IDENTIFICATION OF EPILEPSY RELATED PATHWAYS FROM METHYLOME ANALYSIS

by

ECE EGEMEN

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University

August 2013

IDENTIFICATION OF EPILEPSY RELATED PATHWAYS
FROM METHYLOME ANALYSIS

APPROVED BY:

Prof. Dr. Osman Uğur Sezerman
(Thesis Supervisor) .....

Assoc. Prof. Dr. Berrin Yanıkoğlu .....

Assoc. Prof. Dr. Yücel Saygın .....

Assoc. Prof. Dr. Devrim Gözüaçık .....

Asst. Prof. Dr. Kemal Kılıç .....

DATE OF APPROVAL: 13/08/2013

IDENTIFICATION OF EPILEPSY RELATED PATHWAYS
FROM METHYLOME ANALYSIS

Ece Egemen

Computer Science and Engineering, Master Thesis, 2013

Thesis Advisor: Prof