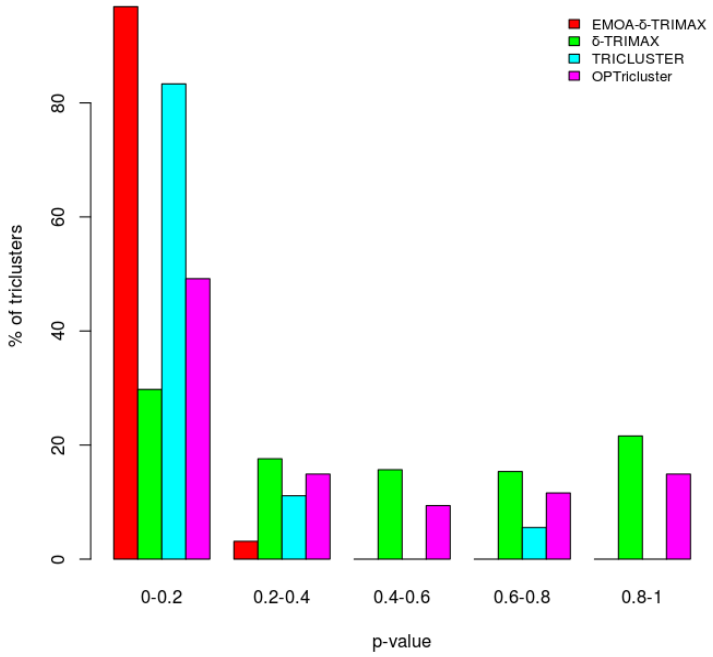

Figure 3.12: Hit scores using KEGG pathway enrichment for Dataset 3. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of Hit scores for Dataset 3. The figure has been taken from [7].

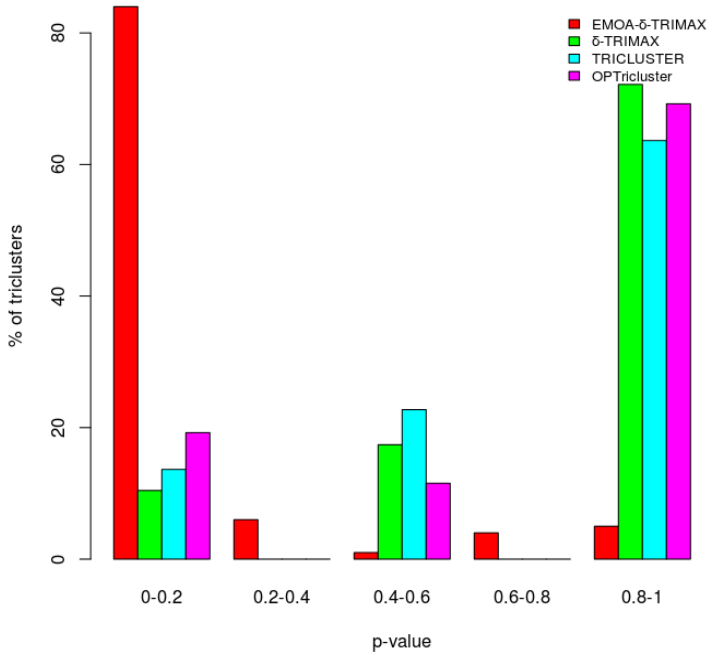

Figure 3.13: Hit scores using TFBS enrichment for Dataset 1. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of Hit scores for Dataset 1. The figure has been taken from [7].

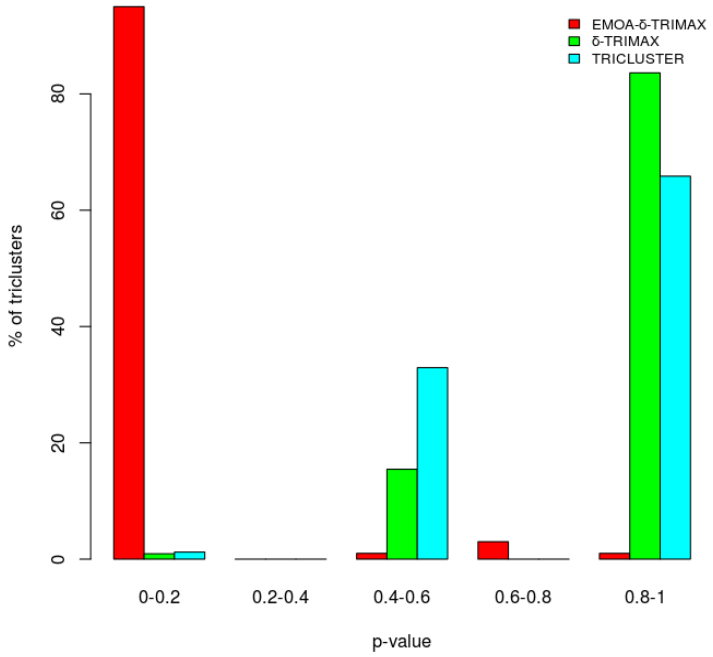

Figure 3.14: Hit scores using TFBS enrichment for Dataset 2. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX and TRICLUSTER in terms of Hit scores for Dataset 2. The figure has been taken from [7].

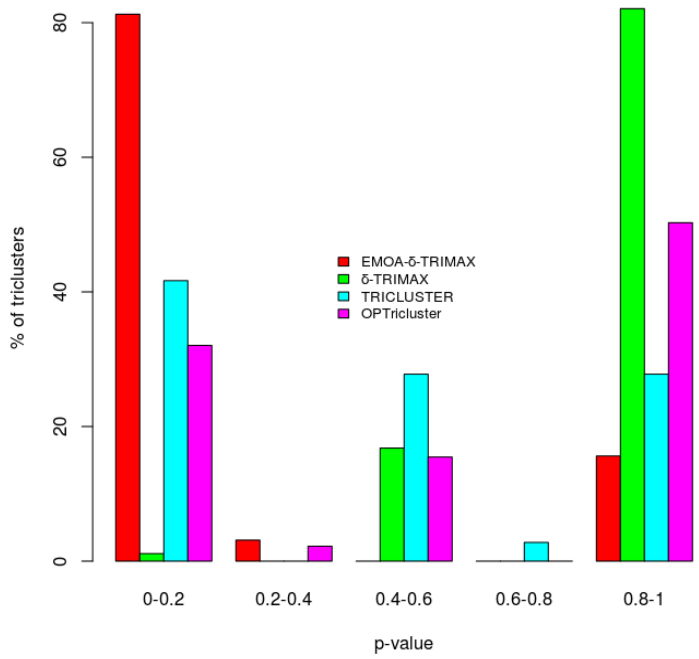

Figure 3.15: Hit scores using TFBS enrichment for Dataset 3. Performance comparison between EMOA-δ-TRIMAX, δ-TRIMAX, TRICLUSTER and OPTricluster in terms of Hit scores for Dataset 3. The figure has been taken from [7].

3.4.4 Identifying Key Genes of Triclusters and Analyzing Their Roles During hiPSC Differentiation into Cardiomyocytes

To validate the proposed triclustering algorithm EMOA-δ-TRIMAX, we have also identified the key genes of each resultant tricluster in the same way that we did in the previous chapter i.e. first we computed the eigen-gene vector for each tricluster using the expression profiles of clustered probe-ids, replicates and time points. In the next step, we ranked the probe-ids in the descending order of the Pearson correlation coefficient between the eigen-gene vector and the expression profiles of probe-ids. Then, we considered the genes of 10 topmost probe-ids of each tricluster as key genes and performed the gene ontology biological process enrichment analysis to detect the specific biological processes. Understanding the specific biological processes might facilitate to gain knowledge about the differentiation of a stem cell into cardiomyocytes. Moreover, the clustered time points may provide

insights into the different stages of differentiation during which the processes remain active. From Figure 3.16, we can deduce that the identified key genes exhibit strong positive correlation with eigen-gene vector of the corresponding tricluster. Furthermore, we have summarized the lists of biological processes and signaling pathways in Figure 3.17 across different time points and most of these identified processes such as lipoprotein, naphthalene, S-adenosylhomocysteine, serotonin, fucose, putrescine, ketone, prostanoid, fatty acid, carbohydrate, spermidine etc. and amine, putrescine, folate biosynthetic processes, canonical Wnt receptor signaling pathway, histone H3 acetylation, the hippo signaling, processes associated with smooth, cardiac and skeletal muscle cell development, etc. are inferred to be associated with the cardiac development in previous studies [25–41].

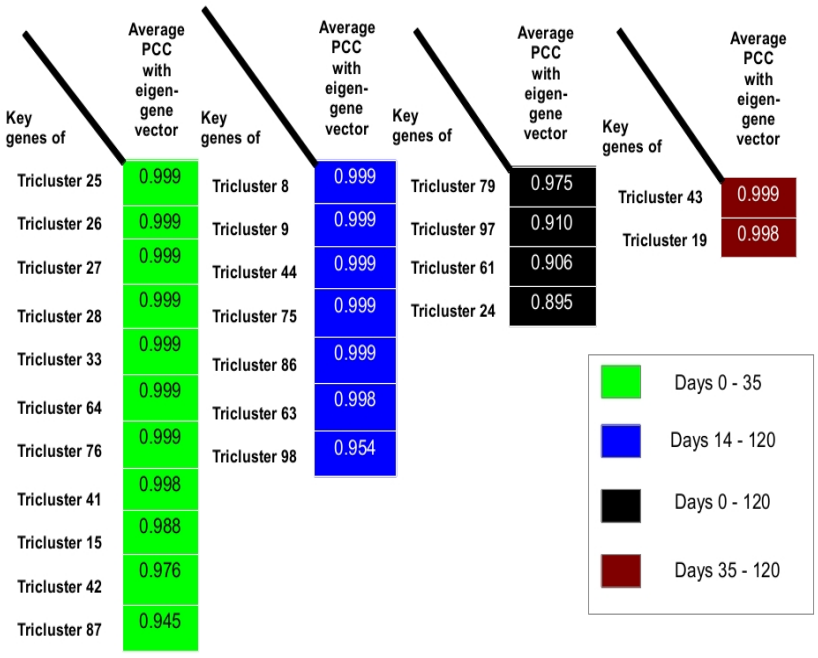

Figure 3.16: Average Pearson correlation coefficient between 10 top-most probe-ids of triclusters and the corresponding eigen-gene vectors during different phases of cardiomyocyte differentiation. The figure has been taken from [7].

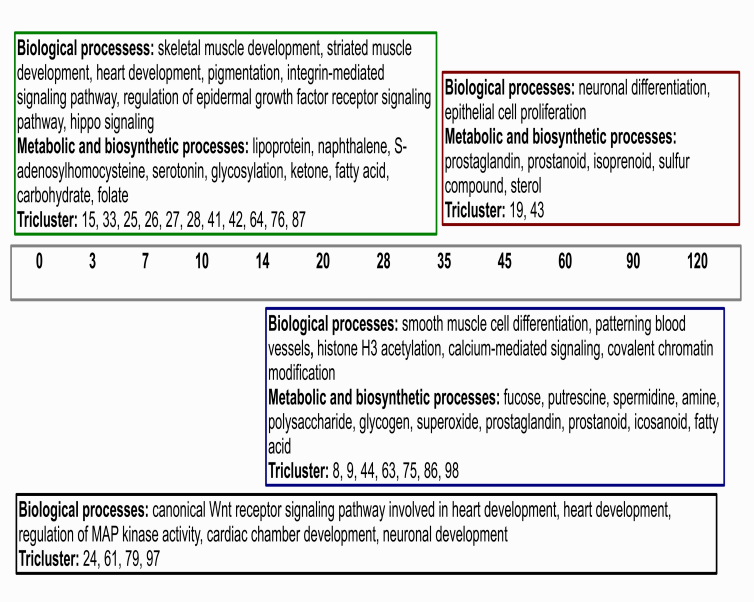

Figure 3.17: Summarization of enriched GOBPs and metabolic pathways during hiPSC differentiation to cardiomyocytes. Green, red, blue and black colored boxes represent the time points Days 0 to 35, Days 35 to 120, Days 14 to 120 and Days 0 to 120, respectively. The figure has been taken from [7].

3.5 Conclusion

In this work, we have presented an improved version (Evolutionary Multiobjective Optimization Algorithm for δ-TRIMAX) of our previously proposed triclustering algorithm δ-TRIMAX and showed that EMOA-δ-TRIMAX outperforms the existing triclustering algorithms in case of one artificial dataset and three real life datasets. Furthermore, we have applied our proposed algorithm to one of the aforementioned real life datasets monitoring the expression profiles of genes during stem cell differentiation into cardiomyocytes in order to unveil the key genes and potential biological processes and/ or KEGG pathways involved at different stages of the differentiation. Through our analysis, we have observed that most of the identified biological processes are known to play instrumental roles in the context of cardiomyocytes development. Furthermore, the roles of most of the identified key genes in cardiac development were already inferred by previous studies. The rest of the genes have been hypothesized as instrumental ones by interpreting and associating their functions with cardiac development or disorders and thus, need to be experimentally

verified. Altogether, our results may provide new insights into the potential therapeutic strategies to the treatment of cardiovascular diseases.

3.6 Bibliography

[1] Burridge, P.W., Keller, G., Gold, J.D., Wu, J.C.: Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. Cell Stem Cell **10**(1), 16–28 (2012). doi:10.1016/j.stem.2011.12.013

[2] Jacot, J.G., Martin, J.C., Hunt, D.L.: Mechanobiology of cardiomyocyte development. Journal of biomechanics **43**(1), 93–98 (2010). doi:10.1016/j.jbiomech.2009.09.014

[3] Ezzeldin, M., Borrego-Diaz, E., Taha, M., Esfandyari, T., Wise, A.L., Peng, W., Rouyanian, A., Asvadi, A. Kermani, Soleimani, M., Patrad, E., Lialyte, K., Wang, K., Williamson, S., Abdulkarim, B., Olyaee, M., Farassati, F.: RalA signaling pathway as a therapeutic target in hepatocellular carcinoma (HCC). Molecular Oncology **8**(5), 1043–1053 (2014). doi:10.1016/j.molonc.2014.03.020

[4] Anastas, J.N., Moon, R.T.: WNT signalling pathways as therapeutic targets in cancer. Nature Reviews Cancer **13**, 11–26 (2013). doi:10.1038/nrc3419

[5] Bhar, A., Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: δ-TRIMAX: Extracting Triclusters and Analysing Coregulation in Time Series Gene Expression Data. In: Raphael, B., Tang, J. (eds.) Algorithms in Bioinformatics. Lecture Notes in Computer Science, vol. 7534, pp. 165–177. Springer, Berlin Heidelberg (2012). doi:10.1007/978-3-642-33122-$0_1$3.

[6] Bhar, A., Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: Coexpression and coregulation analysis of time-series gene expression data in estrogen-induced breast cancer cell. Algorithms for Molecular Biology **8**(1), 9 (2013). doi:10.1186/1748-7188-8-9

[7] Bhar, A., Haubrock, M., Mukhopadhyay, A., Wingender, E.: Multiobjective Triclustering Of Time-Series Transcriptome Data Reveals Key Genes Of Biological Processes. BMC Bioinformatics, (in press) (2015)

[8] Carroll, J., Meyer, C., Song, J., Li, W., Geistlinger, T., Eeckhoute, J., Brodsky, A., Keeton, E., Fertuck, K., Hall, G., Wang, Q., Bekiranov, S., Sementchenko, V., FOX, E., Silver, P., Gingeras, T., Liu, X., Brown, M.: Genome-wide analysis of estrogen receptor binding sites. Nature Genetics **38**(11), 1289–1297 (2006)

[9] Babiarz, J., Ravon, M., Sridhar, S., Ravindran, P., Swanson, B., Bitter, H., Weiser, T., Chiao, E., Certa, U., Kolaja, K.: Determination of the human cardiomyocyte mRNA and miRNA differentiation network by fine-scale profiling. Stem Cells and Development **21**(11), 1956–1965 (2012). doi:10.1089/scd.2011.0357

[10] Hecker, M., Thamilarasan, M., Koczan, D., Schroeder, I., Flechtner, K., Freiesleben, S., Fuellen, G., Thiesen, H.J., Zettl, U.K.: MicroRNA Expression Changes during Interferon-Beta Treatment in the Peripheral Blood of Multiple Sclerosis Patients. International Journal of Molecular Sciences **14**(8), 16087–16110 (2013). doi:10.3390/ijms140816087

[11] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. Evolutionary Computation, IEEE Transactions on **6**(2), 182–197 (2002). doi:10.1109/4235.996017

[12] Maulik, U., Mukhopadhyay, A., Bhattacharyya, M., Kaderali, L., Brors, B., Bandyopadhyay, S., Eils, R.: Mining Quasi-Bicliques from HIV-1-Human Protein Interaction Network: A Multiobjective Biclustering Approach. IEEE/ACM Transactions on Computational Biology and Bioinformatics **10**(2), 423–435 (2013). doi:10.1109/TCBB.2012.139

[13] Spearman, C.: The proof and measurement of association between two things. The American journal of psychology **100**(3-4), 441–471 (1987)

[14] Fogel, D.B., Atmar, J.W.: Comparing Genetic Operators with Gaussian Mutations in Simulated Evolutionary Processes Using Linear Systems. Biological Cybernetics **63**, 111–114 (1990)

[15] Ishibuchi, H., Tsukamoto, N., Nojima, Y.: Evolutionary many-objective optimization: A short review (2008)

[16] Cheng, Y., Church, G.M.: Biclustering of expression data. Proceedings of International Conference on Intelligent Systems for Molecular Biology **8**, 93–103 (2000)

[17] Falcon, S., Gentleman, R.: Using GOstats to test gene lists for GO term association. Bioinformatics **23**(2), 257–258 (2007)

[18] Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) **57**(1), 289–300 (1995). doi:10.2307/2346101

[19] Chen, L., Wang, H., Zhang, L., Li, W., Wang, Q., Shang, Y., He, Y., He, W., Li, X., Tai, J., Li, X.: Uncovering packaging features of co-regulated modules based on human protein interaction and transcriptional regulatory networks. BMC Bioinformatics **11**, 392 (2010). doi:10.1186/1471-2105-11-392

[20] Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhaeuser, R., Pruess, M., Schacherer, F., Thiele, S., Urbach, S.: The TRANSFAC system on gene expression regulation. Nucleic Acids Research **29**(1), 281–283 (2001)

[21] Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., Sherlock, G.: GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics **20**(18), 3710–3715 (2004)

[22] Benjamini, Y., Yekutieli, D.: The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics **29**(4), 1165–1188 (2001). doi:10.1214/aos/1013699998

[23] Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., Kellis, M.: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature **434**(7031), 338–345 (2005)

[24] Mann, H.B., Whitney, D.R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics **18**(1), 50–60 (1947). doi:10.1214/aoms/1177730491

[25] Schaefer, E.J., Asztalos, B.F.: Cholesteryl ester transfer protein inhibition, high-density lipoprotein metabolism and heart disease risk reduction. Current Opinion in Lipidology **17**(4), 394–398 (2006)

[26] Harmon, H.J., Sanborn, M.R.: Effect of naphthalene on respiration in heart mitochondria and intact cultured cells. Environmental Research **29**(1), 160–173 (1982)

[27] Deussen, A., Lloyd, H.G., Schrader, J.: Contribution of S-adenosylhomocysteine to cardiac adenosine formation. Journal of Molecular and Cellular Cardiology **21**(8), 773–782 (1989)

[28] Tian, R., Ingwall, J.S.: How does folic acid cure heart attacks? Circulation **117**(14), 1772–1774 (2008). doi:10.1161/CIRCULATIONAHA.108.766105

[29] Herrmann, H., Kram, D.: Incorporation of fucose in the intact heart and dissociated heart cells of the chick embryo. Experimental Cell Research **107**(2), 455–456 (1977)

[30] Tagliavini, S., Genedani, S., Bertolini, A., Bazzani, C.: Ischemia- and reperfusion-induced arrhythmias are prevented by putrescine. European Journal of Pharmacology **194**(1), 7–10 (1991)

[31] Pegg, A.E., Hibasami, H.: Polyamine metabolism during cardiac hypertrophy. The American Journal of Physiology **239**(5), 372–378 (1980)

[32] Lopaschuk, G.D., Barr, R.L.: Measurements of fatty acid and carbohydrate metabolism in the isolated working rat heart. Molecular and Cellular Biochemistry **172**(1-2), 137–147 (1997)

[33] Krishnamurthy, M., Selvaraju, M., Tamilarasan, M.: Turbinaria conoides (J. Agardh) sulfated polysaccharide protects rat's heart against myocardial injury. International Journal of Biological Macromolecules **50**(5), 1275–1279 (2012). doi:10.1016/j.ijbiomac.2012.03.012

[34] Schaefer, S., Ramasamy, R.: Glycogen utilization and ischemic injury in the isolated rat heart. Cardiovascular Research **35**(1), 90–98 (1997)

[35] Qiu, H., Liu, J.Y., Wei, D., Li, N., Yamoah, E.N., Hammock, B.D., Chiamvimonvat, N.: Cardiac-generated prostanoids mediate cardiac myocyte apoptosis after myocardial ischaemia. Cardiovascular Research **95**(3), 336–345 (2012). doi:10.1093/cvr/cvs191

[36] Nebigil, C.G., Maroteaux, L.: A novel role for serotonin in heart. Trends in cardiovascular medicine **11**(8), 329–335 (2001)

[37] Cole, A.G., Meinertzhagen, I.A.: The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic Ciona intestinalis. Developmental Biology **271**(2), 239–262 (2004)

[38] Christoforou, N., Liau, B., Chakraborty, S., Chellapan, M., Bursac, N., Leong, K.W.: Induced pluripotent stem cell-derived cardiac progenitors differentiate to cardiomyocytes and form biosynthetic tissues. PLoS One **8**(6), e65963 (2013). doi:10.1371/journal.pone.0065963

[39] Lian, X., Hsiao, C., Wilson, G., Zhu, K., Hazeltine, L.B., Azarin, S.M., Raval, K.K., Zhang, J., Kamp, T.J., Palecek, S.P.: Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. PNAS **109**(27), 1848–1857 (2012). doi:10.1073/pnas.1200250109

[40] Otsuji, T.G., Kurose, Y., Suemori, H., Tada, M., Nakatsuji, N.: Dynamic link between histone H3 acetylation and an increase in the functional characteristics of human ESC/iPSC-derived cardiomyocytes. PLos One **7**(9), e45010, (2012). doi:10.1371/journal.pone.0045010

[41] Heallen, T., Morikawa, Y., Leach, J., Tao, G., Willerson, J.T., Johnson, R.L., Martin, J.F.: Hippo signaling impedes adult heart regeneration. Development **140**(23), 4683–4690 (2013). doi:10.1242/dev.102798

[42] Gergs, U., Boknik, P., Buchwalow, I.B., Fabritz, L., Grndker, N., Kucerova, D., Matus, M., Werner, F., Schmitz, W., Neumann, J.: Modulation of cardiac contractility by serine/threonine protein phosphatase type 5. International Journal of Cardiology **154**(2), 116–121 (2012). doi:10.1016/j.ijcard.2010.09.009

[43] Brinckmann, M., Kaschina, E., Altarche-Xifro, W., Curato, C., Timm, M., Grzesiak, A., Dong, J., Kappert, K., Kintscher, U., Unger, T., Li, J.: Estrogen receptor alpha supports cardiomyocytes indirectly through post-infarct cardiac c-kit+ cells. Journal of Molecular and Cellular Cardiology **47**(1), 66–75 (2009). doi:10.1016/j.yjmcc.2009.03.014

[44] Esler, M., Lambert, G., Brunner-La Rocca, H.P., Vaddadi, G., Kaye, D.: Sympathetic nerve activity and neurotransmitter release in humans: translation from pathophysiology into clinical practice. Acta Physiologica Scandinavica **177**(3), 275–284 (2003)

[45] Yang, J., Nie, Y., Wang, F., Hou, J., Cong, X., Hu, S., Chen, X.: Reciprocal regulation of miR-23a and lysophosphatidic acid receptor signaling in cardiomyocyte hypertrophy. Biochimica et Biophysica Acta **1831**(8), 1386–1394 (2013). doi:10.1016/j.bbalip.2013.05.005

[46] Samuels, M.A.: The BrainHeart Connection. Circulation **116**, 77–84 (2007)

[47] Landgren, H., Carlsson, P.: FoxJ3, a novel mammalian forkhead gene expressed in neuroectoderm, neural crest, and myotome. Developmental Dynamics **231**(2), 396–401 (2004)

[48] Consortium, T.U.: Update on activities at the Universal Protein Resource (UniProt) in 2013. Nucleic Acids Research **41**(D1), 43–47 (2013). doi:10.1093/nar/gks1068

[49] Zygmunt, A.C., Gibbons, W.R.: Calcium-activated chloride current in rabbit ventricular myocytes. Circulation Research **68**(2), 424–437 (1991)

[50] Kapur, N.K., Wilson, S., Yunis, A., Qiao, X., Mackey, E., Paruchuri, V., Baker, C., Aronovitz, M.J., Karumanchi, S.A., Letarte, M., Kass, D.A., Mendelsohn, M.E., Karas, R.H.: Reduced endoglin activity limits cardiac fibrosis and improves survival in heart failure. Circulation **125**(22), 2728–2738 (2012). doi:10.1161/CIRCULATIONAHA.111.080002

[51] Ma, Z., Yang, H., Liu, H., Xu, M., Runyan, R.B., Eisenberg, C.A., Markwald, R.R., Borg, T.K., Gao, B.Z.: Mesenchymal stem cell-cardiomyocyte interactions under defined contact modes on laser-patterned biochips. PLoS ONE **8**(2), e56554 (2013). doi:10.1371/journal.pone.0056554

[52] Nohl, H., Gille, L., Staniek, K.: The exogenous NADH dehydrogenase of heart mitochondria is the key enzyme responsible for selective cardiotoxicity of anthracyclines. Zeitschrift fuer Naturforschung. C, journal of biosciences **53**(3-4), 279–285 (1998)

[53] Liu, C.C., Jung, S.M., Orlandi, A., Yeh, T.S., Lin, Y.S., Shiu, T.F., Wu, H.H., Chu, J.J., Lin, P.J., Chu, P.H.: The Fas-mediated apoptotic pathway in cardiac myxoma. International Journal of Surgical Pathology **18**(6), 493–498 (2010). doi:10.1177/1066896910379480

[54] Maier, L.S., Bers, D.M., Brown, J.H.: Calmodulin and Ca^{2+}/calmodulin kinases in the heart Physiology and pathophysiology. Cardiovascular Research **73**(4), 629–630 (2007)

[55] Kliment, C.R., Suliman, H.B., Tobolewski, J.M., Reynolds, C.M., Day, B.J., Zhu, X., McTiernan, C.F., McGaffin, K.R., Piantadosi, C.A., Oury, T.D.: Extracellular superoxide dismutase regulates cardiac function and fibrosis. Journal of Molecular and Cellular Cardiology **47**(5), 730–742 (2009). doi:10.1016/j.yjmcc.2009.08.010

[56] Nithipongvanitch, R., Ittarat, W., Velez, J.M., Zhao, R., St Clair, D.K., Oberley, T.D.: Evidence for p53 as guardian of the cardiomyocyte mitochondrial genome following acute adriamycin treatment. The Journal of Histochemistry and Cytochemistry **55**(6), 629–639 (2007)

[57] Burke, J.A., Levi, R., Guo, Z.G., Corey, E.J.: Leukotrienes C4, D4 and E4: effects on human and guinea-pig cardiac preparations in vitro. The Journal of Pharmacology and Experimental Therapeutics **221**(1), 235–241 (1982)

[58] Golan, O., Issan, Y., Isak, A., Leipziger, J., Robaye, B., Shainberg, A.: Extracellular nucleotide derivatives protect cardiomyocytes against hypoxic stress. Biochemical Pharmacology **81**(10), 1219–1227 (2011). doi:10.1016/j.bcp.2011.02.007

[59] Zandi, S., Farajzadeh, S.: A new cardiac manifestation associated with woolly hair: report of two cases of woolly hair, palmoplantar keratoderma, and mitral valve regurgitation. International Journal of Dermatology **46**(9), 952–954 (2007)

[60] Makki, N., Thiel, K.W., Miller, F.J.J.: The epidermal growth factor receptor and its ligands in cardiovascular disease. International Journal of Molecular Sciences **14**(10), 20597–20613 (2013). doi:10.3390/ijms141020597

[61] McGhie, A.I., Radovancevic, B., Capek, P., Moore, W.H., Kasi, L., Lamki, L., Clubb, F.J.J., Frazier, O.H., Willerson, J.T.: Major histocompatibility complex class II antigen expression in rejecting cardiac allografts: detection using in vivo imaging with radiolabeled monoclonal antibody. Circulation **96**(5), 1605–1611 (1997)

[62] Lopez, B., Gonzalez, A., Querejeta, R., Larman, M., Diez, J.: Alterations in the pattern of collagen deposition may contribute to the deterioration of systolic function in hypertensive patients with heart failure. Journal of The American College of Cardiology **48**(1), 89–96 (2006)

[63] Minamino, T., Komuro, I., Kitakaze, M.: Endoplasmic reticulum stress as a therapeutic target in cardiovascular disease. Circulation Research **107**(9), 1071–1082 (2010). doi:10.1161/CIRCRESAHA.110.227819

[64] Casini, S., Tan, H.L., Demirayak, I., Remme, C.A., Amin, A.S., Scicluna, B.P., Chatyan, H., Ruijter, J.M., Bezzina, C.R., van Ginneken, A.C., Veldkamp, M.W.: Tubulin polymerization modifies cardiac sodium channel expression and gating. Cardiovascular Research **85**(4), 691–700 (2010). doi:10.1093/cvr/cvp352

[65] Massinen, S., Tammimies, K., Tapia-Paez, I., Matsson, H., Hokkanen, M.E., Soederberg, O., Landegren, U., Castren, E., Gustafsson, J.A., Treuter, E., Kere, J.: Functional interaction of DYX1C1 with estrogen receptors suggests involvement of hormonal pathways in dyslexia. Human Molecular Genetics **18**(15), 2802–2812 (2009). doi:10.1093/hmg/ddp215

[66] Du, X.J.: Clues to understanding the role of estrogen receptors in mediating cardiovascular protection. Cardiovascular Research **56**(1), 4–7 (2002)

[67] Eishi, Y., Takemura, T., Sone, R., Yamamura, H., Narisawa, K., Ichinohasama, R., Tanaka, M., Hatakeyama, S.: Glycogen storage disease confined to the heart with deficient activity of cardiac phosphorylase kinase: a new type of glycogen storage disease. Human Pathology **16**(2), 193–197 (1985)

[68] Park, T.S., Yamashita, H., Blaner, W.S., Goldberg, I.J.: Lipids in the heart: a source of fuel and a source of toxins. Current Opinion in Lipidology **18**(3), 277–282 (2007)

[69] Zolk, O., Schenke, C., Sarikas, A.: The ubiquitin-proteasome system: focus on the heart. Cardiovascular Research **70**(3), 410–421 (2006)

[70] Lian, X., Hsiao, C., Wilson, G., Zhu, K., Hazeltine, L., Azarin, S.M., Raval, K.K., Zhang, J., Kamp, T.J., Palecek, S.P.: Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. PNAS **109**(27), 1848–1857 (2012). doi:10.1073/pnas.1200250109

[71] Slachta, C.A., Jeevanandam, V., Goldman, B., Lin, W.L., Platsoucas, C.D.: Coronary arteries from human cardiac allografts with chronic rejection contain oligoclonal T cells: persistence of identical clonally expanded TCR transcripts from the early post-transplantation period (endomyocardial biopsies) to chronic rejection (coronary arteries). Journal of Immunology **165**(6), 3469–3483 (2000)

[72] Bechtold, D.A., Sidibe, A., Saer, B.R., Li, J., Hand, L.E., Ivanova, E.A., Darras, V.M., Dam, J., Jockers, R., Luckman, S.M., Loudon, A.S.: A role for the melatonin-related receptor GPR50 in leptin signaling, adaptive thermogenesis, and torpor. Current Biology **22**(1), 70–77 (2012). doi:10.1016/j.cub.2011.11.043

[73] McGaffin, K.R., Sun, C.K., Rager, J.J., Romano, L.C., Zou, B., Mathier, M.A., O'Doherty, R.M., McTiernan, C.F., O'Donnell, C.P.: Leptin signalling reduces the severity of cardiac dysfunction and remodelling after chronic ischaemic injury. Cardiovascular Research **77**(1), 54–63 (2008)

[74] Manukyan, M.C., Alvernaz, C.H., Poynter, J.A., Wang, Y., Brewster, B.D., Weil, B.R., Abarbanell, A.M., Herrmann, J.L., Crowe, B.J., Keck, A.C., Meldrum, D.R.: Interleukin-10 protects the ischemic heart from reperfusion injury via the STAT3 pathway. Surgery **150**(2), 231–239 (2011). doi:10.1016/j.surg.2011.05.017

[75] Ring, C., Ginsberg, M.H., Haling, J., Pendergast, A.M.: Abl-interactor-1 (Abi1) has a role in cardiovascular and placental development and is a binding partner of the $\alpha 4$ integrin. PNAS **108**(1), 149–154 (2011). doi:10.1073/pnas.1012316108

[76] Zelarayan, L., Renger, A., Noack, C., Zafiriou, M.P., Gehrke, C., van der Nagel, R., Dietz, R., de Windt, L., Bergmann, M.W.: NF-kappaB activation is required for adaptive cardiac hypertrophy. Cardiovascular Research **84**(3), 416–424 (2009). doi:10.1093/cvr/cvp237

[77] Thorpe, S.J., Hunt, B., Yacoub, M.: Expression of ABH blood group antigens in human heart tissue and its relevance to cardiac transplantation. Transplantation **51**(6), 1290–1295 (1991)

[78] Malik, F.I., Hartman, J.J., Elias, K.A., Morgan, B.P., Rodriguez, H., Brejc, K., Anderson, R.L., Sueoka, S.H., Lee, K.H., Finer, J.T., Sakowicz, R., Baliga, R., Cox, D.R., Garard, M., Godinez, G., Kawas, R., Kraynack, E., Lenzi, D., Lu, P.P., Muci, A., Niu, C., Qian, X., Pierce, D.W., Pokrovskii, M., Suehiro, I., Sylvester, S., Tochimoto, T., Valdez, C., Wang, W., Katori, T., Kass, D.A., Shen, Y.T., Vatner, S.F., Morgans, D.J.: Cardiac myosin activation: a potential therapeutic approach for systolic heart failure. Science **331**(6023), 1439–1443 (2011). doi:10.1126/science.1200113

[79] Parissis, J.T., Farmakis, D., Fountoulaki, K., Rigas, A., Nikolaou, M., Paraskevaidis, I.A., Bistola, V., Venetsanou, K., Ikonomidis, I., Anastasiou-Nana, M., Kremastinos,

D.T., Filippatos, G.: Clinical and neurohormonal correlates and prognostic value of serum prolactin levels in patients with chronic heart failure. European Journal of Heart Failure **15**(10), 1122–1130 (2013). doi:10.1093/eurjhf/hft070

[80] Willis, M.S., Schisler, J.C., Patterson, C.: Appetite for destruction: E3 ubiquitin-ligase protection in cardiac disease. Future Cardiology **4**(1), 65–75 (2008). doi:10.2217/14796678.4.1.65

[81] Hermansson, C., Lundqvist, A., Wasslavik, C., Palmqvist, L., Jeppsson, A., Hulten, L.M.: Reduced expression of NLRP3 and MEFV in human ischemic heart tissue. Biochemical and Biophysical Research Communication **430**(1), 425–428 (2013). doi:10.1016/j.bbrc.2012.11.070

[82] Bushdid, P.B., Osinska, H., Waclaw, R.R., Molkentin, J.D., Yutzey, K.E.: NFATc3 and NFATc4 are required for cardiac development and mitochondrial function. Circulation Research **92**(12), 1305–1313 (2003)

[83] Sun, S., Gui, Y., Jiang, Q., Song, H.: Dihydrofolate reductase is required for the development of heart and outflow tract in zebrafish. Acta Biochim Biophys Sin **43**(12), 957–969 (2011). doi:10.1093/abbs/gmr098

[84] Kim, S., Zaghloul, N.A., Bubenshchikova, E., Oh, E.C., Rankin, S., Katsanis, N., Obara, T., Tsiokas, L.: Nde1-mediated inhibition of ciliogenesis affects cell cycle re-entry. Nature Cell Biology **13**(4), 351–360 (2011). doi:10.1038/ncb2183

[85] DiMicco, J.A., Gale, K., Hamilton, B., Gillis, R.A.: GABA receptor control of parasympathetic outflow to heart: characterization and brainstem localization. Science **204**(4397), 1106–1109 (1979)

[86] Wang, J., Irnaten, M., Neff, R.A., Venkatesan, P., Evans, C., Loewy, A.D., Mettenleiter, T.C., Mendelowitz, D.: Synaptic and neurotransmitter activation of cardiac vagal neurons in the nucleus ambiguus. Annals of the New York Academy of Sciences **940**, 237–246 (2001)

[87] Lopez-Sanchez, C., Bartulos, O., Martinez-Campos, E., Ganan, C., Valenciano, A.I., Garcia-Martinez, V., De, F. Pablo, Hernandez-Sanchez, C.: Tyrosine hydroxylase is expressed during early heart development and is required for cardiac chamber formation. Cardiovascular Research **88**(1), 111–120 (2010). doi:10.1093/cvr/cvq179

[88] Tsika, R.W., Ma, L., Kehat, I., Schramm, C., Simmer, G., Morgan, B., Fine, D.M., Hanft, L.M., McDonald, K.S., Molkentin, J.D., Krenz, M., Yang, S., Ji, J.: TEAD-1 overexpression in the mouse heart promotes an age-dependent heart dysfunction. The Journal of Biological Chemistry **285**(18), 13721–13735 (2010). doi:10.1074/jbc.M109.063057

[89] Attramadal, H.: Prokineticins and the heart: diverging actions elicited by signalling through prokineticin receptor-1 or -2. Cardiovascular Research **81**(1), 3–4 (2009). doi:10.1093/cvr/cvn306

[90] Lesch, K.P., Araragi, N., Waider, J., van den Hove, D., Gutknecht, L.: Targeting brain serotonin synthesis: insights into neurodevelopmental disorders with long-term outcomes related to negative emotionality, aggression and antisocial behaviour. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **367**(1601), 2426–2443 (2012). doi:10.1098/rstb.2012.0039

[91] Nebigil, C.G., Maroteaux, L.: A novel role for serotonin in heart. Trends in cardiovascular medicine **11**(8), 329–335 (2001)

[92] Liu, H., Pedram, A., Kim, J.K.: Oestrogen prevents cardiomyocyte apoptosis by suppressing p38α-mediated activation of p53 and by down-regulating p53 inhibition on p38β. Cardiovascular Research **89**, 119–128 (2011). doi:10.1093/cvr/cvq265

[93] Deschamps, A.M., Murphy, E.: Activation of a novel estrogen receptor, GPER, is cardioprotective in male and female rats. American Journal of Physiology - Heart and Circulatory Physiology **297**(5), 1806–1813 (2009). doi:10.1152/ajpheart.00283.2009

[94] Akhter, T., Sawada, N., Yamaguchi, M.: Regucalcin increases Ca2+-ATPase activity in the heart mitochondria of normal and regucalcin transgenic rats. International Journal of Molecular Medicine **18**(1), 171–176 (2006)

[95] Kwon, C., Arnold, J., Hsiao, E.C., Taketo, M.M., Conklin, B.R., Srivastava, D.: Canonical Wnt signaling is a positive regulator of mammalian cardiac progenitors. Proceedings of The National Academy of Sciences of the United States of America **104**(26), 10894–10899 (2007)

[96] Francis, G.A., Annicotte, J.S., Auwerx, J.: PPAR-alpha effects on the heart and other vascular tissues. The American Journal of Physiology. Heart and Circulatory Physiology **285**(1), 1–9 (2003)

[97] Cowan, C.A., Yokoyama, N., Saxena, A., Chumley, M.J., Silvany, R.E., Baker, L.A., Srivastava, D., Henkemeyer, M.: Ephrin-B2 reverse signaling is required for axon pathfinding and cardiac valve formation but not early vascular development. Developmental Biology **271**(2), 263–271 (2004)

[98] Schmitteckert, S., Ziegler, C., Kartes, L., Rolletschek, A.: Transcription Factor Lbx1 Expression in Mouse Embryonic Stem Cell-Derived Phenotypes. Stem Cells International **2011**, 130970 (2011). doi:10.4061/2011/130970

[99] Field, S., Riley, K.L., Grimes, D.T., Hilton, H., Simon, M., Powles-Glover, N., Siggers, P., Bogani, D., Greenfield, A., Norris, D.P.: Pkd1l1 establishes left-right asymmetry and physically interacts with Pkd2. Development **138**(6), 1131–1142 (2011). doi:10.1242/dev.058149

[100] Chen, C.M., Norris, D., Bhattacharya, S.: Transcriptional Control of LeftRight Patterning in Cardiac Development. Pediatric Cardiology **31**(3), 371–377 (2010). doi:10.1007/s00246-009-9610-3

[101] Fouladkou, F., Lu, C., Jiang, C., Zhou, L., She, Y., Walls, J.R., Kawabe, H., Brose, N., Henkelman, R.M., Huang, A., Bruneau, B.G., Rotin, D.: The ubiquitin ligase Nedd4-1 is required for heart development and is a suppressor of thrombospondin-1. The Journal of Biological Chemistry **285**(9), 6770–6780 (2010). doi:10.1074/jbc.M109.082347

[102] den Hollander, A.I., Ghiani, M., de Kok, Y.J., Wijnholds, J., Ballabio, A., Cremers, F.P., Broccoli, V.: Isolation of Crb1, a mouse homologue of Drosophila crumbs, and analysis of its expression pattern in eye and brain. Mechanisms of Development **110**(1-2), 203–207 (2002)

[103] Cole, L.K., Dolinsky, V.W., Dyck, J.R., Vance, D.E.: Impaired phosphatidylcholine biosynthesis reduces atherosclerosis and prevents lipotoxic cardiac dysfunction in ApoE-/- Mice. Circulation Research **108**(6), 686–694 (2011). doi:10.1161/CIRCRESAHA.110.238691

[104] Fiore, R., Rahim, B., Christoffels, V.M., Moorman, A.F., Pschel, A.W.: Inactivation of the Sema5a gene results in embryonic lethality and defective remodeling of the cranial vascular system. Molecular and Cellular Biology **25**(6), 2310–2319 (2005)

[105] Hainfellner, J.A., Voigtlnder, T., Strbel, T., Mazal, P.R., Maddalena, A.S., Aguzzi, A., Budka, H.: Fibroblasts can express glial fibrillary acidic protein (GFAP) in vivo. Journal of Neuropathology and Experimental Neurology **60**(5), 449–461 (2001)

[106] Dell'Era, P., Ronca, R., Coco, L., Nicoli, S., Metra, M., Presta, M.: Fibroblast growth factor receptor-1 is essential for in vitro cardiomyocyte development. Circulation Research **93**(5), 414–420 (2003)

[107] Bushdid, P.B., Osinska, H., Waclaw, R.R., Molkentin, J.D., Yutzey, K.E.: NFATc3 and NFATc4 are required for cardiac development and mitochondrial function. Circulation Research **92**(12), 1305–1313 (2003)

[108] Pan, L., Li, Y., Jia, L., Qin, Y., Qi, G., Cheng, J., Qi, Y., Li, H., Du, J.: Cathepsin S deficiency results in abnormal accumulation of autophagosomes in macrophages and enhances Ang II-induced cardiac inflammation. PLosOne **7**(4), e35315 (2012). doi:10.1371/journal.pone.0035315

[109] Gustafsson, A.B., Gottlieb, R.A.: Autophagy in ischemic heart disease. Circulation Research **104**(2), 150–158 (2009). doi:10.1161/CIRCRESAHA.108.187427

[110] Peal, D.S., Burns, C.G., Macrae, C.A., Milan, D.: Chondroitin sulfate expression is required for cardiac atrioventricular canal formation. Developmental Dynamics **238**(12), 3103–3110 (2009). doi:10.1002/dvdy.22154

[111] Blaschke, R.J., Hahurij, N.D., Kuijper, S., Just, S., Wisse, L.J., Deissler, K., Maxelon, T., K, A., Spitzer, J., Hardt, S.E., Schler, H., Feitsma, H., Rottbauer, W., Blum, M., Meijlink, F., Rappold, G., Gittenberger-de, A.C. Groot: Targeted mutation reveals essential functions of the homeodomain transcription factor Shox2 in sinoatrial and pacemaking development. Circulation **115**(14), 1830–1838 (2007)

[112] Norden, J., Kispert, A.: Wnt/Ctnnb1 signaling and the mesenchymal precursor pools of the heart. Trends in Cardiovascular Medicine **22**(5), 118–122 (2012). doi:10.1016/j.tcm.2012.07.006

[113] Rui, Y., Xu, Z., Xiong, B., Cao, Y., Lin, S., Zhang, M., Chan, S.C., Luo, W., Han, Y., Lu, Z., Ye, Z., Zhou, H.M., Han, J., Meng, A., Lin, S.C.: A beta-catenin-independent dorsalization pathway activated by Axin/JNK signaling and antagonized by Aida. Developmental Cell **13**(2), 268–282 (2007)

3.7 Appendix

Table 3.7: Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during cardiomyocytes differentiation. Functions of bold gene symbols have been described in Tables 3.12-3.16. The contents of the table have been taken from [7].

Tricluster	GOBP/ KEGG Pathway	Time Points	Samples	Key Genes
15	BP: GO:0045843: negative regulation of striated muscle tissue development (LUC7L), GO:0048635: negative regulation of muscle organ development (LUC7L), GO:0048634: regulation of muscle organ development (LUC7L), GO:0014706: striated muscle tissue development (LUC7L) ; KP: KEGG:00790: folate biosynthesis (DHFR), KEGG: 00670: one carbon pool by folate (DHFR)	3day-7day-10day-28day-35day	S1-S2-S3	**DHFR/ DHFRP1**, **GREB1**, ACTR3BP3, **ZNF365**, **SYN1**, PRM2, **LUC7L**
25	BP: GO:0035329: hippo signaling (TEAD1), GO:0070085: glycosylation (ABO), GO:0043413: macromolecule glycosylation (ABO); KP: KEGG:00601: Glycosphingolipid biosynthesis - lacto and neolacto series (ABO), KEGG:04270: Vascular smooth muscle contraction (GNA12)	0day-20day-35day	S1-S2	SNHG14, FABP5P11, **TEAD1**, EEF1A1P15, **PROKR2**, ABO, EVI2A, **BEST3**, GNA12
26	BP: GO:0042427: serotonin biosynthetic process (TPH2), GO:0042428: serotonin metabolic process (TPH2), GO:0007623: circadian rhythm (TPH2), GO:0043627: response to estrogen stimulus (TPH2)	3day-10day-14day	S2-S3	**TCTEX1D4**, ZGPAT, OR9A4, **TPH2**, ENTPD5, MPP7, TMEM245, HTR5A
27	BP: GO:0045742: positive regulation of epidermal growth factor receptor signaling pathway (GPER), GO:0019722: calcium-mediated signaling (RGN), GO:0070374: positive regulation of ERK1 and ERK2 cascade (GPER), GO:0070371: ERK1 and ERK2 cascade (GPER), GO:0007173: epidermal growth factor receptor signaling pathway (GPER)	3day-14day-20day	S2-S3	ZRANB3, **NDUFAF6**, **RGN**, FAM163B, C18orf62, **GPER**, **FAIM2**, LETM2, NUB1, **EMR1**
28	BP: GO:0072366: regulation of cellular ketone metabolic process by positive regulation of transcription from RNA polymerase II promoter (PPARA), GO:0045923: positive regulation of fatty acid metabolic process (PPARA), GO:0006109: regulation of carbohydrate metabolic process (PPARA), GO:0042157: lipoprotein metabolic process (PPARA); KP: KEGG:04340: Hedgehog signaling pathway (WNT8A)	0day-3day-20day	S1-S2	MFSD6L, TTTY16, **WNT8A**, PAK7, **PPARA**, **EPHB2**, MRPL49, DCHS2, **CALML4**, ASB4

Table 3.8: Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during cardiomyocytes differentiation. Functions of bold gene symbols have been described in Tables 3.12-3.16. The contents of the table have been taken from [7].

Tricluster	GOBP/ KEGG Pathway	Time Points	Samples	Key Genes
33	BP: GO:0060537: muscle tissue development (COL19A1), GO:0060538: skeletal muscle organ development (COL19A1), GO:0014706: striated muscle tissue development (COL19A1)	0day-10day-20day	S1-S2	ITGB5-AS1, SLC25A51P1, CLDN18, FAM90A6P, FAM90A7P, TTTY11, MAGEB6P1, COL19A1, C1R, UBXN7
41	BP: GO:0007229: integrin-mediated signaling pathway (ADAM33), GO:0043473: pigmentation (SZT2)	0day-10day-14day-20day	S1-S2	OR12D3, ADAM33, **SZT2**, ABCA13, ERI2, FANK1
42	BP: GO:0035050: embryonic heart tube development (LBX1), GO:0007368: determination of left/right symmetry (LBX1, PKD1L1), GO:0001947: heart looping (LBX1), GO:0003143: embryonic heart tube morphogenesis (LBX1), GO:0061371: determination of heart left/right asymmetry (LBX1), GO:0006508: proteolysis (CUL9, DPEP1)	0day-3day-7day-10day-20day-35day	S1-S2	**PKD1L1**, **KRTAP5-4**, **DPEP1**, CYTIP, **LBX1**, PRR20A/ B/ E, **CUL9**, **DNPH1**, RSPH10B/ B2
64	BP: GO:0042157: lipoprotein metabolic process (PEMT), GO:0018931: naphthalene metabolic process (CYP2F1), GO:0046498: S-adenosylhomocysteine metabolic process (PEMT), GO:0033146: regulation of estrogen receptor signaling pathway (DYX1C1)	10day-14day-20day-28day	S1-S2-S3	DYX1C1-CCPG1, YSK4, CYP2F1, **PEMT**, TMEM59L, RGS9BP, OR10G9, **IAPP**, **DYX1C1**, PDE11A
76	BP: GO:0014002: astrocyte development (GFAP), GO:0014010: Schwann cell proliferation (GFAP), GO:0009566: fertilization (TDRD9)	0day-3day-20day-28day	S1-S2-S3	LINC00330, **LAT**, TDRD9, BARHL2, **GIPC2**, HIST1H3J, FAM115C, KRT78, **GFAP**, **GPR50**
87	BP: GO:0048743: positive regulation of skeletal muscle fiber development (SHOX2), GO:0060272: embryonic skeletal joint morphogenesis (SHOX2), GO:0045844: positive regulation of striated muscle tissue development (SHOX2), GO:0003230: cardiac atrium development (SHOX2), GO:0003170: heart valve development (SHOX2), GO:0003205: cardiac chamber development (SHOX2). GO:0051147: regulation of muscle cell differentiation (SHOX2)	0day-3day-7day-10day-20day-28day-35day	S1-S2	MUC6, DPEP1, CUL9, **MYO5B**, SPACA5/ 5B, ZNF407, **SHOX2**

Table 3.9: Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways since the first heart beat was observed. Functions of bold gene symbols have been described in Tables 3.12-3.16. The contents of the table have been taken from [7].

Tricluster	GOBP/ KEGG Pathway	Time Points	Samples	Key Genes
8	BP: GO:0043966: histone H3 acetylation (TADA2A), GO:0016569: covalent chromatin modification (DNMT3A, TADA2A), GO:0043414: macromolecule methylation (DNMT3A)	14day-20day-28day-60day	S2-S3	LRRC37BP1, EPB41, DNMT3A, **NFATC4**, UVSSA, PTGDR2, SLC5A10, IGHV4-31, LRRC37B, TADA2A
9	BP: GO:0060263: regulation of respiratory burst (NOXO1), GO:0006801: superoxide metabolic process (NOXO1)	14day-60day-120day	S2-S3	FTH1P6, GRTP1, SPNS3, ZNF445, CEACAM5, RUNX2, RTBDN, HTR3D, NOXO1, ZNF41
44	BP: GO:0019722: calcium-mediated signaling (SELE), GO:0002687: positive regulation of leukocyte migration (SELE), GO:0005977: glycogen metabolic process (GYG2), GO:0005976: polysaccharide metabolic process (GYG2), GO:0044042: glucan metabolic process (GYG2), GO:0010517: regulation of phospholipase activity (SELE)	14day-20day-45day	S1-S2-S3	PART1, RGPD5, RGPD6, **CTHRC1**, RGPD8, GYG2, SHC1P1, SYT12, SELE, **CASP4**
63	BP: GO:0050850: positive regulation of calcium-mediated signaling (TRAT1), GO:0050851: antigen receptor-mediated signaling pathway (TRAT1)	14day-28day-60day-120day	S2-S3	ZFHX2, **AGBL3**, TRAT1, GABRQ, ACRV1, TTLL9, CMIP, SLC22A10
75	BP: GO:0006801: superoxide metabolic process (DUOX1, PREX1), GO:0006693: prostaglandin metabolic process (PDPN), GO:0006690: icosanoid metabolic process (PDPN), GO:0048286: lung alveolus development (PDPN), GO:0006692: prostanoid metabolic process (PDPN)	14day-28day-90day-120day	S2-S3	**ASB12**, **MTMR8**, CD200R1, **LGR6**, DUOX1, ZNF442, PDPN, PREX1, TGIF1, SLC22A9
86	BP: GO:0051145: smooth muscle cell differentiation (NFATC3), GO:0001569: patterning of blood vessels (NFATC3), GO:0055001: muscle cell development (NFATC3), GO:0006004: fucose metabolic process (FUT2); KP: KEGG:04370: VEGF signaling pathway (NFATC3)	14day-28day-35day	S1-S2-S3	FTH1P6, INTS6-AS1, **NFATC3**, **OLFML2B**, TTLL3, RPL14P5, **LRR1**, KIAA1328, **CTSS**, **FUT2**
98	BP: GO:0009445: putrescine metabolic process (AGMAT), GO:0009446: putrescine biosynthetic process (AGMAT), GO:0008295: spermidine biosynthetic process (AGMAT), GO:0008216: spermidine metabolic process (AGMAT), GO:0033146: regulation of estrogen receptor signaling pathway (DYX1C1), GO:0009309: amine biosynthetic process(AGMAT)	14day-20day-28day-45day-60day-90day-120day	S2-S3	DYX1C1-CCPG1, **NLRP11**, CETN4P, AWAT2, ANO9, AGMAT, SRGAP2B, UVSSA, DYX1C1, **AIDA**

Table 3.10: Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during all stages of hiPSC-derived cardiomyocyte differentiation. Functions of bold gene symbols have been described in Tables 3.12-3.16. The contents of the table have been taken from [7].

Tricluster	GOBP/ KEGG Pathway	Time Points	Samples	Key Genes
24	BP: GO:0060047: heart contraction (TH, ADRBK1), GO:0003015: heart process (TH, ADRBK1), GO:0003007: heart morphogenesis (LBX1, TH), GO:0007507: heart development (LBX1, TH), GO:0035050: embryonic heart tube development (LBX1), GO:0001947: heart looping (LBX1), GO:0002026: regulation of the force of heart contraction (ADRBK1)	3day-10day-14day-35day-45day-60day-90day-120day	S1-S2-S3	**FOXJ3**, **PCDHA9**, **TH**, TRIM29, ANO9, LBX1, BOD1L1, TSSK3, **ADRBK1**, SPEF2
61	BP: GO:0000093: mitotic telophase (MAD1L1), GO:0000089: mitotic metaphase (MAD1L1), GO:0000090: mitotic anaphase (MAD1L1), GO:0001654: eye development (CRB1), GO:0001754: eye photoreceptor cell differentiation (CRB1)	0day-3day-7day-10day-14day-20day-28day-35day-45day	S1-S2	HERC2P9, CYTIP, **CRB1**, CCT8L1P, MAD1L1
79	BP: GO:0003279: cardiac septum development (FRS2), GO:0003205: cardiac chamber development (FRS2), GO:0003231: cardiac ventricle development (FRS2), GO:0003281: ventricular septum development (FRS2), GO:0000187: activation of MAPK activity (FRS2), GO:0043406: positive regulation of MAP kinase activity (FRS2)	7day-20day-28day-35day-60day-90day	S1-S2-S3	OR2T27, OR2T7, **FRS2**, NAA35, **EPS8L3**, **ICOS**, SPAG16
97	BP: GO:0044336: canonical Wnt receptor signaling pathway involved in negative regulation of apoptosis (CTNNB1), GO:0044334: canonical Wnt receptor signaling pathway involved in positive regulation of epithelial to mesenchymal transition (CTNNB1), GO:0060912: cardiac cell fate specification (CTNNB1), GO:0003306: Wnt receptor signaling pathway involved in heart development (CTNNB1), GO:0061316: canonical Wnt receptor signaling pathway involved in heart development (CTNNB1), GO:0003129: heart induction (CTNNB1), GO:0061311: cell surface receptor linked signaling pathway involved in heart development (CTNNB1); KP: KEGG:04530: tight junction (CTNNB1), KEGG:5412: Arrhythmogenic right ventricular cardiomyopathy (ARVC) (CTNNB1)	0day-3day-10day-14day-28day-45day-60day-120day	S2-S3	LRRC37A4P, RPLP0, ZNF709, **PRLH**, LRRC37A2, WDR89, OR11H6, **CCNB1IP1**, **CTNNB1**

Table 3.11: Time points, samples, key genes of triclusters and the enriched gene ontology biological processes and/ or KEGG pathways during maturation. Functions of bold gene symbols have been described in Tables 3.12-3.16. The contents of the table have been taken from [7].

Tricluster	GOBP/ KEGG Pathway	Time Points	Samples	Key Genes
19	BP: GO:0048934: peripheral nervous system neuron differentiation (RUNX3), GO:0050680: negative regulation of epithelial cell proliferation (RUNX3), GO:0050678: regulation of epithelial cell proliferation (RUNX3)	35day-45day-60day-90day	S1-S2-S3	MYNN, LCN8, **GABRG1**, KRTAP5-5, RUNX3
43	GO:0042357: thiamine diphosphate metabolic process (TPK1), GO:0031958: corticosteroid receptor signaling pathway (NEDD4), GO:0046457: prostanoid biosynthetic process (CD74), GO:0048010: vascular endothelial growth factor receptor signaling pathway (NEDD4); KP: KEGG:00430: Taurine and hypotaurine metabolism (CSAD), KEGG:00900: Terpenoid backbone biosynthesis (IDI2)	45day-60day-90day	S1-S2-S3	KRTAP5-1, IZUMO4, MED12L, CSAD, **ANKRD13D**, **NEDD4**, IDI2, TPK1, ADAM18, **CD74**

Table 3.12: Genes that might be associated with cardiac disorders, cardiovascular development. The contents of the table have been taken from [7].

Tricluster	Genes	Functions
15	*LUC7L*	Serine or arginine proteins are known to be associated with cardiac functions [42]. As *LUC7L* binds to Arg/Ser-rich domain, we hypothesize *LUC7L* as a potential drug target to impede cardiac disorders.
15	*GREB1*	It is known as estrogen responsive protein and can be used as a potential drug target as ERalpha has been reported to play an pivotal role in cardioprotection against cardiac injury [43].
15	*SYN1*	It is a well known regulator of neurotransmitter release. Hence, this finding is quite surprising to us. However a previous study reports that the concentration of the sympathetic nervous neurotransmitter is associated with failing human heart [44]. Thus, *SYN1* might be considered as a potential drug target to prevent cardiac failure.
24	*ADRBK1*	It is known to be a key regulator of LPAR1 signaling which is reported to be involved in mediating cardiomyocyte hypertrophy by a previous work [45].
24	*PCDHA9*	Protocadherin alpha-9 plays a key role in the development and maintenance of specific neuronal connections in the brain. Though it's quite surprising to us that *PCDHA9* has been identified as one of the hubs in a coexpression network, a previous study inferred the brain's effects on the heart for instance neurogenic heart diseases [46].
24	*FOXJ3*	A previous study reports that *FOXJ3* expression is associated with cardiac neural crest and thus it might provide insights into cardiovascular system development [47].
25	*BEST3*	It is known to form calcium-sensitive chloride channel which is known to be associated with heart [48, 49].
26	*TCTEX1D4*	It is known to be an interaction partner of endoglin which is reported to be a potential drug target to prevent cardiac disorders [50]. Hence, we hypothesize targeting *TCTEX1D4* might provide a promising novel therapeutic approach for individuals with cardiac dysfunctions.
27	*EMR1*	It is known to be involved in cell-cell interactions which facilitates in understanding cell based therapies for restoring cardiomyocyte loss during cardiac diseases [51].
27	*NDUFAF6*	It participates in assembly of mitochondrial NADH, dehydrogenase of which plays a key role in cardiotoxicity [52]. Hence, *NDUFAF6* can be hypothesized as potential drug target to prevent heart dysfunctions or muscle damage.
27	*FAIM2*	It is a well-known regulator of Fas-mediated apoptosis in neurons and this pathway is reported to be involved in cardiac myxoma which may originate from sensory nerve tissue [52, 53]. So we hypothesize *FAIM2* as a potential drug target to prevent cardiac myxoma.
28	*CALML4*	It belongs to calmodulin family which may play pivotal role in cardiac function [48, 54].
41	*SZT2*	It may take part in superoxide dismutase which may be involved in regulating cardiac functions [48, 55].
42	*CUL9*	It is known to regulate subcellular localization of p53. A previous study inferred that mitochondrial p53 plays an important role in repairing mitochondrial DNA as a response to oxidative damage of cardiomyocyte [56]. Thus, targeting *CUL9* might prevent oxidative damage of cardiomyocyte.

Table 3.13: Genes that might be associated with cardiac disorders, cardiovascular development. The contents of the table have been taken from [7].

Tricluster	Genes	Functions
42	*DPEP1*	It is known to transform leukotriene D4 to leukotriene E4 and thus involved in regulating activities of leukotrienes myocardial, sedative effects of which play a key role in cardiac dysfunction [57]. Hence, *DPEP1* can be used as a potential drug target to prevent cardiac abnormality associated with leukotrienes reactions.
42	*DNPH1*	It is known to generate purine that has cardioprotective effects against hypoxic stress [58]. Thus, *DNPH1* can be used as a potential drug target to impede hypoxic stress.
42	*KRTAP5-4*	It is quite surprising that keratin-associated protein 5-4 has been identified as a hub gene in coexpression network as it is known to be involved in forming resistant hair shaft. However, an association between mitral valve regurgitation and woolly hair (a hair shaft disorder) was inferred by a previous study [59]. Thus, we hypothesize *KRTAP5-4* as a potential drug target to prevent mitral valve regurgitation.
43	*ANKRD13D*	It is known to be involved in positive regulation of ligand-activated epidermal growth factor receptor that may be associated with cardiac diseases [48, 60]. Thus, *ANKRD13D* can be hypothesized as a potential drug target to prevent cardiac disease.
43	*CD74*	It is inferred to play a crucial role in processing MHC class II antigen which is capable of detecting cardiac allograft rejection [48, 61]. Thus, *CD74* may provide insights into the context of allograft rejection.
44	*CTHRC1*	It is known to be involved in collagen deposition modification of which may result in deterioration of systolic function of heart patients [48, 62].
44	*CASP4*	It plays a pivotal role in endoplasmic reticulum (ER) stress induced apoptosis and thus can be hypothesized as a potential target to protect heart disease [48, 63].
63	*AGBL3*	A previous study reports that *AGBL3* plays an important role to process tubulin which has has been inferred to have a cardioprotective effect [48, 64].
64	*DYX1C1*	It was inferred to be a potential regulator of estrogen receptor stability; thus can be hypothesized to play an instrumental role in cardiovascular development. [65, 66].
64	*IAPP*	It can be hypothesized to play a crucial role in preventing cardiomegaly by inhibiting glycogen deposition. [67].
75	*MTMR8*	It is inferred to act on lipid which is an important enzyme in heart [48, 68].
75	*ASB12*	It is reported to be a mediator of ubiquitination and subsequent proteasomal degradation to play a cardioprotective role [48, 69].
75	*LGR6*	It enhances canonical Wnt signaling pathway which is one of the most important regulators of cardiomyocyte differentiation [48, 70].

Table 3.14: Genes that might be associated with cardiac disorders, cardiovascular development. The contents of the table have been taken from [7].

Tricluster	Genes	Functions
76	*LAT*	Linker for activation of T-cells family member 1 (*LAT*) is known to be needed for T-cell antigen receptor. A previous study reports the presence of T-cells in coronary arteries of cardiac allografts from patients with graft arteriosclerosis (GA) [71]. Hence we hypothesize *LAT* as a potential drug target to prevent graft arteriosclerosis.
76	*GPR50*	*GPR50* known as melatonin-related receptor, was inferred to play a key role in leptin signaling which reduces the severity of cardiac disorders [72, 73]. Hence, *GPR50* can be hypothesized as a potential drug target to prevent cardiac dysfunctions.
79	*ICOS*	Inducible T-cell costimulator promotes the synthesis of interleukin-10 which is inferred to protect ischemic heart from reperfusion injury [74]. Hence we can hypothesize *ICOS* as a potential drug target for individuals with cardiac reperfusion injury.
79	*EPS8L3*	It is a well known interaction partner of ABI1 which has a key role in cardiovascular development [75].
86	*LRR1*	It is known to activate NF-kappaB by negatively regulating 4-1BB-mediated signaling cascades. A previous study reports that NF-kappaB is needed for adaptive cardiac hypertrophy [76]. Thus targeting *LRR1* might provide a novel insight into the treatment of cardiac hypertrophy.
86	*FUT2*	It is known to create H-antigen which is found to be pertinent to cardiac transplantation [77]. Thus, targeting *FUT2* might provide new insights into the treatment for heart transplantation.
87	*MYO5B*	It belongs to myosin family. One of the primary structural proteins of cardiac muscle is myosin, activation which plays an instrumental role in preventing systolic heart failure [78]. Hence, *MYO5B* can be hypothesized as a potential drug target to prevent cardiovascular diseases.
97	*PRLH*	Prolactin-releasing peptide is known to release prolactin (PRL) which is inferred to provide insights into prognosis of advanced chronic heart failure [79]. Hence we hypothesize *PRLH* as a potential target for patients with chronic heart failure.
97	*CCNB1IP1*	A previous study reports E3 ubiquitin-ligase family of proteins as promising targets to impede cardiac diseases [80]. As *CCNB1IP1* is a E3 ubiquitin-protein ligase, targeting this protein might provide new insights into therapeutics of heart diseases.
98	*NLRP11*	It belongs to NLRP family, member of which (NLRP3) is found to be associated with cardiac disorders [81].

Table 3.15: Genes that are known to be associated with cardiovascular development. The contents of the table have been taken from [7].

Tricluster	Genes	Functions
8	*NFATC4*	A previous study infers *NFATC4* as a potential regulator in heart development [82].
15	*DHFR*	It has been identified as key gene that is known to be a key enzyme in folate mediated metabolism, playing a pivotal role in heart development [83].
15	*ZNF365*	It is known to interact with NDE1 depletion of which causes a smaller Kupffer's vesicle [84].
19	*GABRG1*	Gamma-aminobutyric acid (GABA) A receptor, gamma 1 (GABRG1) has proven effects in cardiac vagal neurons, which play a crucial role in controlling the heart rate and cardiac function in previous works [85, 86].
24	*TH*	It has been inferred in a previous study that tyrosine hydroxylase (TH) plays a key role in cardiac differentiation and primitive heart tube formation [87].
25	*TEAD1*	It is a well known regulator of proper heart development [88].
25	*PROKR2*	Prokineticin receptor-2 is reported to be associated with cardiovascular system development [89].
26	*TPH2*	It might play important roles in heart development as it generates serotonin ([5-hydroxytryptamine (5-HT)]) that is involved in regulating differentiation, proliferation during development and cardiac function [90, 91].
27	*GPER*	The role of G-protein-coupled estrogen receptor (GPER) as cardioprotective has been inferred in previous studies [92, 93].
27	*RGN*	It has been suggested in a previous study that RGN, known as regucalcin, plays an essential role in regulating heart mitochondrial function by increasing Ca2+ ATPase activity [94].
28	*WNT8A*	It is inferred to play an essential role in heart tube development [95].
28	*PPARA*	Peroxisome proliferator-activated receptor-alpha plays a critical role in heart through activating fatty acid oxidation [96].
28	*EPHB2*	Ephrin-B2 is reported to play an important role in cardiac valve maturation [97].
42	*LBX1*	It acts as a marker of embryonic stem cell induced cardiomyocyte differentiation [98].
42	*PKD1L1*	It is known to play an essential role in left/ right determination and in cardiac development [99, 100].
43	*NEDD4*	The Hect domain E3 ubiquitin ligase, *NEDD4* is known to play an instrumental role in cardiac development [101].
61	*CRB1*	It is reported to be specifically expressed in central nervous system [102].
64	*PEMT*	Inhibition of phosphatidylethanolamine N-methyltransferase (PEMT) is inferred to play an important role in preventing cardiac dysfunctions [103].

Table 3.16: Genes that are known to be associated with cardiovascular development. The contents of the table have been taken from [7].

Tricluster	Genes	Functions
76	*GIPC2*	It is known to be an interaction partner of SEMA5A inactivation of which leads to imperfect remodeling of the cranial vascular system [104].
76	*GFAP*	Its role in cardiac valve has been established in a previous study [105].
79	*FRS2*	FRS2 phosphorylation is required for activation of MAPK pathway in cardiomyocyte development [106].
86	*NFATC3*	It plays an essential role in cardiac development [107].
86	*CTSS*	It has been suggested by a recent report that cathepsin S (CTSS) takes part in normal accumulation of autophagysomes [108]. Deficiency of autophagy has been inferred to cause heart disease [109].
86	*OLFML2B*	It is known to bind chondroitin sulfate which has been inferred to be involved in cardiac AV canal formation [110].
87	*SHOX2*	The role of SHOX2 in posterior heart field formation has been reported in a previous work [111].
97	*CTNNB1*	It is inferred to play an important role in heart development [112].
98	*AIDA*	A previous study reported that AIDA is found to be highly expressed in heart tissue [113].

4 Speculating about the Role of ZEB2 During Stem Cell Differentiation into Cardiomyocytes

4.1 Introduction

Stem cells are capable of differentiating into a specialized cell types during embryogenesis such as cardiomyocytes, neural progenitors etc. During differentiation into such specific cell types, the cellular proliferation usually occurs at a higher level at early stage where the rate becomes relatively low with the progression of the development [1, 2]. Thus the proper balance between proliferation and regeneration is crucial in the context of stem cell differentiation into a certain cell type. Moreover, depending on the ability of differentiating into three germ layers i.e. mesoderm, ectoderm and endoderm, a stem cell can be categorized into several types such as, pluripotent, multipotent, oligopotent and unipotent. A pluripotent stem cell can differentiate into any of these germ layers, whereas a unipotent stem cell has the ability to differentiate into one of these three germ layers only. Stem cell differentiation into cardiomyocytes is achieved by several steps such as epithelial to mesenchymal transition, mesoderm differentiation, mesoderm specification, cardiac specification, cardiomyocytes differentiation and electrical maturation [3]. Differentiation into ectoderm and mesoderm are crucial for nervous system and muscle (smooth, skeletal, cardiac) development. A previous study inferred the fact that the crosstalk between neuronal and cardiac cells are instrumental for the proper development of both the systems [4]. Furthermore, the role of autonomic nervous system was established to be important for proper heart development by a previous study [5], whereas the deformity in skeletal and smooth muscle cell development may trigger the central nervous system disorders [6, 7]. Thus, differentiation into ectoderm and mesoderm are instrumental during cardiomyocytes development and controlled by several transcription factors (TFs). Although unveiling the regulatory mechanisms involved in the cardiomyocytes differentiation is the subject of substantial number of research over the last decade, how TFs governs the phenotypic changes of the cell during such differentiation remains poorly understood. Furthermore, the roles of key genes identified in the previous chapter show an association with the proper cardiovascular and nervous system development and being motivated by this finding, in this chapter, we aim at elucidating the transcriptional regulation during the differentiation of a human induced pluripotent stem cell (hiPSC) into cardiomyocyte by identifying transcriptional regulatory modules (TRMs) which can be defined as a set of genes regulated by a common set of transcription factors [8]. Though in a complex gene regulatory network, one TF regulates hundreds or thousands of genes, in this chapter we focus on unveiling the transcriptional regulatory network (TRN) during different stages of cardiomyocyte differentiation and in such TRN each node and edge represents one transcription factor and the

regulatory interaction between any two TFs or one TF regulating itself, respectively [9]. Moreover, we have decomposed the inferred regulatory networks into a bow-tie structure as in biological sciences, bow-tie representation helps us to understand how a large set of TFs (IN part) regulates a large set of target genes (OUT part) through a relatively small set of TFs (core part) [10] in order to govern the phenotypic changes of an organism.

4.2 Materials and Methods

4.2.1 Dataset

In order to reprogram hiPSC generated from a human fibroblast, retroviral expression of *SOX2*, *OCT4*, *NANOG* and *LIN28* was used. Afterwards, genetic programming approaches were employed to a single clone from the reprogrammed cell to increase the production of cardiomyocyte population through drug selection such as Blasticidin S deaminase (BSD) [11].

4.2.2 Methods

Figure 4.1 shows the workflow applied in this chapter in order to accomplish the goal mentioned in the previous section. First, we have applied EMOA-δ-TRIMAX triclustering algorithm to retrieve the groups of co-expressed genes which exhibits similar expression profiles over subsets of replicates and time points. Afterwards, we have performed the transcription factor binding site (TFBS) enrichment analysis in order to identify the potential transcriptional regulatory modules (TRMs) during different stages of cardiomyocytes differentiation. As one particular transcription factor subfamily may have more than one transcription factors [12], the genes belonging to the corresponding transcriptional regulatory module are not necessarily to be co-regulated by each of these paralogous TFs [13]. Moreover, the regulators are not necessarily to be co-expressed with their target genes because the TFs may also be post-transcriptionally regulated [13]. Hence, to identify the potential TFs among such paralogous ones, we have computed time lag non-linear distance correlation coefficient between the expression profiles of the target genes belong to each of the TRMs over the clustered time points and replicates and the corresponding TFs expression values either at the clustered time points or at the preceding ones [14, 15]. In the next step, we have used hyper-geometric test [16–18] in order to perform Gene Ontology Biological Process (GOBP) enrichment analysis for reveling the roles of identified TFs in governing cardiomyocytes differentiation.

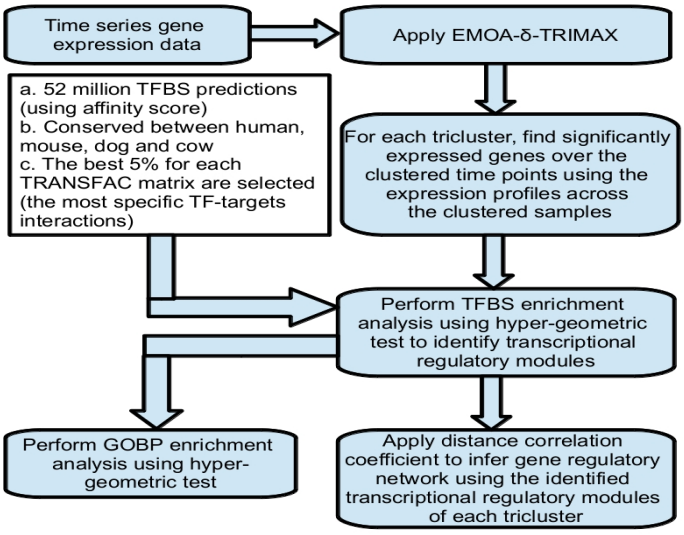

Figure 4.1: Workflow applied in this chapter.

4.3 Results and Discussion

Through our analysis we have identified five triclusters time points of which apparently show different stages of cardiomyocyte differentiation and we named them Module 1, Module 2, Module 3, Module 4 and Module 5. Figure 4.2 shows the time points and replicates of the aforementioned identified modules, whereas from Figure 4.3 we can observe the divergence of the median expression profiles of the genes belonging to each of these identified modules. After reconstructing the transcriptional regulatory networks (TRNs) for each of these identified modules, we have crumbled the TRNs into three parts of bow-tie structure i.e. IN, strongly connected component (SCC) and OUT. Afterwards, we have analyzed the roles of TFs found to be the members of SCCs as they are supposed to drive the phenotypic changes during cardiac development.

Module	Day 00	Day 03	Day 07	Day 10	Day 14	Day 20	Day 28	Day 35	Day 45	Day 60	Day 90	Day 120	Sample 1	Sample 2	Sample 3
Module 1	*	*	*	*	*	*	*	*	*				*	*	
Module 2		*	*										*	*	*
Module 3				*	*	*	*						*	*	*
Module 4								*	*	*	*		*	*	*
Module 5									*	*	*		*	*	*

Figure 4.2: Clustered replicates and time points of the identified modules, namely Module 1, Module 2, Module 3, Module 4 and Module 5.

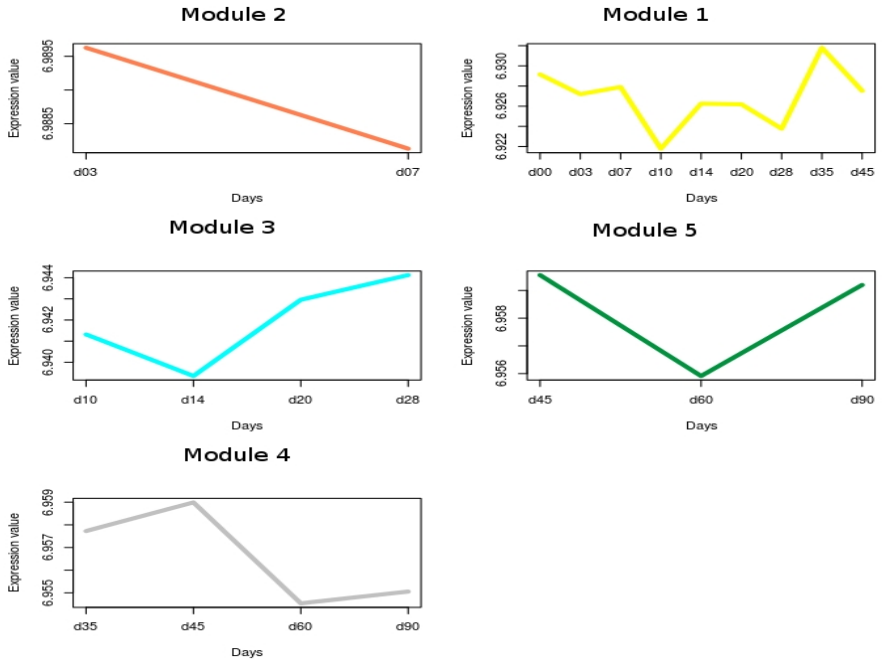

Figure 4.3: Median expression profiles of the genes belonging to the modules, namely Module 1, Module 2, Module 3, Module 4 and Module 5.

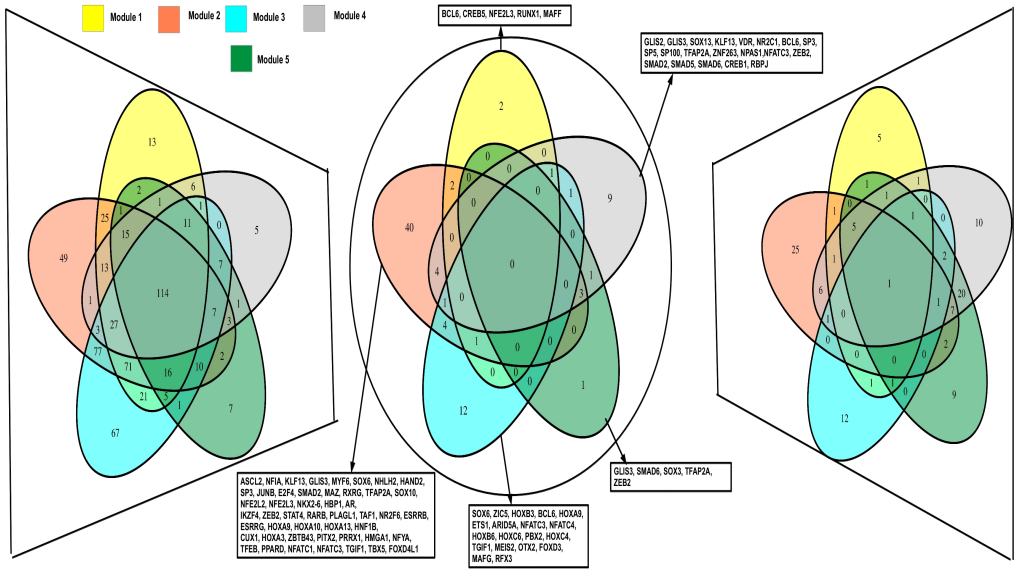

Figure 4.4: The bow-tie structure of transcriptional networks inferred from the Module 1, Module 2, Module 3, Module 4 and Module 5 and the TFs belong to the SCC of each of these modules.

4.3.1 SCCs of Module 1

RUNX1 is found to be an instrumental player in averting skeletal muscle disorders [19]. Depletion of *BCL6* expression may trigger inflammatory diseases in heart [20]. The role *MAFF* has been reported to be crucial during embryonic development of many tissues [21]. *NFE2L3* also known as *NRF3* plays crucial role in smooth muscle cell differentiation during early embryo development [22]. Moreover, our results indicate the role of *CREB5* in regulating systemic arterial blood pressure which was already inferred to be associated with cardiac output by a previous study [23].

4.3.2 SCCs of Module 2

ASCL2 is known to play an instrumental role in driving nervous system differentiation [24]. A previous study inferred the role of *NFIA* in regulating neuronal differentiation [25]. Moreover, the involvement of *KLF13*, *SOX6*, *E2F4*, *MAZ*, *NFE2L1* have been established in governing cardiomyocyte development by previous studies [26–29]. *GLIS3* may also play

an instrumental role in muscle cell differentiation [30]. Furthermore, *MYF6* and *AR* are known to be involved in skeletal muscle cell differentiation [31,32]. The roles of *SOX10* and *NR2F6* were established to be associated with neuronal development [33, 34]. Deficiency of the expression of *NFE2L2*, *RARB*, RXRG, *TFEB* may trigger cardiac disorder [35–37]. Though the role of *NHLH2* has not been established to prevent cardiac disorders, depletion of one of the similar transcription factor *NHLH1* expression may cause cardiac diseases [38]. Additionally, the roles of *HAND2*, *STAT4* and *TBX5* have been established in proper heart development [39–41]. Deletion of the expressions of *SP3*, *JUNB*, *TFAP2A*, *PLAGL1*, *PPARD* and *HMGA1* transcription factors may trigger the occurrence of cardiac disorders [42–47]. Furthermore, the involvement of *IKZF4*, *TAF1* and *HNF1B* TFs have been elucidated in governing neuronal development [48–50]. Moreover, a previous study reports that the interaction between *PITX2* and *PRRX1* may be instrumental in proper heart development [51]. The role of *ESRRB* and *ESRRG* have also been inferred in maintaining the pluripotent state of stem cell [52, 53], whereas *NFYA* play an important role in embryonic stem cell [54]. Moreover, the abundant expression of *NFAT* transcription factors are found in myocardium [55] whereas *TGIF1* is known to play an instrumental role in governing heart looping [56]. *NKX2-6* may play an instrumental role in proper heart development and its expression was observed in each of the three germ layers [57–59].

4.3.3 SCCs of Module 3

Besides the above findings, our results also reveal the roles of *PBX2* and *MEIS2* in governing the heart development [60, 61]. Deficiency of the expression of *RFX3* may trigger cardiac disorders [62]. Moreover, *ETS1* and *FOXD3* may play instrumental roles in neural crest development which supports proper cardiac development [63, 64]. The abundant expression of *ARID5A* and *MAFG* were also observed by previous studies [21, 65]. Furthermore, deletion of the expression of *ZIC5* may cause neural crest disorders [66].

4.3.4 SCCs of Module 4

Additionally, our results infer the roles of *SP5* and *NPAS1* in the neuronal development [67, 68]. *VDR* is known to be associated with sustaining the proper functions of skeletal muscle [69]. The role of *RBPJ* was also inferred to be instrumental in neuronal progenitor

cells [70]. Moreover, *SOX13* may play an important role in neurogenesis and its role was observed in artery wall [71, 72].

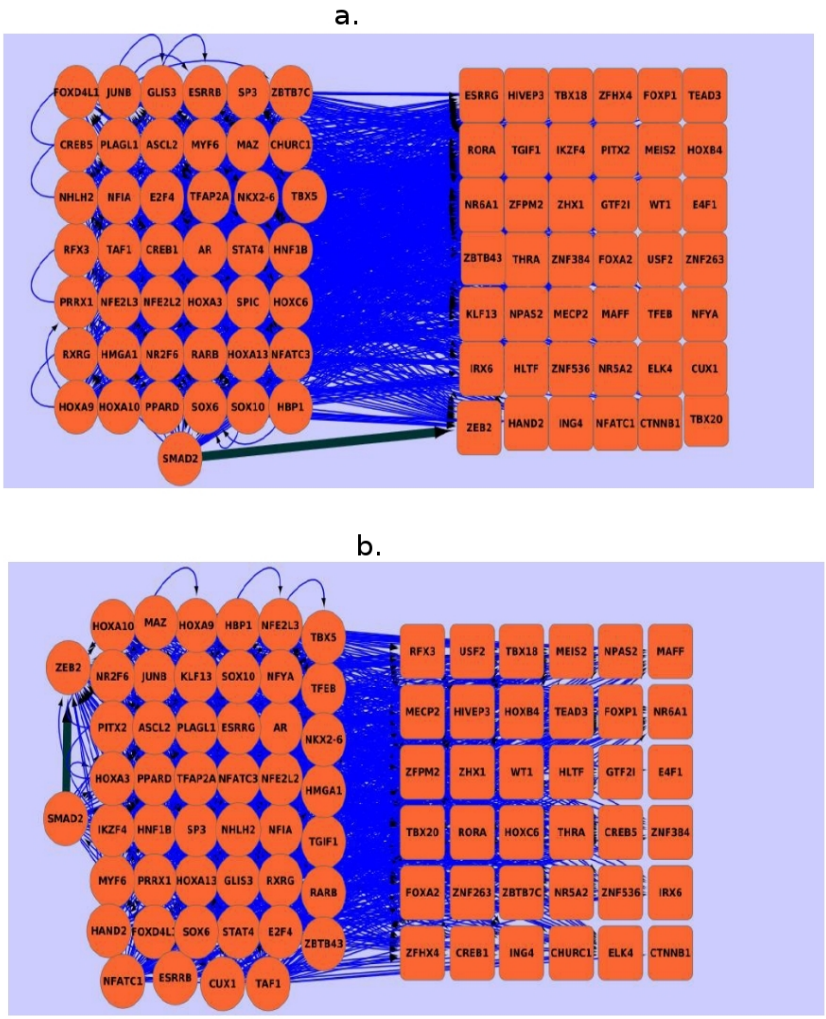

Figure 4.5: SCC and OUT parts of the transcriptional regulatory networks for Module 2 on day 03 (a) and day 07 (b). Circular nodes represent the members of the core part, whereas round-rectangular nodes denote their targets. Green colored edge shows the interaction between *ZEB2* and SMADs.

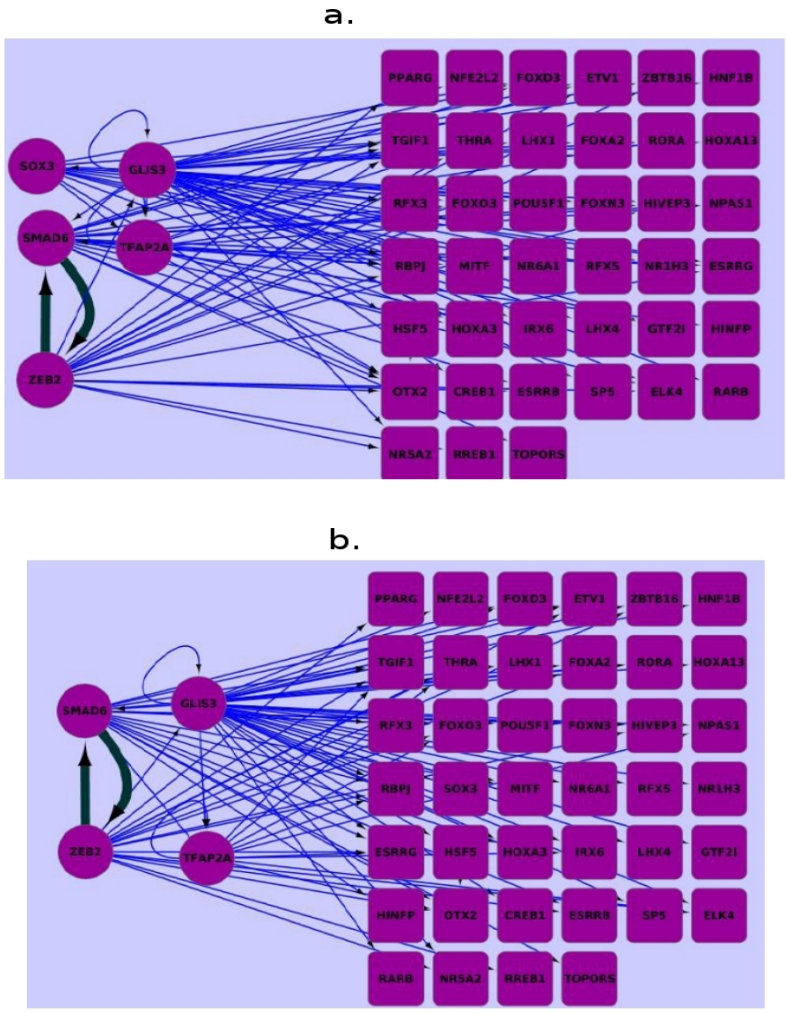

Figure 4.6: SCC and OUT parts of the transcriptional regulatory networks for Module 4 on day 45 (a) and day 60 (b). Circular nodes represent the members of the core part, whereas round-rectangular nodes denote their targets. Green colored edge shows the interaction between *ZEB2* and SMADs.

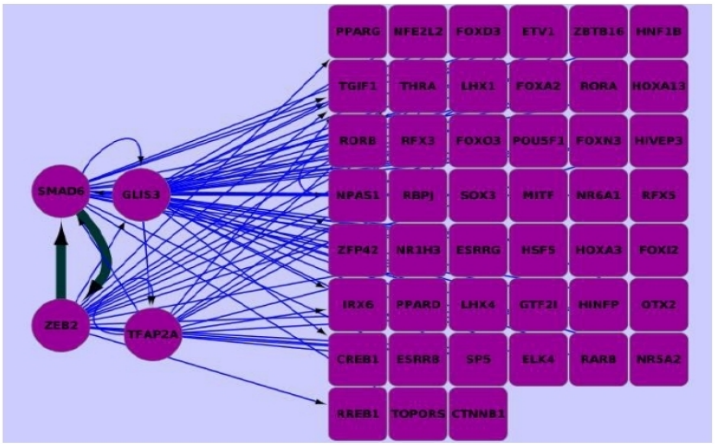

Figure 4.7: SCC and OUT parts of the transcriptional regulatory networks for Module 4 on day 90. Circular nodes represent the members of the core part, whereas round-rectangular nodes denote their targets. Green colored edge shows the interaction between *ZEB2* and SMADs.

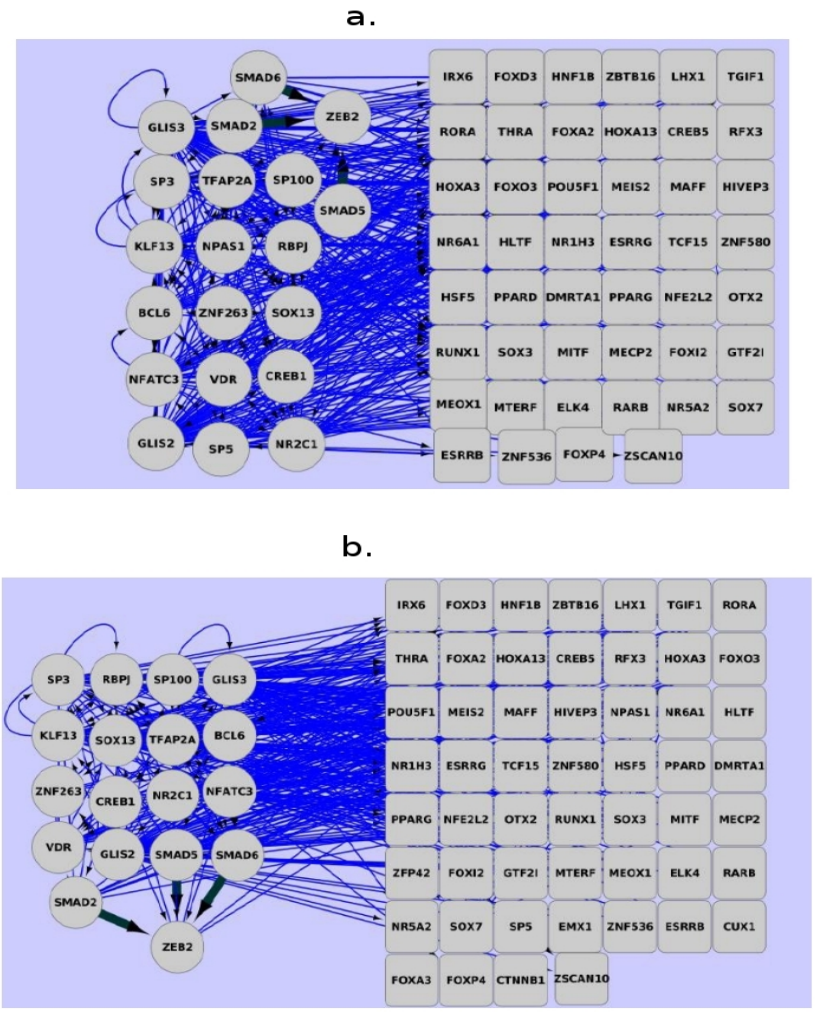

Figure 4.8: SCC and OUT parts of the transcriptional regulatory networks for Module 5 on day 35 (a) and day 45 (b). Circular nodes represent the members of the core part, whereas round-rectangular nodes denote their targets. Green colored edge shows the interaction between *ZEB2* and SMADs.

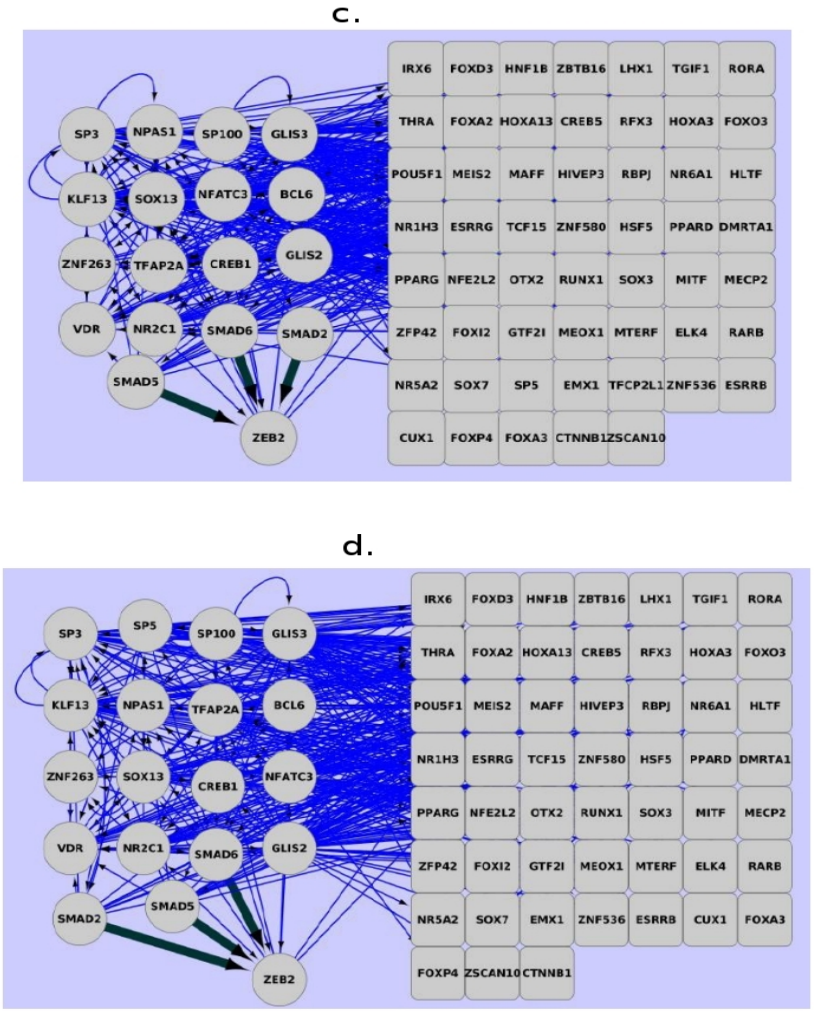

Figure 4.9: SCC and OUT parts of the transcriptional regulatory networks for Module 5 on day 60 (a) and day 90 (b). Circular nodes represent the members of the core part, whereas round-rectangular nodes denote their targets. Green colored edge shows the interaction between *ZEB2* and SMADs.

4.3.5 Elucidating the Roles of ZEB2 and SMADs Transcription Factors During Cardiac Development

Figure 4.4 suggests an important role of *ZEB2*, also known as SMAD interacting protein 1(*SIP1*), in governing the changes of phenotypic behavior of the cell during the hiPSC

differentiation into cardiomyocytes as *ZEB2* is found to be a part of SCC of the transcriptional regulatory networks inferred from module 1, 4 and 5. To further investigate the regulatory interactions between ZEB2 and other TFs belonging to the core part of these transcriptional regulatory networks, we have observed the regulatory interactions between *ZEB2* and different members of SMAD family during the early and late stages of differentiation (Figures 4.5-4.9).

Figure 4.5 shows the interaction between *ZEB2* and *SMAD2* transcription factors during very early stage of differentiation and this finding is not too surprising as the role of *SMAD2* signaling in maintaining the pluripotent stem cell state has already been inferred by a previous study [73]. In order to further elucidate the involvement of *ZEB2* at the early stage of differentiation, we have observed the potential biological processes associated with the regulation of both myocytes and neuronal development (Figure 4.10). Although, mutations in the *ZEB2* gene result in the occurrence of MowatWilsonSyndrome, a genetic disorder that often cooccurs with neurodegenerative malfunctions [74, 75], its role in developing proper heart still remains unclear.

Additionally, Figures 4.6-4.7 evince the significance of the regulatory interactions between SMAD proteins *SMAD2*, *SMAD5*, *SMAD6* and *ZEB2* at the late stage of differentiation, where as Figures 4.8-4.9 show the involvement of a two-node feed-back loop comprising *SMAD6* and *ZEB2* during days 45, 60 and 90. Figures 4.11-4.12 suggest that in case of these two modules, *ZEB2* is also involved in governing both neural and myocytes development. This finding is not too surprising as *SMAD6* and *SMAD5* are known to be important players in regulating proper neural and cardiac development [76–78].

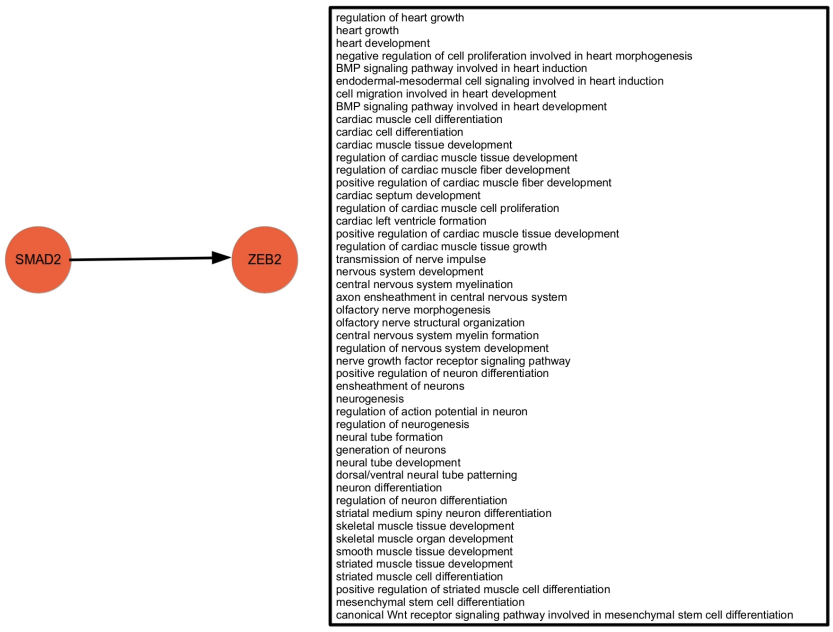

Figure 4.10: Interactions between ZEB2 and SMADs TFs and their roles during the early stage of stem cell differentiation into cardiomyocytes.

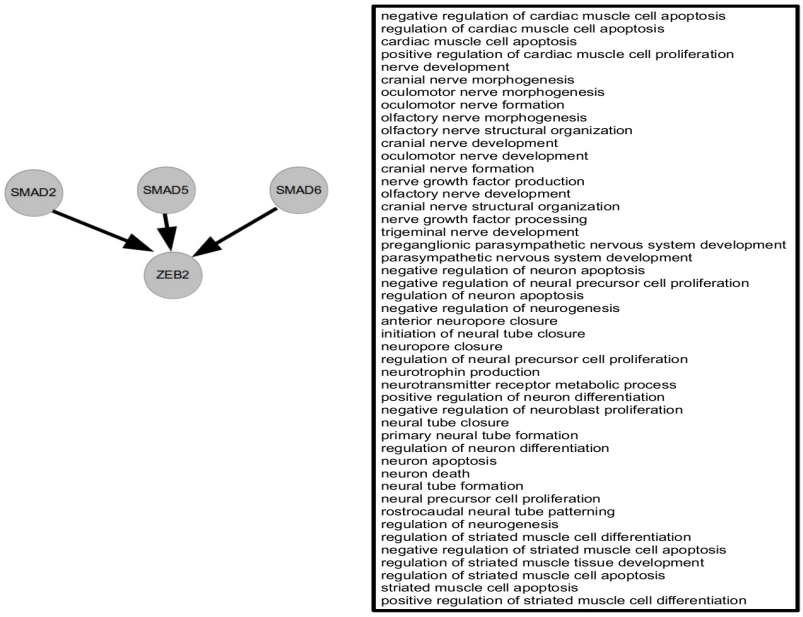

Figure 4.11: Interactions between ZEB2 and SMADs TFs and their roles during the late stage of stem cell differentiation into cardiomyocytes.

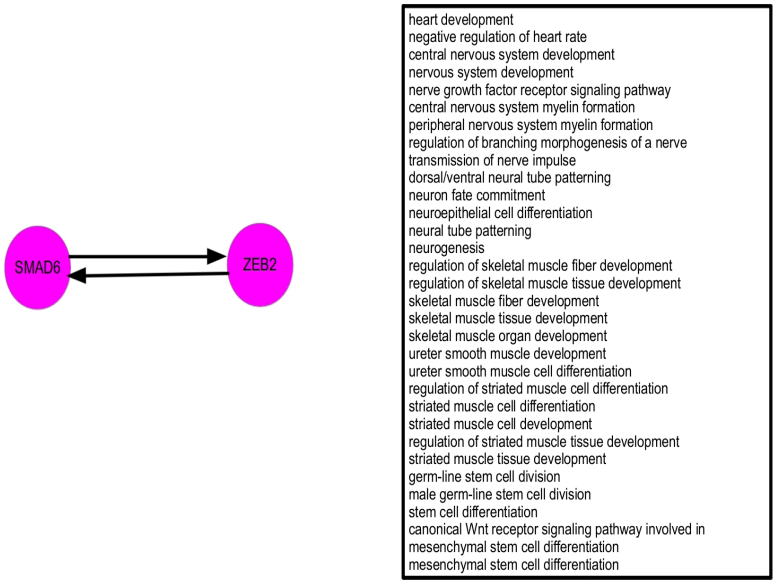

Figure 4.12: Interactions between ZEB2 and SMADs TFs and their roles during the maturation stage of stem cell differentiation into cardiomyocytes.

4.4 Conclusion

In this chapter, we applied the proposed triclustering algorithm EMOA-δ-TRIMAX to a time series gene expression data which contains the expression profiles of genes during stem cell differentiation into cardiomyocytes in order to retrieve the regulatory modules during different stages of cardiomyocytes development. Afterwards, the identified modules have been used to infer the regulatory networks which have then been decomposed into three parts i.e. IN, SCC and OUT. Our results infer the roles of those TFs that were found to be members of the SCCs of the identified modules. Most of these TFs were already known to be associated with the proper development of cardiomyocytes. Moreover, our findings also indicate a role of *ZEB2* in the accomplishment of stem cell differentiation into cardiomyocytes. Altogether, our results provide new insights into the regulatory mechanisms involved in driving the phenotypic changes during adolescence of cardiomyocytes.

4.5 Bibliography

[1] Cooper, G.M.: The Cell Cycle. The Cell, 2nd edn. Sinauer Associates, Sunderland (2000)

[2] Snchez Alvarado, A., Yamanaka, S.: Rethinking differentiation: stem cells, regeneration, and plasticity. Cell **157**(1), 110–119 (2014). doi:10.1016/j.cell.2014.02.041

[3] Burridge, P.W., Keller, G., Gold, J.D., Wu, J.C.: Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. Cell Stem Cell **10**(1), 16–28 (2012). doi:10.1016/j.stem.2011.12.013

[4] Murashov, A.K., Pak, E.S., Katwa, L.C.: Parallel development of cardiomyocytes and neurons in embryonic stem cell culture. Biochemical and Biophysical Research Communications **332**(3), 653–656 (2005)

[5] Kimura, K., Ieda, M., Fukuda, K.: Development, maturation, and transdifferentiation of cardiac sympathetic nerves. Circulation Research **110**(2), 325–336 (2012). doi:10.1161/CIRCRESAHA.111.257253

[6] Kajbafzadeh, A.M., Payabvash, S., Salmasi, A.H., Monajemzadeh, M., Tavangar, S.M.: Smooth muscle cell apoptosis and defective neural development in congenital ureteropelvic junction obstruction. The Journal of Urology **176**(2), 718–723 (2006)

[7] Kablar, B., Rudnicki, M.A.: Information provided by the skeletal muscle and associated neurons is necessary for proper brain development. International Journal of Developmental Neuroscience **20**(7), 573–584 (2002)

[8] Yan, X., Mehan, M.R., Huang, Y., Waterman, M.S., Yu, P.S., Zhou, X.J.: A graph-based approach to systematically reconstruct human transcriptional regulatory modules. Bioinformatics **23**(13), 577–586 (2007)

[9] Neph, S., Stergachis, A.B., Reynolds, A., Sandstrom, R., Borenstein, E., Stamatoyannopoulos, J.A.: Circuitry and dynamics of human transcription factor regulatory networks. Cell **150**(6), 1276–1286 (2012). doi:10.1016/j.cell.2012.04.040

[10] Csete, M., Doyle, J.: Bow ties, metabolism and disease. Trends in Biotechnology **22**(9), 446–450 (2004)

[11] Babiarz, J.E., Ravon, M., Sridhar, S., Ravindran, P., Swanson, B., Bitter, H., Weiser, T., Chiao, E., Certa, U., Kolaja, K.L.: Determination of the human cardiomyocyte mRNA and miRNA differentiation network by fine-scale profiling. Stem Cells and Development **21**(11), 1956–1965 (2012). doi:10.1089/scd.2011.0357

[12] Wingender, E., Schoeps, T., Doenitz, J.: TFClass: an expandable hierarchical classification of human transcription factors. Nucleic Acids Research **41**(Database issue), 165–170 (2013). doi:10.1093/nar/gks1123

[13] Ma, S., Shah, S., Bohnert, H.J., Snyder, M., Dinesh-Kumar, S.P.: Incorporating motif analysis into gene co-expression networks reveals novel modular expression pattern and new signaling pathways. PLoS Genetics **9**(10), 1003840 (2013). doi:10.1371/journal.pgen.1003840

[14] Szkely, G.J., Rizzo, M.L., Bakirov, N.K.: Measuring and testing dependence by correlation of distances. The Annals of Statistics **35**(6), 2769–2794 (2007). doi:10.1214/009053607000000505

[15] de Siqueira Santos, S., Takahashi, D.Y., Nakata, A., Fujita, A.: A comparative study of statistical methods used to identify dependencies between gene expression signals. Briefings in Bioinformatics (2013). doi:10.1093/bib/bbt051

[16] Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., Sherlock, G.: GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics **20**(18), 3710–3715 (2004)

[17] Benjamini, Y., Yekutieli, D.: The Control of the False Discovery Rate in Multiple Testing under Dependency. The Annals of Statistics **29**(4), 1165–1188 (2001)

[18] Falcon, S., Gentleman, R.: Using GOstats to test gene lists for GO term association. Bioinformatics **23**(2), 257–258 (2007)

[19] Mondal, A., Sawant, D., Dent, A.L.: Transcriptional repressor BCL6 controls Th17 responses by controlling gene expression in both T cells and macrophages. Journal of Immunology **184**(8), 4123–4132 (2010). doi:10.4049/jimmunol.0901242

[20] Wang, X., Blagden, C., Fan, J., Nowak, S.J., Taniuchi, I., Littman, D.R., Burden, S.J.: Runx1 prevents wasting, myofibrillar disorganization, and autophagy of skeletal muscle. Genes and Development **19**(14), 1715–1722 (2005)

[21] Onodera, K., Shavit, J.A., Motohashi, H., Katsuoka, F., Akasaka, J.E., Engel, J.D., Yamamoto, M.: Characterization of the murine mafF gene. The Journal of Biological Chemistry **274**(30), 21162–21169 (1999)

[22] Pepe, A.E., Xiao, Q., Zampetaki, A., Zhang, Z., Kobayashi, A., Hu, Y., Xu, Q.: Crucial role of nrf3 in smooth muscle cell differentiation from stem cells. Circulation Research **106**(5), 870–879 (2010). doi:10.1161/CIRCRESAHA.109.211417

[23] Guyton, A.C.: The relationship of cardiac output and arterial pressure control. Circulation **64**(6), 1079–1088 (1981)

[24] Kwon, O.H., Park, J.L., Baek, S.J., Noh, S.M., Song, K.S., Kim, S.Y., Kim, Y.S.: Aberrant upregulation of ASCL2 by promoter demethylation promotes the growth and resistance to 5-fluorouracil of gastric cancer cells. Cancer Science **104**(3), 391–397 (2013). doi:10.1111/cas.12076

[25] Piper, M., Barry, G., Hawkins, J., Mason, S., Lindwall, C., Little, E., Sarkar, A., Smith, A.G., Moldrich, R.X., Boyle, G.M., Tole, S., Gronostajski, R.M., Bailey, T.L., Richards, L.J.: NFIA controls telencephalic progenitor cell differentiation through repression of the Notch effector Hes1. The Journal of Neuroscience **30**(27), 9127–9139 (2010). doi:10.1523/JNEUROSCI.6167-09.2010

[26] Lavalle, G., Andelfinger, G., Nadeau, M., Lefebvre, C., Nemer, G., Horb, M., Nemer, M.: The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. The EMBO Journal **25**(21), 5201–5213 (2006)

[27] Cohen-Barak, O., Yi, Z., Hagiwara, N., Monzen, K., Komuro, I., Brilliant, M.H.: Sox6 regulation of cardiac myocyte development. Nucleic Acids Research **31**(20), 5941–5948 (2003)

[28] van Amerongen, M.J., Diehl, F., Novoyatleva, T., Patra, C., Engel, F.B.: E2F4 is required for cardiomyocyte proliferation. Cardiovascular Research **86**(1), 92–102 (2010). doi:10.1093/cvr/cvp383

[29] Xia, Y., Buja, L.M., McMillin, J.B.: Activation of the cytochrome c gene by electrical stimulation in neonatal rat cardiac myocytes. Role of NRF-1 and c-Jun. Journal of Biological Chemistry **273**(20), 12593–12598 (1998)

[30] Kim, Y.S., Nakanishi, G., Lewandoski, M., Jetten, A.M.: GLIS3, a novel member of the GLIS subfamily of Krppel-like zinc finger proteins with repressor and activation functions. Nucleic Acids Research **31**(19), 5513–5525 (2003)

[31] Hinits, Y., Osborn, D.P., Carvajal, J.J., Rigby, P.W., Hughes, S.M.: Mrf4 (myf6) is dynamically expressed in differentiated zebrafish skeletal muscle. Gene Expression Patterns **7**(7), 738–754 (2007)

[32] Altuwaijri, S., Lee, D.K., Chuang, K.H., Ting, H.J., Yang, Z., Xu, Q., Tsai, M.Y., Yeh, S., Hanchett, L.A., Chang, H.C., Chang, C.: Androgen receptor regulates expression of skeletal muscle-specific proteins and muscle cell types. Endocrine **25**(1), 27–32 (2004)

[33] Mongera, A., Singh, A.P., Levesque, M.P., Chen, Y.Y., Konstantinidis, P., Nuesslein-Volhard, C.: Genetic lineage labeling in zebrafish uncovers novel neural crest contributions to the head, including gill pillar cells. Development **140**(4), 916–925 (2013). doi:10.1242/dev.091066

[34] Warnecke, M., Oster, H., Revelli, J.P., Alvarez-Bolado, G., Eichele, G.: Abnormal development of the locus coeruleus in Ear2(Nr2f6)-deficient mice impairs the functionality of the forebrain clock and affects nociception. Genes and Development **19**(5), 614–625 (2005)

[35] Zhou, S., Sun, W., Zhang, Z., Zheng, Y.: The Role of Nrf2-Mediated Pathway in Cardiac Remodeling and Heart Failure. Oxidative Medicine and Cellular Longevity **2014**, 260429 (2014). doi:10.1155/2014/260429

[36] Guleria, R.S., Choudhary, R., Tanaka, T., Baker, K.M., Pan, J.: Retinoic acid receptor-mediated signaling protects cardiomyocytes from hyperglycemia induced apoptosis: role of the renin-angiotensin system. Journal of Cellular Physiology **226**(5), 1292–1307 (2011). doi:10.1002/jcp.22457

[37] Ma, X., Godar, R.J., Liu, H., Diwan, A.: Enhancing lysosome biogenesis attenuates BNIP3-induced cardiomyocyte death. Autophagy **8**(3), 297–309 (2012). doi:10.4161/auto.18658

[38] Good, D.J., Coyle, C.A., Fox, D.L.: Nhlh2: a basic helix-loop-helix transcription factor controlling physical activity. Exercise and Sport Sciences Review **36**(4), 187–192 (2008). doi:10.1097/JES.0b013e31818782dd

[39] Schindler, Y.L., Garske, K.M., Wang, J., Firulli, B.A., Firulli, A.B., Poss, K.D., Yelon, D.: Hand2 elevates cardiomyocyte production during zebrafish heart development and regeneration. Development **141**(16), 3112–3122 (2014). doi:10.1242/dev.106336

[40] Afanasyeva, M., Wang, Y., Kaya, Z., Stafford, E.A., Dohmen, K.M., Sadighi, A.A. Akha, Rose, N.R.: Interleukin-12 receptor/STAT4 signaling is required for the development of autoimmune myocarditis in mice by an interferon-gamma-independent pathway. Circulation **104**(25), 3145–3151

[41] Hatcher, C.J., Kim, M.S., Pennisi, D., Song, Y., Diman, N., Goldstein, M., Mikawa, T., Basson, C.T.: TBX5 Regulates Cardiac Cell Behavior During Cardiogenesis, pp. 27–30. Blackwell Publishing Ltd, Malden, Massachusetts (2007). doi:10.1002/9780470988664.ch7. http://dx.doi.org/10.1002/9780470988664.ch7

[42] van Loo, P.F., Mahtab, E.A., Wisse, L.J., Hou, J., Grosveld, F., Suske, G., Philipsen, S., Gittenberger-de Groot, A.C.: Transcription factor Sp3 knockout mice display serious cardiac malformations. Molecular and Cellular Biology **27**(24), 8571–8582 (2007)

[43] Meder, B., Just, S., Vogel, B., Rudloff, J., Gaertner, L., Dahme, T., Huttner, I., Zankl, A., Katus, H.A., Rottbauer, W.: JunB-CBFbeta signaling is essential to maintain sarcomeric Z-disc structure and when defective leads to heart failure. Journal od Cell Science **123**(Pt 15), 2613–2620 (2010). doi:10.1242/jcs.067967

[44] Hammer, S., Toenjes, M., Lange, M., Fischer, J.J., Dunkel, I., Mebus, S., Grimm, C.H., Hetzer, R., Berger, F., Sperling, S.: Characterization of TBX20 in human hearts and its regulation by TFAP2. Journal of Cellular Biochemistry **104**(3), 1022–1033 (2008). doi:10.1002/jcb.21686

[45] Xuan, C., Wang, B.B., Gao, G., Bai, X.Y., Yang, Q., Liu, X.C., Jing, W.B., Ma, X., He, G.W.: A novel variation of PLAGL1 in Chinese patients with isolated ventricular septal defect. Genetic Testing and Molecular Biomarkers **16**(8), 984–987 (2012). doi:10.1089/gtmb.2012.0003

[46] Finck, B.N.: The PPAR regulatory system in cardiac physiology and disease. Cardiovascular Research **73**(2), 269–277 (2007)

[47] Fedele, M., Fidanza, V., Battista, S., Pentimalli, F., Klein-Szanto, A.J., Visone, R., De Martino, I., Curcio, A., Morisco, C., Del Vecchio, L., Baldassarre, G., Arra, C., Viglietto, G., Indolfi, C., Croce, C.M., Fusco, A.: Haploinsufficiency of the Hmga1 gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. Cancer Research **66**(5), 2536–2543 (2006)

[48] Honma, Y., Kiyosawa, H., Mori, T., Oguri, A., Nikaido, T., Kanazawa, K., Tojo, M., Takeda, J., Tanno, Y., Yokoya, S., Kawabata, I., Ikeda, H., Wanaka, A.: Eos: a novel member of the Ikaros gene family expressed predominantly in the developing nervous system. FEBS Letters **447**(1), 76–80 (1999)

[49] Rashid, D., Newell, K., Shama, L., Bradley, R.: A requirement for NF-protocadherin and TAF1/Set in cell adhesion and neural tube formation. Developmental Biology **291**(1), 170–181 (2006)

[50] Makki, N., Capecchi, M.R.: Cardiovascular defects in a mouse model of HOXA1 syndrome. Human Molecular Genetics **21**(1), 26–31 (2012). doi:10.1093/hmg/ddr434

[51] Dolmatova, E., Cooper, R.R., Ellinor, P.T.: Prrx1 Interacts with and Inhibits Pitx2 Function. Circulation **126**, 19452 (2012)

[52] Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., Chambers, I.: Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell **11**(4), 477–490 (2012). doi:10.1016/j.stem.2012.08.002

[53] Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H., Han, J., Vega, V.B., Cacheux-Rataboul, V., Lim, B., Lufkin, T., Ng, H.H.: Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nature Cell Biology **11**(2), 197–203 (2009). doi:10.1038/ncb1827

[54] Dolfini, D., Minuzzo, M., Pavesi, G., Mantovani, R.: The short isoform of NF-YA belongs to the embryonic stem cell transcription factor circuitry. Stem Cells **30**(11), 2450–2459 (2012). doi:10.1002/stem.1232

[55] Molkentin, J.D.: Calcineurin-NFAT signaling regulates the cardiac hypertrophic response in coordination with the MAPKs. Cardiovasc Research **63**(3), 467–475 (2004)

[56] Powers, S.E., Taniguchi, K., Yen, W., Melhuish, T.A., Shen, J., Walsh, C.A., Sutherland, A.E., Wotton, D.: Tgif1 and Tgif2 regulate Nodal signaling and are required for gastrulation. Development **137**(2), 249–259 (2010). doi:10.1242/dev.040782

[57] Nikolova, M., Chen, X., Lufkin, T.: Nkx2.6 expression is transiently and specifically restricted to the branchial region of pharyngeal-stage mouse embryos. Mechanisms of Development **69**(1-2), 215–218 (1997)

[58] Heathcote, K., Braybrook, C., Abushaban, L., Guy, M., Khetyar, M.E., Patton, M.A., Carter, N.D., Scambler, P.J., Syrris, P.: Common arterial trunk associated with a homeodomain mutation of NKX2.6. Human Molecular Genetics **14**(5), 585–593 (2005)

[59] Wang, J., Mao, J.H., Ding, K.K., Xu, W.J., Liu, X.Y., Qiu, X.B., Li, R.G., Qu, X.K., Xu, Y.J., Huang, R.T., Xue, S., Yang, Y.Q.: A Novel NKX2.6 Mutation Associated with Congenital Ventricular Septal Defect. Pediatric Cardiology, 1–11 (2014). doi:10.1007/s00246-014-1060-x

[60] Stankunas, K., Shang, C., Twu, K.Y., Kao, S.C., Jenkins, N.A., Copeland, N.G., Sanyal, M., Selleri, L., Cleary, M.L., Chang, C.P.: Pbx/Meis deficiencies demonstrate multigenetic origins of congenital heart disease. Circulation Research **103**(7), 702–709 (2008). doi:10.1161/CIRCRESAHA.108.175489

[61] Paige, S.L., Thomas, S., Stoick-Cooper, C.L., Wang, H., Maves, L., Sandstrom, R., Pabon, L., Reinecke, H., Pratt, G., Keller, G., Moon, R.T., Stamatoyannopoulos, J., Murry, C.E.: A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. Cell **151**(1), 221–232 (2012). doi:10.1016/j.cell.2012.08.027

[62] Bonnafe, E., Touka, M., AitLounis, A., Baas, D., Barras, E., Ucla, C., Moreau, A., Flamant, F., Dubruille, R., Couble, P., Collignon, J., Durand, B., Reith, W.: The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. Molecular and Cellular Biology **24**(10), 4417–4427 (2004)

[63] Gao, Z., Kim, G., Mackinnon, A.C., Flagg, A.E., Bassett, B., Earley, J.U., Svensson, E.C.: Ets1 is required for proper migration and differentiation of the cardiac neural crest. Development **137**(9), 1543–1551 (2010). doi:10.1242/dev.047696

[64] Mundell, N.A., Labosky, P.A.: Neural crest stem cell multipotency requires Foxd3 to maintain neural potential and repress mesenchymal fates. Development **138**(4), 641–652 (2011). doi:10.1242/dev.054718

[65] Amano, K., Hata, K., Muramatsu, S., Wakabayashi, M., Takigawa, Y., Ono, K., Nakanishi, M., Takashima, R., Kogo, M., Matsuda, A., Nishimura, R., Yoneda, T.: Arid5a cooperates with Sox9 to stimulate chondrocyte-specific transcription. Molecular Biology of the Cell **22**(8), 1300–1311 (2011). doi:10.1091/mbc.E10-07-0566

[66] Inoue, T., Hatayama, M., Tohmonda, T., Itohara, S., Aruga, J., Mikoshiba, K.: Mouse Zic5 deficiency results in neural tube defects and hypoplasia of cephalic neural crest derivatives. Developmental Biology **270**(1), 146–162 (2004)

[67] Park, D.S., Seo, J.H., Hong, M., Bang, W., Han, J.K., Choi, S.C.: Role of Sp5 as an essential early regulator of neural crest specification in xenopus. Developmental Dynamics **242**(12), 1382–1394 (2013). doi:10.1002/dvdy.24034

[68] Ohsawa, S., Hamada, S., Kakinuma, Y., Yagi, T., Miura, M.: Novel function of neuronal PAS domain protein 1 in erythropoietin expression in neuronal cells. Journal od Neuroscience Research **79**(4), 451–458 (2005)

[69] Girgis, C.M., Mokbel, N., Cha, K.M., Houweling, P.J., Abboud, M., Fraser, D.R., Mason, R.S., Clifton-Bligh, R.J., Gunton, J.E.: The vitamin D receptor (VDR) is expressed in skeletal muscle of male mice and modulates 25-hydroxyvitamin D (25OHD) uptake in myofibers. Endocrinology **155**(9), 3227–3237 (2014). doi:10.1210/en.2014-1016

[70] Hori, K., Cholewa-Waclaw, J., Nakada, Y., Glasgow, S., Masui, T., Henke, R.M., Wildner, H., Martarelli, B., Beres, T.M., Epstein, J.A., Magnuson, M.A., Macdonald, R.J., Birchmeier, C., Johnson, J.E.: A nonclassical bHLH Rbpj transcription factor complex is required for specification of GABAergic neurons independent of Notch signaling. Genes and Development **22**(2), 166–178 (2008). doi:10.1101/gad.1628008

[71] Wang, Y., Bagheri-Fam, S., Harley, V.R.: SOX13 is up-regulated in the developing mouse neuroepithelium and identifies a sub-population of differentiating neurons. Brain Research. Developmental Brain Research **157**(2), 201–208 (2005)

[72] Roose, J., Korver, W., Oving, E., Wilson, A., Wagenaar, G., Markman, M., Lamers, W., Clevers, H.: High expression of the HMG box factor sox-13 in arterial walls during embryonic development. Nucleic Acids Research **26**(2), 469–476 (1998)

[73] Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N., Derynck, R.: Smad2 is essential for maintenance of the human and mouse primed pluripotent stem cell state. The Journal of Biological Chemistry **288**(25), 18546–18560 (2013). doi:10.1074/jbc.M112.446591

[74] Ghoumid, J., Drevillon, L., Alavi-Naini, S., Bondurand, N., Rio, M., Briand-Suleau, A., Nasser, M., Goodwin, L., Raymond, P., Yanicostas, C., Goossens, M., Lyonnet, S., Mowat, D., Amiel, J., Soussi-Yanicostas, N., Giurgea, I.: ZEB2 zinc-finger missense mutations lead to hypomorphic alleles and a mild Mowat-Wilson syndrome. Human Molecular Genetics **22**(13), 2652–2661 (2013). doi:10.1093/hmg/ddt114

[75] Mowat, D.R., Croaker, G.D., Cass, D.T., Kerr, B.A., Chaitow, J., Ades, L.C., Chia, N.L., Wilson, M.J.: Hirschsprung disease, microcephaly, mental retardation, and characteristic facial features: delineation of a new syndrome and identification of a locus at chromosome 2q22-q23. Journal of Medical Genetics **35**(8), 617–623 (1998)

[76] Xie, Z., Chen, Y., Li, Z., Bai, G., Zhu, Y., Yan, R., Tan, F., Chen, Y.G., Guillemot, F., Li, L., Jing, N.: Smad6 promotes neuronal differentiation in the intermediate zone of the dorsal neural tube by inhibition of the Wnt/beta-catenin pathway. Proceedings of the National Academy of Sciences of the United States of America **108**(29), 12119–12124 (2011). doi:10.1073/pnas.1100160108

[77] Galvin, K.M., Donovan, M.J., Lynch, C.A., Meyer, R.I., Paul, R.J., Lorenz, J.N., Fairchild-Huntress, V., Dixon, K.L., Dunmore, J.H., Gimbrone, M.A.J., Falb, D., Huszar, D.: A role for smad6 in development and homeostasis of the cardiovascular system. Nature Genetics **24**(2), 171–174 (2000)

[78] Hegarty, S.V., O'Keeffe, G.W., Sullivan, A.M.: BMP-Smad 1/5/8 signalling in the development of the nervous system. Progress in Neurobiology **109**, 28–41 (2013). doi:10.1016/j.pneurobio.2013.07.002

5 Co-regulation Analysis of Time Series Transcriptomics Data Unveils the Roles of Three Node Feed-Forward Loops in Regulating Genes of Signaling Pathways of Breast Cancer Progression

5.1 Introduction

Extracellular signaling molecules trigger the activation of specific receptors, which are usually confined to a certain cell type. As a response, these receptors engender cellular changes by altering gene expressions, metabolic events etc. In the context of breast cancer, after invading into the cell the estrogen molecule binds to the estrogen receptor which thereafter binds to specific DNA sites known as estrogen responsive elements (EREs). Afterwards, EREs regulate their target genes by recruiting co-activators or co-repressors to prompt assembly of the transcriptoin complex and synthesize messenger RNAs (mRNAs) from which specific proteins are synthesized. These proteins then cause changes of the cellular behavior for instance, enhancing the proliferation etc. Targeting these signaling pathways may provide insightful knowledge about therapeutic strategies as these pathways trigger the progression of the breast cancer cells [1]. Though over the last decade elucidating the roles of transcription factors in (i) responding to these signaling pathways and (ii) specifically enhancing the synthesis of components of the same or other pathways is the topic of considerable research in the context of breast cancer, the transcriptional regulatory circuitry still remains poorly understood.

Due to the advancement of microarray technology, over the last decade, it has become easier to monitor the expression profiles of thousands of genes not only over a set of replicates but also at a set of time points. The changes of the gene expression profiles are controlled by transcription factors (TFs) which escalate or diminish specific target gene expression by interacting with either co-activators or co-repressors, respectively [2]. Moreover, due to the paralogous expansion of the transcription factor subfamily [3], each of the paralogous transcription factors may not necessarily regulate their target genes. Gene regulatory networks (GRNs) can be viewed as a collection of interactions between transcription factors and their target genes. In graph theory, GRNs can be delineated as a mixed graph $G = (N, DE, UE)$, where N represents a set of vertices, each of which can be either a gene or a TF (or the gene of a TF). DE denotes the set of directed edges, each of which represents regulatory relationships between a gene and a TF. UE refers to the set of undirected edges which can be expounded as the associations between genes deduced from the degree of similarities in their expression profiles [4]. A gene regulatory network can further be decomposed as the components of a bow-tie structure i.e. IN, OUT, strongly connected component (SCC), tendrils and tubes (Figure 5.1). Nodes belonging to the IN part can act as regulators of the nodes belonging to the SCC and OUT components. Nodes in the OUT part which are reachable from the IN part without having any connection to

the SCC part form tube. Nodes are said to be the members of SCC if there is a path from each of these nodes to every other nodes. Tendrils comprise nodes that are reachable from IN part, but do not have any link to SCC or the OUT component, and nodes that can reach OUT component, but are neither linked to the SCC or to the IN component [4]. In the context of developmental biology, TFs in the giant component play crucial roles in driving the phenotypic changes in an adolescent cell, whereas in the context of disease biology, these strongly connected transcription factors play instrumental role in disease progression. Moreover, due to the complexity of the gene regulatory networks, recurrent network motifs such as, feed-forward or feed-back loops can be used to easily understand the transcriptional regulatory circuitry and their involvement in governing the disease progression. Feed-forward loops (FFLs), one of the coregulatory motifs, are reported by several studies to provide apparent insights into the governance of cellular functions and in the progression of cancer [5–7]. A three node FFL comprises two transcription factors one of which regulates the other and a gene which is regulated by these two transcription factors. A previous study reported the involvement of a cyclic feed-forward loop comprising *SAF-1*, *RAS* and *VEGF* in governing angiogenesis in breast cancer [8].

In this chapter, we have applied EMOA-δ-TRIMAX algorithm [11] to a time series gene expression dataset to provide insights into the regulatory mechanisms of signaling during MCF-7 breast cancer cell exposure to estrogen.

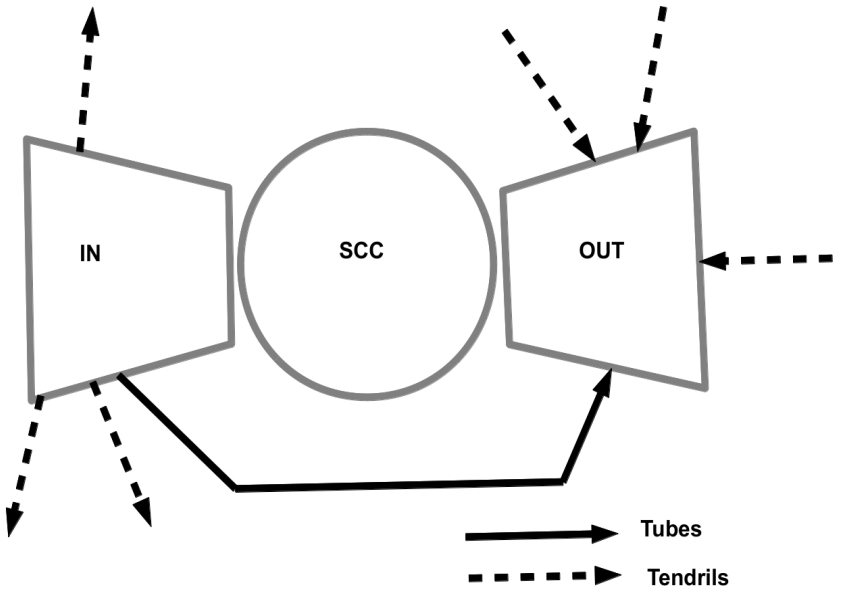

Figure 5.1: Schematic diagram of bow-tie structure.

5.2 Materials and Method

5.2.1 Dataset

The description of this dataset was provided in chapter 2. The experiment was carried out to unravel the transcriptional regulatory mechanisms during the exposure of MCF-7 breast cancer cell to estrogen [12].

5.2.2 Methods

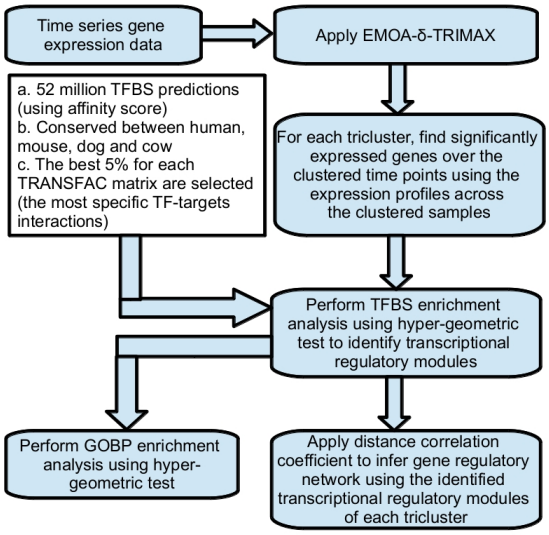

Figure 5.2: Workflow applied in this work.

Figure 5.2 demonstrates the workflow applied in this work. First we have applied the EMOA-δ-TRIMAX [11] algorithm to the aforementioned dataset to extract the groups of genes exhibiting similar expression profiles over a subset of replicates and time points. In the next step, we have used the Limma package [13] to find out the significantly expressed genes from each group of the co-expressed genes using clustered time points and replicates, as we are interested in elucidating the regulatory mechanisms of estrogen induced genes. Afterwards, performing the transcription factor binding site (TFBS) enrichment analysis on the groups of co- and significantly expressed genes using TRANSFAC library version 12.2 [14], we have detected the corresponding potential regulators based on the statistical enrichment of TRANSFAC matrices. The target genes of each of the identified potential regulators are not necessarily to be co-expressed with TF itself as a transcription factor can also be post-transcriptionally regulated. Being motivated by this fact, we have computed the distance correlation coefficient [15] between the median expression profiles of the target genes and that of the corresponding transcription factors to identify the most specific regulators. Distance correlation is used to compute the non-linear dependence between two random variables (V_1 and V_2) and its value ranges from 0 to 1, where 0 represents the

distance correlation between two independent variables.

5.3 Results and Discussion

Through our analysis, we have found two modules comprising the time points that clearly depict the early and late stages of exposure of MCF-7 cells to estrogen. After reconstructing the gene regulatory network using the aforementioned regulatory modules, we fragmented the network into IN, OUT and SCC parts. TFs that belong to the SCC are identified using the igraph package [16]. Figure 5.3 shows the overlap between the members belonging to each of the components of the bow-tie structures of the corresponding gene regulatory networks. We can observe that inspite of having a higher number (Jaccard similarity coefficient is 0.55) of common transcription factors belonging to the IN parts, the overlap in case of both the SCC and OUT parts (Jaccard similarity coefficients are 0.15 and 0.19,respectively) are comparatively low. The transcription factors that switch from each part of the bow-tie structure to every other part are shown in Figure 5.4.
To further investigate the commonalities between the results obtained in this chapter and chapter 2, we have observed that in case of the early module 64% of all transcription factors have also been found to be in chapter 2 using the δ-TRIMAX algorithm, whereas this percentage is 65% in case of the late module. And this finding indicate the robustness of the two methods used in this work.

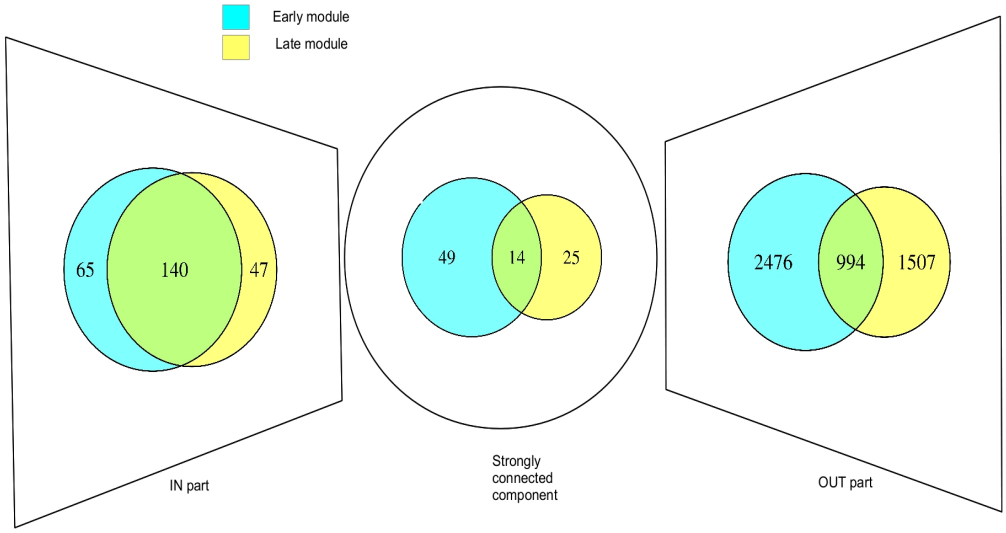

Figure 5.3: Overlap between the members belonging to each part of the bow-tie structure of the regulatory networks inferred from early and late modules.

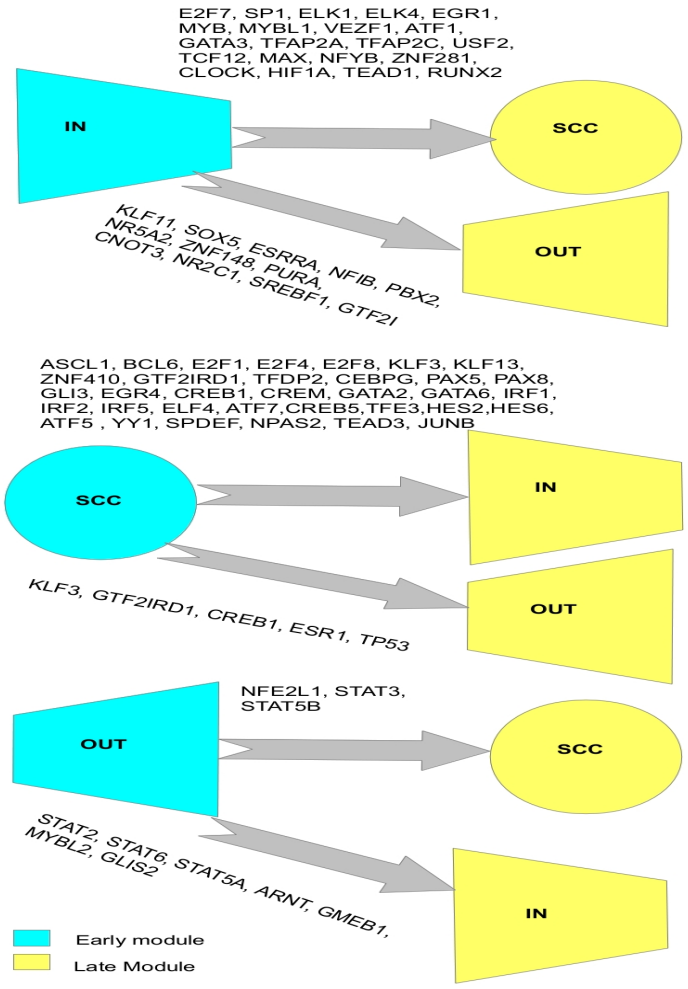

Figure 5.4: Inter-modular switch of transcription factors from each part of the bow-tie structure to every other part.

Moreover, the identified strongly connected components as shown in Figure 5.5, will provide insights into several feed-forward or feed-back loops.

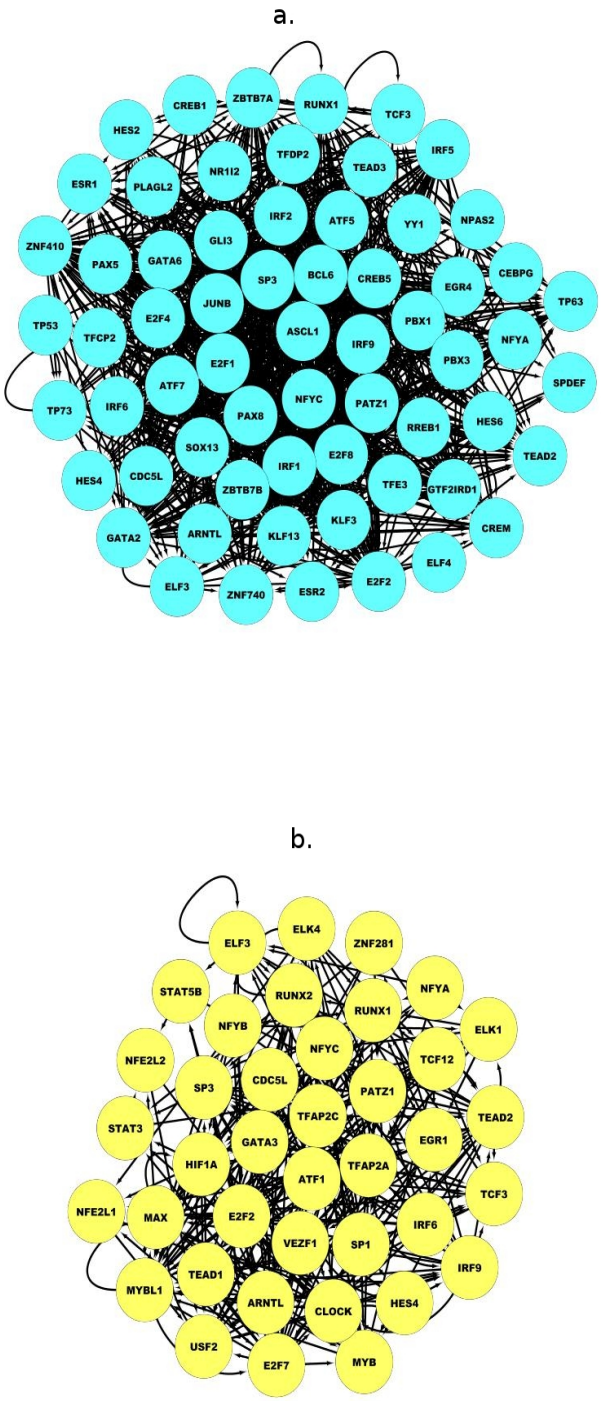

Figure 5.5: The strongly connected components of early (a.) and late (b.) modules.

In this chapter, we aim at elucidating the roles of 3-node FFL motifs identified through performing an exhaustive search, in order to unveil how TFs cooperatively accelerate estrogen induced breast cancer cell growth and proliferation. From Figure 5.5a, we can observe that *ESR1* also known as estrogen receptor alpha belongs to the strongly connected component of the regulatory network at early time points. Analyzing its targets in the SCC, we have found the regulatory interactions between *ESR1* and *E2F2*, *BCL6* and analyzed the roles of FFLs where *E2F2* and *BCL6* act as co-regulators with *ESR1* in governing the expression of genes involved in signaling pathways which may enhance breast cancer risks. Table 5.1 shows the lists of signaling pathways which are mediated by the TFs *ESR1*, *E2F2* and *ESR1*, *BCL6*, respectively. A comprehensive list of signaling pathway genes regulated by two TFs in the identified FFLs are enlisted in Table 5.3 (see appendix). It is of interest to see that the FFL comprising *ESR1*, *E2F2* and *HDAC1* is involved in diminishing the rate of androgen receptor signaling which is known to block the ER-alpha signaling [17]. Table 5.2 shows those signaling pathways which are governed by the transcription factors of the FFLs comprising *E2F2* during the late stage of breast cancer cell exposure to estrogen. A comprehensive list of signaling pathway genes regulated by two TFs in the identified FFLs during the late stage of cellular exposure to estrogen are enlisted in Table 5.4 (see appendix). At this stage, we can see the reduction of the rates of pathways such as bone morphogenetic protein (BMP) signaling, non-canonical Wnt (β-catenin independent pathway) signaling, prolactin signaling, interleukin-6 (IL-6)-mediated signaling, macrophage colony-stimulating factor signaling, tumor necrosis factor-mediated signaling, apoptotic signaling, protein kinase C signaling, transforming growth factor beta (TGF-β) receptor signaling which may in turn enhance the risks of breast cancer. For instance, inhibition of BMP signaling may cause mitotic arrest defect which is known to be involved in chromosomal instability observed in human breast cancer [18, 19]. Non-canonical Wnt signaling mediates the maintenance of cellular polarity the loss of which may initiate tumorigenesis [20, 21]. Moreover, a previous study reported that the blockage of prolactin signaling may also be observed in human breast cancer [22]. IL-6 signaling has been inferred to be an inhibitor of the cellular proliferation in estrogen receptor positive breast cancer cell [23]. The inhibition of colony-stimulating factor signaling may also trigger breast cancer metastasis [24]. Protein kinase C and transforming growth factor beta signaling may play instrumental roles in inhibiting breast cancer cell proliferation; thus diminishing the rate of these two signaling pathways may positively influence the progression of breast cancer [25, 26]. Tumor necrosis factor (TNF)-mediated signaling was inferred to

act as an antagonist of ER-alpha signaling in MCF7 breast cancer cell [27].
Besides, targeting both extrinsic and intrinsic apoptotic signaling pathway has been reported to be therapeutic strategies in the context of breast cancer treatment. TNF-related apoptosis inducing ligand (TRAIL) triggers the activation of extrinsic apoptotic signaling which exhibits anti-tumor activities in breast cancer cell [28]. It is of interest to see the regulation of "extrinsic apoptotic signaling pathway in the absence of ligand" (Table 5.4) also treated as an example of intrinsic apoptotic pathway during the late stage of exposure of MCF-7 cell to estrogen. Additionally, in case of breast cancer cell, mutations in the genes involved in the regulation of mitochondrial pathways may cause oncogenesis. For instance, it has been reported by a previous study that the chromosomal translocation of Bcl-2 gene may cause the over-expression of Bcl-2 family members exhibit the anti-apoptotic behavior in breast cancer cells [28]. From our results (Table 5.4), the involvement of *YWHAQ* also known as *14-3-3*, *BCL2L11*, *BCL2L1*, *SEPT4* also known as *ARTS* and *BCLAF1* becomes apparent in the regulation of intrinsic apoptotic signaling during the late stage of cellular exposure to estrogen. *14-3-3* is known to be an instrumental mediator of anti-apoptotic signal [29], whereas a previous study reported that the over expression of *XIAP* may inhibit the activities of caspase-3 which is known to be instrumental in mediating apoptosis in breast cancer cell [28, 30]. Furthermore, another previous study inferred that mutation in *ARTS* may result in promoting the tumor growth [31]. Though BCL2-associated transcription factor 1 (*BCLAF1*) and *BCL2L11* are well known apoptosis inducers, the co-expression of the former one with other members of BCL-2 family may inhibit its apoptotic behavior in breast cancer cell, whereas the interaction of the later with *GRB10* may suppress apoptosis [32].

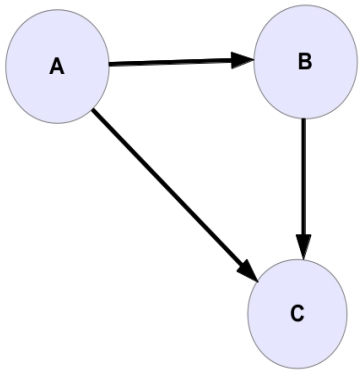

Figure 5.6: Feed-forward loop comprising three nodes.

Table 5.1: Signaling pathways regulated by three node feed-forward loop in the early module. "x" means yes.

GO Term	ESR1 (A)/ E2F2 (B)	ESR1 (A)/ BCL6 (B)
androgen receptor signaling pathway	x	
regulation of opioid receptor signaling pathway	x	x
fibroblast growth factor receptor signaling pathway	x	
intracellular steroid hormone receptor signaling pathway	x	x
neurotrophin TRK receptor signaling pathway	x	x
cell surface receptor signaling pathway	x	
negative regulation of androgen receptor signaling pathway	x	

Table 5.2: Signaling pathways regulated by three node feed-forward loop in the late module. "x" means yes.

GO Term	E2F2 (A)/ ARNTL (B)	E2F2 (A)/ ATF1 (B)	E2F2 (A)/ CLOCK (B)	E2F2 (A)/ EGR1 (B)	E2F2 (A)/ GATA3 (B)	E2F2 (A)/ HIF1A (B)	E2F2 (A)/ MAX (B)	E2F2 (A)/MYB or MYBL1 (B)	E2F2 (A)/NFYB or NFYC (B)	E2F2 (A)/ PATZ1 (B)	E2F2 (A)/RUNX1 or RUNX2 (B)	E2F2 (A)/ SP1 or SP3 (B)	E2F2 (A)/STAT5B or STAT6 (B)	E2F2 (A)/ TEAD1 or TEAD2 (B)	E2F2 (A)/TFAP2A or TFAP2C (B)	E2F2 (A)/ USF2 (B)
regulation of peroxisome proliferator activated receptor signaling pathway	x		x			x		x							x	
transmembrane receptor protein serine/threonine kinase signaling pathway	x		x			x						x				
neurotrophin TRK receptor signaling pathway	x		x	x		x		x	x		x	x		x	x	
transforming growth factor beta receptor signaling pathway	x	x	x	x		x	x	x		x	x	x	x	x	x	x
Notch signaling involved in heart development	x		x			x								x		
Wnt signaling pathway involved in forebrain neuroblast division	x		x			x	x	x				x	x		x	x
canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	x		x			x	x	x				x	x		x	x
positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	x		x			x	x	x				x	x		x	x
canonical Wnt signaling pathway involved in neural crest cell differentiation	x		x			x	x	x				x	x		x	x
canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	x		x			x	x	x				x	x		x	x
canonical Wnt signaling pathway involved in regulation of cell proliferation	x		x			x	x					x	x			x
canonical Wnt signaling pathway	x		x			x			x					x	x	
cell surface receptor signaling pathway	x	x	x	x		x				x		x			x	
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x			x						x			x	
Wnt signaling pathway involved in dorsal/ventral axis specification	x		x			x							x			
negative regulation of non-canonical Wnt signaling pathway	x		x			x							x			
Wnt signaling pathway involved in somitogenesis	x		x	x		x						x	x			
intracellular receptor signaling pathway	x	x	x	x	x	x		x				x		x	x	
G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	x		x			x										
Notch signaling pathway	x		x			x	x		x					x	x	x
negative regulation of BMP signaling pathway	x		x			x										

Table 5.2 continued...

GO Term	E2F2 (A)/ ARNTL (B)	E2F2 (A)/ ATF1 (B)	E2F2 (A)/ CLOCK (B)	E2F2 (A)/ EGR1 (B)	E2F2 (A)/ GATA3 (B)	E2F2 (A)/ HIF1A (B)	E2F2 (A)/ MAX (B)	E2F2 (A)/MYB or MYBL1 (B)	E2F2 (A)/NFYB or NFYC (B)	E2F2 (A)/ PATZ1 (B)	E2F2 (A)/RUNX1 or RUNX2 (B)	E2F2 (A)/ SP1 or SP3 (B)	E2F2 (A)/STAT5B or STAT6 (B)	E2F2 (A)/ TEAD1 or TEAD2 (B)	E2F2 (A)/TFAP2A or TFAP2C (B)	E2F2 (A)/ USF2 (B)
toll-like receptor 10 signaling pathway		x		x					x			x		x	x	
toll-like receptor 5 signaling pathway		x		x					x			x		x	x	
toll-like receptor TLR1:TLR2 signaling pathway		x		x					x							
toll-like receptor TLR6:TLR2 signaling pathway		x		x					x							
toll-like receptor 9 signaling pathway		x		x					x							
TRIF-dependent toll-like receptor signaling pathway		x		x					x			x				
toll-like receptor 2 signaling pathway		x		x					x							
MyD88-independent toll-like receptor signaling pathway		x		x					x			x				
toll-like receptor 3 signaling pathway		x		x								x				
MyD88-dependent toll-like receptor signaling pathway		x		x												
toll-like receptor 4 signaling pathway		x		x								x				
toll-like receptor signaling pathway		x		x												
insulin receptor signaling pathway		x		x							x	x		x	x	
regulation of Wnt signaling pathway		x						x				x		x	x	
Wnt signaling pathway		x				x	x	x	x		x	x		x	x	x
BMP signaling pathway		x				x						x		x		
intracellular steroid hormone receptor signaling pathway		x		x												
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway		x				x		x			x	x			x	
positive regulation of intracellular estrogen receptor signaling pathway		x		x												
fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development				x	x							x		x	x	
insulin receptor signaling pathway via phosphatidylinositol 3-kinase				x								x		x	x	

Table 5.2 continued...

GO Term	E2F2 (A)/ ARNTL (B)	E2F2 (A)/ ATF1 (B)	E2F2 (A)/ CLOCK (B)	E2F2 (A)/ EGR1 (B)	E2F2 (A)/ GATA3 (B)	E2F2 (A)/ HIF1A (B)	E2F2 (A)/ MAX (B)	E2F2 (A)/MYB or MYBL1 (B)	E2F2 (A)/NFYB or NFYC (B)	E2F2 (A)/ PATZ1 (B)	E2F2 (A)/RUNX1 or RUNX2 (B)	E2F2 (A)/ SP1 or SP3 (B)	E2F2 (A)/STAT5B or STAT6 (B)	E2F2 (A)/ TEAD1 or TEAD2 (B)	E2F2 (A)/TFAP2A or TFAP2C (B)	E2F2 (A)/ USF2 (B)
negative regulation of Wnt signaling pathway				x	x	x		x				x				
fibroblast growth factor receptor signaling pathway				x					x	x	x	x		x	x	
regulation of endocannabinoid signaling pathway				x			x	x				x		x	x	x
extrinsic apoptotic signaling pathway				x				x				x			x	
positive regulation of intrinsic apoptotic signaling pathway				x		x			x				x		x	
androgen receptor signaling pathway				x				x						x	x	
glucocorticoid receptor signaling pathway				x				x				x				
epidermal growth factor receptor signaling pathway				x						x		x		x	x	
negative regulation of interleukin-6-mediated signaling pathway				x						x		x	x			
Fc-epsilon receptor signaling pathway				x						x		x		x	x	
regulation of transforming growth factor beta receptor signaling pathway				x										x	x	
intracellular estrogen receptor signaling pathway				x	x	x		x				x		x	x	
negative regulation of cGMP-mediated signaling				x	x							x		x		
activation of prostate induction by androgen receptor signaling pathway				x								x			x	
negative regulation of interleukin-2-mediated signaling pathway				x						x		x				
negative regulation of prolactin signaling pathway				x						x		x				
negative regulation of interleukin-4-mediated signaling pathway				x						x		x				
negative regulation of macrophage colony-stimulating factor signaling pathway				x						x		x				
negative regulation of tumor necrosis factor-mediated signaling pathway				x						x		x				
extrinsic apoptotic signaling pathway in absence of ligand				x		x			x			x		x	x	
activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway				x								x			x	

Table 5.2 continued...

GO Term	E2F2 (A)/ ARNTL (B)	E2F2 (A)/ ATF1 (B)	E2F2 (A)/ CLOCK (B)	E2F2 (A)/ EGR1 (B)	E2F2 (A)/ GATA3 (B)	E2F2 (A)/ HIF1A (B)	E2F2 (A)/ MAX (B)	E2F2 (A)/MYB or MYBL1 (B)	E2F2 (A)/NFYB or NFYC (B)	E2F2 (A)/ PATZ1 (B)	E2F2 (A)/RUNX1 or RUNX2 (B)	E2F2 (A)/ SP1 or SP3 (B)	E2F2 (A)/STAT5B or STAT6 (B)	E2F2 (A)/ TEAD1 or TEAD2 (B)	E2F2 (A)/TFAP2A or TFAP2C (B)	E2F2 (A)/ USF2 (B)
adenylate cyclase-inhibiting G-protein coupled glutamate receptor signaling pathway					x			x							x	
autocrine signaling					x	x										
canonical Wnt signaling pathway involved in mesenchymal stem cell differentiation					x	x										
canonical Wnt signaling pathway involved in osteoblast differentiation					x	x										
regulation of fibroblast growth factor receptor signaling pathway					x							x		x	x	
regulation of phosphatidylinositol 3-kinase signaling					x	x						x			x	
positive regulation of ephrin receptor signaling pathway						x										
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway						x		x							x	
Wnt signaling pathway, calcium modulating pathway						x										
SREBP signaling pathway						x										
cytokine-mediated signaling pathway						x						x				
apolipoprotein A-I-mediated signaling pathway							x	x								x
apoptotic signaling pathway								x	x			x			x	
negative regulation of apoptotic signaling pathway								x				x				
Fc-gamma receptor signaling pathway involved in phagocytosis								x	x							
intrinsic apoptotic signaling pathway								x	x			x			x	
nucleotide-binding oligomerization domain containing signaling pathway								x								
hippo signaling								x			x	x			x	
nerve growth factor signaling pathway									x							
positive regulation of apoptotic signaling pathway									x							
positive regulation of vascular endothelial growth factor receptor signaling pathway									x			x			x	x

Table 5.2 continued...

GO Term	E2F2 (A)/ ARNTL (B)	E2F2 (A)/ ATF1 (B)	E2F2 (A)/ CLOCK (B)	E2F2 (A)/ EGR1 (B)	E2F2 (A)/ GATA3 (B)	E2F2 (A)/ HIF1A (B)	E2F2 (A)/ MAX (B)	E2F2 (A)/MYB or MYBL1 (B)	E2F2 (A)/NFYB or NFYC (B)	E2F2 (A)/ PATZ1 (B)	E2F2 (A)/RUNX1 or RUNX2 (B)	E2F2 (A)/ SP1 or SP3 (B)	E2F2 (A)/STAT5B or STAT6 (B)	E2F2 (A)/ TEAD1 or TEAD2 (B)	E2F2 (A)/TFAP2A or TFAP2C (B)	E2F2 (A)/ USF2 (B)
positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway									x							
steroid hormone mediated signaling pathway												x				
phospholipase C-activating G-protein coupled glutamate receptor signaling pathway												x				
interleukin-6-mediated signaling pathway												x	x			
negative regulation of cytokine-mediated signaling pathway												x				
VEGF-activated neuropilin signaling pathway												x			x	
interleukin-27-mediated signaling pathway												x	x			
positive regulation of protein kinase C signaling												x				
negative regulation of protein kinase C signaling														x		
interleukin-15-mediated signaling pathway														x		
positive regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway															x	
positive regulation of I-kappaB kinase/NF-kappaB signaling															x	
canonical Wnt signaling pathway involved in metanephric kidney development															x	
semaphorin-plexin signaling pathway															x	
semaphorin-plexin signaling pathway involved in neuron projection guidance															x	
regulation of epidermal growth factor receptor signaling pathway															x	
I-kappaB kinase/NF-kappaB signaling															x	
negative regulation of transforming growth factor beta receptor signaling pathway															x	

5.4 Conclusion

In this chapter, we have applied the proposed triclustering algorithm EMOA-δ-TRIMAX to a time-series gene expression data set monitoring the exposure of MCF-7 breast cancer to estrogen, in order to reveal the involvement of FFLs comprising two TFs and one common target gene, in mediating the potential signaling pathways which support uncontrolled cellular proliferation of breast cancer cell. Most of the identifying signaling pathways, genes of which are regulated by the TFs of the identified FFLs are known to be involved in promoting breast tumor growth.

5.5 Bibliography

[1] Malinowsky, K., Wolff, C., Ergin, B., Berg, D., Becker, K.F.: Deciphering signaling pathways in clinical tissues for personalized medicine using protein microarrays. Journal of Cellular Physiology **225**(2), 364–370 (2010). doi:10.1002/jcp.22307

[2] Macneil, L.T., Walhout, A.J.: Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome Research **21**(5), 645–657 (2011). doi:10.1101/gr.097378.109

[3] Wingender, E., Schoeps, T., Doenitz, J.: TFClass: an expandable hierarchical classification of human transcription factors. Nucleic Acids Research **41**(Database issue), 165–170 (2013). doi:10.1093/nar/gks1123

[4] de la Fuente, A.: What are Gene Regulatory Networks? IGI Global (2010). doi:10.4018/978-1-60566-685-3.ch001

[5] Widder, S., Sole, R., Macia, J.: Evolvability of feed-forward loop architecture biases its abundance in transcription networks. BMC Systems Biology **6**, 7 (2012). doi:10.1186/1752-0509-6-7

[6] Yan, Z., Shah, P., Amin, S., Samur, M., Huang, N., Wang, X., Misra, V., Ji, H., Gabuzda, D., Li, C.: Integrative analysis of gene and miRNA expression profiles with transcription factor-miRNA feed-forward loops identifies regulators in human cancers. Nucleic Acids Research **40**(17), 135 (2012)

[7] Peng, C., Wang, M., Shen, Y., Feng, H., Li, A.: Reconstruction and analysis of transcription factor-miRNA co-regulatory feed-forward loops in human

cancers using filter-wrapper feature selection. PLoS One **8**(10), e78197 (2013). doi:10.1371/journal.pone.0078197

[8] Ray, A., Ray, B.K.: Induction of Ras by SAF-1/MAZ through a feed-forward loop promotes angiogenesis in breast cancer. Cancer Medicine (2014). doi:10.1002/cam4.362

[9] Zhou, W., Srinivasan, S., Nawaz, Z., Slingerland, J.M.: ERα, SKP2 and E2F-1 form a feed forward loop driving late ERα targets and G1 cell cycle progression. Oncogene **33**(18), 2341–2353 (2014). doi:10.1038/onc.2013.197

[10] Alvarez, J.V., Pan, T.C., Ruth, J., Feng, Y., Zhou, A., Pant, D., Grimley, J.S., Wandless, T.J., Demichele, A., Chodosh, L.A.: Par-4 downregulation promotes breast cancer recurrence by preventing multinucleation following targeted therapy. Cancer Cell **24**(1), 30–44 (2013). doi:10.1016/j.ccr.2013.05.007

[11] Bhar, A., Haubrock, M., Mukhopadhyay, A., Wingender, E.: Multiobjective Triclustering Of Time-Series Transcriptome Data Reveals Key Genes Of Biological Processes. BMC Bioinformatics, (in press) (2015)

[12] Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E., Silver, P., Gingeras, T.R., Liu, X.S., Brown, M.: Genome-wide analysis of estrogen receptor binding sites. Nature Genetics **38**(11), 1289–1297 (2006)

[13] Wettenhall, J.M., Smyth, G.K.: limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics **20**(18), 3705–3706 (2004)

[14] Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhaeuser, R., Pruess, M., Schacherer, F., Thiele, S., Urbach, S.: The TRANSFAC system on gene expression regulation. Nucleic Acids Research **29**(1), 281–283 (2001)

[15] Szekely, G.J., Rizzo, M.L., Bakirov, N.K.: Measuring and testing dependence by correlation of distances. The Annals of Statistics **35**(6), 2769–2794 (2007)

[16] Csardi, G., Nepusz, T.: The igraph software package for complex network research. InterJournal **Complex Systems**, 1695 (2006)

[17] Rizza, P., Barone, I., Zito, D., Giordano, F., Lanzino, M., De Amicis, F., Mauro, L., Sisci, D., Catalano, S., Wright, K.D., Gustafsson, J.A., Ando, S.: Estrogen receptor beta as a novel target of androgen receptor action in breast cancer cell lines. Breast Cancer Research **16**(1), 21 (2014). doi:10.1186/bcr3619

[18] Yan, H., Zhu, S., Song, C., Liu, N., Kang, J.: Bone morphogenetic protein (BMP) signaling regulates mitotic checkpoint protein levels in human breast cancer cells. Cellular Signalling **24**(4), 961–968 (2012). doi:10.1016/j.cellsig.2011.12.019

[19] Fung, M.K., Cheung, H.W., Wong, H.L., Yuen, H.F., Ling, M.T., Chan, K.W., Wong, Y.C., Cheung, A.L., Wang, X.: MAD2 expression and its significance in mitotic checkpoint control in testicular germ cell tumour. Biochimica et Biophysica Acta **1773**(6), 821–832 (2007)

[20] Vandenberg, A.L., Sassoon, D.A.: Non-canonical Wnt signaling regulates cell polarity in female reproductive tract development via van gogh-like 2. Development **136**(9), 1559–1570 (2009). doi:10.1242/dev.034066

[21] Royer, C., Lu, X.: Epithelial cell polarity: a major gatekeeper against cancer? Cell Death and Differentiation **18**(9), 1470–1477 (2011). doi:10.1038/cdd.2011.60

[22] Yang, N., Liu, C., Peck, A.R., Girondo, M.A., Yanac, A.F., Tran, T.H., Utama, F.E., Tanaka, T., Freydin, B., Chervoneva, I., Hyslop, T., Kovatich, A.J., Hooke, J.A., Shriver, C.D., Rui, H.: Prolactin-Stat5 signaling in breast cancer is potently disrupted by acidosis within the tumor microenvironment. Breast Cancer Research **15**(5), 73 (2013)

[23] Knuepfer, H., Preiss, R.: Significance of interleukin-6 (IL-6) in breast cancer (review). Breast Cancer Research and Treatment **102**(2), 129–135 (2007)

[24] Swierczak, A., Cook, A.D., Lenzo, J.C., Restall, C.M., Doherty, J.P., Anderson, R.L., Hamilton, J.A.: The promotion of breast cancer metastasis caused by inhibition of CSF-1R/CSF-1 signaling is blocked by targeting the G-CSF receptor. Cancer Immunology Research **2**(8), 1–12 (2014). doi:10.1158/2326-6066.CIR-13-0190

[25] Rubis, B., Grodecka-Gazdecka, S., Lecybyl, R., Ociepa, M., Krozowski, Z., Trzeciak, W.H.: Contribution of protein kinase A and protein kinase C signalling pathways to the regulation of HSD11B2 expression and proliferation of MCF-7 cells. Acta Biochimica Polonica **51**(4), 919–924 (2004)

[26] Scollen, S., Luccarini, C., Baynes, C., Driver, K., Humphreys, M.K., Garcia-Closas, M., Figueroa, J., Lissowska, J., Pharoah, P.D., Easton, D.F., Hesketh, R., Metcalfe, J.C., Dunning, A.M.: TGF-β signaling pathway and breast cancer susceptibility. Cancer Epidemiology, Biomarkers and Prevention **20**(6), 1112–1119 (2011). doi:10.1158/1055-9965.EPI-11-0062

[27] Lee, S.H., Nam, H.S.: TNF alpha-induced down-regulation of estrogen receptor alpha in MCF-7 breast cancer cells. Molecules and Cells **26**(3), 285–290 (2008)

[28] Fulda, S., Debatin, K.M.: Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy. Oncogene **25**(34), 4798–4811 (2006)

[29] Masters, S.C., Fu, H.: 14-3-3 proteins mediate an essential anti-apoptotic signal. The Journal of Biological Chemistry **276**(48), 45193–45200 (2001)

[30] Silke, J., Hawkins, C.J., Ekert, P.G., Chew, J., Day, C.L., Pakusch, M., Verhagen, A.M., Vaux, D.L.: The anti-apoptotic activity of XIAP is retained upon mutation of both the caspase 3- and caspase 9-interacting sites. The Journal of Cell Biology **157**(1), 115–124 (2002)

[31] Lee, J.W., Soung, Y.H., Young Kim, S., Woo Nam, S., Sang Park, W., Young Lee, J., Jin Yoo, N., Lee, S.H.: Mutational analysis of proapoptotic ARTS P-loop domain in common human cancers. Pathology, Research and Practice **202**(2), 67–70 (2006)

[32] Hu, Z.Q., Zhang, J.Y., Ji, C.N., Xie, Y., Chen, J.Z., Mao, Y.M.: Grb10 interacts with Bim L and inhibits apoptosis. Molecular Biology Reports **37**(7), 3547–3552 (2010). doi:10.1007/s11033-010-0002-9

5.6 Appendix

Table 5.3: Signaling pathways regulated by three node feed-forward loop in the early module. A, B and C represent the positions in FFL depicted by Figure 5.6.

GOID	GOBPTerm regulated by *ESR1* (A) and *E2F2* (B)	Gene Symbol (C) in the identified FFLs
GO:0030521	androgen receptor signaling pathway	CCNE1
GO:2000474	regulation of opioid receptor signaling pathway	SYP
GO:0008543	fibroblast growth factor receptor signaling pathway	ERBB2
GO:0030518	intracellular steroid hormone receptor signaling pathway	THRAP3
GO:0048011	neurotrophin TRK receptor signaling pathway	ERBB2
GO:0007166	cell surface receptor signaling pathway	ERBB2
GO:0060766	negative regulation of androgen receptor signaling pathway	HDAC1

Table 5.3 continued...

GOID	GOBPTerm regulated by *ESR1* (A) and *BCL6* (B)	Gene Symbol (C) in the identified FFLs
GO:2000474	regulation of opioid receptor signaling pathway	SYP
GO:0030518	intracellular steroid hormone receptor signaling pathway	THRAP3
GO:0048011	neurotrophin TRK receptor signaling pathway	FGFR4

Table 5.4: Signaling pathways regulated by three node feed-forward loop in the late module. A, B and C represent the positions in FFL depicted by Figure 5.6.

GOID	GOBPTerm regulated by *E2F2* (A) and *ARNTL* (B)	Gene Symbol (C) in the identified FFLs
GO:0035358	regulation of peroxisome proliferator activated receptor signaling pathway	SIRT1
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	TRPS1
GO:0048011	neurotrophin TRK receptor signaling pathway	TIAM1
GO:0007179	transforming growth factor beta receptor signaling pathway	CDK8
GO:0061314	Notch signaling involved in heart development	JAG1
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0060070	canonical Wnt signaling pathway	KDM6A
GO:0007166	cell surface receptor signaling pathway	FZD3
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	MTSS1
GO:0044332	Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:2000051	negative regulation of non-canonical Wnt signaling pathway	LRP6
GO:0090244	Wnt signaling pathway involved in somitogenesis	LRP6
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0007199	G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	FZD3
GO:0007219	Notch signaling pathway	CDK8
GO:0030514	negative regulation of BMP signaling pathway	TRIM33

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *ATF1* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	MEF2A
GO:0034146	toll-like receptor 5 signaling pathway	MEF2A
GO:0038123	toll-like receptor TLR1:TLR2 signaling pathway	MEF2A
GO:0038124	toll-like receptor TLR6:TLR2 signaling pathway	MEF2A
GO:0034162	toll-like receptor 9 signaling pathway	MEF2A
GO:0035666	TRIF-dependent toll-like receptor signaling pathway	MEF2A
GO:0034134	toll-like receptor 2 signaling pathway	MEF2A
GO:0002756	MyD88-independent toll-like receptor signaling pathway	MEF2A
GO:0034138	toll-like receptor 3 signaling pathway	MEF2A
GO:0002755	MyD88-dependent toll-like receptor signaling pathway	MEF2A
GO:0034142	toll-like receptor 4 signaling pathway	MEF2A
GO:0002224	toll-like receptor signaling pathway	MEF2A
GO:0007179	transforming growth factor beta receptor signaling pathway	CGN
GO:0008286	insulin receptor signaling pathway	AKT2
GO:0030111	regulation of Wnt signaling pathway	SENP2
GO:0007166	cell surface receptor signaling pathway	STC2
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	AHI1
GO:0016055	Wnt signaling pathway	MARK2
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0030509	BMP signaling pathway	EGR1
GO:0030518	intracellular steroid hormone receptor signaling pathway	MED14
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	YWHAQ
GO:0033148	positive regulation of intracellular estrogen receptor signaling pathway	DDX17

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *CLOCK* (B)	Gene Symbol (C) in the identified FFLs
GO:0035358	regulation of peroxisome proliferator activated receptor signaling pathway	SIRT1
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	TRPS1
GO:0048011	neurotrophin TRK receptor signaling pathway	TIAM1
GO:0007179	transforming growth factor beta receptor signaling pathway	CDK8
GO:0061314	Notch signaling involved in heart development	JAG1
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0060070	canonical Wnt signaling pathway	KDM6A
GO:0007166	cell surface receptor signaling pathway	FZD3
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	MTSS1
GO:0044332	Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:2000051	negative regulation of non-canonical Wnt signaling pathway	LRP6
GO:0090244	Wnt signaling pathway involved in somitogenesis	LRP6
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0007199	G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	FZD3
GO:0007219	Notch signaling pathway	CDK8
GO:0030514	negative regulation of BMP signaling pathway	TRIM33

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *EGR1* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	TAB2
GO:0034146	toll-like receptor 5 signaling pathway	TAB2
GO:0038123	toll-like receptor TLR1:TLR2 signaling pathway	TAB2
GO:0038124	toll-like receptor TLR6:TLR2 signaling pathway	TAB2
GO:0034162	toll-like receptor 9 signaling pathway	TAB2
GO:0035666	TRIF-dependent toll-like receptor signaling pathway	TAB2
GO:0034134	toll-like receptor 2 signaling pathway	TAB2
GO:0002756	MyD88-independent toll-like receptor signaling pathway	TAB2
GO:0034138	toll-like receptor 3 signaling pathway	TAB2
GO:0002755	MyD88-dependent toll-like receptor signaling pathway	TAB2
GO:0034142	toll-like receptor 4 signaling pathway	TAB2
GO:0002224	toll-like receptor signaling pathway	TAB2
GO:0048011	neurotrophin TRK receptor signaling pathway	FGFR1
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0008286	insulin receptor signaling pathway	FGFR1
GO:0035607	fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	FGFR1
GO:0038028	insulin receptor signaling pathway via phosphatidylinositol 3-kinase	PIK3CA
GO:0007166	cell surface receptor signaling pathway	IL13RA1
GO:0030178	negative regulation of Wnt signaling pathway	NLK
GO:0090244	Wnt signaling pathway involved in somitogenesis	PPP2R3A
GO:0030522	intracellular receptor signaling pathway	NR2F2

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *EGR1* (B)	Gene Symbol (C) in the identified FFLs
GO:0008543	fibroblast growth factor receptor signaling pathway	FGFR1
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0097191	extrinsic apoptotic signaling pathway	SMAD3
GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	BCLAF1
GO:0030521	androgen receptor signaling pathway	MED14
GO:0042921	glucocorticoid receptor signaling pathway	ARID1A
GO:0007173	epidermal growth factor receptor signaling pathway	FGFR1
GO:0070104	negative regulation of interleukin-6-mediated signaling pathway	PTPN2
GO:0030518	intracellular steroid hormone receptor signaling pathway	MED14
GO:0038095	Fc-epsilon receptor signaling pathway	FGFR1
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	LTBP4
GO:0030520	intracellular estrogen receptor signaling pathway	ARID1A
GO:0010754	negative regulation of cGMP-mediated signaling	THBS1
GO:0060520	activation of prostate induction by androgen receptor signaling pathway	AR
GO:1902206	negative regulation of interleukin-2-mediated signaling pathway	PTPN2
GO:1902212	negative regulation of prolactin signaling pathway	PTPN2
GO:1902215	negative regulation of interleukin-4-mediated signaling pathway	PTPN2
GO:1902227	negative regulation of macrophage colony-stimulating factor signaling pathway	PTPN2
GO:0010804	negative regulation of tumor necrosis factor-mediated signaling pathway	RFFL
GO:0033148	positive regulation of intracellular estrogen receptor signaling pathway	AR
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	BCL2L11
GO:0097296	activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	SMAD3

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *GATA3* (B)	Gene Symbol (C) in the identified FFLs
GO:0007196	adenylate cyclase-inhibiting G-protein coupled glutamate receptor signaling pathway	GRM8
GO:0035607	fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	FGFR1
GO:0030178	negative regulation of Wnt signaling pathway	NLK
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0035425	autocrine signaling	FZD1
GO:0044338	canonical Wnt signaling pathway involved in mesenchymal stem cell differentiation	FZD1
GO:0044339	canonical Wnt signaling pathway involved in osteoblast differentiation	FZD1
GO:0040036	regulation of fibroblast growth factor receptor signaling pathway	FAM20C
GO:0014066	regulation of phosphatidylinositol 3-kinase signaling	C3orf58
GO:0030520	intracellular estrogen receptor signaling pathway	RARA
GO:0010754	negative regulation of cGMP-mediated signaling	THBS1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *HIF1A* (B)	Gene Symbol (C) in the identified FFLs
GO:0035358	regulation of peroxisome proliferator activated receptor signaling pathway	SIRT1
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	TRPS1
GO:0048011	neurotrophin TRK receptor signaling pathway	TIAM1
GO:0007179	transforming growth factor beta receptor signaling pathway	CDK8
GO:0061314	Notch signaling involved in heart development	JAG1
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:1901189	positive regulation of ephrin receptor signaling pathway	RBPJ
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0060070	canonical Wnt signaling pathway	TCF7L1
GO:0007166	cell surface receptor signaling pathway	FZD1
GO:0030178	negative regulation of Wnt signaling pathway	LRP6
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	MTSS1
GO:0016055	Wnt signaling pathway	RTF1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *HIF1A* (B)	Gene Symbol (C) in the identified FFLs
GO:0044332	Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:2000051	negative regulation of non-canonical Wnt signaling pathway	LRP6
GO:0090244	Wnt signaling pathway involved in somitogenesis	LRP6
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0030509	BMP signaling pathway	EGR1
GO:0035425	autocrine signaling	FZD1
GO:0044338	canonical Wnt signaling pathway involved in mesenchymal stem cell differentiation	FZD1
GO:0044339	canonical Wnt signaling pathway involved in osteoblast differentiation	FZD1
GO:0014066	regulation of phosphatidylinositol 3-kinase signaling	C3orf58
GO:0035872	nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	TAB2
GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	BCL2L11
GO:0007223	Wnt signaling pathway, calcium modulating pathway	FZD1
GO:0007199	G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	FZD1
GO:0007219	Notch signaling pathway	CDK8
GO:0030514	negative regulation of BMP signaling pathway	FZD1
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	BCL2L11
GO:0030520	intracellular estrogen receptor signaling pathway	RBFOX2
GO:0032933	SREBP signaling pathway	INSIG1
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	FOXO3
GO:0019221	cytokine-mediated signaling pathway	NUP188

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *MAX* (B)	Gene Symbol (C) in the identified FFLs
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0016055	Wnt signaling pathway	NLK
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0007219	Notch signaling pathway	CDK8
GO:0038027	apolipoprotein A-I-mediated signaling pathway	ABCA1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *MYB/ MYBL1* (B)	Gene Symbol (C) in the identified FFLs
GO:0007196	adenylate cyclase-inhibiting G-protein coupled glutamate receptor signaling pathway	GRM8
GO:0035358	regulation of peroxisome proliferator activated receptor signaling pathway	SIRT1
GO:0048011	neurotrophin TRK receptor signaling pathway	BRAF
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0097190	apoptotic signaling pathway	FASTK
GO:2001234	negative regulation of apoptotic signaling pathway	PCGF2
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0030111	regulation of Wnt signaling pathway	LEF1
GO:0030178	negative regulation of Wnt signaling pathway	LZTS2
GO:0016055	Wnt signaling pathway	LZTS2
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0097191	extrinsic apoptotic signaling pathway	SMAD3
GO:0038096	Fc-gamma receptor signaling pathway involved in phagocytosis	NCK1
GO:0035872	nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	TAB3
GO:0097193	intrinsic apoptotic signaling pathway	YWHAQ
GO:0070423	nucleotide-binding oligomerization domain containing signaling pathway	TAB3
GO:0030521	androgen receptor signaling pathway	UBE3A
GO:0042921	glucocorticoid receptor signaling pathway	NR3C1
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	YWHAQ
GO:0035329	hippo signaling	STK3
GO:0030520	intracellular estrogen receptor signaling pathway	ARID1A
GO:0038027	apolipoprotein A-I-mediated signaling pathway	ABCA1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *NFYB/ NFYC* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	NFKB2
GO:0034146	toll-like receptor 5 signaling pathway	NFKB2
GO:0038123	toll-like receptor TLR1:TLR2 signaling pathway	NFKB2
GO:0038124	toll-like receptor TLR6:TLR2 signaling pathway	NFKB2
GO:0034162	toll-like receptor 9 signaling pathway	NFKB2
GO:0035666	TRIF-dependent toll-like receptor signaling pathway	NFKB2
GO:0034134	toll-like receptor 2 signaling pathway	NFKB2
GO:0002756	MyD88-independent toll-like receptor signaling pathway	NFKB2
GO:0048011	neurotrophin TRK receptor signaling pathway	CASP2
GO:0097190	apoptotic signaling pathway	CASP2
GO:0060070	canonical Wnt signaling pathway	GSK3B
GO:0016055	Wnt signaling pathway	GSK3B
GO:0038180	nerve growth factor signaling pathway	SORT1
GO:0008543	fibroblast growth factor receptor signaling pathway	GSK3B
GO:0038096	Fc-gamma receptor signaling pathway involved in phagocytosis	WASF2
GO:0097193	intrinsic apoptotic signaling pathway	BCL2L1
GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	SEPT4
GO:2001235	positive regulation of apoptotic signaling pathway	CASP2
GO:0007219	Notch signaling pathway	SEL1L
GO:0030949	positive regulation of vascular endothelial growth factor receptor signaling pathway	HIF1A
GO:1901030	positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	GSK3B
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	CASP2

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *PATZ1* (B)	Gene Symbol (C) in the identified FFLs
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0007166	cell surface receptor signaling pathway	STC2
GO:0008543	fibroblast growth factor receptor signaling pathway	FOXO3
GO:0007173	epidermal growth factor receptor signaling pathway	REPS2
GO:0070104	negative regulation of interleukin-6-mediated signaling pathway	PTPN2
GO:0038095	Fc-epsilon receptor signaling pathway	FOXO3
GO:1902206	negative regulation of interleukin-2-mediated signaling pathway	PTPN2
GO:1902212	negative regulation of prolactin signaling pathway	PTPN2
GO:1902215	negative regulation of interleukin-4-mediated signaling pathway	PTPN2
GO:1902227	negative regulation of macrophage colony-stimulating factor signaling pathway	PTPN2
GO:0010804	negative regulation of tumor necrosis factor-mediated signaling pathway	PTPN2

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *RUNX1/ RUNX2* (B)	Gene Symbol (C) in the identified FFLs
GO:0048011	neurotrophin TRK receptor signaling pathway	PPP2R5D
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0008286	insulin receptor signaling pathway	PHIP
GO:0016055	Wnt signaling pathway	NLK
GO:0008543	fibroblast growth factor receptor signaling pathway	YWHAB
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	YWHAE
GO:0035329	hippo signaling	YWHAE

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *SP1/ SP3* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	TAB2
GO:0034146	toll-like receptor 5 signaling pathway	TAB2
GO:0035666	TRIF-dependent toll-like receptor signaling pathway	TAB2
GO:0002756	MyD88-independent toll-like receptor signaling pathway	TAB2
GO:0034138	toll-like receptor 3 signaling pathway	TAB2
GO:0034142	toll-like receptor 4 signaling pathway	TAB2
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	TRPS1
GO:0048011	neurotrophin TRK receptor signaling pathway	FRS2
GO:0007179	transforming growth factor beta receptor signaling pathway	SMAD2
GO:0097190	apoptotic signaling pathway	FASTK
GO:2001234	negative regulation of apoptotic signaling pathway	PCGF2
GO:0043401	steroid hormone mediated signaling pathway	BMP7
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0008286	insulin receptor signaling pathway	FRS2
GO:0007206	phospholipase C-activating G-protein coupled glutamate receptor signaling pathway	HOMER1
GO:0035607	fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	FGFR1
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0038028	insulin receptor signaling pathway via phosphatidylinositol 3-kinase	PIK3CA
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0030111	regulation of Wnt signaling pathway	LEF1
GO:0007166	cell surface receptor signaling pathway	STC2
GO:0030178	negative regulation of Wnt signaling pathway	LZTS2
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	FRS2
GO:0016055	Wnt signaling pathway	RTF1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *SP1/ SP3* (B)	Gene Symbol (C) in the identified FFLs
GO:0090244	Wnt signaling pathway involved in somitogenesis	PPP2R3A
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0030509	BMP signaling pathway	BMP7
GO:0008543	fibroblast growth factor receptor signaling pathway	FRS2
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0097191	extrinsic apoptotic signaling pathway	SMAD3
GO:0040036	regulation of fibroblast growth factor receptor signaling pathway	RUNX2
GO:0014066	regulation of phosphatidylinositol 3-kinase signaling	C3orf58
GO:0097193	intrinsic apoptotic signaling pathway	XIAP
GO:0042921	glucocorticoid receptor signaling pathway	ARID1A
GO:0007173	epidermal growth factor receptor signaling pathway	FRS2
GO:0070104	negative regulation of interleukin-6-mediated signaling pathway	PTPN2
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	BCL2L11
GO:0038095	Fc-epsilon receptor signaling pathway	FRS2
GO:0070102	interleukin-6-mediated signaling pathway	CTR9
GO:0035329	hippo signaling	STK3
GO:0030520	intracellular estrogen receptor signaling pathway	ARID1A
GO:0030949	positive regulation of vascular endothelial growth factor receptor signaling pathway	VEGFA
GO:0001960	negative regulation of cytokine-mediated signaling pathway	SIGIRR
GO:0010754	negative regulation of cGMP-mediated signaling	THBS1
GO:0038190	VEGF-activated neuropilin signaling pathway	VEGFA
GO:0060520	activation of prostate induction by androgen receptor signaling pathway	AR
GO:0070106	interleukin-27-mediated signaling pathway	IL6ST
GO:1902206	negative regulation of interleukin-2-mediated signaling pathway	PTPN2
GO:1902212	negative regulation of prolactin signaling pathway	PTPN2
GO:1902215	negative regulation of interleukin-4-mediated signaling pathway	PTPN2
GO:1902227	negative regulation of macrophage colony-stimulating factor signaling pathway	PTPN2
GO:0010804	negative regulation of tumor necrosis factor-mediated signaling pathway	RFFL
GO:0090037	positive regulation of protein kinase C signaling	VEGFA
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	FOXO3
GO:0097296	activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	SMAD3
GO:0019221	cytokine-mediated signaling pathway	KPNB1

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *STAT5B/ STAT6* (B)	Gene Symbol (C) in the identified FFLs
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0044332	Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:2000051	negative regulation of non-canonical Wnt signaling pathway	LRP6
GO:0090244	Wnt signaling pathway involved in somitogenesis	LRP6
GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	SEPT4
GO:0070104	negative regulation of interleukin-6-mediated signaling pathway	IL6ST
GO:0070102	interleukin-6-mediated signaling pathway	IL6ST
GO:0070106	interleukin-27-mediated signaling pathway	IL6ST

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *TEAD1/ TEAD2* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	TAB2
GO:0034146	toll-like receptor 5 signaling pathway	TAB2
GO:0048011	neurotrophin TRK receptor signaling pathway	FGFR1
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0061314	Notch signaling involved in heart development	JAG1
GO:0008286	insulin receptor signaling pathway	FGFR1
GO:0035607	fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	FGFR1
GO:0038028	insulin receptor signaling pathway via phosphatidylinositol 3-kinase	PIK3CA
GO:0090038	negative regulation of protein kinase C signaling	GPD1L
GO:0060070	canonical Wnt signaling pathway	LEF1
GO:0030111	regulation of Wnt signaling pathway	LEF1
GO:0016055	Wnt signaling pathway	NLK
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0030509	BMP signaling pathway	LEF1
GO:0008543	fibroblast growth factor receptor signaling pathway	FGFR1
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0040036	regulation of fibroblast growth factor receptor signaling pathway	RUNX2
GO:0007219	Notch signaling pathway	MIB1
GO:0030521	androgen receptor signaling pathway	ARID1A
GO:0007173	epidermal growth factor receptor signaling pathway	FGFR1
GO:0038095	Fc-epsilon receptor signaling pathway	FGFR1
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	LTBP4
GO:0030520	intracellular estrogen receptor signaling pathway	ARID1A
GO:0010754	negative regulation of cGMP-mediated signaling	THBS1
GO:0035723	interleukin-15-mediated signaling pathway	PLCB1
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	FOXO3

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *TFAP2A/ TFAP2C* (B)	Gene Symbol (C) in the identified FFLs
GO:0034166	toll-like receptor 10 signaling pathway	TAB2
GO:0034146	toll-like receptor 5 signaling pathway	TAB2
GO:0007196	adenylate cyclase-inhibiting G-protein coupled glutamate receptor signaling pathway	GRM8
GO:0035358	regulation of peroxisome proliferator activated receptor signaling pathway	SIRT1
GO:0048011	neurotrophin TRK receptor signaling pathway	TIAM1
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0097190	apoptotic signaling pathway	TIAM1
GO:2001269	positive regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	TFAP4
GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	TAB2
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0008286	insulin receptor signaling pathway	FRS2
GO:0035607	fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	FGFR1
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0038028	insulin receptor signaling pathway via phosphatidylinositol 3-kinase	PIK3CA
GO:0061290	canonical Wnt signaling pathway involved in metanephric kidney development	GATA3
GO:0060070	canonical Wnt signaling pathway	KDM6A
GO:0030111	regulation of Wnt signaling pathway	SENP2

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *TFAP2A/ TFAP2C* (B)	Gene Symbol (C) in the identified FFLs
GO:0007166	cell surface receptor signaling pathway	IL13RA1
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	FRS2
GO:0016055	Wnt signaling pathway	LZTS2
GO:0030522	intracellular receptor signaling pathway	NR2F2
GO:0008543	fibroblast growth factor receptor signaling pathway	FRS2
GO:0071526	semaphorin-plexin signaling pathway	SEMA4C
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0097191	extrinsic apoptotic signaling pathway	SMAD3
GO:0040036	regulation of fibroblast growth factor receptor signaling pathway	RUNX2
GO:0014066	regulation of phosphatidylinositol 3-kinase signaling	C3orf58
GO:1902285	semaphorin-plexin signaling pathway involved in neuron projection guidance	SEMA3F
GO:0035872	nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	TAB2
GO:0097193	intrinsic apoptotic signaling pathway	XIAP
GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	BCLAF1
GO:0007219	Notch signaling pathway	MIB1
GO:0030521	androgen receptor signaling pathway	UBE3A
GO:0042058	regulation of epidermal growth factor receptor signaling pathway	RHBDF1
GO:0007173	epidermal growth factor receptor signaling pathway	FRS2
GO:1900740	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	YWHAQ
GO:0038095	Fc-epsilon receptor signaling pathway	FRS2
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	LTBP4
GO:0035329	hippo signaling	TEAD2
GO:0007249	I-kappaB kinase/NF-kappaB signaling	TAB2
GO:0030520	intracellular estrogen receptor signaling pathway	ARID1A
GO:0030949	positive regulation of vascular endothelial growth factor receptor signaling pathway	VEGFA
GO:0030512	negative regulation of transforming growth factor beta receptor signaling pathway	SMAD3
GO:0038190	VEGF-activated neuropilin signaling pathway	VEGFA
GO:0060520	activation of prostate induction by androgen receptor signaling pathway	AR
GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	FOXO3
GO:0097296	activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	SMAD3

Table 5.4 continued...

GOID	GOBPTerm regulated by *E2F2* (A) and *USF2* (B)	Gene Symbol (C) in the identified FFLs
GO:0007179	transforming growth factor beta receptor signaling pathway	NLK
GO:0021874	Wnt signaling pathway involved in forebrain neuroblast division	LRP6
GO:0061310	canonical Wnt signaling pathway involved in cardiac neural crest cell differentiation involved in heart development	LRP6
GO:2000055	positive regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	LRP6
GO:0044335	canonical Wnt signaling pathway involved in neural crest cell differentiation	LRP6
GO:0061324	canonical Wnt signaling pathway involved in positive regulation of cardiac outflow tract cell proliferation	LRP6
GO:0044340	canonical Wnt signaling pathway involved in regulation of cell proliferation	LRP6
GO:0016055	Wnt signaling pathway	NLK
GO:2000124	regulation of endocannabinoid signaling pathway	MGLL
GO:0007219	Notch signaling pathway	CDK8
GO:0030949	positive regulation of vascular endothelial growth factor receptor signaling pathway	HIF1A
GO:0038027	apolipoprotein A-I-mediated signaling pathway	ABCA1

6 Unraveling Potential Signaling Pathways During the Exposure of Several Tissues to Different Toxicants in Different Species

6.1 Introduction

Maintaining proper cellular functions usually requires the crosstalk between several signaling pathways. Several toxic agents may impinge on such signaling pathways in a dose-dependent manner which in turn alters the activities of the signaling pathways and triggers the activation or inhibition of cellular apoptosis [1]. Studying toxicology may provide insights into the effects of several toxicants on living organisms. Advancement of microarray technology over the last decade not only provides ample opportunity to measure the gene expression over a set of time points during the exposure of tissues to certain chemical substances such as naphthalene, 1,2,3-trichloropropane etc. but also facilitates to monitor the gene expression changes over a set of doses of a particular toxicant. Furthermore, these substances may be potential in causing cancer and work in a dose-dependent manner. Exploratory analysis, such as gene co-expression studies, may help us to elucidate the relevant signaling pathways perturbed by those substances. In this chapter, we have applied the proposed triclustering algorithm EMOA-δ-TRIMAX [2] to four microarray gene expression datasets to reveal the signaling pathways which may be perturbed by the toxicants either over different time points or at different dosages.

6.2 Materials and Methods

6.2.1 Dataset 1

In the study of J.C. Snyder et al. and A.C. Zemke et al. [3, 4], experiments were carried out to investigate the role of imperfect epithelial repair in extracellular matrix (ECM) displacement and causing pulmonary fibrosis. Expression profiles of 18675 probe ids were measured over 5 time points (control, Days 1, 2, 3 and 6) and 4 replicates during the exposure of mouse lung tissue to naphthalene (GSE17693) [3, 4].

6.2.2 Dataset 2

Expression profiles of 45101 probe ids were monitored across 4 replicates during the exposure of female mouse lung tumor cell to the air and naphthalene toxicant with different doses (0.5, 3, 10, 20 and 30 parts per million (ppm)) to unveil potential biomarkers involved in causing lung cancer (GSE17993) [5].

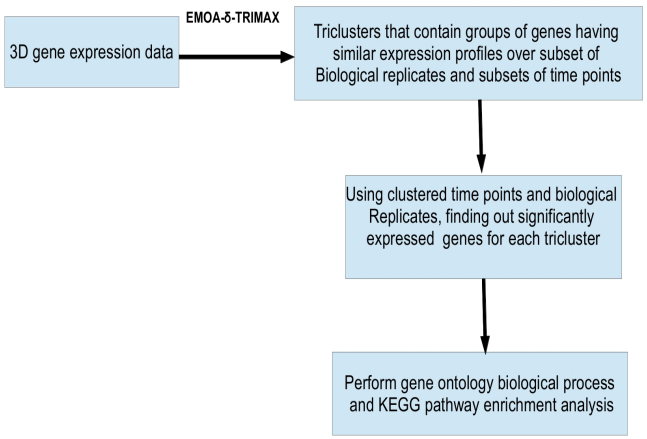

Figure 6.1: Workflow applied in this work.

6.2.3 Dataset 3

Expression profiles of 45101 probe ids were monitored across 4 replicates during the exposure of female mouse liver tumor cell to distilled water, air, naphthalene (0.5, 3, 10, 20 and 30 ppm) and 1,2,3-trichloropropane (2, 6, 20, 40 and 60 mg/kg) substances with different doses to unveil potential biomarkers involved in causing hepatic cancer (GSE18858) [6].

6.2.4 Dataset 4

To infer the role of matrix metalloproteinase 7 (MMP7) in diminishing ciliogenesis during wound repair, expression profiles of 45101 probe ids were measured over 4 replicates using MMP7-null and wild type mice exposed to naphthalene (GSE38513) [7].

6.2.5 Workflow

Figure 6.1 shows the workflow used in this chapter.

6.3 Results and Discussion

Table 6.1 shows the values assigned to the input parameters of EMOA-δ-TRIMAX algorithm for each of the datasets used in this chapter.

Table 6.1: Values of input parameters of EMOA-δ-TRIMAX algorithm.

Dataset	Value of λ	Value of δ	Size of population	Number of generations
Dataset 1	1.2	0.063	100	100
Dataset 2	1.2	0.00253	100	100
Dataset 3	1.2	110.216	100	100
Dataset 4	1.2	0.0041	100	100

We have used the *minSum* metric delineated in chapter 2 to examine whether solutions converge towards the Pareto optimal front around its center region. From Figure 6.2, we can see the convergence of solutions towards the Pareto optimal front for each of the aforementioned datasets. For instance, for dataset 1 and 2 the solutions converge after 78th generation, whereas in case of dataset 3, the convergence of the solutions occurs after 30th generation.

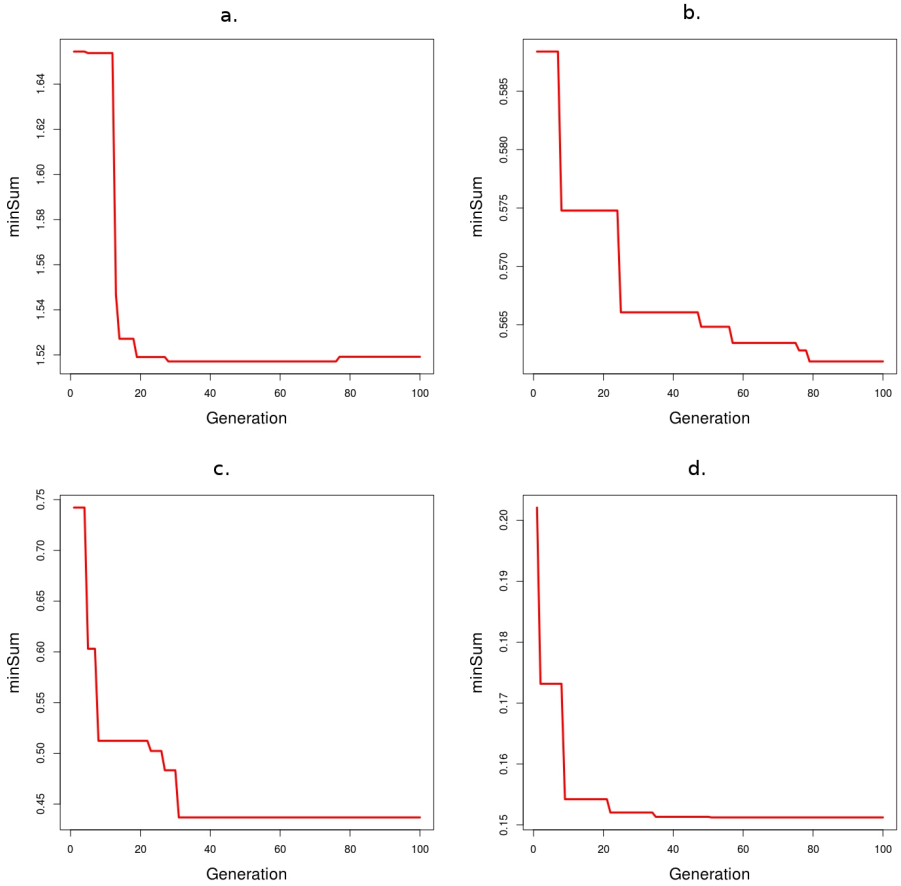

Figure 6.2: Convergence of solutions towards the Pareto optimal front. *minSum* values are plotted for dataset 1 (a), dataset 2 (b), dataset 3 (c) and dataset 4 (d)

6.3.1 Results on Dataset 1

From Tables 6.3-6.4 (see Appendix), we can see the list of signaling pathways that are triggered by naphthalene at different stages of the treatment. Moreover, the activation of ERBB, epidermal growth factor receptor signaling, unfolded protein response signaling protein, Fc-receptor mediated signaling, endoplasmic reticulum (ER) to nucleus signaling, corticosteroid receptor signaling, glucocorticoid receptor signaling, NF-kappaB signaling, arginine metabolism, thromboxane A2 signaling, neurotrophin signaling, PPAR signaling, Wnt signaling, sphingosine-1-phosphate signaling, JAK-STAT pathway, Notch signaling have been inferred to be associated with pulmonary fibrosis, supported by previously published evidences [8–23]. Moreover, the co-occurrence of diseases such as pulmonary fibrosis,

diabetes mellitus, *Vibrio cholerae infection*, acute myeloid leukemia, Alzheimers disease have also been established by previous studies [24–27]. Furthermore, type I interferon and MDA-5 signaling both jointly play crucial roles in the context of pulmonary fibrosis [28]. In addition to the aforementioned signaling pathways, finding the inhibition of BMP, activin receptor and protein kinase C signaling pathways is not astonishing to us as deletion of these pathways may trigger the pulmonary fibrosis [29–31]. From Figure 6.3a, we can see that those that start early, which are most of the triggered pathways (41%), persist for at least 3 days, additional 30% persist for the whole period of exposure to the toxic substance. Among the very late ones (activated on days 3-6) are those that are specific for lung cancer (either small-cell or non-small cell lung cancer) and several other cancer types. From Figure 6.3b, we can obtain the similar trend i.e. 58% of the triggered pathways start early and persist for at least 3 days, whereas additional 12% persist for the whole period of exposure to the toxic substance. Additional 3% of the triggered pathways are found to persist on days 3-6 and these pathways include somatostatin receptor, insulin-like growth factor receptor signaling which were also inferred to be associated with the pulmonary fibrosis in previous studies [32, 33].

Furthermore, to investigated whether there is a overlap between the pathways found to be significant for the up- and down-regulated genes, we have computed Jaccard similarity coefficient which is 0.14 in case of this dataset. This overlap is due to the molecules constituting it.

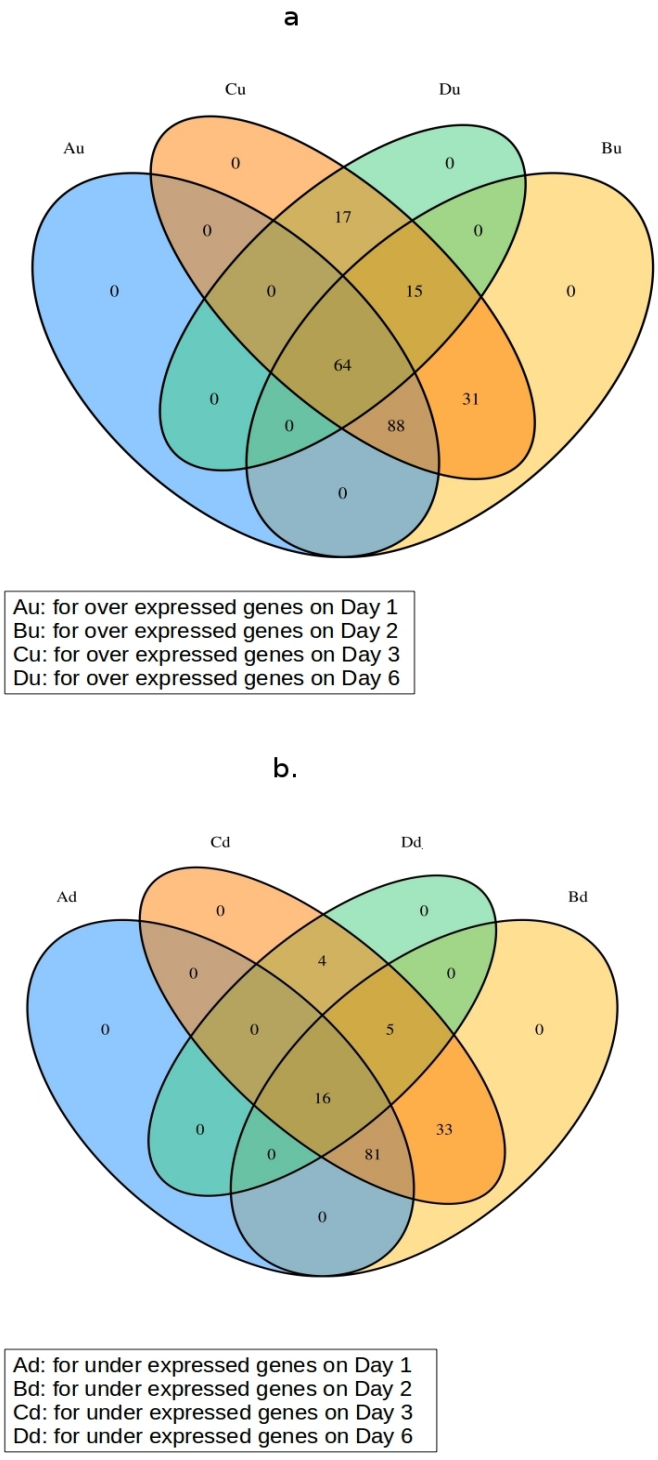

Figure 6.3: Venn diagram showing the overlap between the lists of pathways found to be perturbed for over- (a.) and under (b.)-expressed genes across different days of cellular exposure to naphthalene (NPTH) for dataset 1.

6.3.2 Results on Dataset 2

From Table 6.5 and Table 6.6 (see Appendix), we can see that estrogen receptor plays an important role in driving the growth of lung tumor through stimulating steroid hormone receptor signaling; this fact has already been established by a previous study [34]. Moreover, from the lists of signaling pathways, we can observe the inverse relationships between ERBB signaling and estrogen receptor signaling and this finding was also supported by a previous study [35, 36]. Our results also indicate that Shigellosis, renal cell carcinoma, depression are known to co-occur with lung cancer [37–39]. Furthermore, insulin receptor signaling, mTOR signaling, lipopolysaccharide-mediated signaling, GnRH signaling, *Helicobacter pylori infection*, roundabout signaling, apoptotic signaling, natural killer cell, calcium signaling, gamma-aminobutyric acid signaling have already been inferred to play crucial roles in driving lung tumor growth [40–50]. Moreover, finding the involvement of the genes expressed at lower levels in purinergic receptor signaling, may indicate the fact that dysregulation of purinergic signaling may be associated with the risk for lung infection [51]. Figure 6.4a shows that most of the triggered pathways (45%) are found to be perturbed by the two lowest doses of naphthalene, whereas additional 10% pathways are found to be activated by each of the doses of the toxic substance. Additional 16% of the pathways are perturbed by the three highest doses of the toxicant. Moreover, it is of interest to see that additional 4%, 3% and 8% pathways are perturbed by the two lowest doses and the highest dose of the toxic substance, respectively. From Figure 6.4b, we can see that 34% of the triggered pathways are found to be activated by the two lowest doses of naphthalene, whereas additional 8% pathways are found to be perturbed by the two highest doses only. Additional 6% pathways are found to be perturbed by all doses of the toxic substance, whereas additional 14% of the triggered pathways are activated by the three highest doses of naphthalene. Furthermore, this figure also indicates the fact that perturbation of the pathways may occur in a dose-dependent manner as additional 9%, 5% and 8% pathways are uniquely activated by the two lowest doses and the highest dose of naphthalene, respectively.

Furthermore, to investigated whether there is a overlap between the pathways found to be significant for the up- and down-regulated genes, we have computed Jaccard similarity coefficient which is 0.40 in case of this dataset. This overlap is due to the molecules constituting it.

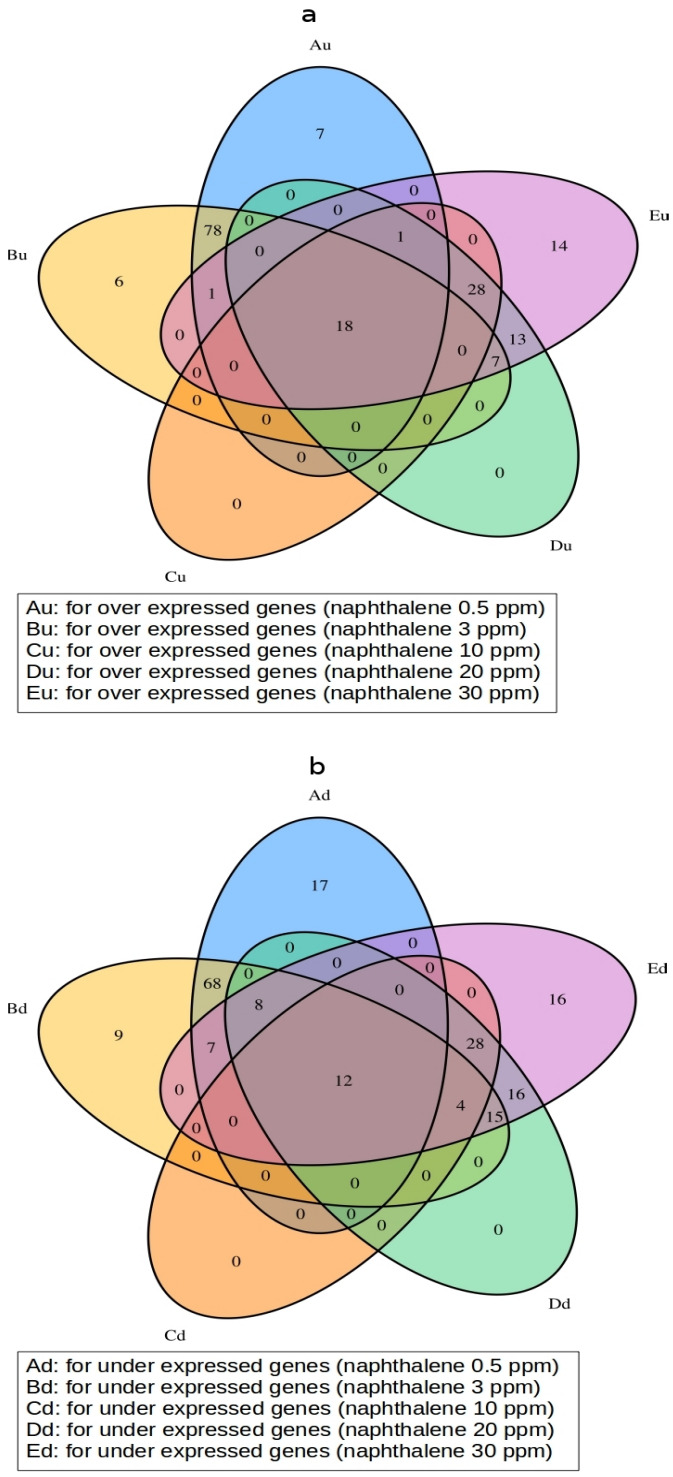

Figure 6.4: Venn diagram showing the overlap between the lists of pathways found to be perturbed for over- (a.) and under (b.)-expressed genes at different doses of naphthalene (NPTH) for dataset 2.

6.3.3 Results on Dataset 3

From Tables 6.7-6.8 (see Appendix), we can observe the inhibition of transforming growth factor beta receptor signaling and this finding may indicate the fact that hormone-mediated signaling play crucial roles in liver cancer growth as TGF-beta signaling was inferred to be a suppressor in hormone-induced cancers [52]. Co-enrichment of PPAR signaling and fatty acid metabolism is also biologically meaningful as PPAR is known to trigger the fatty acid catabolism in the context of liver cancer [53]. Moreover, adiponectin and porphyrin were inferred to be major players in governing the liver tumor growth [54, 55]. Figures 6.5-6.6 show the overlap between the lists of pathways found to be activated by two toxic substances namely, naphthalene and 1,2,3-trichloropropane (TCPN), respectively. Figures 6.5a and 6.6a show that 18% and 30% of the triggered pathways are perturbed by all doses of naphthalene and 1,2,3-trichloropropane, respectively. Moreover, from Figure 6.5a, we can see that 4% of the triggered pathways are activated by the four highest doses of naphthalene, whereas additional 9% pathways are perturbed by the first four doses of the toxic substance. From figure 6.5a it becomes apparent that naphthalene activates the pathways in a dose-dependent manner as additional 7%, 1%, 3% and 6% pathways are found to be perturbed by the first three doses and the highest dose of naphthalene, respectively. From Figure 6.5b, we can observe that most of the triggered pathways (36%) are activated by all doses of naphthalene, whereas additional 21 % pathways are found to be perturbed by the first and last three doses of the toxic substance. Another 4% pathways pathways are found to be perturbed by the last four doses of naphthalene, whereas additional 5% pathways are activated by the first four doses of the toxic substance. Furthermore, additional 2%, 2%, 1% and 2% pathways are found to be unique at 0.5, 3, 10 and 30 parts per million of naphthalene, respectively. Besides, from Figure 6.6a, we can see that 8% pathways are triggered by the three highest doses of TCPN, whereas another 6% pathways persist at the two lowest doses of the toxic substance. Additional 6% of the triggered pathways persist at the two highest doses of TCPN, whereas another 1% and 6% pathways are found to be unique at 2 and 60 mg/kg of 1,2,3-trichloropropane, respectively. From figure 6.6b, it becomes apparent that most of the triggered pathways persist at all doses of TCPN, additional 20% pathways persist only the two lowest doses of the toxic substance. Moreover, another 9% pathways are found to be triggered by the three highest doses, whereas additional 4% pathways persist at the first four doses of TCPN. Furthermore, additional 1% pathways are found to be unique at 6 and 20mg/kg of 1,2,3-trichloropropane. Altogether, from figures 6.5-6.6, it becomes apparent that the pathways are perturbed by naphthalene

and 1,2,3-trichloropropane in a dose-dependent manner.
Additionally, Figure 6.7 shows to what extent the pathways activated by the toxic substances namely, naphthalene and 1,2,3-trichloropropane overlap or differ. From this figure, we can conclude the fact that there is a large overlap between the pathways perturbed by these two toxic substances in case of both the up- and down-regulated genes (Jaccard similarity coefficient is 0.81 and 0.58, respectively).
Furthermore, to investigated whether there is a overlap between the pathways found to be significant for the up- and down-regulated genes, we have computed Jaccard similarity coefficient which is 0.46 in case of this dataset. This overlap is due to the molecules constituting it.

6.3.4 Results on Dataset 4

The aim of generating this data was to unveil the role of matrix metalloproteinase 7 (MMP7) in injured airway epithelial cells. Overexpression of MMP7 may facilitate profibrotic effects in pulmonary fibrosis by enhancing wound healing and tissue inflammation, whereas diminishing its expression may result in differentiation of ciliated cells in the epithelial barrier [7]. Tables 6.9-6.10 (see Appendix) indicate that the involvement of insulin receptor signaling, G-protein coupled receptor signaling, adrenergic receptor signaling, TGF-beta signaling induced by a bile compound, lipopolysaccharide-mediated signaling (LPS), protein kinase A signaling in both matrix metalloproteinase 7 (MMP7)-null and wild type mice injured airway epithelial cells. Each of these aforementioned pathways have been reported to be associated with injured airway epithelial cells [56–62]. Furthermore, it is of interest to see the pathways such as thyroid hormone mediated signaling, androgen receptor signaling, vitamin D receptor signaling, estrogen receptor signaling, oncostatin-M-mediated signaling, purine metabolism are coming up in the MMP7-null mice only in case of the up-regulated genes as these signaling pathways were inferred to exhibit anti-fibrotic or anti-inflammatory activities by the previously published literatures [63–68].
Furthermore, to investigated whether there is a overlap between the pathways found to be significant for the up- and down-regulated genes, we have computed Jaccard similarity coefficient which is 0 in case of this dataset.

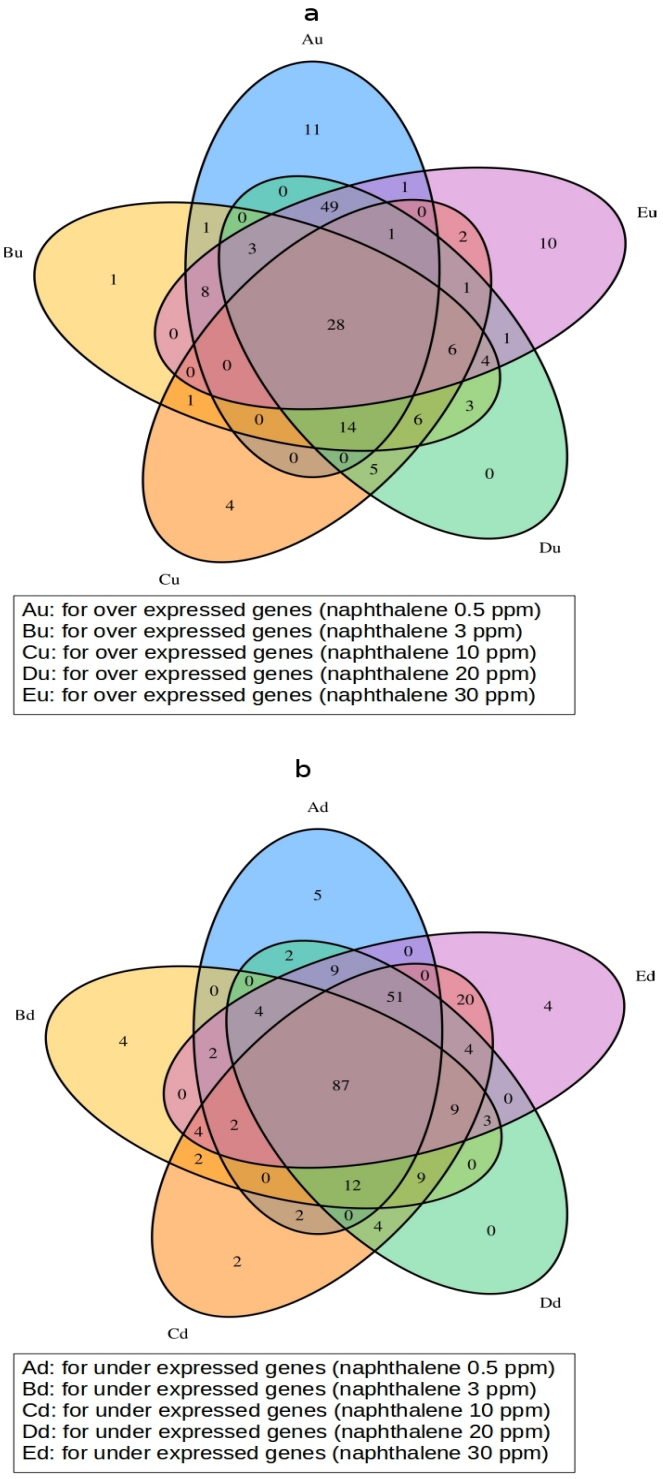

Figure 6.5: Venn diagram showing the overlap between the lists of pathways found to be perturbed for over- (a.) and under (b.)-expressed genes at different doses of naphthalene (NPTH) for dataset 3.

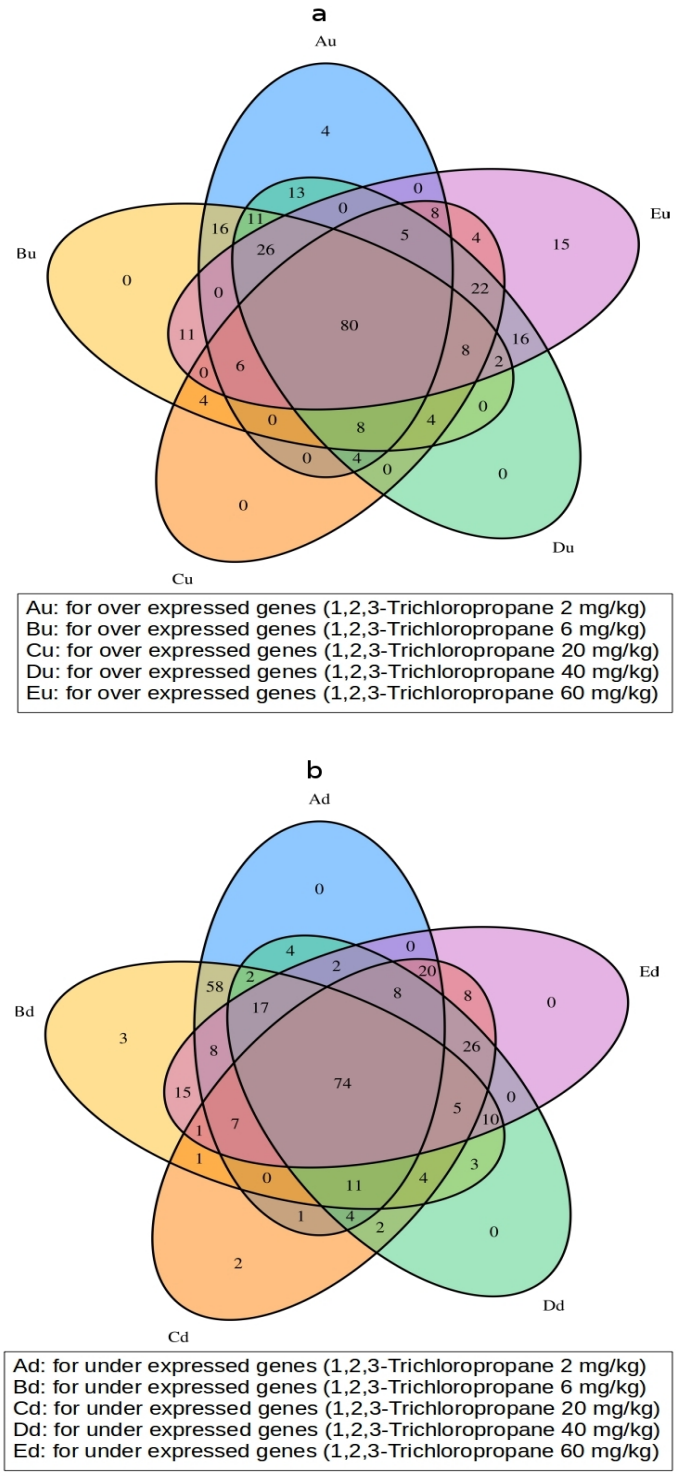

Figure 6.6: Venn diagram showing the overlap between the lists of pathways found to be perturbed for over- (a.) and under (b.)-expressed genes at different doses of 1,2,3-trichloropropane (TCPN) for dataset 3.

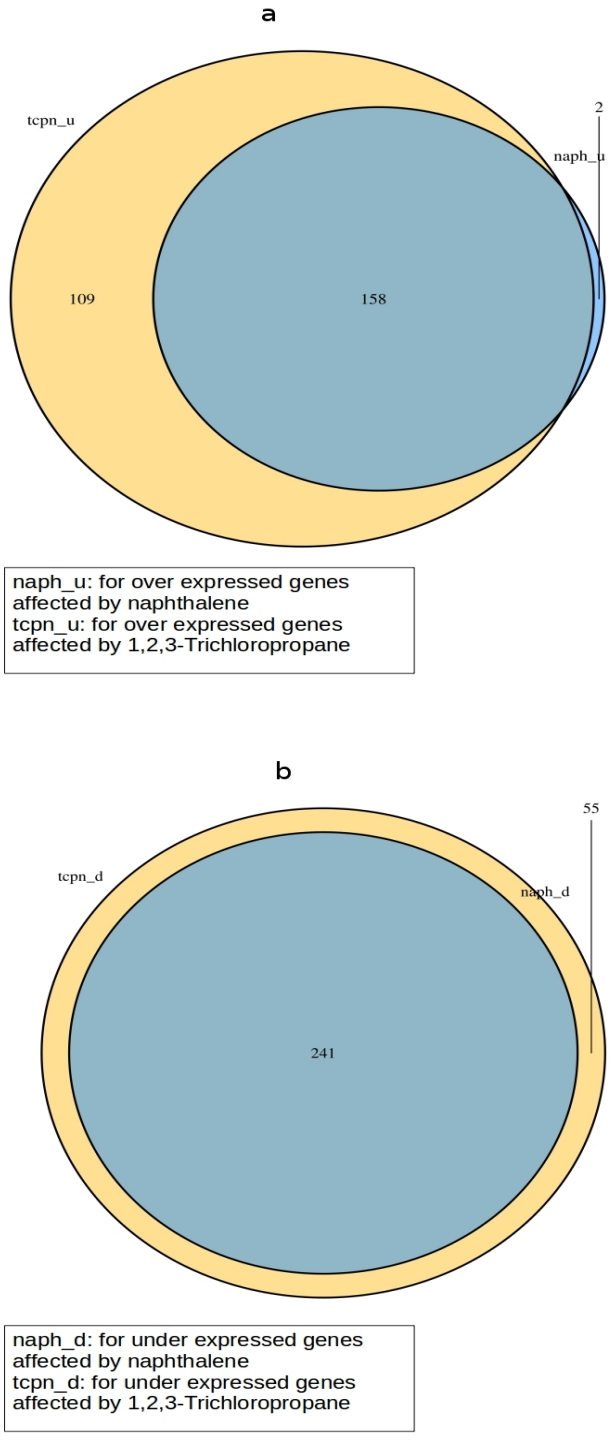

Figure 6.7: Venn diagram showing the overlap between the lists of pathways found to be perturbed for over- (a.) and under (b.)-expressed genes by naphthalene and 1,2,3-trichloropropane for dataset 3.

Tables 6.11-6.12 (see Appendix) provide a comprehensive list of pathways that are found to be perturbed by several chemical substances in case of all datasets. To further investigate the degree of similarity and diversity between the lists of pathways for every pair of datasets, we have computed the Jaccard similarity coefficient. From Table 6.2, we can observe a low Jaccard similarity coefficient in case of each pair of the datasets and this is not too surprising as the experimental backgrounds differ from each other either in terms of tissues or doses of the toxicants used in the studies.

Table 6.2: Jaccard similarity coefficient (JSC) between the pathways found to be significant for every pair of datasets.

Datasets	For over expressed genes	For under expressed genes
Dataset 1 - Dataset 2	0.2125	0.2889
Dataset 1 - Dataset 3	0.213	0.2357
Dataset 1 - Dataset 4	0.0305	0.0728
Dataset 2 - Dataset 3	0.2664	0.3191
Dataset 2 - Dataset 4	0.0374	0.0372
Dataset 3 - Dataset 4	0.0469	0.0598

6.4 Conclusion

This chapter emphasizes the application of EMOA-δ-TRIMAX triclustering algorithm in the analysis of 3D gene expression datasets which consist of expression profiles of thousands of genes during the exposure of specific tissues to several toxicants either in a time- or dose dependent manner in order to identify the signaling pathways triggered by the activation of different toxicants. Our results indicate the lists of signaling pathways that can either be activated or blocked by different toxicants in order to affect several tissues such as lung, liver etc. Most the identified pathways have already been found to be associated with diseases caused by toxicants by several studies over the last decade and hence, the others need to be experimentally verified. Thus altogether, this chapter may provide new insights into the mechanisms of inhalation toxicology.

6.5 Bibliography

[1] Nicotera, P.: Alteration of Cell Signalling in Chemical Toxicity. In: Seiler, J., Kroftov, O., Eybl, V. (eds.) Toxicology - From Cells to Man. Archives of Toxicology, vol. 18, pp. 3–11. Springer, Berlin Heidelberg (1996). doi:10.1007/978-3-642-61105-6_1.http://dx.doi.org/10.1007/978-3-642-61105-6_1

[2] Bhar, A., Haubrock, M., Mukhopadhyay, A., Wingender, E.: Multiobjective Triclustering Of Time-Series Transcriptome Data Reveals Key Genes Of Biological Processes. BMC Bioinformatics, (in press) (2015)

[3] Snyder, J.C., Zemke, A.C., Stripp, B.R.: Reparative capacity of airway epithelium impacts deposition and remodeling of extracellular matrix. American Journal of Respiratory Cell and Molecular Biology **40**(6), 633–642 (2009). doi:10.1165/rcmb.2008-0334OC

[4] Zemke, A.C., Snyder, J.C., Brockway, B.L., Drake, J.A., Reynolds, S.D., Kaminski, N., Stripp, B.R.: Molecular staging of epithelial maturation using secretory cell-specific genes as markers. American Journal of Respiratory Cell and Molecular Biology **40**(3), 340–358 (2009). doi:10.1165/rcmb.2007-0380OC

[5] Thomas, R.S., Bao, W., Chu, T.M., Bessarabova, M., Nikolskaya, T., Nikolsky, Y., Andersen, M.E., Wolfinger, R.D.: Use of short-term transcriptional profiles to assess the long-term cancer-related safety of environmental and industrial chemicals. Toxicological Sciences **112**(2), 311–321 (2009). doi:10.1093/toxsci/kfp233

[6] Thomas, R.S., Clewell, H.J.r., Allen, B.C., Wesselkamper, S.C., Wang, N.C., Lambert, J.C., Hess-Wilson, J.K., Zhao, Q.J., Andersen, M.E.: Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicological Sciences **120**(1), 194–205 (2011). doi:10.1093/toxsci/kfq355

[7] Gharib, S.A., Altemeier, W.A., VanWinkle, L.S., Plopper, C.G., Schlesinger, S.Y., Buell, C.A., Brauer, R., Lee, V., Parks, W.C., Chen, P.: Matrix metalloproteinase-7 coordinates airway epithelial injury response and differentiation of ciliated cells. American Journal of Respiratory Cell and Molecular Biology **48**(3), 390–396 (2013). doi:10.1165/rcmb.2012-0083OC

[8] Andrianifahanana, M., Wilkes, M.C., Repellin, C.E., Edens, M., Kottom, T.J., Rahimi, R.A., Leof, E.B.: ERBB receptor activation is required for profibrotic responses to transforming growth factor beta. Cancer Research **70**(19), 7421–7430 (2010). doi:10.1158/0008-5472.CAN-10-0232

[9] Suzuki, H., Aoshiba, K., Yokohori, N., Nagai, A.: Epidermal growth factor receptor tyrosine kinase inhibition augments a murine model of pulmonary fibrosis. Cancer Research **63**(16), 5054–5059 (2003)

[10] Kelsen, S.G., Duan, X., Ji, R., Perez, O., Liu, C., Merali, S.: Cigarette Smoke Induces an Unfolded Protein Response in the Human Lung. American Journal of Respiratory Cell and Molecular Biology **38**(5), 541–550 (2008). doi:10.1165/rcmb.2007-0221OC

[11] Bournazos, S., Grinfeld, J., Alexander, K.M., Murchison, J.T., Wallace, W.A., McFarlane, P., Hirani, N., Simpson, A.J., Dransfield, I., Hart, S.P.: Association of FcRIIa R131H polymorphism with idiopathic pulmonary fibrosis severity and progression. BMC Pulmonary Medicine **10**, 51 (2010). doi:10.1186/1471-2466-10-51

[12] Martino, M.B., Jones, L., Brighton, B., Ehre, C., Abdulah, L., Davis, C.W., Ron, D., O'Neal, W.K., Ribeiro, C.M.P.: The ER stress transducer IRE1β is required for airway epithelial mucin production. Mucosal Immunology **6**(3), 639–654 (2013). doi:10.1038/mi.2012.105

[13] Richeldi, L., Davies, H.R., Ferrara, G., Franco, F.: Corticosteroids for idiopathic pulmonary fibrosis. The Cochrane Database of Systematic Reviews **3**, 002880 (2003). doi:10.1002/14651858.CD002880

[14] Li, L.R., Sydenham, E., Chaudhary, B., You, C.: Glucocorticoid with cyclophosphamide for paraquat-induced lung fibrosis. The Cochrane Database of Systematic Reviews **7**, 008084 (2012). doi:10.1002/14651858.CD008084.pub3

[15] Saadane, A., Masters, S., DiDonato, J., Li, J., Berger, M.: Parthenolide inhibits IkappaB kinase, NF-kappaB activation, and inflammatory response in cystic fibrosis cells and mice. American Journal of Respiratory Cell and Molecular Biology **36**(6), 728–736 (2007)

[16] Kitowska, K., Zakrzewicz, D., Koenigshoff, M., Chrobak, I., Grimminger, F., Seeger, W., Bulau, P., Eickelberg, O.: Functional role and species-specific contribution of

arginases in pulmonary fibrosis. American Journal of Physiology. Lung Cellular and Molecular Physiology **294**(1), 34–35 (2008)

[17] Jankov, R.P., Belcastro, R., Ovcina, E., Lee, J., Massaeli, H., Lye, S.J., Tanswell, A.K.: Thromboxane A2 Receptors Mediate Pulmonary Hypertension in 60% Oxygenexposed Newborn Rats by a Cyclooxygenase-independent Mechanism. American Journal of Respiratory and Critical Care Medicine **166**(2), 208–214 (2002). doi:10.1164/rccm.200112-124OC

[18] Lakatos, H.F., Thatcher, T.H., Kottmann, R.M., Garcia, T.M., Phipps, R.P., Sime, P.J.: The Role of PPARs in Lung Fibrosis. PPAR Research **2007**, 71323 (2007). doi:10.1155/2007/71323

[19] Prakash, Y., Thompson, M.A., Meuchel, L., Pabelick, C.M., Mantilla, C.B., Zaidi, S., Martin, R.J.: Neurotrophins in lung health and disease. Expert Review of Respiratory Medicine **4**(3), 395–411 (2010). doi:10.1586/ers.10.29

[20] Milara, J., Navarro, R., Juan, G., Peir, T., Serrano, A., Ramn, M., Morcillo, E., Cortijo, J.: Sphingosine-1-phosphate is increased in patients with idiopathic pulmonary fibrosis and mediates epithelial to mesenchymal transition. Thorax **67**(2), 147–156 (2012). doi:10.1136/thoraxjnl-2011-200026

[21] Morrisey, E.E.: Wnt Signaling and Pulmonary Fibrosis. The American Journal of Pathology **162**(5), 1393–1397 (2003). doi:10.1016/S0002-9440(10)64271-X

[22] Eschenbrenner, J., Janssen, W., Kojonazarov, B., Murmann, K., Ghofrani, A., Weissmann, N., Grimminger, F., Seeger, W., Schermuly, R.T.: Role of JAK-STAT pathway in pulmonary fibrosis. Pneumologie **67**(05), 15 (2013). doi:10.1055/s-0033-1345044

[23] Kavian, N., Servettaz, A., Weill, B., Batteux, F.: New Insights into the Mechanism of Notch Signalling in Fibrosis. The Open Rheumatology Journal **6**, 96–102 (2012). doi:10.2174/1874312901206010096

[24] Enomoto, T., Usuki, J., Azuma, A., Nakagawa, T., Kudoh, S.: Diabetes mellitus may increase risk for idiopathic pulmonary fibrosis. Chest **123**(6), 2007–2011 (2003)

[25] Lyczak, J.B., Cannon, C.L., Pier, G.B.: Lung Infections Associated with Cystic Fibrosis. Clinical Microbiology Reviews **15**(2), 194–222 (2002). doi:10.1128/CMR.15.2.194-222.2002

[26] Gansner, J.M., Rosas, I.O., Ebert, B.L.: Pulmonary Fibrosis, Bone Marrow Failure, and Telomerase Mutation. New England Journal of Medicine **366**(16), 1551–1553 (2012). doi:10.1056/NEJMc1200999

[27] Study Links Pulmonary Fibrosis And Heart Disease. `www.sciencedaily.com/releases/2004/03/040310080526.htm` (2004)

[28] Chaisson, N.F., Paik, J., Orbai, A.M., Casciola-Rosen, L., Fiorentino, D., Danoff, S., Rosen, A.: A novel dermato-pulmonary syndrome associated with MDA-5 antibodies: report of 2 cases and review of the literature. Medicine **91**(4), 220–228 (2012). doi:10.1097/MD.0b013e3182606f0b

[29] Rafii, R., Juarez, M.M., Albertson, T.E., Chan, A.L.: A review of current and novel therapies for idiopathic pulmonary fibrosis. Journal of Thoracic Disease **5**(1) (2013)

[30] Ma, B., Zhou, P.Y., Ni, W., Wei, W., Ben, D.F., Lu, W., Xia, Z.F.: Inhibition of activin receptor-like kinase 5 induces matrix metallopeptidase 9 expression and aggravates lipopolysaccharide-induced pulmonary injury in mice. European Review for Medical and Pharmacological Sciences **17**(8), 1051–1059 (2013)

[31] Meier, M., Menne, J., Park, J.K., Holtz, M., Gueler, F., Kirsch, T., Schiffer, M., Mengel, M., Lindschau, C., Leitges, M., Haller, H.: Deletion of protein kinase C-epsilon signaling pathway induces glomerulosclerosis and tubulointerstitial fibrosis in vivo. Journal of the American Society of Nephrology **18**(4), 1190–1198 (2007)

[32] Borie, R., Fabre, A., Prost, F., Marchal-Somme, J., Lebtahi, R., Marchand-Adam, S., Aubier, M., Soler, P., Crestani, B.: Activation of somatostatin receptors attenuates pulmonary fibrosis. Thorax **63**(3), 251–258 (2008)

[33] Allen, J.T., Spiteri, M.: Growth factors in idiopathic pulmonary fibrosis: relative roles. Respiratory Research **3**, 13 (2002)

[34] Siegfried, J.M., Stabile, L.P.: Estrongenic steroid hormones in lung cancer. Seminars in Oncology **41**, 2014 (1). doi:10.1053/j.seminoncol.2013.12.009

[35] Siegfried, J.M., Hershberger, P.A., Stabile, L.P.: Estrogen receptor signaling in lung cancer. Seminars in Oncology **36**(6), 524–531 (2009). doi:10.1053/j.seminoncol.2009.10.004

[36] Yarden, Y., Sliwkowski, M.X.: Untangling the ErbB signalling network. Nature Reviews. Molecular Cell Biology **2**(2), 127–137 (2001)

[37] Liu, C.Y., Huang, Y.T., Liao, C.H., Chang, S.C., Hsueh, P.R.: Rapidly fatal bacteremia caused by Shigella sonnei without preceding gastrointestinal symptoms in an adult patient with lung cancer. Clinical Infectious Diseases **48**(11), 1635–1636 (2009). doi:10.1086/598994

[38] Assouad, J., Petkova, B., Berna, P., Dujon, A., Foucault, C., Riquet, M.: Renal cell carcinoma lung metastases surgery: pathologic findings and prognostic factors. The Annals of Thoracic Surgery **84**(4), 1114–1120 (2007)

[39] Hopwood, P., Stephens, R.J.: Depression in patients with lung cancer: prevalence and risk factors derived from quality-of-life data. Journal of Clinical Oncology **18**(4), 893–903 (2000)

[40] Ekman, S., Wynes, M.W., Hirsch, F.R.: The mTOR pathway in lung cancer and implications for therapy and biomarker analysis. Journal of Thoracic Oncology **7**(6), 947–953 (2012). doi:10.1097/JTO.0b013e31825581bd

[41] Dziadziuszko, R., Camidge, D.R., Hirsch, F.R.: The insulin-like growth factor pathway in lung cancer. Journal of Thoracic Oncology **3**(8), 815–818 (2008). doi:10.1097/JTO.0b013e31818180f5

[42] Melkamu, T., Qian, X., Upadhyaya, P., O'Sullivan, M.G., Kassie, F.: Lipopolysaccharide enhances mouse lung tumorigenesis: a model for inflammation-driven lung cancer. Veterinary Pathology **50**(5), 895–902 (2013). doi:10.1177/0300985813476061

[43] Limonta, P., Montagnani-Marelli, M., Mai, S., Motta, M., Martini, L., Moretti, R.M.: GnRH receptors in cancer: from cell biology to novel targeted therapeutic strategies. Endocrine Reviews **33**(5), 784–811 (2012). doi:10.1210/er.2012-1014

[44] Ece, F., F Hatabay, N., Erdal, N., Gedik, C., Guney, C., Aksoy, F.: Does Helicobacter pylori infection play a role in lung cancer? Respiratory Medicine **99**(10), 1258–1262 (2005)

[45] Wen, P., Kong, R., Liu, J., Zhu, L., Chen, X., Li, X., Nie, Y., Wu, K., Wu, J.Y.: USP33, a new player in lung cancer, mediates Slit-Robo signaling. Protein and Cell **5**(9), 704–713 (2014). doi:10.1007/s13238-014-0070-z

[46] Chen, S., Li, X., Lu, D., Xu, Y., Mou, W., Wang, L., Chen, Y., Liu, Y., Li, X., Li, L., Liu, L., Stupack, D., Reisfeld, R.A., Xiang, R., Li, N.: SOX2 regulates apoptosis through MAP4K4-survivin signaling pathway in human lung cancer cells. Carcinogenesis **35**(3), 613–623 (2014). doi:10.1093/carcin/bgt371

[47] He, S., Yin, T., Li, D., Gao, X., Wan, Y., Ma, X., Ye, T., Guo, F., Sun, J., Lin, Z., Wang, Y.: Enhanced interaction between natural killer cells and lung cancer cells: involvement in gefitinib-mediated immunoregulation. Journal of Translational Medicine **11**, 186 (2013). doi:10.1186/1479-5876-11-186

[48] Yamada, M., Doi, Y., Yorioka, H., Yonezawa, Y., Ozawa, T.: Dilated cardiomyopathy associated with tumor pulmonary microembolism of small cell lung cancer, an elderly case report. Nihon Ronen Igakkai Zasshi **26**(5), 521–522 (1989)

[49] Yang, H., Zhang, Q., He, J., Lu, W.: Regulation of calcium signaling in lung cancer. Journal of Thoracic Disease **2**(1), 52–56 (2010)

[50] Schuller, H.M., Al-Wadei, H.A., Majidi, M.: Gamma-aminobutyric acid, a potential tumor suppressor for small airway-derived lung adenocarcinoma. Carcinogenesis **29**(10), 1979–1985 (2008). doi:10.1093/carcin/bgn041

[51] Geary, C., Akinbi, H., Korfhagen, T., Fabre, J.E., Boucher, R., Rice, W.: Increased susceptibility of purinergic receptor-deficient mice to lung infection with Pseudomonas aeruginosa. American Journal of Physiology. Lung Cellular and Molecular Physiology **289**(5), 890–895 (2005)

[52] Crowley, M.R., Frost, A., Chen, D.T., Baffi, M.O., Nicola, T., Serra, R.: Transforming growth factor-beta signaling helps specify tumor type in DMBA and hormone-induced mammary cancers. Differentiation **74**(1), 40–52 (2006)

[53] Tachibana, K., Yamasaki, D., Ishimoto, K., Doi, T.: The Role of PPARs in Cancer. PPAR Research **2008**, 102737 (2008). doi:10.1155/2008/102737

[54] Arano, T., Nakagawa, H., Tateishi, R., Ikeda, H., Uchino, K., Enooku, K., Goto, E., Masuzaki, R., Asaoka, Y., Kondo, Y., Goto, T., Shiina, S., Omata, M., Yoshida, H., Koike, K.: Serum level of adiponectin and the risk of liver cancer development in chronic hepatitis C patients. International Journal of Cancer **129**(9), 2226–2235 (2011). doi:10.1002/ijc.25861

[55] Kaczynski, J., Hansson, G., Wallerstedt, S.: Increased porphyrins in primary liver cancer mainly reflect a parallel liver disease. Gastroenterology Research and Practice **2009**, 402394 (2009). doi:10.1155/2009/402394

[56] White, S.R., Martin, L.D., Abe, M.K., Marroquin, B.A., Stern, R., Fu, X.: Insulin receptor substrate-1/2 mediates IL-4-induced migration of human airway epithelial cells. American Journal of Physiology. Lung Cellular and Molecular Physiology **297**(1), 164–173 (2009). doi:10.1152/ajplung.90453.2008

[57] Chang, W., Chen, J., Schlueter, C.F., Hoyle, G.W.: Common pathways for activation of proinflammatory gene expression by G protein-coupled receptors in primary lung epithelial and endothelial cells. Experimental Lung Research **35**(4), 324–343 (2009). doi:10.1080/01902140802712738

[58] Mutlu, G.M., Factor, P.: Alveolar epithelial beta2-adrenergic receptors. American Journal of cell and Respiratory Cell and Molecular Biology **38**(2), 127–134 (2008)

[59] Zemke, A.C.: Roles for TGF-beta/BMP and beta-catenin signaling pathways in lung development and repair (2007)

[60] Foureau, D.M., Mielcarz, D.W., Menard, L.C., Schulthess, J., Werts, C., Vasseur, V., Ryffel, B., Kasper, L.H., Buzoni-Gatel, D.: TLR9-dependent induction of intestinal alpha-defensins by Toxoplasma gondii. Journal of Immunology **184**(12), 7022–7029 (2010). doi:10.4049/jimmunol.0901642

[61] Camelo, A., Dunmore, R., Sleeman, M.A., Clarke, D.L.: The epithelium in idiopathic pulmonary fibrosis: breaking the barrier. Frontiers of Pharmacology **4**, 173 (2014). doi:10.3389/fphar.2013.00173

[62] Picher, M., Boucher, R.C. (eds.): Purinergic Regulation of Respiratory Diseases vol. 55. Springer, Netherlands (2011). doi:10.1007/978-94-007-1217-1

[63] Lynch, C.C.: The expression and regulation of matrilysin (MMP-7) in human colon cancer and leukaemia cell lines. PhD thesis, School of Biotechnology, Dublin City University (2001)

[64] Huang, C.K., Lee, S.O., Lai, K.P., Ma, W.L., Lin, T.H., Tsai, M.Y., Luo, J., Chang, C.: Targeting androgen receptor in bone marrow mesenchymal stem cells leads to

better transplantation therapy efficacy in liver cirrhosis. Hepatology **57**(4), 1550–1563 (2013). doi:10.1002/hep.26135

[65] Ding, N., Yu, R.T., Subramaniam, N., Sherman, M.H., Wilson, C., Rao, R., Leblanc, M., Coulter, S., He, M., Scott, C., Lau, S.L., Atkins, A.R., Barish, G.D., Gunton, J.E., Liddle, C., Downes, M., Evans, R.M.: A vitamin D receptor/SMAD genomic circuit gates hepatic fibrotic response. Cell **153**(3), 601–613 (2013). doi:10.1016/j.cell.2013.03.028

[66] Spence, R.D., Wisdom, A.J., Cao, Y., Hill, H.M., Mongerson, C.R., Stapornkul, B., Itoh, N., Sofroniew, M.V., Voskuhl, R.R.: Estrogen mediates neuroprotection and anti-inflammatory effects during EAE through ER signaling on astrocytes but not through ER signaling on astrocytes or neurons. The Journal of Neuroscience **33**(26), 10924–10933 (2013). doi:10.1523/JNEUROSCI.0886-13.2013

[67] Wallace, P.M., MacMaster, J.F., Rouleau, K.A., Brown, T.J., Loy, J.K., Donaldson, K.L., Wahl, A.F.: Regulation of inflammatory responses by oncostatin M. Journal of Immunology **162**(9), 5547–5555 (1999)

[68] Marwaha, S.S., Kaur, J., Sodhi, G.S.: Structure Determination and Anti-Inflammatory Activity of some Purine Complexes. Metal-Based Drugs **2**(1), 13–17 (1995)

6.6 Appendix

Table 6.3: Enriched pathways for over expressed genes for dataset 1. "x" stands for "found to be significant". NPTH represents naphthalene.

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
apoptotic signaling pathway	x	x	x	x
regulation of apoptotic signaling pathway	x	x	x	x
intrinsic apoptotic signaling pathway in response to DNA damage	x	x	x	x
regulation of cytokine-mediated signaling pathway	x	x	x	x
adenosine receptor signaling pathway	x	x	x	
G-protein coupled purinergic receptor signaling pathway	x	x	x	
intrinsic apoptotic signaling pathway	x	x	x	x
positive regulation of I-kappaB kinase/NF-kappaB signaling	x	x	x	x
positive regulation of apoptotic signaling pathway	x	x	x	x
extrinsic apoptotic signaling pathway via death domain receptors	x	x	x	x
regulation of interferon-gamma-mediated signaling pathway	x	x	x	x
growth hormone receptor signaling pathway	x	x	x	x
intrinsic apoptotic signaling pathway by p53 class mediator	x	x	x	x
regulation of Fc receptor mediated stimulatory signaling pathway	x	x	x	x
ER-nucleus signaling pathway	x	x	x	x
JAK-STAT cascade involved in growth hormone signaling pathway	x	x	x	x
platelet-derived growth factor receptor-beta signaling pathway	x	x	x	
extrinsic apoptotic signaling pathway	x	x	x	x
regulation of intrinsic apoptotic signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway	x	x	x	
negative regulation of cytokine-mediated signaling pathway	x	x	x	x
positive regulation of metanephric mesenchymal cell migration by platelet-derived growth factor receptor-beta signaling pathway	x	x	x	
regulation of metanephric mesenchymal cell migration by platelet-derived growth factor receptor-beta signaling pathway	x	x	x	
positive regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	x	x	x	
activation of signaling protein activity involved in unfolded protein response	x	x	x	x
negative regulation of protein kinase B signaling	x	x	x	
regulation of I-kappaB kinase/NF-kappaB signaling	x	x	x	
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x	x	x	
I-kappaB kinase/NF-kappaB signaling	x	x	x	
adenylate cyclase-activating dopamine receptor signaling pathway	x	x	x	x
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x
androgen receptor signaling pathway	x	x	x	

Table 6.3 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
MyD88-independent toll-like receptor signaling pathway	x	x	x	
toll-like receptor 3 signaling pathway	x	x	x	
positive regulation of intrinsic apoptotic signaling pathway	x	x	x	
cytokine-mediated signaling pathway		x	x	x
interferon-gamma-mediated signaling pathway		x	x	x
type I interferon signaling pathway		x	x	x
Fc receptor mediated stimulatory signaling pathway	x	x	x	x
dopamine receptor signaling pathway	x	x	x	x
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis	x	x	x	x
Fc-gamma receptor signaling pathway involved in phagocytosis	x	x	x	x
positive regulation of extrinsic apoptotic signaling pathway in absence of ligand	x	x	x	x
positive regulation of cytokine-mediated signaling pathway		x	x	x
interleukin-4-mediated signaling pathway		x	x	x
MDA-5 signaling pathway	x	x	x	x
positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x	x	x	x
Fc-gamma receptor signaling pathway	x	x	x	x
toll-like receptor 4 signaling pathway	x	x	x	
toll-like receptor 2 signaling pathway	x	x	x	
regulation of BMP signaling pathway		x	x	
negative regulation of BMP signaling pathway		x	x	
BMP signaling pathway		x	x	
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway		x	x	
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway		x	x	
negative regulation of activin receptor signaling pathway		x	x	
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x	x	
transmembrane receptor protein serine/threonine kinase signaling pathway		x	x	
regulation of adenosine receptor signaling pathway		x	x	
negative regulation of adenosine receptor signaling pathway		x	x	
neurotrophin TRK receptor signaling pathway	x	x	x	
neurotrophin signaling pathway	x	x	x	
intrinsic apoptotic signaling pathway in response to oxidative stress	x	x	x	
thromboxane A2 signaling pathway		x	x	x
negative regulation of protein kinase C signaling			x	x
regulation of opioid receptor signaling pathway			x	x
sphingosine-1-phosphate signaling pathway			x	x
sphingolipid mediated signaling pathway			x	x
Wnt signaling pathway, calcium modulating pathway		x	x	
regulation of intrinsic apoptotic signaling pathway in response to DNA damage		x	x	
regulation of toll-like receptor 4 signaling pathway			x	x
positive regulation of toll-like receptor 4 signaling pathway			x	x
positive regulation of toll-like receptor signaling pathway			x	x

Table 6.3 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
immune response-regulating signaling pathway			x	x
regulation of toll-like receptor signaling pathway			x	x
regulation of tumor necrosis factor-mediated signaling pathway	x	x	x	
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	
toll-like receptor 5 signaling pathway	x	x	x	
toll-like receptor 10 signaling pathway	x	x	x	
insulin receptor signaling pathway	x	x	x	
toll-like receptor 9 signaling pathway	x	x	x	
negative regulation of tumor necrosis factor-mediated signaling pathway	x	x	x	
negative regulation of fibroblast growth factor receptor signaling pathway	x	x	x	
tumor necrosis factor-mediated signaling pathway	x	x	x	
regulation of extrinsic apoptotic signaling pathway	x	x	x	
toll-like receptor TLR1:TLR2 signaling pathway	x	x	x	
toll-like receptor TLR6:TLR2 signaling pathway	x	x	x	
negative regulation of apoptotic signaling pathway	x	x	x	
fibroblast growth factor receptor signaling pathway	x	x	x	
regulation of signaling	x	x	x	
negative regulation of signaling	x	x	x	
vascular endothelial growth factor receptor signaling pathway	x	x	x	
phospholipase C-activating dopamine receptor signaling pathway	x	x	x	
regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	x	x	x	
intracellular steroid hormone receptor signaling pathway	x	x	x	
transforming growth factor beta receptor signaling pathway	x	x	x	
intracellular receptor signaling pathway	x	x	x	
retinoic acid receptor signaling pathway	x	x	x	
positive regulation of Wnt signaling pathway, planar cell polarity pathway	x	x	x	
regulation of androgen receptor signaling pathway	x	x	x	
regulation of transforming growth factor beta receptor signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage		x	x	
activin receptor signaling pathway	x	x	x	
positive regulation of signaling	x	x	x	
negative regulation of phosphatidylinositol 3-kinase signaling	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x	x	x	
regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x	x	x	
glucocorticoid receptor signaling pathway	x	x	x	x
rhodopsin mediated signaling pathway	x	x	x	x
corticosteroid receptor signaling pathway	x	x	x	x
positive regulation of epidermal growth factor receptor signaling pathway	x	x	x	x
positive regulation of ERBB signaling pathway	x	x	x	x
canonical Wnt signaling pathway involved in mesenchymal stem cell differentiation	x	x	x	x
canonical Wnt signaling pathway involved in osteoblast differentiation	x	x	x	x

Table 6.3 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
Amoebiasis		x	x	x
Non-small cell lung cancer			x	x
PPAR signaling pathway			x	x
Chronic myeloid leukemia	x	x	x	x
Pancreatic cancer	x	x	x	x
Axon guidance			x	x
ECM-receptor interaction		x	x	x
Small cell lung cancer			x	x
Prostate cancer			x	x
Biotin metabolism			x	x
Pyrimidine metabolism		x	x	
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	x	x	x	x
DNA replication		x	x	
Toxoplasmosis		x	x	
Antigen processing and presentation		x	x	
Colorectal cancer			x	x
Amyotrophic lateral sclerosis (ALS)	x	x	x	
Prion diseases	x	x	x	
Lysosome		x	x	
Vibrio cholerae infection	x	x	x	x
Phagosome		x	x	
Chagas disease (American trypanosomiasis)	x	x	x	
MAPK signaling pathway	x	x	x	
Focal adhesion			x	x
Bladder cancer	x	x	x	
Melanoma	x	x	x	
Glycerolipid metabolism	x	x	x	
mTOR signaling pathway	x	x	x	
Glycerophospholipid metabolism	x	x	x	
Fat digestion and absorption	x	x	x	
Glyoxylate and dicarboxylate metabolism	x	x	x	x
Ether lipid metabolism	x	x	x	
Mucin type O-Glycan biosynthesis	x	x	x	x
Vasopressin-regulated water reabsorption	x	x	x	x
Cell cycle		x	x	
Oocyte meiosis		x	x	
p53 signaling pathway		x	x	
Progesterone-mediated oocyte maturation		x	x	
RNA transport	x	x	x	
Fructose and mannose metabolism	x	x	x	
Circadian rhythm - mammal	x	x	x	

Table 6.3 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
Huntington's disease	x	x	x	x
Parkinson's disease	x	x	x	x
Oxidative phosphorylation	x	x	x	x
Protein processing in endoplasmic reticulum	x	x	x	x
Alzheimer's disease	x	x	x	x
Metabolic pathways	x	x	x	x
Ribosome	x	x	x	x
Arginine and proline metabolism	x	x	x	
mRNA surveillance pathway	x	x	x	x
Base excision repair	x	x	x	
Ubiquitin mediated proteolysis	x	x	x	x
Proteasome	x	x	x	
Pathogenic Escherichia coli infection	x	x	x	
Insulin signaling pathway	x	x	x	x
Bacterial invasion of epithelial cells	x	x	x	
Shigellosis	x	x	x	
Fc gamma R-mediated phagocytosis	x	x	x	
Alanine, aspartate and glutamate metabolism	x	x	x	
RNA polymerase	x	x	x	
Neurotrophin signaling pathway	x	x	x	x
Apoptosis	x	x	x	x
Ribosome biogenesis in eukaryotes	x	x	x	x
Adipocytokine signaling pathway	x	x	x	
N-Glycan biosynthesis	x	x	x	x
Other types of O-glycan biosynthesis		x	x	x
Glycosphingolipid biosynthesis - globo series		x	x	x
Glycosphingolipid biosynthesis - lacto and neolacto series	x	x	x	x
Type II diabetes mellitus	x	x	x	x
Notch signaling pathway		x	x	x
Peroxisome	x	x	x	x
Thiamine metabolism		x	x	x
Jak-STAT signaling pathway		x	x	x
Acute myeloid leukemia	x	x	x	x
Amino sugar and nucleotide sugar metabolism		x	x	x
Renal cell carcinoma	x	x	x	
Glutathione metabolism	x	x	x	
Protein digestion and absorption		x	x	x
Glioma		x	x	
Galactose metabolism	x	x	x	
Citrate cycle (TCA cycle)	x	x	x	x

Table 6.3 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
hepatocyte growth factor receptor signaling pathway	x	x	x	
positive regulation of non-canonical Wnt signaling pathway	x	x	x	
thrombopoietin-mediated signaling pathway	x	x	x	x
negative regulation of dopamine receptor signaling pathway		x	x	
positive regulation of dopamine receptor signaling pathway		x	x	
adenylate cyclase-inhibiting dopamine receptor signaling pathway		x	x	
regulation of dopamine receptor signaling pathway		x	x	
desensitization of G-protein coupled receptor protein signaling pathway		x	x	
negative adaptation of signaling pathway		x	x	
adaptation of signaling pathway		x	x	
positive regulation of G-protein coupled receptor protein signaling pathway		x	x	
protein kinase C-activating G-protein coupled receptor signaling pathway	x	x	x	
positive regulation of extrinsic apoptotic signaling pathway	x	x	x	
extrinsic apoptotic signaling pathway in absence of ligand	x	x	x	
regulation of extrinsic apoptotic signaling pathway in absence of ligand	x	x	x	
positive regulation of type I interferon-mediated signaling pathway	x	x	x	
TRIF-dependent toll-like receptor signaling pathway	x	x	x	
regulation of type I interferon-mediated signaling pathway	x	x	x	x

Table 6.4: Enriched pathways for under expressed genes for dataset 1. "x" stands for "found to be significant". NPTH represents naphthalene.

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
glutamate receptor signaling pathway	x	x	x	
negative regulation of TOR signaling	x	x	x	
regulation of extrinsic apoptotic signaling pathway via death domain receptors	x	x	x	
extrinsic apoptotic signaling pathway via death domain receptors	x	x	x	
cell-cell signaling involved in cardiac conduction	x	x	x	
Fas signaling pathway	x	x	x	
glucocorticoid mediated signaling pathway	x	x	x	
regulation of glucocorticoid mediated signaling pathway	x	x	x	
regulation of Fas signaling pathway	x	x	x	
multicellular organismal signaling	x	x	x	
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	
regulation of cardiac muscle contraction by calcium ion signaling	x	x	x	
regulation of glutamate receptor signaling pathway	x	x	x	
negative regulation of transforming growth factor beta receptor signaling pathway	x	x	x	
enzyme linked receptor protein signaling pathway	x	x	x	
reelin-mediated signaling pathway	x	x	x	
semaphorin-plexin signaling pathway involved in neuron projection guidance	x	x	x	x
transmembrane receptor protein tyrosine phosphatase signaling pathway	x	x	x	
interferon-gamma-mediated signaling pathway	x	x	x	
interleukin-2-mediated signaling pathway	x	x	x	
calcineurin-NFAT signaling cascade	x	x	x	x
regulation of calcineurin-NFAT signaling cascade	x	x	x	x
type I interferon signaling pathway	x	x	x	
TRAM-dependent toll-like receptor signaling pathway		x	x	x
TRAM-dependent toll-like receptor 4 signaling pathway		x	x	x
semaphorin-plexin signaling pathway		x	x	x
positive regulation of transforming growth factor beta receptor signaling pathway		x	x	x
cytokine-mediated signaling pathway		x	x	
negative regulation of stress-activated protein kinase signaling cascade		x	x	
regulation of interferon-gamma-mediated signaling pathway		x	x	
regulation of type I interferon-mediated signaling pathway		x	x	
positive regulation of Wnt signaling pathway		x	x	
interleukin-15-mediated signaling pathway	x	x	x	
canonical Wnt signaling pathway involved in positive regulation of apoptotic process	x	x	x	x
interleukin-12-mediated signaling pathway	x	x	x	
nerve growth factor signaling pathway	x	x	x	x
G-protein coupled acetylcholine receptor signaling pathway	x	x	x	
smoothened signaling pathway involved in regulation of cerebellar granule cell precursor cell proliferation	x	x	x	
canonical Wnt signaling pathway involved in negative regulation of apoptotic process	x	x	x	x
regulation of lipopolysaccharide-mediated signaling pathway	x	x	x	x

Table 6.4 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
lipopolysaccharide-mediated signaling pathway			x	x
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis		x	x	
Fc-gamma receptor signaling pathway involved in phagocytosis		x	x	
Fc-gamma receptor signaling pathway		x	x	
Fc receptor mediated stimulatory signaling pathway		x	x	
immune response-activating cell surface receptor signaling pathway		x	x	
Fc receptor signaling pathway		x	x	
regulation of signaling		x	x	
immune response-regulating cell surface receptor signaling pathway		x	x	
cell surface receptor signaling pathway		x	x	
regulation of stress-activated protein kinase signaling cascade	x	x	x	
immune response-regulating signaling pathway		x	x	
fibroblast growth factor receptor signaling pathway		x	x	
signaling		x	x	
single organism signaling		x	x	
epidermal growth factor receptor signaling pathway		x	x	
stress-activated protein kinase signaling cascade		x	x	
ERBB signaling pathway		x	x	
negative regulation of insulin receptor signaling pathway		x	x	
regulation of insulin receptor signaling pathway		x	x	
neurotrophin TRK receptor signaling pathway		x	x	
Tie signaling pathway	x	x	x	
negative regulation of epidermal growth factor receptor signaling pathway	x	x	x	
negative regulation of ERBB signaling pathway	x	x	x	
regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x	x	
regulation of epidermal growth factor receptor signaling pathway	x	x	x	
regulation of ERBB signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x	x	x	
regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x	x	x	
phosphatidylinositol-mediated signaling	x	x	x	
inositol lipid-mediated signaling	x	x	x	
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x	x	
collagen-activated tyrosine kinase receptor signaling pathway	x	x	x	
collagen-activated signaling pathway	x	x	x	
leukemia inhibitory factor signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x	x	x	
phosphatidylinositol 3-kinase signaling	x	x	x	
cAMP-mediated signaling		x	x	
cyclic-nucleotide-mediated signaling		x	x	
vascular endothelial growth factor receptor signaling pathway		x	x	

Table 6.4 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
Ribosome biogenesis in eukaryotes	x	x	x	x
Circadian rhythm - mammal	x	x	x	
Phosphatidylinositol signaling system	x	x	x	
Ubiquitin mediated proteolysis	x	x	x	
Biotin metabolism	x	x	x	
Renal cell carcinoma	x	x	x	
Focal adhesion	x	x	x	
VEGF signaling pathway	x	x	x	
TGF-beta signaling pathway	x	x	x	
Vascular smooth muscle contraction	x	x	x	
Toxoplasmosis	x	x	x	
Axon guidance	x	x	x	
Leukocyte transendothelial migration	x	x	x	
Huntington's disease	x	x	x	
RNA transport	x	x	x	
Endocytosis	x	x	x	
ABC transporters		x	x	x
Gastric acid secretion	x	x	x	x
T cell receptor signaling pathway	x	x	x	
Basal transcription factors		x	x	
Leishmaniasis		x	x	
Pancreatic cancer		x	x	
Adherens junction	x	x	x	
Wnt signaling pathway	x	x	x	x
Basal cell carcinoma	x	x	x	
Viral myocarditis			x	x
Complement and coagulation cascades			x	x
Hematopoietic cell lineage			x	x
Long-term potentiation	x	x	x	
Fc gamma R-mediated phagocytosis		x	x	
Calcium signaling pathway		x	x	
Pancreatic secretion	x	x	x	
Oocyte meiosis	x	x	x	
Jak-STAT signaling pathway	x	x	x	
Glioma	x	x	x	
Melanoma	x	x	x	
Acute myeloid leukemia	x	x	x	
Prostate cancer	x	x	x	
Drug metabolism - cytochrome P450	x	x	x	
Bladder cancer	x	x	x	

Table 6.4 continued...

Pathways	NPTH Day1	NPTH Day2	NPTH Day3	NPTH Day6
negative regulation of signaling	x	x	x	
regulation of nucleotide-binding oligomerization domain containing signaling pathway	x	x	x	x
regulation of viral-induced cytoplasmic pattern recognition receptor signaling pathway	x	x	x	x
regulation of RIG-I signaling pathway	x	x	x	x
positive regulation of protein kinase A signaling	x	x	x	
BMP signaling pathway	x	x	x	
Chronic myeloid leukemia	x	x	x	
Staphylococcus aureus infection		x	x	
Non-small cell lung cancer	x	x	x	
Cell cycle	x	x	x	
Pathways in cancer	x	x	x	
Spliceosome	x	x	x	
Regulation of actin cytoskeleton	x	x	x	x
Bile secretion	x	x	x	x
Lysine biosynthesis	x	x	x	
Long-term depression	x	x	x	
Notch signaling pathway	x	x	x	x
Inositol phosphate metabolism	x	x	x	

Table 6.5: Enriched pathways for over expressed genes for dataset 2. "x" stands for "found to be significant". NPTH represents naphthalene (ppm).

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Lysine degradation	x				
Dorso-ventral axis formation	x				
Endocytosis	x	x	x	x	x
ECM-receptor interaction	x				
Leishmaniasis	x				
mRNA surveillance pathway	x				
Jak-STAT signaling pathway	x				
Pancreatic cancer	x	x			
Notch signaling pathway	x	x			
Small cell lung cancer	x	x			
RNA transport	x	x			
Viral myocarditis	x	x			
Fc epsilon RI signaling pathway	x	x			
Base excision repair	x	x			
Wnt signaling pathway	x	x			
Chronic myeloid leukemia	x	x			
Ubiquitin mediated proteolysis	x	x			
ErbB signaling pathway	x	x			
Acute myeloid leukemia	x	x			
Fc gamma R-mediated phagocytosis	x	x			
Shigellosis	x	x			
Prostate cancer	x	x			
Non-small cell lung cancer	x	x			
Renal cell carcinoma	x	x			
Axon guidance	x	x	x	x	x
Glioma	x	x			
Pathogenic Escherichia coli infection	x	x			
Pathways in cancer	x	x			
T cell receptor signaling pathway	x	x			
VEGF signaling pathway	x	x			

Table 6.5 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Type II diabetes mellitus	x	x	x	x	x
Epithelial cell signaling in Helicobacter pylori infection	x	x			
Protein processing in endoplasmic reticulum	x	x			
B cell receptor signaling pathway	x	x			
Tight junction	x	x			
Cell cycle	x	x			
Adherens junction	x	x			
GnRH signaling pathway	x	x			
Phosphatidylinositol signaling system	x	x			
Fatty acid elongation in mitochondria	x	x			
Neurotrophin signaling pathway	x	x			
Amino sugar and nucleotide sugar metabolism	x	x			
mTOR signaling pathway	x	x			
Oocyte meiosis	x	x			
Natural killer cell mediated cytotoxicity	x	x			
Insulin signaling pathway	x	x			
Long-term potentiation	x	x			
transforming growth factor beta receptor signaling pathway	x	x			
negative regulation of transforming growth factor beta receptor signaling pathway	x	x			
transmembrane receptor protein serine/threonine kinase signaling pathway	x	x			
enzyme linked receptor protein signaling pathway	x	x	x	x	x
regulation of transforming growth factor beta receptor signaling pathway	x	x			
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x			
positive regulation of intracellular steroid hormone receptor signaling pathway	x	x	x	x	x
positive regulation of intracellular estrogen receptor signaling pathway	x	x	x	x	x
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x			
signaling	x	x	x	x	x
single organism signaling	x	x	x	x	x
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	x	x
intracellular receptor signaling pathway	x	x			
thyroid hormone mediated signaling pathway	x		x	x	x

Table 6.5 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
mesenchymal-epithelial cell signaling involved in prostate gland development	x				
cell surface receptor signaling pathway	x	x	x	x	x
regulation of intracellular estrogen receptor signaling pathway	x	x			
lipopolysaccharide-mediated signaling pathway	x	x			
neurotrophin signaling pathway	x	x			
apoptotic signaling pathway	x	x			
regulation of G-protein coupled receptor protein signaling pathway	x	x			
neurotrophin TRK receptor signaling pathway	x	x			
regulation of rhodopsin mediated signaling pathway	x	x			
regulation of lipopolysaccharide-mediated signaling pathway	x	x			
intrinsic apoptotic signaling pathway in response to DNA damage	x	x			
rhodopsin mediated signaling pathway	x	x			
TOR signaling	x	x			
positive regulation of Notch signaling pathway	x	x			
adenylate cyclase-activating dopamine receptor signaling pathway	x	x			
regulation of epidermal growth factor receptor signaling pathway	x	x	x	x	x
negative regulation of G-protein coupled receptor protein signaling pathway	x	x			
cAMP-mediated signaling	x	x	x	x	x
regulation of ERBB signaling pathway	x	x	x	x	x
negative regulation of epidermal growth factor receptor signaling pathway	x	x	x	x	x
epidermal growth factor receptor signaling pathway	x	x	x	x	x
ERBB signaling pathway	x	x	x	x	x
negative regulation of ERBB signaling pathway	x	x	x	x	x
Fc receptor signaling pathway	x	x			
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis	x	x			
Fc-gamma receptor signaling pathway involved in phagocytosis	x	x			
immune response-regulating cell surface receptor signaling pathway	x	x			
Fc-gamma receptor signaling pathway	x	x			
Fc receptor mediated stimulatory signaling pathway	x	x			
regulation of signaling	x	x	x	x	x
Fc-epsilon receptor signaling pathway	x	x			

Table 6.5 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Roundabout signaling pathway	x	x			
negative regulation of signaling	x	x			
immune response-activating cell surface receptor signaling pathway	x	x			
nucleotide-binding oligomerization domain containing signaling pathway	x	x			
regulation of chemokine-mediated signaling pathway	x	x			
intracellular estrogen receptor signaling pathway	x	x			
toll-like receptor 5 signaling pathway	x	x			
toll-like receptor 10 signaling pathway	x	x			
MyD88-independent toll-like receptor signaling pathway	x	x			
stress-activated protein kinase signaling cascade	x	x			
fibroblast growth factor receptor signaling pathway	x	x			
regulation of stress-activated protein kinase signaling cascade	x	x			x
insulin receptor signaling pathway	x	x			
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis		x		x	x
Phenylalanine, tyrosine and tryptophan biosynthesis		x		x	x
Biosynthesis of unsaturated fatty acids		x		x	x
Phenylalanine metabolism		x		x	x
Maturity onset diabetes of the young		x		x	x
negative regulation of BMP signaling pathway		x			
regulation of BMP signaling pathway		x			
BMP signaling pathway		x			
negative regulation of retinoic acid receptor signaling pathway		x		x	x
regulation of retinoic acid receptor signaling pathway		x		x	x
activation of signaling protein activity involved in unfolded protein response		x			
ER-nucleus signaling pathway		x			
protein kinase B signaling		x			
Neuroactive ligand-receptor interaction			x	x	x
Alanine, aspartate and glutamate metabolism			x	x	x
Cell adhesion molecules (CAMs)			x	x	x
Other glycan degradation			x	x	x
cell-cell signaling			x	x	x

Table 6.5 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
gamma-aminobutyric acid signaling pathway			x	x	x
positive regulation of I-kappaB kinase/NF-kappaB signaling			x	x	x
cyclic-nucleotide-mediated signaling			x	x	x
regulation of I-kappaB kinase/NF-kappaB signaling			x	x	x
positive regulation of signaling			x	x	x
glutamate receptor signaling pathway			x	x	x
I-kappaB kinase/NF-kappaB signaling			x	x	x
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress			x	x	x
protein kinase C-activating G-protein coupled receptor signaling pathway			x	x	x
regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway			x	x	x
positive regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway			x	x	x
TRAM-dependent toll-like receptor signaling pathway			x	x	x
TRAM-dependent toll-like receptor 4 signaling pathway			x	x	x
glucocorticoid mediated signaling pathway			x	x	x
regulation of glucocorticoid mediated signaling pathway			x	x	x
regulation of cardiac muscle contraction by calcium ion signaling			x	x	x
platelet-derived growth factor receptor signaling pathway			x	x	x
retinoic acid receptor signaling pathway			x	x	x
phospholipase C-activating G-protein coupled glutamate receptor signaling pathway			x	x	x
thyroid-stimulating hormone signaling pathway			x	x	x
positive regulation of retinoic acid receptor signaling pathway			x	x	x
positive regulation of interferon-gamma-mediated signaling pathway			x	x	x
regulation of vitamin D receptor signaling pathway			x	x	x
Purine metabolism				x	x
Gastric acid secretion				x	x
Tyrosine metabolism				x	x
negative regulation of extrinsic apoptotic signaling pathway				x	x
mesodermal-endodermal cell signaling				x	x
somatostatin receptor signaling pathway				x	x
somatostatin signaling pathway				x	x
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation				x	x

Table 6.5 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
stromal-epithelial cell signaling involved in prostate gland development				x	x
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation				x	x
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation				x	x
regulation of insulin-like growth factor receptor signaling pathway				x	x
negative regulation of apoptotic signaling pathway				x	x
Metabolic pathways					x
Butirosin and neomycin biosynthesis					x
Sphingolipid metabolism					x
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification					x
canonical Wnt signaling pathway involved in regulation of cell proliferation					x
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification					x
Wnt signaling pathway involved in dorsal/ventral axis specification					x
Wnt signaling pathway involved in somitogenesis					x
negative regulation of non-canonical Wnt signaling pathway					x
insulin-like growth factor receptor signaling pathway					x
regulation of extrinsic apoptotic signaling pathway					x
positive regulation of insulin-like growth factor receptor signaling pathway					x
negative regulation of insulin-like growth factor receptor signaling pathway					x
positive regulation of non-canonical Wnt signaling pathway					x

Table 6.6: Enriched pathways for under expressed genes for dataset 2. "x" stands for "found to be significant". NPTH represents naphthalene (ppm).

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Steroid hormone biosynthesis	x				
Glycosphingolipid biosynthesis - globo series	x	x			
Wnt signaling pathway	x	x		x	x
Lysosome	x	x			
Chagas disease (American trypanosomiasis)	x	x			
ErbB signaling pathway	x	x	x	x	x
Bile secretion	x	x			
Other glycan degradation	x	x			
Gastric acid secretion	x	x			
Glycosphingolipid biosynthesis - lacto and neolacto series	x	x			
Renal cell carcinoma	x	x			
Calcium signaling pathway	x	x			
Phosphatidylinositol signaling system	x	x			
Adherens junction	x	x			
Pathways in cancer	x	x	x	x	x
Endometrial cancer	x	x	x	x	x
Axon guidance	x	x	x	x	x
Melanogenesis	x	x			
Dilated cardiomyopathy	x	x			x
Jak-STAT signaling pathway	x	x			
Long-term depression	x	x			x
Inositol phosphate metabolism	x	x			
Olfactory transduction	x	x			
Glioma	x	x	x	x	x
Acute myeloid leukemia	x	x			
Hypertrophic cardiomyopathy (HCM)	x	x	x	x	x

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Type II diabetes mellitus	x	x	x	x	x
Gap junction	x	x			
Toxoplasmosis	x	x			
Long-term potentiation	x	x			x
Fc gamma R-mediated phagocytosis	x	x			
Alanine, aspartate and glutamate metabolism	x	x			
Circadian rhythm - mammal	x	x			
mTOR signaling pathway	x	x			
Neuroactive ligand-receptor interaction	x	x			
Regulation of actin cytoskeleton	x	x			
MAPK signaling pathway	x	x			x
Melanoma	x	x	x	x	x
Type I diabetes mellitus	x	x			
Mucin type O-Glycan biosynthesis	x	x			
osmosensory signaling pathway	x				
somatostatin receptor signaling pathway	x				
somatostatin signaling pathway	x				
glucocorticoid mediated signaling pathway	x				
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation	x				
stromal-epithelial cell signaling involved in prostate gland development	x				
canonical Wnt signaling pathway involved in cardiac muscle cell fate commitment	x				
regulation of glucocorticoid mediated signaling pathway	x				
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation	x				
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation	x				
negative regulation of extrinsic apoptotic signaling pathway	x				
negative regulation of apoptotic signaling pathway	x				
canonical Wnt signaling pathway involved in neural crest cell differentiation	x				
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	x				
intracellular receptor signaling pathway	x	x			
signaling	x	x		x	x
single organism signaling	x	x		x	x

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
intracellular estrogen receptor signaling pathway	x	x			
positive regulation of retinoic acid receptor signaling pathway	x	x			
regulation of retinoic acid receptor signaling pathway	x	x			
intracellular steroid hormone receptor signaling pathway	x	x			
regulation of nucleotide-binding oligomerization domain containing signaling pathway	x	x			
cAMP-mediated signaling	x	x			
second-messenger-mediated signaling	x	x	x	x	x
cyclic-nucleotide-mediated signaling	x	x			
NIK/NF-kappaB signaling	x	x			
regulation of cardiac muscle contraction by calcium ion signaling	x	x	x	x	x
regulation of intracellular estrogen receptor signaling pathway	x	x			
transmembrane receptor protein tyrosine kinase signaling pathway	x	x		x	x
positive regulation of intracellular steroid hormone receptor signaling pathway	x	x			
positive regulation of intracellular estrogen receptor signaling pathway	x	x			
retinoic acid receptor signaling pathway	x	x			
regulation of signaling	x	x		x	x
negative regulation of signaling	x	x			
cell surface receptor signaling pathway	x	x		x	x
negative regulation of epidermal growth factor receptor signaling pathway	x	x			x
cell-cell signaling	x	x			
purinergic receptor signaling pathway	x	x			
negative regulation of ERBB signaling pathway	x	x			x
multicellular organismal signaling	x	x			
regulation of intracellular steroid hormone receptor signaling pathway	x	x			
enzyme linked receptor protein signaling pathway	x	x		x	x
phospholipase C-activating G-protein coupled glutamate receptor signaling pathway	x	x			
bile acid signaling pathway	x	x			
regulation of epidermal growth factor receptor signaling pathway	x	x			
regulation of ERBB signaling pathway	x	x			
extrinsic apoptotic signaling pathway in absence of ligand	x	x			
purinergic nucleotide receptor signaling pathway	x	x			

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
steroid hormone mediated signaling pathway	x	x			
Fc receptor signaling pathway	x	x			
gamma-aminobutyric acid signaling pathway	x	x			
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress	x	x			
negative regulation of stress-activated protein kinase signaling cascade	x	x			
positive regulation of TOR signaling	x	x			
neurotrophin signaling pathway	x	x			
negative regulation of intracellular steroid hormone receptor signaling pathway	x	x			
G-protein coupled glutamate receptor signaling pathway	x	x			
Fc-epsilon receptor signaling pathway	x	x			
stress-activated protein kinase signaling cascade	x	x			
glutamate receptor signaling pathway	x	x			
fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	x	x			
neurotrophin TRK receptor signaling pathway	x	x			
G-protein coupled receptor signaling pathway	x	x			
insulin-like growth factor receptor signaling pathway	x	x			
epidermal growth factor receptor signaling pathway	x	x	x	x	x
positive regulation of signaling	x	x			
regulation of stress-activated protein kinase signaling cascade	x	x		x	x
ERBB signaling pathway	x	x	x	x	x
Base excision repair		x			
RNA transport		x		x	x
Mismatch repair		x		x	x
Taste transduction		x		x	x
Staphylococcus aureus infection		x			
intrinsic apoptotic signaling pathway by p53 class mediator		x		x	x
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis		x	x	x	x
Fc-gamma receptor signaling pathway involved in phagocytosis		x	x	x	x
Fc-gamma receptor signaling pathway		x	x	x	x
Fc receptor mediated stimulatory signaling pathway		x	x	x	x
intrinsic apoptotic signaling pathway in response to DNA damage		x		x	x

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator		x		x	x
intrinsic apoptotic signaling pathway		x		x	x
platelet-derived growth factor receptor signaling pathway		x		x	x
immune response-activating cell surface receptor signaling pathway		x		x	x
adiponectin-activated signaling pathway		x		x	x
reelin-mediated signaling pathway		x		x	x
G-protein coupled acetylcholine receptor signaling pathway		x		x	x
cytokine-mediated signaling pathway		x		x	x
adenylate cyclase-activating dopamine receptor signaling pathway		x		x	x
regulation of G-protein coupled receptor protein signaling pathway		x		x	x
interleukin-6-mediated signaling pathway		x			
negative regulation of insulin-like growth factor receptor signaling pathway		x			
regulation of insulin-like growth factor receptor signaling pathway		x			
positive regulation of Notch signaling pathway		x			
JAK-STAT cascade involved in growth hormone signaling pathway		x			
growth hormone receptor signaling pathway		x			
regulation of Notch signaling pathway		x			
Non-small cell lung cancer			x	x	x
Small cell lung cancer			x	x	x
Prostate cancer			x	x	x
Chronic myeloid leukemia			x	x	x
Pancreatic cancer			x	x	x
Cell adhesion molecules (CAMs)			x	x	x
Citrate cycle (TCA cycle)			x	x	x
Neurotrophin signaling pathway			x	x	x
Proximal tubule bicarbonate reclamation			x	x	x
Apoptosis			x	x	x
Shigellosis			x	x	x
Cell cycle			x	x	x
Arrhythmogenic right ventricular cardiomyopathy (ARVC)			x	x	x
Alzheimer's disease			x	x	x

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Cyanoamino acid metabolism			x	x	x
Viral myocarditis			x	x	x
TRAM-dependent toll-like receptor signaling pathway			x	x	x
TRAM-dependent toll-like receptor 4 signaling pathway			x	x	x
regulation of smoothened signaling pathway involved in dorsal/ventral neural tube patterning			x	x	x
negative regulation of smoothened signaling pathway involved in dorsal/ventral neural tube patterning			x	x	x
positive regulation of I-kappaB kinase/NF-kappaB signaling			x	x	x
activation of signaling protein activity involved in unfolded protein response			x	x	x
Roundabout signaling pathway			x	x	x
regulation of adrenergic receptor signaling pathway			x	x	x
ER-nucleus signaling pathway			x	x	x
positive regulation of protein kinase A signaling			x	x	x
chemokine-mediated signaling pathway			x	x	x
calcium-mediated signaling			x	x	x
RNA polymerase				x	x
Ribosome biogenesis in eukaryotes				x	x
Glycerolipid metabolism				x	x
Notch signaling pathway				x	x
Proteasome				x	x
lipopolysaccharide-mediated signaling pathway				x	x
positive regulation of hippo signaling				x	x
Fas signaling pathway				x	x
collagen-activated tyrosine kinase receptor signaling pathway				x	x
collagen-activated signaling pathway				x	x
activation of prostate induction by androgen receptor signaling pathway				x	x
regulation of thrombin receptor signaling pathway				x	x
negative regulation of thrombin receptor signaling pathway				x	x
regulation of Fas signaling pathway				x	x
positive regulation of intrinsic apoptotic signaling pathway by p53 class mediator				x	x
negative regulation of smoothened signaling pathway				x	x
Oocyte meiosis					x

Table 6.6 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)
Ubiquitin mediated proteolysis					x
Pathogenic Escherichia coli infection					x
Metabolic pathways					x
Fatty acid elongation in mitochondria					x
transforming growth factor beta receptor signaling pathway					x
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway					x
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway					x
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway					x
hippo signaling					x
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway					x
regulation of hippo signaling					x
regulation of skeletal muscle contraction by calcium ion signaling					x
negative regulation of chemokine-mediated signaling pathway					x
positive regulation of lipopolysaccharide-mediated signaling pathway					x
fibroblast growth factor receptor signaling pathway					x

Table 6.7: Enriched pathways for over expressed genes for dataset 3. "x" stands for "found to be significant". TCPN and NPTH represent naphthalene (ppm) and 1,2,3-Trichloropropane (mg/kg), respectively.

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Lysosome	x	x	x	x	x	x	x	x	x	x
Terpenoid backbone biosynthesis	x	x	x	x	x	x		x		x
Glycosaminoglycan degradation	x	x	x	x	x	x	x	x		x
Bladder cancer	x			x	x	x	x	x	x	x
Proteasome	x			x	x	x	x			
Non-small cell lung cancer	x			x	x	x	x	x	x	x
Inositol phosphate metabolism	x			x	x	x	x		x	x
Pathways in cancer	x	x	x	x	x	x	x	x	x	x
RNA transport	x	x		x	x	x	x	x	x	x
Leishmaniasis	x			x	x	x	x		x	
Phosphatidylinositol signaling system	x			x	x	x	x		x	
Protein digestion and absorption	x	x	x	x	x	x	x	x	x	x
Small cell lung cancer	x	x	x	x	x	x	x	x	x	x
Pancreatic secretion	x			x	x	x	x		x	x
Toxoplasmosis	x					x	x	x	x	x
Jak-STAT signaling pathway	x					x	x	x	x	x
Ribosome biogenesis in eukaryotes	x	x	x	x	x	x	x	x	x	x
Protein processing in endoplasmic reticulum	x			x	x			x	x	x
mRNA surveillance pathway	x			x	x			x	x	x
Ether lipid metabolism	x	x	x	x	x	x	x	x	x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Complement and coagulation cascades	x	x	x	x	x	x	x	x	x	x
Notch signaling pathway	x	x	x	x	x	x	x	x	x	x
Neurotrophin signaling pathway	x	x	x	x		x	x	x	x	x
Alzheimer's disease	x	x	x	x			x			x
Wnt signaling pathway	x	x	x	x			x			x
Basal transcription factors	x	x	x	x	x			x		x
Nucleotide excision repair	x	x	x	x	x			x		x
adenylate cyclase-activating adrenergic receptor signaling pathway	x	x	x	x	x	x		x		x
adenylate cyclase-activating dopamine receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	x	x	x	x	x	x	x	x		x
positive regulation of cAMP-mediated signaling	x	x	x	x	x	x		x		x
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger	x	x	x	x	x	x		x	x	x
regulation of cAMP-mediated signaling	x	x	x	x	x	x		x		x
adrenergic receptor signaling pathway	x	x	x	x	x	x		x	x	x
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x			x	x	x	x	x	x	x
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x			x	x	x	x	x	x	x
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x			x	x	x	x	x	x	x
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x			x	x	x	x			
intrinsic apoptotic signaling pathway by p53 class mediator	x			x	x	x	x			
intrinsic apoptotic signaling pathway in response to DNA damage	x			x	x	x	x			
regulation of interleukin-6-mediated signaling pathway	x			x	x	x	x		x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
negative regulation of interleukin-6-mediated signaling pathway	x			x	x	x	x		x	x
regulation of interleukin-2-mediated signaling pathway	x			x	x	x	x		x	x
negative regulation of interleukin-2-mediated signaling pathway	x			x	x	x	x		x	x
regulation of prolactin signaling pathway	x			x	x	x	x		x	x
negative regulation of prolactin signaling pathway	x			x	x	x	x		x	x
regulation of interleukin-4-mediated signaling pathway	x			x	x	x	x		x	x
negative regulation of interleukin-4-mediated signaling pathway	x			x	x	x	x		x	x
regulation of macrophage colony-stimulating factor signaling pathway	x			x	x	x	x		x	x
negative regulation of macrophage colony-stimulating factor signaling pathway	x			x	x	x	x		x	x
prolactin signaling pathway	x			x	x	x	x		x	x
negative regulation of type I interferon-mediated signaling pathway	x			x	x	x	x		x	x
positive regulation of chemokine-mediated signaling pathway	x	x	x	x	x	x	x	x	x	x
regulation of platelet-derived growth factor receptor-beta signaling pathway	x			x	x	x	x		x	x
negative regulation of platelet-derived growth factor receptor-beta signaling pathway	x			x	x	x	x		x	x
transmembrane receptor protein tyrosine phosphatase signaling pathway	x	x		x	x	x	x	x	x	x
macrophage colony-stimulating factor signaling pathway	x			x	x	x	x		x	x
regulation of hepatocyte growth factor receptor signaling pathway	x			x	x	x	x		x	x
negative regulation of tumor necrosis factor-mediated signaling pathway	x			x	x	x	x		x	x
interleukin-4-mediated signaling pathway	x			x	x	x	x		x	x
interleukin-2-mediated signaling pathway	x			x	x	x	x	x	x	x
negative regulation of interferon-gamma-mediated signaling pathway	x			x	x	x	x		x	

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
interleukin-6-mediated signaling pathway	x			x	x	x	x		x	
atrial cardiac muscle cell to AV node cell signaling	x	x		x	x	x	x	x	x	x
negative regulation of platelet-derived growth factor receptor signaling pathway	x			x	x	x	x		x	
platelet-derived growth factor receptor-beta signaling pathway	x			x	x	x	x		x	
negative regulation of insulin-like growth factor receptor signaling pathway	x			x	x	x	x		x	
hepatocyte growth factor receptor signaling pathway	x			x	x	x	x		x	
regulation of chemokine-mediated signaling pathway	x	x	x	x	x	x	x	x	x	x
cytokine-mediated signaling pathway	x					x	x	x	x	x
negative regulation of TOR signaling	x	x	x	x		x	x	x	x	x
cell-cell signaling involved in cardiac conduction	x				x	x	x	x	x	x
positive regulation of vascular endothelial growth factor receptor signaling pathway	x					x	x		x	x
regulation of lipopolysaccharide-mediated signaling pathway	x					x	x	x	x	x
ionotropic glutamate receptor signaling pathway	x					x	x		x	x
positive regulation of signaling	x					x	x	x	x	x
positive regulation of cytokine-mediated signaling pathway	x					x	x		x	x
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress	x			x	x	x	x	x	x	x
negative regulation of transforming growth factor beta receptor signaling pathway	x			x	x			x	x	x
activation of signaling protein activity involved in unfolded protein response	x			x	x	x	x	x	x	x
SREBP signaling pathway	x			x	x	x	x	x	x	x
regulation of transforming growth factor beta receptor signaling pathway	x			x	x			x	x	x
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x			x	x			x	x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
ER-nucleus signaling pathway	x			x	x			x	x	x
enzyme linked receptor protein signaling pathway	x			x	x			x	x	x
cell surface receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
brain-derived neurotrophic factor receptor signaling pathway	x	x			x	x	x			
ciliary receptor clustering involved in smoothened signaling pathway	x	x			x	x	x	x	x	
smoothened signaling pathway involved in ventral spinal cord patterning	x	x			x	x	x	x	x	
smoothened signaling pathway involved in regulation of cerebellar granule cell precursor cell proliferation	x	x			x	x	x	x	x	
cerebellar Purkinje cell-granule cell precursor cell signaling involved in regulation of granule cell precursor cell proliferation	x	x			x	x	x	x	x	
epithelial-mesenchymal cell signaling	x	x			x	x	x	x	x	
calcium-mediated signaling using intracellular calcium source	x	x			x	x	x			
positive regulation of smoothened signaling pathway	x	x			x	x	x			
apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
negative regulation of epidermal growth factor receptor signaling pathway	x	x	x	x	x	x	x	x		x
negative regulation of ERBB signaling pathway	x	x	x	x	x	x	x	x		x
regulation of epidermal growth factor receptor signaling pathway	x	x	x	x	x	x	x	x		x
regulation of ERBB signaling pathway	x	x	x	x	x	x	x	x		x
extrinsic apoptotic signaling pathway via death domain receptors	x	x	x	x		x	x	x	x	x
T cell receptor signaling pathway	x	x	x	x		x	x	x	x	x
antigen receptor-mediated signaling pathway	x	x	x	x			x			x
negative regulation of apoptotic signaling pathway	x	x	x	x		x	x	x	x	x
immune response-activating cell surface receptor signaling pathway	x	x	x	x			x			x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
epidermal growth factor receptor signaling pathway	x	x	x	x			x			x
ERBB signaling pathway	x	x	x	x			x			x
extrinsic apoptotic signaling pathway	x	x	x	x			x	x	x	x
positive regulation of phosphatidylinositol 3-kinase signaling	x	x				x	x			
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	x	x			x		x
gamma-aminobutyric acid signaling pathway	x						x	x		
positive regulation of transforming growth factor beta receptor signaling pathway	x		x	x	x		x	x		
insulin-like growth factor receptor signaling pathway	x			x	x		x	x	x	x
adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway	x						x	x		
positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x						x	x		
intrinsic apoptotic signaling pathway	x	x	x	x		x	x	x	x	x
negative regulation of intrinsic apoptotic signaling pathway	x	x	x	x		x	x	x	x	x
Cell adhesion molecules (CAMs)		x		x	x	x	x	x	x	x
Phagosome		x		x	x	x	x	x	x	x
Pathogenic Escherichia coli infection		x		x	x	x	x	x	x	x
Olfactory transduction		x	x	x	x	x		x		x
RNA polymerase		x	x	x	x	x		x		x
Cytosolic DNA-sensing pathway		x	x	x	x	x		x		x
Long-term potentiation		x	x	x	x	x	x	x	x	x
Arrhythmogenic right ventricular cardiomyopathy (ARVC)		x	x	x		x	x	x	x	x
ECM-receptor interaction		x	x	x		x	x	x	x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Hypertrophic cardiomyopathy (HCM)		x	x	x		x	x	x	x	x
Dilated cardiomyopathy		x	x	x		x	x	x	x	x
adiponectin-activated signaling pathway		x		x	x	x	x	x	x	
interferon-gamma-mediated signaling pathway		x	x	x	x	x		x		x
regulation of intrinsic apoptotic signaling pathway		x	x					x	x	x
regulation of apoptotic signaling pathway		x				x		x	x	x
integrin-mediated signaling pathway		x	x	x	x	x	x	x	x	x
chemokine-mediated signaling pathway		x	x	x		x	x	x	x	x
negative regulation of signaling		x	x	x		x	x	x	x	x
transforming growth factor beta receptor signaling pathway		x		x		x	x			
phosphatidylinositol-mediated signaling		x		x		x	x	x	x	x
inositol lipid-mediated signaling		x		x		x	x	x	x	x
Tight junction			x	x		x	x	x	x	x
Porphyrin and chlorophyll metabolism			x	x	x	x	x	x	x	x
Other glycan degradation			x					x	x	x
Long-term depression			x				x	x	x	x
Vascular smooth muscle contraction			x				x	x	x	x
regulation of TOR signaling			x		x		x	x	x	x
TOR signaling			x		x		x	x	x	x
TRAM-dependent toll-like receptor signaling pathway			x	x				x	x	x
TRAM-dependent toll-like receptor 4 signaling pathway			x	x				x	x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
neurotrophin TRK receptor signaling pathway			x	x		x	x	x	x	x
neurotrophin signaling pathway			x	x		x	x	x	x	x
adenylate cyclase-modulating G-protein coupled receptor signaling pathway			x			x		x	x	x
Wnt signaling pathway involved in wound healing, spreading of epidermal cells				x	x			x		x
Intestinal immune network for IgA production					x				x	x
Leukocyte transendothelial migration					x	x	x			
activation of MAPK activity by adrenergic receptor signaling pathway					x				x	x
epidermal growth factor-activated receptor transactivation by G-protein coupled receptor signaling pathway					x				x	x
positive regulation of epidermal growth factor receptor signaling pathway					x				x	x
positive regulation of ERBB signaling pathway					x				x	x
regulation of intrinsic apoptotic signaling pathway in response to oxidative stress					x	x	x	x	x	x
intrinsic apoptotic signaling pathway in response to oxidative stress					x	x		x	x	x
cAMP-mediated signaling					x					
cyclic-nucleotide-mediated signaling					x					
VEGF signaling pathway						x	x	x	x	x
B cell receptor signaling pathway						x	x	x	x	x
Carbohydrate digestion and absorption						x	x	x	x	x
Endometrial cancer						x	x	x	x	x
mTOR signaling pathway						x	x	x	x	x
Acute myeloid leukemia						x	x	x	x	x
Colorectal cancer						x	x	x	x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Glioma						x	x	x	x	x
Chronic myeloid leukemia						x	x	x	x	x
Renal cell carcinoma						x	x	x	x	x
Pancreatic cancer						x	x	x	x	x
Fc epsilon RI signaling pathway						x	x	x	x	x
Apoptosis						x	x	x	x	x
ErbB signaling pathway						x	x	x	x	x
Melanoma						x	x	x	x	x
Progesterone-mediated oocyte maturation						x	x	x	x	x
Fc gamma R-mediated phagocytosis						x	x	x	x	x
Natural killer cell mediated cytotoxicity						x	x			
Prostate cancer						x	x	x	x	x
Toll-like receptor signaling pathway						x	x	x	x	x
Chagas disease (American trypanosomiasis)						x	x	x	x	x
Osteoclast differentiation						x	x	x	x	x
Insulin signaling pathway						x	x	x	x	x
Hepatitis C						x	x	x	x	x
Vitamin digestion and absorption						x	x			
Fatty acid metabolism						x	x	x	x	
PPAR signaling pathway						x		x	x	
Adipocytokine signaling pathway						x		x	x	

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Peroxisome						x		x	x	
Steroid biosynthesis						x	x			
Glycosphingolipid biosynthesis - lacto and neolacto series						x	x		x	x
Alanine, aspartate and glutamate metabolism						x	x		x	
Amino sugar and nucleotide sugar metabolism						x	x		x	
Glycosphingolipid biosynthesis - globo series						x				
Galactose metabolism						x				
Spliceosome						x	x		x	
Staphylococcus aureus infection						x	x			
Systemic lupus erythematosus						x	x			
negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress						x	x	x	x	x
Fc-epsilon receptor signaling pathway						x	x			
hormone-mediated signaling pathway						x		x	x	
negative regulation of extrinsic apoptotic signaling pathway						x				
regulation of extrinsic apoptotic signaling pathway						x				
G-protein coupled acetylcholine receptor signaling pathway						x	x		x	x
epiblast cell-extraembryonic ectoderm cell signaling involved in anterior/posterior axis specification						x			x	
transforming growth factor beta receptor signaling pathway involved in primitive streak formation						x			x	
positive regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry						x			x	
adenylate cyclase-activating G-protein coupled receptor signaling pathway						x			x	
lipopolysaccharide-mediated signaling pathway						x	x	x	x	

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
regulation of nodal signaling pathway involved in determination of left/right asymmetry						x			x	
regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry						x			x	
regulation of signaling						x	x		x	x
positive regulation of activin receptor signaling pathway						x			x	
nodal signaling pathway involved in determination of left/right asymmetry						x			x	
regulation of nodal signaling pathway						x			x	
nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry						x			x	
negative regulation of androgen receptor signaling pathway						x			x	
calcium-mediated signaling						x			x	
nodal signaling pathway						x			x	
Regulation of actin cytoskeleton							x	x	x	x
Biosynthesis of unsaturated fatty acids							x	x	x	
alpha-Linolenic acid metabolism							x	x	x	
Viral myocarditis							x	x	x	
Taste transduction							x			x
Phototransduction							x			x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
regulation of smoothened signaling pathway							x	x	x	x
smoothened signaling pathway							x	x	x	
glucocorticoid mediated signaling pathway							x		x	x
regulation of glucocorticoid mediated signaling pathway							x		x	x
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway							x			x
dopamine receptor signaling pathway							x			x
regulation of rhodopsin mediated signaling pathway							x			x
Steroid hormone biosynthesis								x	x	x
RIG-I-like receptor signaling pathway								x	x	x
Chemokine signaling pathway								x	x	x
Focal adhesion								x	x	x
MAPK signaling pathway								x	x	x
protein kinase B signaling								x	x	x
positive regulation of protein kinase B signaling								x	x	x
regulation of protein kinase B signaling								x	x	x
negative regulation of stress-activated protein kinase signaling cascade								x	x	x
G-protein coupled receptor signaling pathway								x	x	x
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand								x	x	x
Melanogenesis									x	x
Riboflavin metabolism									x	x
positive regulation of nucleotide-binding oligomerization domain containing signaling pathway									x	x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
regulation of nucleotide-binding oligomerization domain containing 1 signaling pathway									x	x
positive regulation of nucleotide-binding oligomerization domain containing 1 signaling pathway									x	x
regulation of nucleotide-binding oligomerization domain containing 2 signaling pathway									x	x
positive regulation of nucleotide-binding oligomerization domain containing 2 signaling pathway									x	x
regulation of growth hormone receptor signaling pathway									x	x
positive regulation of growth hormone receptor signaling pathway									x	x
regulation of nucleotide-binding oligomerization domain containing signaling pathway									x	x
nucleotide-binding oligomerization domain containing 1 signaling pathway									x	x
signaling										x
single organism signaling										x
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation										x
stromal-epithelial cell signaling involved in prostate gland development										x
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation										x
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation										x
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification										x
non-canonical Wnt signaling pathway										x
canonical Wnt signaling pathway involved in regulation of cell proliferation										x
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification										x
positive regulation of canonical Wnt signaling pathway										x

Table 6.7 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
negative regulation of non-canonical Wnt signaling pathway										x
androgen receptor signaling pathway										x
Wnt signaling pathway involved in dorsal/ventral axis specification										x
Wnt signaling pathway involved in somitogenesis										x

Table 6.8: Enriched pathways for under expressed genes for dataset 3. "x" stands for "found to be significant". TCPN and NPTH represent naphthalene (ppm) and 1,2,3-Trichloropropane (mg/kg), respectively.

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Alzheimer's disease	x	x	x	x	x	x	x	x	x	x
Protein processing in endoplasmic reticulum	x	x	x	x	x	x	x	x	x	x
Biosynthesis of unsaturated fatty acids	x	x	x	x	x	x	x	x	x	x
alpha-Linolenic acid metabolism	x	x	x	x	x	x	x	x	x	x
Taste transduction	x	x	x	x	x	x	x	x	x	x
RNA transport	x	x	x	x	x	x	x	x	x	x
Small cell lung cancer	x	x	x	x	x	x	x	x	x	x
Ribosome biogenesis in eukaryotes	x	x	x	x	x	x	x	x	x	x
Hematopoietic cell lineage	x			x	x	x	x		x	
Lysosome	x	x	x	x	x	x	x	x	x	x
RNA degradation	x	x	x	x	x	x	x	x	x	x
mRNA surveillance pathway	x	x	x	x	x	x		x	x	x
Tyrosine metabolism	x					x	x	x	x	x
Proteasome	x					x	x		x	x
Nucleotide excision repair	x		x	x	x	x	x	x	x	x
Regulation of actin cytoskeleton	x			x	x			x	x	x
Glycosaminoglycan degradation	x			x	x			x	x	x
Other glycan degradation	x			x	x	x	x	x	x	x
Renal cell carcinoma	x	x	x	x	x	x	x	x	x	
TGF-beta signaling pathway	x	x	x	x	x	x	x	x	x	x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
MAPK signaling pathway	x	x	x	x	x	x	x	x	x	x
Carbohydrate digestion and absorption	x	x	x	x	x	x	x	x	x	x
Endometrial cancer	x		x	x	x	x	x		x	x
Non-small cell lung cancer	x		x	x	x	x	x		x	x
mTOR signaling pathway	x		x	x	x	x	x		x	x
Acute myeloid leukemia	x		x	x	x	x	x		x	x
Colorectal cancer	x		x	x	x	x	x		x	x
Adipocytokine signaling pathway	x		x	x	x	x	x			
Long-term depression	x		x	x	x	x	x			
Glioma	x		x	x	x	x	x			
VEGF signaling pathway	x	x	x	x	x	x	x	x		x
Chronic myeloid leukemia	x		x	x	x	x	x			
B cell receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
Pancreatic cancer	x		x	x	x	x	x			
Fc epsilon RI signaling pathway	x		x	x	x	x	x			
Apoptosis	x		x	x	x	x	x			
ErbB signaling pathway	x		x	x	x	x	x			
Melanoma	x		x	x	x	x	x			
Progesterone-mediated oocyte maturation	x		x	x	x	x	x			
Fc gamma R-mediated phagocytosis	x		x	x	x	x	x			
Prostate cancer	x		x	x	x	x	x			

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Toll-like receptor signaling pathway	x		x	x	x	x	x			
Vascular smooth muscle contraction	x		x	x	x	x	x			
Chagas disease (American trypanosomiasis)	x		x	x	x	x	x			
Osteoclast differentiation	x		x	x	x	x	x			
Tight junction	x		x	x	x	x	x			
T cell receptor signaling pathway	x		x	x	x	x	x			
Toxoplasmosis	x		x	x	x	x	x			
Ether lipid metabolism	x	x	x	x	x	x	x			x
Basal cell carcinoma	x	x	x	x	x	x	x	x	x	x
Hedgehog signaling pathway	x	x	x	x	x	x	x	x	x	x
Pathways in cancer	x		x	x	x	x		x	x	x
DNA replication	x			x	x	x	x		x	x
Cardiac muscle contraction	x	x		x	x	x		x		
Oxidative phosphorylation	x	x		x	x	x	x	x	x	
Parkinson's disease	x	x		x	x	x		x	x	
Huntington's disease	x	x	x	x	x	x		x	x	x
Vibrio cholerae infection	x						x		x	
Epithelial cell signaling in Helicobacter pylori infection	x						x		x	
Rheumatoid arthritis	x						x		x	
Phagosome	x	x	x	x	x	x	x	x	x	x
Protein digestion and absorption	x	x	x	x	x	x	x	x		x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Glycosphingolipid biosynthesis - lacto and neolacto series	x	x	x	x	x	x	x	x	x	x
cGMP-mediated signaling	x	x	x	x	x	x	x	x	x	x
G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger	x	x	x	x	x	x	x	x	x	x
antigen receptor-mediated signaling pathway	x	x	x	x	x	x	x	x	x	x
ionotropic glutamate receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
smoothened signaling pathway	x	x	x	x	x	x	x	x	x	x
intracellular steroid hormone receptor signaling pathway	x	x	x	x	x	x		x		x
regulation of smoothened signaling pathway	x	x	x	x	x	x	x	x	x	x
androgen receptor signaling pathway	x	x	x	x	x	x		x		x
mesodermal-endodermal cell signaling	x	x	x	x	x	x		x		x
TRAM-dependent toll-like receptor signaling pathway	x	x	x	x	x	x	x	x		x
TRAM-dependent toll-like receptor 4 signaling pathway	x	x	x	x	x	x	x	x		x
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation	x	x	x	x	x	x		x		x
stromal-epithelial cell signaling involved in prostate gland development	x	x	x	x	x	x		x		x
semaphorin-plexin signaling pathway involved in bone trabecula morphogenesis	x	x	x	x	x	x		x		x
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation	x	x	x	x	x	x		x		x
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation	x	x	x	x	x	x		x		x
negative regulation of intracellular steroid hormone receptor signaling pathway	x	x	x	x	x	x		x		x
integrin-mediated signaling pathway	x	x	x	x	x	x	x	x		x
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x	x	x	x	x		x		x
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	x	x	x	x	x	x		x		x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
negative regulation of BMP signaling pathway	x	x	x	x	x	x		x		x
canonical Wnt signaling pathway involved in regulation of cell proliferation	x	x	x	x	x	x		x		x
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification	x	x	x	x	x	x		x		x
regulation of intracellular steroid hormone receptor signaling pathway	x	x	x	x	x	x		x		x
Wnt signaling pathway involved in dorsal/ventral axis specification	x	x	x	x	x	x		x		x
Wnt signaling pathway involved in somitogenesis	x	x	x	x	x	x		x		x
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress	x	x	x	x	x	x	x	x	x	x
semaphorin-plexin signaling pathway involved in outflow tract morphogenesis	x	x	x	x	x	x	x	x	x	x
positive regulation of semaphorin-plexin signaling pathway involved in outflow tract morphogenesis	x	x	x	x	x	x	x	x	x	x
regulation of semaphorin-plexin signaling pathway	x	x	x	x	x	x	x	x	x	x
positive regulation of semaphorin-plexin signaling pathway	x	x	x	x	x	x	x	x	x	x
cholecystokinin signaling pathway	x	x		x	x	x	x	x	x	
negative regulation of insulin-like growth factor receptor signaling pathway	x			x	x	x	x		x	
negative regulation of retinoic acid receptor signaling pathway	x		x	x	x	x	x	x	x	x
regulation of retinoic acid receptor signaling pathway	x		x	x	x	x	x	x	x	x
retinoic acid receptor signaling pathway	x		x	x	x	x	x	x	x	x
semaphorin-plexin signaling pathway	x	x	x	x	x	x		x	x	x
cell surface receptor signaling pathway involved in heart development	x	x	x	x	x	x		x	x	x
Notch signaling pathway	x	x	x		x	x	x	x	x	
positive regulation of chemokine-mediated signaling pathway	x	x			x	x	x			
regulation of signaling	x	x	x	x	x	x	x			

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
regulation of chemokine-mediated signaling pathway	x	x			x	x	x			
negative regulation of TOR signaling	x	x	x		x	x	x	x		x
negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x	x	x	x	x	x	x	x	x	x
regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x	x	x	x	x	x	x	x	x	x
negative regulation of stress-activated protein kinase signaling cascade	x	x	x	x	x	x	x	x	x	x
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
regulation of TOR signaling	x	x	x	x	x	x	x	x	x	x
G-protein coupled receptor signaling pathway	x		x	x	x	x	x			
intrinsic apoptotic signaling pathway in response to oxidative stress	x		x	x	x	x	x		x	x
insulin-like growth factor receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand	x		x	x	x	x	x		x	x
TOR signaling	x		x	x	x	x	x			
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x		x	x	x	x	x		x	x
negative regulation of intrinsic apoptotic signaling pathway	x		x	x	x	x	x	x	x	x
regulation of extrinsic apoptotic signaling pathway in absence of ligand	x		x	x	x	x	x			
extrinsic apoptotic signaling pathway in absence of ligand	x		x	x	x	x	x			
regulation of intrinsic apoptotic signaling pathway	x		x	x	x	x	x	x	x	
negative regulation of extrinsic apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
adenylate cyclase-modulating G-protein coupled receptor signaling pathway	x		x	x	x	x	x			

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
protein kinase B signaling	x	x	x	x	x	x	x	x	x	x
negative regulation of apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
regulation of stress-activated protein kinase signaling cascade	x		x	x	x	x	x			
Fc-epsilon receptor signaling pathway	x		x	x	x	x	x			
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger	x	x	x	x	x	x	x	x	x	x
regulation of extrinsic apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
fibroblast growth factor receptor signaling pathway	x		x	x	x	x	x			
phosphatidylinositol-mediated signaling	x	x	x	x	x	x	x	x	x	x
inositol lipid-mediated signaling	x	x	x	x	x	x	x	x	x	x
insulin receptor signaling pathway	x		x	x	x	x	x			
epidermal growth factor receptor signaling pathway	x		x	x	x	x	x			
stress-activated protein kinase signaling cascade	x		x	x	x	x	x			
Fc receptor signaling pathway	x		x	x	x	x	x			
ERBB signaling pathway	x		x	x	x	x	x			
extrinsic apoptotic signaling pathway	x	x	x	x	x	x	x	x	x	x
intrinsic apoptotic signaling pathway	x		x	x	x	x	x	x	x	
cell surface receptor signaling pathway	x		x	x	x	x	x	x		x
neurotrophin TRK receptor signaling pathway	x		x	x	x	x	x		x	x
neurotrophin signaling pathway	x		x	x	x	x	x			
mesenchymal smoothened signaling pathway involved in prostate gland development	x	x	x	x			x			x
epithelial-mesenchymal signaling involved in prostate gland development	x	x	x	x			x			x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
regulation of nodal signaling pathway involved in determination of left/right asymmetry	x	x	x	x			x			x
regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry	x	x	x	x			x			x
smoothened signaling pathway involved in regulation of cerebellar granule cell precursor cell proliferation	x	x	x	x	x	x	x		x	x
nodal signaling pathway involved in determination of left/right asymmetry	x	x	x	x			x			x
regulation of nodal signaling pathway	x	x	x	x			x			x
nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry	x	x	x	x			x			x
cerebellar Purkinje cell-granule cell precursor cell signaling involved in regulation of granule cell precursor cell proliferation	x	x	x	x	x	x	x		x	x
epithelial-mesenchymal cell signaling	x	x	x	x	x	x	x		x	x
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x	x	x			x			x
nodal signaling pathway	x	x	x	x			x			x
regulation of activin receptor signaling pathway	x	x	x	x			x			x
positive regulation of smoothened signaling pathway	x	x	x	x	x	x	x		x	x
transmembrane receptor protein serine/threonine kinase signaling pathway	x	x	x	x			x			x
activin receptor signaling pathway	x	x	x	x			x			x
ciliary receptor clustering involved in smoothened signaling pathway	x			x	x	x			x	x
smoothened signaling pathway involved in ventral spinal cord patterning	x			x	x	x			x	x
cAMP-mediated signaling	x	x	x	x	x	x	x			x
signaling	x		x				x	x	x	x
single organism signaling	x		x				x	x	x	x
regulation of thrombin receptor signaling pathway	x	x	x	x	x			x		x
negative regulation of thrombin receptor signaling pathway	x	x	x	x	x			x		x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
thrombin receptor signaling pathway	x	x	x	x	x			x		x
negative regulation of platelet-derived growth factor receptor signaling pathway	x	x	x	x	x			x		x
regulation of platelet-derived growth factor receptor signaling pathway	x	x	x	x	x			x		x
brain-derived neurotrophic factor receptor signaling pathway	x	x	x	x	x	x	x	x	x	x
calcium-mediated signaling using intracellular calcium source	x	x	x	x	x	x	x	x	x	x
transforming growth factor beta receptor signaling pathway	x			x		x	x		x	x
enzyme linked receptor protein signaling pathway	x			x		x	x		x	x
positive regulation of transforming growth factor beta receptor signaling pathway	x			x	x	x	x	x	x	x
chemokine-mediated signaling pathway	x	x	x	x	x	x		x	x	x
Long-term potentiation		x	x		x	x	x	x	x	
Adherens junction		x	x	x	x	x	x	x	x	x
Neurotrophin signaling pathway		x		x	x	x	x	x	x	x
Glycerophospholipid metabolism		x		x	x	x	x	x	x	
Natural killer cell mediated cytotoxicity		x	x	x	x	x		x		x
ECM-receptor interaction		x	x	x		x	x	x	x	x
Focal adhesion		x	x	x		x	x	x	x	x
regulation of lipopolysaccharide-mediated signaling pathway		x	x	x	x	x	x	x	x	x
regulation of fibroblast growth factor receptor signaling pathway		x	x	x	x	x	x	x	x	x
negative regulation of smoothened signaling pathway		x	x	x	x	x	x	x	x	x
lipopolysaccharide-mediated signaling pathway		x	x	x	x	x	x	x	x	x
positive regulation of insulin-like growth factor receptor signaling pathway		x	x		x	x	x	x	x	

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
apoptotic signaling pathway		x	x	x	x	x	x	x	x	x
regulation of insulin-like growth factor receptor signaling pathway		x	x		x	x	x	x	x	
regulation of apoptotic signaling pathway		x	x		x	x		x	x	x
glutamate receptor signaling pathway		x		x	x	x	x	x	x	
positive regulation of G-protein coupled receptor protein signaling pathway		x	x	x	x	x		x		x
gamma-aminobutyric acid signaling pathway		x	x	x	x	x		x		x
positive regulation of phosphatidylinositol 3-kinase signaling		x	x					x	x	x
regulation of phosphatidylinositol 3-kinase signaling		x						x	x	x
phosphatidylinositol 3-kinase signaling		x						x	x	x
calcium-mediated signaling		x						x	x	x
regulation of protein kinase B signaling		x	x					x	x	x
second-messenger-mediated signaling		x						x	x	x
mesenchymal-epithelial cell signaling involved in prostate gland development		x	x	x		x	x	x	x	x
negative regulation of epidermal growth factor receptor signaling pathway		x	x	x		x	x	x	x	x
negative regulation of ERBB signaling pathway		x	x	x		x	x	x	x	x
regulation of epidermal growth factor receptor signaling pathway		x	x	x		x	x	x	x	x
regulation of ERBB signaling pathway		x	x	x		x	x	x	x	x
steroid hormone mediated signaling pathway		x	x	x		x	x	x	x	x
mesenchymal-epithelial cell signaling		x	x	x		x	x	x	x	x
Fatty acid metabolism			x		x		x			
PPAR signaling pathway			x		x		x			

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Inositol phosphate metabolism			x		x		x	x		x
Complement and coagulation cascades			x		x			x		x
Phosphatidylinositol signaling system			x		x			x		x
Pancreatic secretion			x		x		x	x	x	x
RIG-I-like receptor signaling pathway			x	x	x			x		
Type II diabetes mellitus			x		x	x	x			x
Melanogenesis			x		x	x	x			
Salivary secretion			x				x		x	x
negative regulation of signaling			x					x		x
transmembrane receptor protein tyrosine phosphatase signaling pathway			x	x	x		x	x		
extrinsic apoptotic signaling pathway via death domain receptors			x	x	x			x		
regulation of glutamate receptor signaling pathway			x	x	x	x		x	x	x
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis			x		x	x	x			
Fc-gamma receptor signaling pathway involved in phagocytosis			x		x	x	x			
Fc-gamma receptor signaling pathway			x		x	x	x			
Fc receptor mediated stimulatory signaling pathway			x		x	x	x			
MyD88-independent toll-like receptor signaling pathway			x		x	x	x			
toll-like receptor 4 signaling pathway			x		x	x	x			
positive regulation of I-kappaB kinase/NF-kappaB signaling			x		x	x	x			
toll-like receptor signaling pathway			x		x	x	x			
pattern recognition receptor signaling pathway			x		x	x	x			

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
immune response-activating cell surface receptor signaling pathway			x		x	x	x			
regulation of I-kappaB kinase/NF-kappaB signaling			x		x	x	x			
I-kappaB kinase/NF-kappaB signaling			x		x	x	x			
negative regulation of canonical Wnt signaling pathway			x	x				x	x	x
negative regulation of Wnt signaling pathway			x	x				x	x	x
regulation of canonical Wnt signaling pathway			x	x				x	x	x
regulation of Wnt signaling pathway			x	x				x	x	x
Glycosaminoglycan biosynthesis - keratan sulfate					x	x	x	x	x	x
Glyoxylate and dicarboxylate metabolism					x	x	x	x	x	x
Porphyrin and chlorophyll metabolism					x	x	x	x	x	x
serotonin receptor signaling pathway					x	x		x	x	
Fatty acid elongation in mitochondria						x	x			
Spliceosome						x		x	x	
Ubiquitin mediated proteolysis						x		x	x	
RNA polymerase						x	x			x
Other types of O-glycan biosynthesis						x	x			x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
Metabolic pathways						x	x			x
N-Glycan biosynthesis						x	x			x
Cytosolic DNA-sensing pathway						x	x			x
Aminoacyl-tRNA biosynthesis						x	x			
Antigen processing and presentation						x	x			
Renin-angiotensin system						x			x	
African trypanosomiasis						x			x	
regulation of toll-like receptor 3 signaling pathway						x	x			
positive regulation of toll-like receptor 3 signaling pathway						x	x			
Toll signaling pathway						x	x			
regulation of toll-like receptor 4 signaling pathway						x	x			
positive regulation of toll-like receptor 4 signaling pathway						x	x			
positive regulation of protein kinase B signaling						x	x			
positive regulation of Notch signaling pathway						x			x	
regulation of interferon-gamma-mediated signaling pathway						x			x	
Aldosterone-regulated sodium reabsorption							x			x
Proximal tubule bicarbonate reclamation							x		x	x
Citrate cycle (TCA cycle)							x		x	x
Pyruvate metabolism							x		x	x
Calcium signaling pathway							x		x	x
Dilated cardiomyopathy							x		x	x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
transmembrane receptor protein tyrosine kinase signaling pathway							x			
activation of MAPK activity by adrenergic receptor signaling pathway							x	x	x	
epidermal growth factor-activated receptor transactivation by G-protein coupled receptor signaling pathway							x	x	x	
positive regulation of epidermal growth factor receptor signaling pathway							x	x	x	
positive regulation of ERBB signaling pathway							x	x	x	
adrenergic receptor signaling pathway							x	x	x	x
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator							x	x	x	x
adenylate cyclase-activating adrenergic receptor signaling pathway							x		x	x
adenylate cyclase-activating dopamine receptor signaling pathway							x		x	x
positive regulation of cAMP-mediated signaling							x		x	x
regulation of cAMP-mediated signaling							x		x	x
regulation of cardiac muscle contraction by calcium ion signaling							x			x
interferon-gamma-mediated signaling pathway							x			x
Phenylalanine metabolism								x	x	
One carbon pool by folate								x	x	x
Steroid hormone biosynthesis								x	x	x
intrinsic apoptotic signaling pathway by p53 class mediator								x	x	x
positive regulation of protein kinase A signaling								x	x	x
regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator								x	x	x
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator								x	x	x
positive regulation of lipopolysaccharide-mediated signaling pathway								x	x	x

Table 6.8 continued...

Pathways	NPTH (0.5)	NPTH (3)	NPTH (10)	NPTH (20)	NPTH (30)	TCPN (2)	TCPN (6)	TCPN (20)	TCPN (40)	TCPN (60
negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator								x	x	x
protein kinase A signaling								x	x	x
regulation of protein kinase A signaling								x	x	x
regulation of intrinsic apoptotic signaling pathway by p53 class mediator								x	x	x
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage								x	x	x
regulation of intrinsic apoptotic signaling pathway in response to DNA damage								x	x	x
intrinsic apoptotic signaling pathway in response to DNA damage								x	x	x
positive regulation of signaling								x	x	

Table 6.9: Enriched pathways for over expressed genes for dataset 4. "x" stands for "found to be significant", where "0" represents "not found to be significant". NPTH represents naphthalene.

Pathways	NPTH (MMP7 Null)	NPTH (Wild type)
adrenergic receptor signaling pathway involved in positive regulation of heart rate	x	x
extrinsic apoptotic signaling pathway	x	x
regulation of extrinsic apoptotic signaling pathway	x	x
cell surface receptor signaling pathway	x	x
adenylate cyclase-activating adrenergic receptor signaling pathway	x	x
adrenergic receptor signaling pathway involved in heart process	x	x
regulation of apoptotic signaling pathway	x	x
G-protein coupled receptor signaling pathway involved in heart process	x	x
apoptotic signaling pathway	x	x
positive regulation of insulin receptor signaling pathway	x	x
regulation of signaling	x	x
extrinsic apoptotic signaling pathway in absence of ligand	x	x
negative regulation of extrinsic apoptotic signaling pathway	x	x
T cell receptor signaling pathway	x	x
Bile secretion	x	x
Toxoplasmosis	x	x
Hepatitis C	x	x
Endocytosis	x	x
Vascular smooth muscle contraction	x	x
Steroid biosynthesis	x	x
TGF-beta signaling pathway	x	x
thyroid hormone mediated signaling pathway	x	
oncostatin-M-mediated signaling pathway	x	
positive regulation of interferon-gamma-mediated signaling pathway	x	
regulation of vitamin D receptor signaling pathway	x	
vitamin D receptor signaling pathway	x	
interferon-gamma-mediated signaling pathway	x	
androgen receptor signaling pathway	x	
positive regulation of intracellular steroid hormone receptor signaling pathway	x	
positive regulation of intracellular estrogen receptor signaling pathway	x	
cytokine-mediated signaling pathway	x	
Alanine, aspartate and glutamate metabolism	x	
Purine metabolism	x	

Table 6.10: Enriched pathways for under expressed genes for dataset 4. "x" stands for "found to be significant", whereas "0" represents "not found to be significant". NPTH represents naphthalene.

Pathways	NPTH (MMP7 Null)	NPTH (Wild type)
negative regulation of insulin receptor signaling pathway	x	x
regulation of insulin receptor signaling pathway	x	x
Notch signaling pathway	x	x
positive regulation of protein kinase A signaling	x	x
regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x
positive regulation of lipopolysaccharide-mediated signaling pathway	x	x
negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x	x
protein kinase A signaling	x	x
regulation of protein kinase A signaling	x	x
regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x	x
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x	x
regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x	x
regulation of lipopolysaccharide-mediated signaling pathway	x	x
adiponectin-activated signaling pathway	x	x
androgen receptor signaling pathway	x	x
intracellular steroid hormone receptor signaling pathway	x	x
insulin receptor signaling pathway	x	x
Adherens junction	x	x
Cell adhesion molecules (CAMs)	x	x
Protein processing in endoplasmic reticulum	x	x
Phenylalanine metabolism	x	x
Collecting duct acid secretion	x	x
fibroblast growth factor receptor signaling pathway involved in negative regulation of apoptotic process in bone marrow	x	
fibroblast growth factor receptor signaling pathway involved in hemopoiesis	x	
fibroblast growth factor receptor signaling pathway involved in positive regulation of cell proliferation in bone marrow	x	
fibroblast growth factor receptor signaling pathway involved in mammary gland specification	x	
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	x	
fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development	x	
granzyme-mediated apoptotic signaling pathway	x	
Toll signaling pathway	x	
N-Glycan biosynthesis	x	

Table 6.11: Summarization of the enriched pathways for over expressed genes for all data sets. "x" stands for "found to be significant".

Pathways	D1	D2	D3	D4
apoptotic signaling pathway	x	x	x	x
regulation of apoptotic signaling pathway	x		x	x
intrinsic apoptotic signaling pathway in response to DNA damage	x	x	x	
regulation of cytokine-mediated signaling pathway	x			
adenosine receptor signaling pathway	x			
G-protein coupled purinergic receptor signaling pathway	x			
intrinsic apoptotic signaling pathway	x		x	
positive regulation of I-kappaB kinase/NF-kappaB signaling	x	x		
positive regulation of apoptotic signaling pathway	x			
extrinsic apoptotic signaling pathway via death domain receptors	x		x	
regulation of interferon-gamma-mediated signaling pathway	x			
growth hormone receptor signaling pathway	x			
intrinsic apoptotic signaling pathway by p53 class mediator	x		x	
regulation of Fc receptor mediated stimulatory signaling pathway	x			
ER-nucleus signaling pathway	x	x	x	
JAK-STAT cascade involved in growth hormone signaling pathway	x			
platelet-derived growth factor receptor-beta signaling pathway	x		x	
extrinsic apoptotic signaling pathway	x		x	x
regulation of intrinsic apoptotic signaling pathway	x		x	
negative regulation of intrinsic apoptotic signaling pathway	x		x	
negative regulation of cytokine-mediated signaling pathway	x			
positive regulation of metanephric mesenchymal cell migration by platelet-derived growth factor receptor-beta signaling pathway	x			
regulation of metanephric mesenchymal cell migration by platelet-derived growth factor receptor-beta signaling pathway	x			
positive regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	x			
activation of signaling protein activity involved in unfolded protein response	x	x	x	
negative regulation of protein kinase B signaling	x			
regulation of I-kappaB kinase/NF-kappaB signaling	x	x		
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x		x	
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x		x	
I-kappaB kinase/NF-kappaB signaling	x	x		
adenylate cyclase-activating dopamine receptor signaling pathway	x	x	x	
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x		x	
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway	x		x	
androgen receptor signaling pathway	x		x	x
positive regulation of extrinsic apoptotic signaling pathway	x			

Table 6.11 continued...

Pathways	D1	D2	D3	D4
extrinsic apoptotic signaling pathway in absence of ligand	x			x
regulation of extrinsic apoptotic signaling pathway in absence of ligand	x			
positive regulation of type I interferon-mediated signaling pathway	x			
TRIF-dependent toll-like receptor signaling pathway	x			
regulation of type I interferon-mediated signaling pathway	x			
MyD88-independent toll-like receptor signaling pathway	x	x		
toll-like receptor 3 signaling pathway	x			
positive regulation of intrinsic apoptotic signaling pathway	x			
cytokine-mediated signaling pathway	x		x	x
interferon-gamma-mediated signaling pathway	x		x	x
type I interferon signaling pathway	x			
Fc receptor mediated stimulatory signaling pathway	x	x		
dopamine receptor signaling pathway	x		x	
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis	x	x		
Fc-gamma receptor signaling pathway involved in phagocytosis	x	x		
positive regulation of extrinsic apoptotic signaling pathway in absence of ligand	x			
positive regulation of cytokine-mediated signaling pathway	x		x	
interleukin-4-mediated signaling pathway	x		x	
MDA-5 signaling pathway	x			
positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway	x			
Fc-gamma receptor signaling pathway	x	x		
toll-like receptor 4 signaling pathway	x			
toll-like receptor 2 signaling pathway	x			
regulation of BMP signaling pathway	x	x		
negative regulation of BMP signaling pathway	x	x		
BMP signaling pathway	x	x		
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x		
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	x	x	x	
negative regulation of activin receptor signaling pathway	x			
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x			
transmembrane receptor protein serine/threonine kinase signaling pathway	x	x		
regulation of adenosine receptor signaling pathway	x			
negative regulation of adenosine receptor signaling pathway	x			
neurotrophin TRK receptor signaling pathway	x	x	x	
neurotrophin signaling pathway	x	x	x	
intrinsic apoptotic signaling pathway in response to oxidative stress	x		x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
thromboxane A2 signaling pathway	x			
negative regulation of protein kinase C signaling	x			
regulation of opioid receptor signaling pathway	x			
sphingosine-1-phosphate signaling pathway	x			
sphingolipid mediated signaling pathway	x			
Wnt signaling pathway, calcium modulating pathway	x			
regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x			
regulation of toll-like receptor 4 signaling pathway	x			
positive regulation of toll-like receptor 4 signaling pathway	x			
positive regulation of toll-like receptor signaling pathway	x			
immune response-regulating signaling pathway	x			
regulation of toll-like receptor signaling pathway	x			
regulation of tumor necrosis factor-mediated signaling pathway	x			
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	
toll-like receptor 5 signaling pathway	x	x		
toll-like receptor 10 signaling pathway	x	x		
insulin receptor signaling pathway	x	x		
toll-like receptor 9 signaling pathway	x			
negative regulation of tumor necrosis factor-mediated signaling pathway	x		x	
negative regulation of fibroblast growth factor receptor signaling pathway	x			
tumor necrosis factor-mediated signaling pathway	x			
regulation of extrinsic apoptotic signaling pathway	x	x	x	x
toll-like receptor TLR1:TLR2 signaling pathway	x			
toll-like receptor TLR6:TLR2 signaling pathway	x			
negative regulation of apoptotic signaling pathway	x	x	x	
fibroblast growth factor receptor signaling pathway	x	x		
regulation of signaling	x	x	x	x
negative regulation of signaling	x	x	x	
vascular endothelial growth factor receptor signaling pathway	x			
phospholipase C-activating dopamine receptor signaling pathway	x			
regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	x			
intracellular steroid hormone receptor signaling pathway	x			
transforming growth factor beta receptor signaling pathway	x	x	x	
intracellular receptor signaling pathway	x	x		
retinoic acid receptor signaling pathway	x	x		
positive regulation of Wnt signaling pathway, planar cell polarity pathway	x			

Table 6.11 continued...

Pathways	D1	D2	D3	D4
regulation of androgen receptor signaling pathway	x			
regulation of transforming growth factor beta receptor signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x			
activin receptor signaling pathway	x			
positive regulation of signaling	x	x	x	
negative regulation of phosphatidylinositol 3-kinase signaling	x			
negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x		x	
regulation of intrinsic apoptotic signaling pathway in response to oxidative stress	x		x	
glucocorticoid receptor signaling pathway	x			
rhodopsin mediated signaling pathway	x	x		
corticosteroid receptor signaling pathway	x			
positive regulation of epidermal growth factor receptor signaling pathway	x		x	
positive regulation of ERBB signaling pathway	x		x	
canonical Wnt signaling pathway involved in mesenchymal stem cell differentiation	x			
canonical Wnt signaling pathway involved in osteoblast differentiation	x			
hepatocyte growth factor receptor signaling pathway	x		x	
positive regulation of non-canonical Wnt signaling pathway	x	x		
thrombopoietin-mediated signaling pathway	x			
negative regulation of dopamine receptor signaling pathway	x			
positive regulation of dopamine receptor signaling pathway	x			
adenylate cyclase-inhibiting dopamine receptor signaling pathway	x			
regulation of dopamine receptor signaling pathway	x			
desensitization of G-protein coupled receptor protein signaling pathway	x			
negative adaptation of signaling pathway	x			
adaptation of signaling pathway	x			
positive regulation of G-protein coupled receptor protein signaling pathway	x			
protein kinase C-activating G-protein coupled receptor signaling pathway	x	x		
Huntington's disease	x			
Parkinson's disease	x			
Oxidative phosphorylation	x			
Protein processing in endoplasmic reticulum	x	x	x	
Alzheimer's disease	x		x	
Metabolic pathways	x	x		
Ribosome	x			
Arginine and proline metabolism	x			
mRNA surveillance pathway	x	x	x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
Base excision repair	x	x		
Ubiquitin mediated proteolysis	x	x		
Proteasome	x		x	
Pathogenic Escherichia coli infection	x	x	x	
Insulin signaling pathway	x	x	x	
Bacterial invasion of epithelial cells	x			
Shigellosis	x	x		
Fc gamma R-mediated phagocytosis	x	x	x	
Alanine, aspartate and glutamate metabolism	x	x	x	x
RNA polymerase	x		x	
Neurotrophin signaling pathway	x	x	x	
Apoptosis	x		x	
Ribosome biogenesis in eukaryotes	x		x	
Adipocytokine signaling pathway	x		x	
N-Glycan biosynthesis	x			
Other types of O-glycan biosynthesis	x			
Glycosphingolipid biosynthesis - globo series	x		x	
Glycosphingolipid biosynthesis - lacto and neolacto series	x		x	
Type II diabetes mellitus	x	x		
Notch signaling pathway	x	x	x	
Peroxisome	x		x	
Thiamine metabolism	x			
Jak-STAT signaling pathway	x	x	x	
Acute myeloid leukemia	x	x	x	
Amino sugar and nucleotide sugar metabolism	x	x	x	
Renal cell carcinoma	x	x	x	
Glutathione metabolism	x			
Protein digestion and absorption	x		x	
Glioma	x	x	x	
Galactose metabolism	x		x	
Citrate cycle (TCA cycle)	x			
Amoebiasis	x			
Non-small cell lung cancer	x	x	x	
PPAR signaling pathway	x		x	
Chronic myeloid leukemia	x	x	x	
Pancreatic cancer	x	x	x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
Axon guidance	x	x		
ECM-receptor interaction	x	x	x	
Small cell lung cancer	x	x	x	
Prostate cancer	x	x	x	
Biotin metabolism	x			
Pyrimidine metabolism	x			
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	x	x		
DNA replication	x			
Toxoplasmosis	x		x	x
Antigen processing and presentation	x			
Colorectal cancer	x		x	
Amyotrophic lateral sclerosis (ALS)	x			
Prion diseases	x			
Lysosome	x		x	
Vibrio cholerae infection	x			
Phagosome	x		x	
Chagas disease (American trypanosomiasis)	x		x	
MAPK signaling pathway	x		x	
Focal adhesion	x		x	
Bladder cancer	x		x	
Melanoma	x		x	
Glycerolipid metabolism	x			
mTOR signaling pathway	x	x	x	
Glycerophospholipid metabolism	x			
Fat digestion and absorption	x			
Glyoxylate and dicarboxylate metabolism	x			
Ether lipid metabolism	x		x	
Mucin type O-Glycan biosynthesis	x			
Vasopressin-regulated water reabsorption	x			
Cell cycle	x	x		
Oocyte meiosis	x	x		
p53 signaling pathway	x			
Progesterone-mediated oocyte maturation	x		x	
RNA transport	x	x	x	
Fructose and mannose metabolism	x			
Circadian rhythm - mammal	x			

Table 6.11 continued...

Pathways	D1	D2	D3	D4
Lysine degradation		x		
Dorso-ventral axis formation		x		
Endocytosis		x		x
Leishmaniasis		x	x	
Viral myocarditis		x	x	
Fc epsilon RI signaling pathway		x	x	
Wnt signaling pathway		x	x	
ErbB signaling pathway		x	x	
Pathways in cancer		x	x	
T cell receptor signaling pathway		x	x	x
VEGF signaling pathway		x	x	
Epithelial cell signaling in Helicobacter pylori infection		x		
B cell receptor signaling pathway		x	x	
Tight junction		x	x	
Adherens junction		x		
GnRH signaling pathway		x		
Phosphatidylinositol signaling system		x	x	
Fatty acid elongation in mitochondria		x		
Natural killer cell mediated cytotoxicity		x	x	
Long-term potentiation		x	x	
negative regulation of transforming growth factor beta receptor signaling pathway		x	x	
enzyme linked receptor protein signaling pathway		x	x	
positive regulation of intracellular steroid hormone receptor signaling pathway		x		x
positive regulation of intracellular estrogen receptor signaling pathway		x		x
signaling		x	x	
single organism signaling		x	x	
thyroid hormone mediated signaling pathway		x		x
mesenchymal-epithelial cell signaling involved in prostate gland development		x		
cell surface receptor signaling pathway		x	x	x
regulation of intracellular estrogen receptor signaling pathway		x		
lipopolysaccharide-mediated signaling pathway		x	x	
regulation of G-protein coupled receptor protein signaling pathway		x		
regulation of rhodopsin mediated signaling pathway		x	x	
regulation of lipopolysaccharide-mediated signaling pathway		x	x	
TOR signaling		x	x	
positive regulation of Notch signaling pathway		x		

Table 6.11 continued...

Pathways	D1	D2	D3	D4
regulation of epidermal growth factor receptor signaling pathway		x	x	
negative regulation of G-protein coupled receptor protein signaling pathway		x		
cAMP-mediated signaling		x	x	
regulation of ERBB signaling pathway		x	x	
negative regulation of epidermal growth factor receptor signaling pathway		x	x	
epidermal growth factor receptor signaling pathway		x	x	
ERBB signaling pathway		x	x	
negative regulation of ERBB signaling pathway		x	x	
Fc receptor signaling pathway		x		
immune response-regulating cell surface receptor signaling pathway		x		
Fc-epsilon receptor signaling pathway		x	x	
Roundabout signaling pathway		x		
immune response-activating cell surface receptor signaling pathway		x	x	
nucleotide-binding oligomerization domain containing signaling pathway		x		
regulation of chemokine-mediated signaling pathway		x	x	
intracellular estrogen receptor signaling pathway		x		
stress-activated protein kinase signaling cascade		x		
regulation of stress-activated protein kinase signaling cascade		x		
Phenylalanine, tyrosine and tryptophan biosynthesis		x		
Biosynthesis of unsaturated fatty acids		x	x	
Phenylalanine metabolism		x		
Maturity onset diabetes of the young		x		
negative regulation of retinoic acid receptor signaling pathway		x		
regulation of retinoic acid receptor signaling pathway		x		
protein kinase B signaling		x	x	
Neuroactive ligand-receptor interaction		x		
Cell adhesion molecules (CAMs)		x	x	
Other glycan degradation		x	x	
cell-cell signaling		x		
gamma-aminobutyric acid signaling pathway		x	x	
cyclic-nucleotide-mediated signaling		x	x	
glutamate receptor signaling pathway		x		
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress		x	x	
regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway		x		
positive regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway		x		
TRAM-dependent toll-like receptor signaling pathway		x	x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
TRAM-dependent toll-like receptor 4 signaling pathway		x	x	
glucocorticoid mediated signaling pathway		x	x	
regulation of glucocorticoid mediated signaling pathway		x	x	
regulation of cardiac muscle contraction by calcium ion signaling		x		
platelet-derived growth factor receptor signaling pathway		x		
phospholipase C-activating G-protein coupled glutamate receptor signaling pathway		x		
thyroid-stimulating hormone signaling pathway		x		
positive regulation of retinoic acid receptor signaling pathway		x		
positive regulation of interferon-gamma-mediated signaling pathway		x		x
regulation of vitamin D receptor signaling pathway		x		x
Purine metabolism		x		x
Gastric acid secretion		x		
Tyrosine metabolism		x		
negative regulation of extrinsic apoptotic signaling pathway		x	x	x
mesodermal-endodermal cell signaling		x		
somatostatin receptor signaling pathway		x		
somatostatin signaling pathway		x		
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation		x	x	
stromal-epithelial cell signaling involved in prostate gland development		x	x	
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation		x	x	
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation		x	x	
regulation of insulin-like growth factor receptor signaling pathway		x		
Butirosin and neomycin biosynthesis		x		
Sphingolipid metabolism		x		
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification		x	x	
canonical Wnt signaling pathway involved in regulation of cell proliferation		x	x	
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification		x	x	
Wnt signaling pathway involved in dorsal/ventral axis specification		x	x	
Wnt signaling pathway involved in somitogenesis		x	x	
negative regulation of non-canonical Wnt signaling pathway		x	x	
insulin-like growth factor receptor signaling pathway		x	x	
positive regulation of insulin-like growth factor receptor signaling pathway		x		
negative regulation of insulin-like growth factor receptor signaling pathway		x	x	
Terpenoid backbone biosynthesis			x	
Glycosaminoglycan degradation			x	
Inositol phosphate metabolism			x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
Pancreatic secretion			x	
Complement and coagulation cascades			x	
Basal transcription factors			x	
Nucleotide excision repair			x	
adenylate cyclase-activating adrenergic receptor signaling pathway			x	x
G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger			x	
positive regulation of cAMP-mediated signaling			x	
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger			x	
regulation of cAMP-mediated signaling			x	
adrenergic receptor signaling pathway			x	
regulation of interleukin-6-mediated signaling pathway			x	
negative regulation of interleukin-6-mediated signaling pathway			x	
regulation of interleukin-2-mediated signaling pathway			x	
negative regulation of interleukin-2-mediated signaling pathway			x	
regulation of prolactin signaling pathway			x	
negative regulation of prolactin signaling pathway			x	
regulation of interleukin-4-mediated signaling pathway			x	
negative regulation of interleukin-4-mediated signaling pathway			x	
regulation of macrophage colony-stimulating factor signaling pathway			x	
negative regulation of macrophage colony-stimulating factor signaling pathway			x	
prolactin signaling pathway			x	
negative regulation of type I interferon-mediated signaling pathway			x	
positive regulation of chemokine-mediated signaling pathway			x	
regulation of platelet-derived growth factor receptor-beta signaling pathway			x	
negative regulation of platelet-derived growth factor receptor-beta signaling pathway			x	
transmembrane receptor protein tyrosine phosphatase signaling pathway			x	
macrophage colony-stimulating factor signaling pathway			x	
regulation of hepatocyte growth factor receptor signaling pathway			x	
interleukin-2-mediated signaling pathway			x	
negative regulation of interferon-gamma-mediated signaling pathway			x	
interleukin-6-mediated signaling pathway			x	
atrial cardiac muscle cell to AV node cell signaling			x	
negative regulation of platelet-derived growth factor receptor signaling pathway			x	
negative regulation of TOR signaling			x	
cell-cell signaling involved in cardiac conduction			x	
positive regulation of vascular endothelial growth factor receptor signaling pathway			x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
ionotropic glutamate receptor signaling pathway			x	
SREBP signaling pathway			x	
brain-derived neurotrophic factor receptor signaling pathway			x	
ciliary receptor clustering involved in smoothened signaling pathway			x	
smoothened signaling pathway involved in ventral spinal cord patterning			x	
smoothened signaling pathway involved in regulation of cerebellar granule cell precursor cell proliferation			x	
cerebellar Purkinje cell-granule cell precursor cell signaling involved in regulation of granule cell precursor cell proliferation			x	
epithelial-mesenchymal cell signaling			x	
calcium-mediated signaling using intracellular calcium source			x	
positive regulation of smoothened signaling pathway			x	
antigen receptor-mediated signaling pathway			x	
positive regulation of phosphatidylinositol 3-kinase signaling			x	
positive regulation of transforming growth factor beta receptor signaling pathway			x	
adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway			x	
positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway			x	
Olfactory transduction			x	
Cytosolic DNA-sensing pathway			x	
Arrhythmogenic right ventricular cardiomyopathy (ARVC)			x	
Hypertrophic cardiomyopathy (HCM)			x	
Dilated cardiomyopathy			x	
adiponectin-activated signaling pathway			x	
integrin-mediated signaling pathway			x	
chemokine-mediated signaling pathway			x	
phosphatidylinositol-mediated signaling			x	
inositol lipid-mediated signaling			x	
Porphyrin and chlorophyll metabolism			x	
Long-term depression			x	
Vascular smooth muscle contraction			x	x
regulation of TOR signaling			x	
adenylate cyclase-modulating G-protein coupled receptor signaling pathway			x	
Wnt signaling pathway involved in wound healing, spreading of epidermal cells			x	
Intestinal immune network for IgA production			x	
Leukocyte transendothelial migration			x	
activation of MAPK activity by adrenergic receptor signaling pathway			x	
epidermal growth factor-activated receptor transactivation by G-protein coupled receptor signaling pathway			x	
Carbohydrate digestion and absorption			x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
Endometrial cancer			x	
Toll-like receptor signaling pathway			x	
Osteoclast differentiation			x	
Hepatitis C			x	x
Vitamin digestion and absorption			x	
Fatty acid metabolism			x	
Steroid biosynthesis			x	x
Spliceosome			x	
Staphylococcus aureus infection			x	
Systemic lupus erythematosus			x	
hormone-mediated signaling pathway			x	
G-protein coupled acetylcholine receptor signaling pathway			x	
epiblast cell-extraembryonic ectoderm cell signaling involved in anterior/posterior axis specification			x	
transforming growth factor beta receptor signaling pathway involved in primitive streak formation			x	
positive regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry			x	
adenylate cyclase-activating G-protein coupled receptor signaling pathway			x	
regulation of nodal signaling pathway involved in determination of left/right asymmetry			x	
regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry			x	
positive regulation of activin receptor signaling pathway			x	
nodal signaling pathway involved in determination of left/right asymmetry			x	
regulation of nodal signaling pathway			x	
nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry			x	
negative regulation of androgen receptor signaling pathway			x	
calcium-mediated signaling			x	
nodal signaling pathway			x	
Regulation of actin cytoskeleton			x	
alpha-Linolenic acid metabolism			x	
Taste transduction			x	
Phototransduction			x	
regulation of smoothened signaling pathway			x	
smoothened signaling pathway			x	
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway			x	
Steroid hormone biosynthesis			x	
RIG-I-like receptor signaling pathway			x	
Chemokine signaling pathway			x	
positive regulation of protein kinase B signaling			x	

Table 6.11 continued...

Pathways	D1	D2	D3	D4
regulation of protein kinase B signaling			x	
negative regulation of stress-activated protein kinase signaling cascade			x	
G-protein coupled receptor signaling pathway			x	
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand			x	
Melanogenesis			x	
Riboflavin metabolism			x	
positive regulation of nucleotide-binding oligomerization domain containing signaling pathway			x	
regulation of nucleotide-binding oligomerization domain containing 1 signaling pathway			x	
positive regulation of nucleotide-binding oligomerization domain containing 1 signaling pathway			x	
regulation of nucleotide-binding oligomerization domain containing 2 signaling pathway			x	
positive regulation of nucleotide-binding oligomerization domain containing 2 signaling pathway			x	
regulation of growth hormone receptor signaling pathway			x	
positive regulation of growth hormone receptor signaling pathway			x	
regulation of nucleotide-binding oligomerization domain containing signaling pathway			x	
nucleotide-binding oligomerization domain containing 1 signaling pathway			x	
non-canonical Wnt signaling pathway			x	
positive regulation of canonical Wnt signaling pathway			x	
adrenergic receptor signaling pathway involved in positive regulation of heart rate				x
adrenergic receptor signaling pathway involved in heart process				x
G-protein coupled receptor signaling pathway involved in heart process				x
positive regulation of insulin receptor signaling pathway				x
Bile secretion				x
TGF-beta signaling pathway				x
oncostatin-M-mediated signaling pathway				x
vitamin D receptor signaling pathway				x

Table 6.12: Summarization of the enriched pathways for under expressed genes for all data sets. "x" stands for "found to be significant".

Pathways	D1	D2	D3	D4
glutamate receptor signaling pathway	x	x	x	
negative regulation of TOR signaling	x		x	
regulation of extrinsic apoptotic signaling pathway via death domain receptors	x			
extrinsic apoptotic signaling pathway via death domain receptors	x		x	
cell-cell signaling involved in cardiac conduction	x			
Fas signaling pathway	x	x		
glucocorticoid mediated signaling pathway	x	x		
regulation of glucocorticoid mediated signaling pathway	x	x		
regulation of Fas signaling pathway	x	x		
multicellular organismal signaling	x	x		
transmembrane receptor protein tyrosine kinase signaling pathway	x	x	x	
regulation of cardiac muscle contraction by calcium ion signaling	x	x	x	
regulation of glutamate receptor signaling pathway	x		x	
negative regulation of transforming growth factor beta receptor signaling pathway	x			
enzyme linked receptor protein signaling pathway	x	x	x	
reelin-mediated signaling pathway	x	x		
semaphorin-plexin signaling pathway involved in neuron projection guidance	x			
transmembrane receptor protein tyrosine phosphatase signaling pathway	x		x	
interferon-gamma-mediated signaling pathway	x		x	
interleukin-2-mediated signaling pathway	x			
calcineurin-NFAT signaling cascade	x			
regulation of calcineurin-NFAT signaling cascade	x			
type I interferon signaling pathway	x			
TRAM-dependent toll-like receptor signaling pathway	x	x	x	
TRAM-dependent toll-like receptor 4 signaling pathway	x	x	x	
semaphorin-plexin signaling pathway	x		x	
positive regulation of transforming growth factor beta receptor signaling pathway	x		x	
cytokine-mediated signaling pathway	x	x		
negative regulation of stress-activated protein kinase signaling cascade	x	x	x	
regulation of interferon-gamma-mediated signaling pathway	x		x	
regulation of type I interferon-mediated signaling pathway	x			
positive regulation of Wnt signaling pathway	x			
interleukin-15-mediated signaling pathway	x			
canonical Wnt signaling pathway involved in positive regulation of apoptotic process	x			
interleukin-12-mediated signaling pathway	x			

Table 6.12 continued...

Pathways	D1	D2	D3	D4
nerve growth factor signaling pathway	x			
G-protein coupled acetylcholine receptor signaling pathway	x	x		
smoothened signaling pathway involved in regulation of cerebellar granule cell precursor cell proliferation	x		x	
canonical Wnt signaling pathway involved in negative regulation of apoptotic process	x			
regulation of lipopolysaccharide-mediated signaling pathway	x		x	x
lipopolysaccharide-mediated signaling pathway	x	x	x	
immune response-regulating cell surface receptor signaling pathway involved in phagocytosis	x	x	x	
Fc-gamma receptor signaling pathway involved in phagocytosis	x	x	x	
Fc-gamma receptor signaling pathway	x	x	x	
Fc receptor mediated stimulatory signaling pathway	x	x	x	
immune response-activating cell surface receptor signaling pathway	x	x	x	
Fc receptor signaling pathway	x	x	x	
regulation of signaling	x	x	x	
immune response-regulating cell surface receptor signaling pathway	x			
cell surface receptor signaling pathway	x	x	x	
regulation of stress-activated protein kinase signaling cascade	x	x	x	
immune response-regulating signaling pathway	x			
fibroblast growth factor receptor signaling pathway	x	x	x	
signaling	x	x	x	
single organism signaling	x	x	x	
epidermal growth factor receptor signaling pathway	x	x	x	
stress-activated protein kinase signaling cascade	x	x	x	
ERBB signaling pathway	x	x	x	
negative regulation of insulin receptor signaling pathway	x			x
regulation of insulin receptor signaling pathway	x			x
neurotrophin TRK receptor signaling pathway	x	x	x	
Tie signaling pathway	x			
negative regulation of epidermal growth factor receptor signaling pathway	x	x	x	
negative regulation of ERBB signaling pathway	x	x	x	
regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x		x	x
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x		x	x
regulation of epidermal growth factor receptor signaling pathway	x	x	x	
regulation of ERBB signaling pathway	x	x	x	
negative regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x		x	x
regulation of intrinsic apoptotic signaling pathway by p53 class mediator	x		x	x
phosphatidylinositol-mediated signaling	x		x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
inositol lipid-mediated signaling	x		x	
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	x	x	x	
collagen-activated tyrosine kinase receptor signaling pathway	x	x		
collagen-activated signaling pathway	x	x		
leukemia inhibitory factor signaling pathway	x			
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage	x		x	x
phosphatidylinositol 3-kinase signaling	x		x	
cAMP-mediated signaling	x	x	x	
cyclic-nucleotide-mediated signaling	x	x		
vascular endothelial growth factor receptor signaling pathway	x			
negative regulation of signaling	x	x	x	
regulation of nucleotide-binding oligomerization domain containing signaling pathway	x	x		
regulation of viral-induced cytoplasmic pattern recognition receptor signaling pathway	x			
regulation of RIG-I signaling pathway	x			
positive regulation of protein kinase A signaling	x	x	x	x
BMP signaling pathway	x			
Ribosome biogenesis in eukaryotes	x	x	x	
Circadian rhythm - mammal	x	x		
Phosphatidylinositol signaling system	x	x	x	
Ubiquitin mediated proteolysis	x	x	x	
Biotin metabolism	x			
Renal cell carcinoma	x	x	x	
Focal adhesion	x		x	
VEGF signaling pathway	x		x	
TGF-beta signaling pathway	x		x	
Vascular smooth muscle contraction	x		x	
Toxoplasmosis	x	x	x	
Axon guidance	x	x		
Leukocyte transendothelial migration	x			
Huntington's disease	x		x	
RNA transport	x	x	x	
Endocytosis	x			
ABC transporters	x			
Gastric acid secretion	x	x		
T cell receptor signaling pathway	x		x	
Basal transcription factors	x			

Table 6.12 continued...

Pathways	D1	D2	D3	D4
Leishmaniasis	x			
Pancreatic cancer	x	x	x	
Adherens junction	x	x	x	x
Wnt signaling pathway	x	x		
Basal cell carcinoma	x		x	
Viral myocarditis	x	x		
Complement and coagulation cascades	x		x	
Hematopoietic cell lineage	x		x	
Long-term potentiation	x	x	x	
Fc gamma R-mediated phagocytosis	x	x	x	
Calcium signaling pathway	x	x	x	
Pancreatic secretion	x		x	
Oocyte meiosis	x	x		
Jak-STAT signaling pathway	x	x		
Glioma	x	x	x	
Melanoma	x	x	x	
Acute myeloid leukemia	x	x	x	
Prostate cancer	x	x	x	
Drug metabolism - cytochrome P450	x			
Bladder cancer	x			
Chronic myeloid leukemia	x	x	x	
Staphylococcus aureus infection	x	x		
Non-small cell lung cancer	x	x	x	
Cell cycle	x	x		
Pathways in cancer	x	x	x	
Spliceosome	x		x	
Regulation of actin cytoskeleton	x	x	x	
Bile secretion	x	x		
Lysine biosynthesis	x			
Long-term depression	x	x	x	
Notch signaling pathway	x	x	x	x
Inositol phosphate metabolism	x	x	x	
Selenocompound metabolism		x		
Tyrosine metabolism		x	x	
Phenylalanine metabolism		x	x	x
Glycolysis / Gluconeogenesis		x		

Table 6.12 continued...

Pathways	D1	D2	D3	D4
Steroid hormone biosynthesis		x	x	
Glycosphingolipid biosynthesis - globo series		x		
Lysosome		x	x	
Chagas disease (American trypanosomiasis)		x	x	
ErbB signaling pathway		x	x	
Other glycan degradation		x	x	
Glycosphingolipid biosynthesis - lacto and neolacto series		x	x	
Endometrial cancer		x	x	
Melanogenesis		x	x	
Dilated cardiomyopathy		x	x	
Olfactory transduction		x		
Hypertrophic cardiomyopathy (HCM)		x		
Type II diabetes mellitus		x	x	
Gap junction		x		
Alanine, aspartate and glutamate metabolism		x		
mTOR signaling pathway		x	x	
Neuroactive ligand-receptor interaction		x		
MAPK signaling pathway		x	x	
Type I diabetes mellitus		x		
Mucin type O-Glycan biosynthesis		x		
osmosensory signaling pathway		x		
somatostatin receptor signaling pathway		x		
somatostatin signaling pathway		x		
canonical Wnt signaling pathway involved in regulation of type B pancreatic cell proliferation		x	x	
stromal-epithelial cell signaling involved in prostate gland development		x	x	
canonical Wnt signaling pathway involved in cardiac muscle cell fate commitment		x		
regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation		x	x	
negative regulation of canonical Wnt signaling pathway involved in controlling type B pancreatic cell proliferation		x	x	
negative regulation of extrinsic apoptotic signaling pathway		x	x	
negative regulation of apoptotic signaling pathway		x	x	
canonical Wnt signaling pathway involved in neural crest cell differentiation		x		
negative regulation of Wnt signaling pathway involved in dorsal/ventral axis specification		x	x	
intracellular receptor signaling pathway		x		
intracellular estrogen receptor signaling pathway		x		
positive regulation of retinoic acid receptor signaling pathway		x		
regulation of retinoic acid receptor signaling pathway		x	x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
intracellular steroid hormone receptor signaling pathway		x	x	x
second-messenger-mediated signaling		x	x	
NIK/NF-kappaB signaling		x		
regulation of intracellular estrogen receptor signaling pathway		x		
positive regulation of intracellular steroid hormone receptor signaling pathway		x		
positive regulation of intracellular estrogen receptor signaling pathway		x		
retinoic acid receptor signaling pathway		x	x	
cell-cell signaling		x		
purinergic receptor signaling pathway		x		
regulation of intracellular steroid hormone receptor signaling pathway		x	x	
phospholipase C-activating G-protein coupled glutamate receptor signaling pathway		x		
bile acid signaling pathway		x		
extrinsic apoptotic signaling pathway in absence of ligand		x	x	
purinergic nucleotide receptor signaling pathway		x		
steroid hormone mediated signaling pathway		x	x	
gamma-aminobutyric acid signaling pathway		x	x	
intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress		x	x	
positive regulation of TOR signaling		x		
neurotrophin signaling pathway		x	x	
negative regulation of intracellular steroid hormone receptor signaling pathway		x	x	
G-protein coupled glutamate receptor signaling pathway		x		
Fc-epsilon receptor signaling pathway		x	x	
fibroblast growth factor receptor signaling pathway involved in orbitofrontal cortex development		x		x
G-protein coupled receptor signaling pathway		x	x	
insulin-like growth factor receptor signaling pathway		x	x	
positive regulation of signaling		x	x	
Base excision repair		x		
Mismatch repair		x		
Taste transduction		x	x	
intrinsic apoptotic signaling pathway by p53 class mediator		x	x	
intrinsic apoptotic signaling pathway in response to DNA damage		x	x	
intrinsic apoptotic signaling pathway		x	x	
platelet-derived growth factor receptor signaling pathway		x		
adiponectin-activated signaling pathway		x		x
adenylate cyclase-activating dopamine receptor signaling pathway		x	x	
regulation of G-protein coupled receptor protein signaling pathway		x		

Table 6.12 continued...

Pathways	D1	D2	D3	D4
interleukin-6-mediated signaling pathway		x		
negative regulation of insulin-like growth factor receptor signaling pathway		x	x	
regulation of insulin-like growth factor receptor signaling pathway		x	x	
positive regulation of Notch signaling pathway		x	x	
JAK-STAT cascade involved in growth hormone signaling pathway		x		
growth hormone receptor signaling pathway		x		
regulation of Notch signaling pathway		x		
Small cell lung cancer		x	x	
Cell adhesion molecules (CAMs)		x		x
Citrate cycle (TCA cycle)		x	x	
Neurotrophin signaling pathway		x	x	
Proximal tubule bicarbonate reclamation		x	x	
Apoptosis		x	x	
Shigellosis		x		
Arrhythmogenic right ventricular cardiomyopathy (ARVC)		x		
Alzheimer's disease		x	x	
Cyanoamino acid metabolism		x		
regulation of smoothened signaling pathway involved in dorsal/ventral neural tube patterning		x		
negative regulation of smoothened signaling pathway involved in dorsal/ventral neural tube patterning		x		
positive regulation of I-kappaB kinase/NF-kappaB signaling		x	x	
activation of signaling protein activity involved in unfolded protein response		x		
Roundabout signaling pathway		x		
regulation of adrenergic receptor signaling pathway		x		
ER-nucleus signaling pathway		x		
chemokine-mediated signaling pathway		x	x	
calcium-mediated signaling		x	x	
RNA polymerase		x	x	
Glycerolipid metabolism		x		
Proteasome		x	x	
positive regulation of hippo signaling		x		
activation of prostate induction by androgen receptor signaling pathway		x		
regulation of thrombin receptor signaling pathway		x	x	
negative regulation of thrombin receptor signaling pathway		x	x	
positive regulation of intrinsic apoptotic signaling pathway by p53 class mediator		x		
negative regulation of smoothened signaling pathway		x	x	
Pathogenic Escherichia coli infection		x		

Table 6.12 continued...

Pathways	D1	D2	D3	D4
Metabolic pathways		x	x	
Fatty acid elongation in mitochondria		x	x	
transforming growth factor beta receptor signaling pathway		x	x	
regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway		x	x	
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway		x	x	
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway		x	x	
hippo signaling		x		
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway		x	x	
regulation of hippo signaling		x		
regulation of skeletal muscle contraction by calcium ion signaling		x		
negative regulation of chemokine-mediated signaling pathway		x		
positive regulation of lipopolysaccharide-mediated signaling pathway		x	x	x
Protein processing in endoplasmic reticulum			x	x
Biosynthesis of unsaturated fatty acids			x	
alpha-Linolenic acid metabolism			x	
RNA degradation			x	
mRNA surveillance pathway			x	
Nucleotide excision repair			x	
Glycosaminoglycan degradation			x	
Carbohydrate digestion and absorption			x	
Colorectal cancer			x	
Adipocytokine signaling pathway			x	
B cell receptor signaling pathway			x	
Fc epsilon RI signaling pathway			x	
Progesterone-mediated oocyte maturation			x	
Toll-like receptor signaling pathway			x	
Osteoclast differentiation			x	
Tight junction			x	
Ether lipid metabolism			x	
Hedgehog signaling pathway			x	
DNA replication			x	
Cardiac muscle contraction			x	
Oxidative phosphorylation			x	
Parkinson's disease			x	
Vibrio cholerae infection			x	
Epithelial cell signaling in Helicobacter pylori infection			x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
Rheumatoid arthritis			x	
Phagosome			x	
Protein digestion and absorption			x	
cGMP-mediated signaling			x	
G-protein coupled receptor signaling pathway coupled to cGMP nucleotide second messenger			x	
antigen receptor-mediated signaling pathway			x	
ionotropic glutamate receptor signaling pathway			x	
smoothened signaling pathway			x	
regulation of smoothened signaling pathway			x	
androgen receptor signaling pathway			x	x
mesodermal-endodermal cell signaling			x	
semaphorin-plexin signaling pathway involved in bone trabecula morphogenesis			x	
integrin-mediated signaling pathway			x	
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway			x	
negative regulation of BMP signaling pathway			x	
canonical Wnt signaling pathway involved in regulation of cell proliferation			x	
regulation of Wnt signaling pathway involved in dorsal/ventral axis specification			x	
Wnt signaling pathway involved in dorsal/ventral axis specification			x	
Wnt signaling pathway involved in somitogenesis			x	
semaphorin-plexin signaling pathway involved in outflow tract morphogenesis			x	
positive regulation of semaphorin-plexin signaling pathway involved in outflow tract morphogenesis			x	
regulation of semaphorin-plexin signaling pathway			x	
positive regulation of semaphorin-plexin signaling pathway			x	
cholecystokinin signaling pathway			x	
negative regulation of retinoic acid receptor signaling pathway			x	
cell surface receptor signaling pathway involved in heart development			x	
positive regulation of chemokine-mediated signaling pathway			x	
regulation of chemokine-mediated signaling pathway			x	
negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress			x	
regulation of intrinsic apoptotic signaling pathway in response to oxidative stress			x	
regulation of TOR signaling			x	
intrinsic apoptotic signaling pathway in response to oxidative stress			x	
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand			x	
TOR signaling			x	
negative regulation of intrinsic apoptotic signaling pathway			x	
regulation of extrinsic apoptotic signaling pathway in absence of ligand			x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
regulation of intrinsic apoptotic signaling pathway			x	
adenylate cyclase-modulating G-protein coupled receptor signaling pathway			x	
protein kinase B signaling			x	
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger			x	
regulation of extrinsic apoptotic signaling pathway			x	
insulin receptor signaling pathway			x	x
extrinsic apoptotic signaling pathway			x	
mesenchymal smoothened signaling pathway involved in prostate gland development			x	
epithelial-mesenchymal signaling involved in prostate gland development			x	
regulation of nodal signaling pathway involved in determination of left/right asymmetry			x	
regulation of nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry			x	
nodal signaling pathway involved in determination of left/right asymmetry			x	
regulation of nodal signaling pathway			x	
nodal signaling pathway involved in determination of lateral mesoderm left/right asymmetry			x	
cerebellar Purkinje cell-granule cell precursor cell signaling involved in regulation of granule cell precursor cell proliferation			x	
epithelial-mesenchymal cell signaling			x	
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway			x	
nodal signaling pathway			x	
regulation of activin receptor signaling pathway			x	
positive regulation of smoothened signaling pathway			x	
transmembrane receptor protein serine/threonine kinase signaling pathway			x	
activin receptor signaling pathway			x	
ciliary receptor clustering involved in smoothened signaling pathway			x	
smoothened signaling pathway involved in ventral spinal cord patterning			x	
thrombin receptor signaling pathway			x	
negative regulation of platelet-derived growth factor receptor signaling pathway			x	
regulation of platelet-derived growth factor receptor signaling pathway			x	
brain-derived neurotrophic factor receptor signaling pathway			x	
calcium-mediated signaling using intracellular calcium source			x	
Glycerophospholipid metabolism			x	
Natural killer cell mediated cytotoxicity			x	
ECM-receptor interaction			x	
regulation of fibroblast growth factor receptor signaling pathway			x	
positive regulation of insulin-like growth factor receptor signaling pathway			x	
apoptotic signaling pathway			x	
regulation of apoptotic signaling pathway			x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
positive regulation of G-protein coupled receptor protein signaling pathway			x	
positive regulation of phosphatidylinositol 3-kinase signaling			x	
regulation of phosphatidylinositol 3-kinase signaling			x	
regulation of protein kinase B signaling			x	
mesenchymal-epithelial cell signaling involved in prostate gland development			x	
mesenchymal-epithelial cell signaling			x	
Fatty acid metabolism			x	
PPAR signaling pathway			x	
RIG-I-like receptor signaling pathway			x	
Salivary secretion			x	
MyD88-independent toll-like receptor signaling pathway			x	
toll-like receptor 4 signaling pathway			x	
toll-like receptor signaling pathway			x	
pattern recognition receptor signaling pathway			x	
regulation of I-kappaB kinase/NF-kappaB signaling			x	
I-kappaB kinase/NF-kappaB signaling			x	
negative regulation of canonical Wnt signaling pathway			x	
negative regulation of Wnt signaling pathway			x	
regulation of canonical Wnt signaling pathway			x	
regulation of Wnt signaling pathway			x	
Glycosaminoglycan biosynthesis - keratan sulfate			x	
Glyoxylate and dicarboxylate metabolism			x	
Porphyrin and chlorophyll metabolism			x	
serotonin receptor signaling pathway			x	
Other types of O-glycan biosynthesis			x	
N-Glycan biosynthesis			x	x
Cytosolic DNA-sensing pathway			x	
Aminoacyl-tRNA biosynthesis			x	
Antigen processing and presentation			x	
Renin-angiotensin system			x	
African trypanosomiasis			x	
regulation of toll-like receptor 3 signaling pathway			x	
positive regulation of toll-like receptor 3 signaling pathway			x	
Toll signaling pathway			x	x
regulation of toll-like receptor 4 signaling pathway			x	
positive regulation of toll-like receptor 4 signaling pathway			x	

Table 6.12 continued...

Pathways	D1	D2	D3	D4
positive regulation of protein kinase B signaling			x	
Aldosterone-regulated sodium reabsorption			x	
Pyruvate metabolism			x	
activation of MAPK activity by adrenergic receptor signaling pathway			x	
epidermal growth factor-activated receptor transactivation by G-protein coupled receptor signaling pathway			x	
positive regulation of epidermal growth factor receptor signaling pathway			x	
positive regulation of ERBB signaling pathway			x	
adrenergic receptor signaling pathway			x	
adenylate cyclase-activating adrenergic receptor signaling pathway			x	
positive regulation of cAMP-mediated signaling			x	
regulation of cAMP-mediated signaling			x	
One carbon pool by folate			x	
protein kinase A signaling			x	x
regulation of protein kinase A signaling			x	x
regulation of intrinsic apoptotic signaling pathway in response to DNA damage			x	x
Collecting duct acid secretion				x
fibroblast growth factor receptor signaling pathway involved in negative regulation of apoptotic process in bone marrow				x
fibroblast growth factor receptor signaling pathway involved in hemopoiesis				x
fibroblast growth factor receptor signaling pathway involved in positive regulation of cell proliferation in bone marrow				x
fibroblast growth factor receptor signaling pathway involved in mammary gland specification				x
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway				x
granzyme-mediated apoptotic signaling pathway				x

7 Elucidating the Importance of Clustering Replicates in Three Dimensional Microarray Gene Expression Data

7.1 Introduction

The advancement of microarray technology facilitates monitoring the expression profiles of thousands of genes not only over a set of replicates but also across a set of time points or different doses of several chemical substances. Though the experimental set up remains unchanged to measure the expression values of genes at each time point over a set of replicates, the variations in the expression values over the replicates can still be present due to abnormalities of the experimental protocol when taking the expression profiles (technical replicates) or physiological variations in the population leading to heterogeneity among the samples (biological replicates). Thus, equalizing the effects of the replicates by forcing together those which exhibits dissimilar expression patters may not be beneficial in mining three dimensional dataset to extract the biologically meaningful information. In this chapter, we aim at unraveling the effects of non-clustered replicates on the quality of the resultant triclusters obtained after applying the EMOA-δ-TRIMAX triclustering algorithm to the gene expression datasets used in the previous chapters.

7.2 Materials and Methods

7.2.1 Dataset 1 (Accession no.- GSE11324)

The work of J.S. Carroll et al. aimed at elucidating the transcriptional regulation during the exposure of a MCF-7 breast cancer cell to estrogen across 4 time points (0, 3, 6, and 12 hours) [1]. At each time point, the experiment was performed to monitor the expression profiles of 54675 Affymetrix probe ids for three times keeping the experimental conditions unchanged, thus the replicates can be referred to as biological ones.

7.2.2 Dataset 2 (Accession Number- GSE35671)

This work was carried out to reveal the regulatory mechanisms during the differentiation of a human induced pluripotent stem cell (hiPSC) into cardiomyocytes [2]. This dataset contains the expression profiles of 48803 Illumina probe ids across 12 time points (days 0, 3, 7, 10, 14, 20, 28, 35, 45, 60, 90 and 120) and at each time point the expression profiles were measured for three biological replicates.

7.2.3 Dataset 3 (Accession Number- GSE46280)

The purpose of the work carried out by M. Hecker et al. was to explore the expression dynamics of 54675 Affymetrix probe ids during the exposure of peripheral blood mononuclear cells to subcutaneous IFN-beta-1b treatment across four time points (first, second, third IFN-beta injection and after 1 month) and peripheral blood mononuclear cells were collected from six patients [3]. Thus, in this dataset, the replicates can be considered as biological ones.

7.2.4 Dataset 4 (Accession Number- GSE17693)

The dataset comprises expression profiles of 18675 probe ids across 5 time points (control, day 1, day 2, day 3 and day 6) and 4 biological replicates. The experiment was done in order to elucidate the potential key genes during the productive and unproductive repair of mouse airway epithelial cells exposed to naphthalene [4, 5].

7.2.5 Dataset 5 (Accession Number- GSE17933)

This experiment was carried out by R.S. Thomas et al. to unravel the potential biomarkers during the exposure of mouse lung tumors to naphthalene having five dose levels i.e. 0.5, 3, 10, 20 and 30 parts per million (ppm) [6]. For each of these dose levels, the experiment was performed to monitor the expression profiles 45101 Affymetrix probe ids over four different mice.

7.2.6 Dataset 6 (Accession Number- GSE18858)

The aim of this experiment was to provide insights into the potential biomarkers during the exposure of five-week-old female B6C3F1 mouse liver tumors to the toxicants namely, naphthalene and 1,2,3-trichloropropane having different dose levels i.e. (0.5, 3, 10, 20, 30 ppm for naphthalene and 2, 6, 20, 40, 60 mg/kg for 1,2,3-trichloropropane) [7]. For each of the aforementioned dose levels, this dataset contains the expression profiles 45101 Affymetrix probe ids measured across four liver tumors taken from four different mice.

7.2.7 Dataset 7 (Accession Number- GSE38513)

The experiment was carried out by S.A. Gharib et al. in order to investigate the role of matrix metalloproteinase 7 (MMP7) in governing the injured airway epithelial cell repair [8].

To accomplish the goal, expression profiles of 45101 Affymetrix probe ids were measured over four biological replicates during the exposure of Mmp7-null and wildtype mice airway epithelial cells to naphthalene.

7.2.8 Workflow

Figure 7.1 shows the workflow applied in this chapter. After applying the proposed triclustering algorithm EMOA-δ-TRIMAX to each of the datasets described in the previous subsection, we have evaluated the mean of the pairwise Euclidean distances of the clustered replicates for the triclusters which do not have all replicates. Afterwards, we have calculated the mean of the pairwise Euclidean distances of all replicates for the triclusters used in the previous step. Finally, we have computed the following metric for each of the datasets.

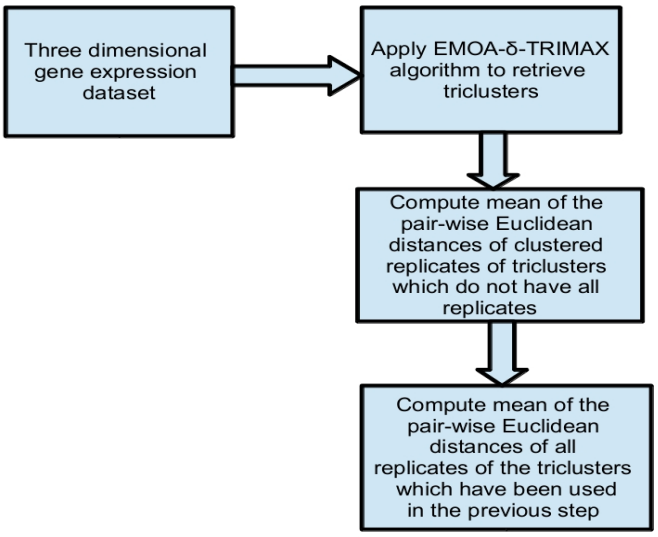

Figure 7.1: Workflow used in this chapter.

$$divergence_i = \frac{\sum_{k=1}^{n} \frac{dist\_non\_clust[k]}{dist\_clust[k]}}{n} \tag{7.1}$$

, where $dist\_non\_clust[k]$ and $dist\_clust[k]$ represent the mean of the pairwise Euclidean distances kth tricluster which does not have all replicates considering all and clustered

replicates, respectively. n represents the total number of triclusters having at least one replicate missing in case of dataset i. A higher value of $divergence_i$ corresponds to a dataset having more divergent replicates.

7.3 Results and Discussion

From each of the Figures 7.2-7.8, we can observe the enhancement of the average intra-cluster Euclidean distances between replicates if we include all non-clustered replicates in the triclusters which do not have all replicates.

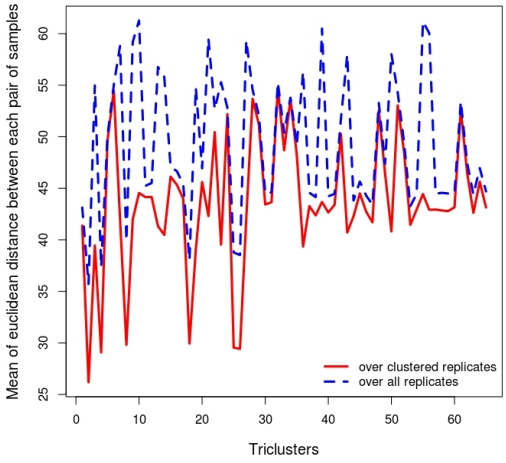

Figure 7.2: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and time points for Dataset 1.

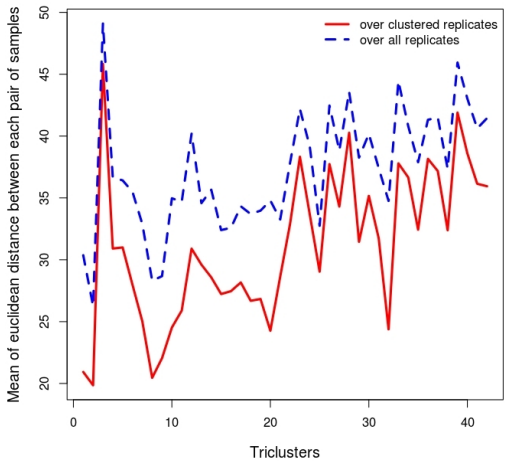

Figure 7.3: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and time points for Dataset 2.

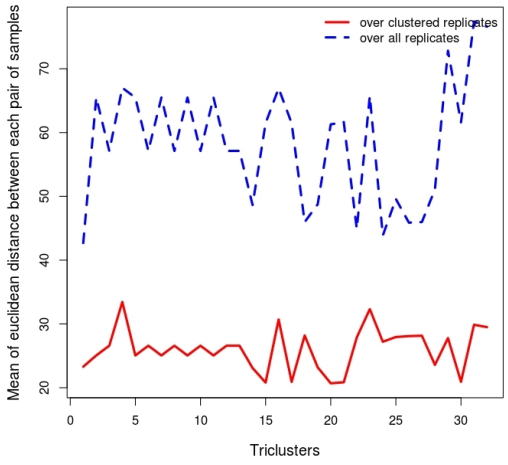

Figure 7.4: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and time points for Dataset 3.

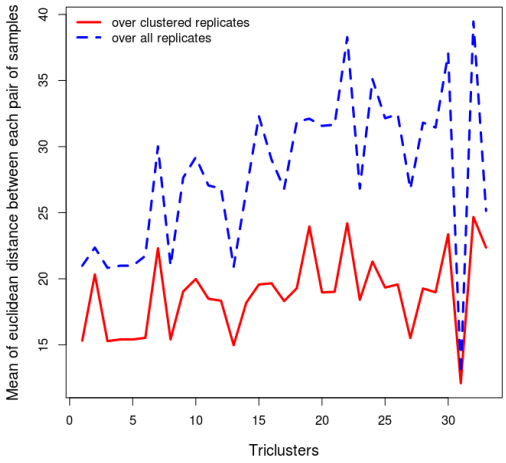

Figure 7.5: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and different doses of toxicants for Dataset 4.

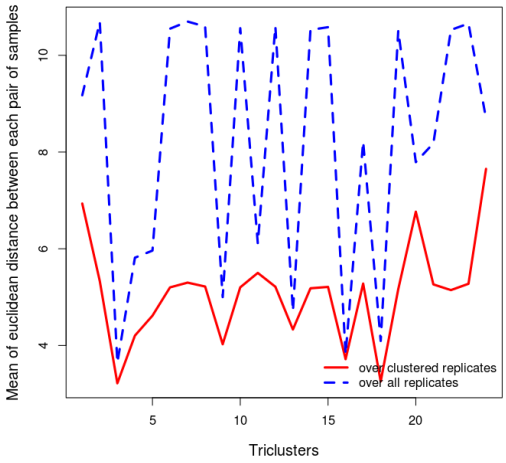

Figure 7.6: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and different doses of toxicants for Dataset 5.

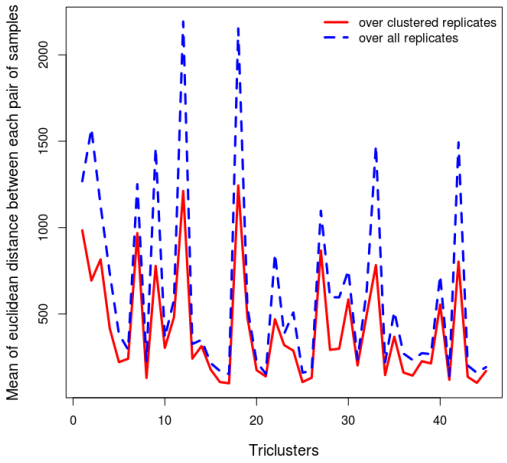

Figure 7.7: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and different doses of toxicants for Dataset 6.

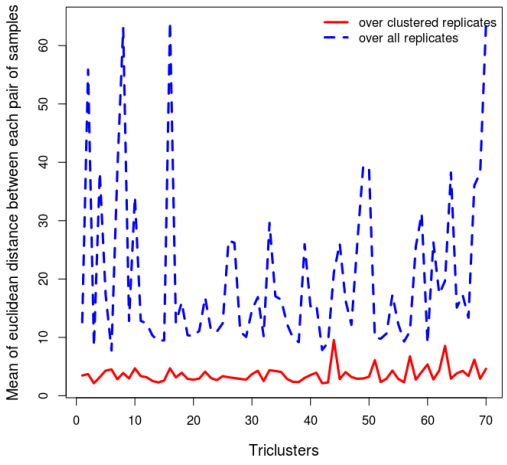

Figure 7.8: Average Euclidean distances between the expression profiles of each pair of clustered (red) and all (blue, dashed line) replicates over the clustered genes and different doses of toxicants for Dataset 7.

Additionally, Table 7.1 enlists the values of the metric $divergence_i$ for each of the datasets described in the previous section. It is of interest to see a relatively small value of $divergence_i$ in case of the datasets 1 and 2, whereas this value becomes higher for the rest of the datasets.

Table 7.1: Values of $divergence_i$ metric for each of the datasets used in this chapter.

Dataset	$divergence_i$
Dataset 1	1.145301
Dataset 2	1.20993
Dataset 3	2.276298
Dataset 4	1.47531
Dataset 5	1.606549
Dataset 6	1.479335
Dataset 7	5.813373

7.4 Conclusion

Altogether, this chapter provides insights into the importance of clustering replicates in the context of mining three dimensional microarray gene expression datasets in order to extract biologically meaningful information. Moreover, our results indicate the fact that if you move more from in vitro towards in vivo, the divergence between the biological replicates increases significantly. For instance, we have found the lowest divergence in case of dataset 1 (cells in culture), whereas the value becomes significantly higher in case of the dataset 3 and dataset 5 where the biological replicates are of more complex nature such as whole animals or patients.

7.5 Bibliography

[1] Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E.A., Silver, P.A., Gingeras, T.R., Liu, X.S., Brown, M.: Genome-wide analysis of estrogen receptor binding sites. Nature Genetics **38**(11), 1289–1297 (2006)

[2] Babiarz, J.E., Ravon, M., Sridhar, S., Ravindran, P., Swanson, B., Bitter, H., Weiser, T., Chiao, E., Certa, U., Kolaja, K.L.: Determination of the human cardiomyocyte

mRNA and miRNA differentiation network by fine-scale profiling. Stem Cells and Development **21**(11), 1956–1965 (2012). doi:10.1089/scd.2011.0357

[3] Hecker, M., Thamilarasan, M., Koczan, D., Schroeder, I., Flechtner, K., Freiesleben, S., Fuellen, G., Thiesen, H.J., Zettl, U.K.: MicroRNA expression changes during interferon-beta treatment in the peripheral blood of multiple sclerosis patients. International Journal of Molecular Sciences **14**(8), 16087–16110 (2013). doi:10.3390/ijms140816087

[4] Snyder, J.C., Zemke, A.C., Stripp, B.R.: Reparative capacity of airway epithelium impacts deposition and remodeling of extracellular matrix. American Journal of Respiratory Cell and Molecular Biology **40**(6), 633–642 (2009). doi:10.1165/rcmb.2008-0334OC

[5] Zemke, A.C., Snyder, J.C., Brockway, B.L., Drake, J.A., Reynolds, S.D., Kaminski, N., Stripp, B.R.: Molecular staging of epithelial maturation using secretory cell-specific genes as markers. American Journal of Respiratory Cell and Molecular Biology **40**(3), 340–348 (2009). doi:10.1165/rcmb.2007-0380OC

[6] Thomas, R.S., Bao, W., Chu, T.M., Bessarabova, M., Nikolskaya, T., Nikolsky, Y., Andersen, M.E., Wolfinger, R.D.: Use of short-term transcriptional profiles to assess the long-term cancer-related safety of environmental and industrial chemicals. Toxicological Sciences **112**(2), 311–321 (2009). doi:10.1093/toxsci/kfp233

[7] Thomas, R.S., Clewell, H.J.r., Allen, B.C., Wesselkamper, S.C., Wang, N.C., Lambert, J.C., Hess-Wilson, J.K., Zhao, Q.J., Andersen, M.E.: Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicological Sciences **120**(1), 194–205 (2011). doi:10.1093/toxsci/kfq355

[8] Gharib, S.A., Altemeier, W.A., Van Winkle, L.S., Plopper, C.G., Schlesinger, S.Y., Buell, C.A., Brauer, R., Lee, V., Parks, W.C., Chen, P.: Matrix metalloproteinase-7 coordinates airway epithelial injury response and differentiation of ciliated cells. American Journal of Respiratory Cell and Molecular Biology **48**(3), 390–396 (2013). doi:10.1165/rcmb.2012-0083OC

8 Conclusion

In this work, I have proposed a triclustering algorithm δ-TRIMAX by introducing a novel coherence measure to mine three-dimensional gene expression datasets. This algorithm has been applied to a time series dataset which monitors expression profiles upon exposure of an MCF-7 breast cancer cell to estrogen. In this application, the algorithm has been found to be useful in identifying the potential biological processes involved in the progression of breast cancer. However, it had clear limitations such as the inability to extract overlapping triclusters and the possibility of getting stuck into local optima due to its greedy search heuristic approach. I therefore developed an improved version of δ-TRIMAX algorithm, EMOA-δ-TRIMAX which effectively remediated the pitfalls of the former one. We have proven the merits of EMOA-δ-TRIMAX using one artificial dataset and three real-life datasets where it outperforms the other algorithms for each of the aforementioned datasets. Moreover, the improved version of δ-TRIMAX algorithm has been found to be beneficial in analyzing a time series gene expression dataset which monitors gene expression profiles during stem cell differentiation into cardiomyocytes for revealing the key genes play instrumental roles in driving the phenotypic changes of the cell and unveil the regulatory mechanisms during different stages of cardiomyocytes development. Moreover, EMOA-δ-TRIMAX has also been used not only to infer gene regulatory networks during the exposure of a MCF-7 breast cancer cell to estrogen but also to elucidate the signaling pathways that may be triggered by different chemical substances or toxicants. Altogether, this work emphasizes the application of the proposed triclustering algorithms in the context of analyzing three dimensional microarray gene expression dataset and provides new insights in the context of transcriptional regulation, developmental biology, disease biology and toxicology. One possible extension of the work would be incorporating the regulatory information into the EMOA-δ-TRIMAX algorithm to ensure both co-expression and co-regulation of the genes belonging to the resultant triclusters. Additionally, in the context of mining time series gene expression data, EMOA-δ-TRIMAX algorithm yields triclusters that may have nonconsecutive time points which may be treated as inconsequential from a biological point of view. Thus representing time intervals instead of the time points in the encoding of chromosomes may solve this problem.

Curriculum Vitae

Name:	Anirban Bhar
Date of birth:	22.03.1984
Place of birth:	Kolkata, India
Address:	Apartment number- 44, Gossler-strasse 33 B, Goettingen- D37075, Germany

Education and Research Experience

2010 - present:	**Institute of bioinformatics, University Medical Center, Georg August University, Goettingen.** Doctoral thesis: Application of A Novel Tricluster-ing Method in Analyzing Three Dimensional Transcriptomics Data.
2006 - 2008:	**University of Kalyani, Kalyani, India.** M.Tech. in Computer Sci-ence and Engineering.
2002 - 2006:	**West bengal University of Technology, Kolkata, India.** B.Tech. in Computer Science and Engineering.

Awards and Scholarships

2010 - 2013:	Fellowship of Erasmus Mundus Ex-ternal Cooperation Window, Lot 13 - India - EURINDIA project.

Conference Presentation

Title	Conference
(I) Application of a Novel Triclustering Method (δ-TRIMAX) to Mine 3D Gene Expression Data of Breast Cancer Cells	German Conference on Bioinformatics, 2013.
(II) Revealing Exclusive Usage Of T-BOX Family Paralogous Transcription Factors Through Identifying Diversity In Expression Profiles During hiPSC-Derived Cardiomyocytes Generation	Regulatory Genomics Special Interest Group - RegGenSig, ISMB/ ECCB, 2013.
(III) δ-TRIMAX: Extracting Triclusters and Analyzing Coregulation in Time Series Gene Expression Data	Workshop on Algorithms in Bioinformatics, 2012.

Publications

(I) **Bhar, A.**, Haubrock, M., Zeidler, S., Wingender, E.: Unraveling The Regulatory Roles of *ZEB2* During The Adolescence of Cardiomyocytes. [In preparation].

(II) **Bhar, A.**, Haubrock, M., Mukhopadhyay, A., Wingender, E.: Multiobjective Triclustering Of Time-Series Transcriptome Data Reveals Key Genes Of Biological Processes. BMC Bioinformatics, in press (2015)

(III) Koschmann, J., **Bhar, A.**, Stegmaier,P., Kel, A.E. and Wingender, E.: Upstream Analysis: An integrated promoter-pathway analysis approach to causal interpretation of microarray data. Microarrays 4, 270-286 (2015).

(IV) **Bhar, A.**, Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: Coexpression and coregulation analysis of time-series gene expression data in estrogen-induced breast cancer cell. Algorithms for molecular biology 8(1), 9 (2013).

(V) **Bhar, A.**, Haubrock, M., Mukhopadhyay, A., Maulik, U., Bandyopadhyay, S., Wingender, E.: δ-TRIMAX: Extracting triclusters and analysing coregulation in time series gene expression data. In: Raphael, B., Tang, J. (eds.) Algorithms in Bioinformatics. Lecture Notes in Computer Science, vol. 7534, pp. 165177. Springer, Berlin Heidelberg

(2012). doi:10.1007/978-3-642-33122-0_13.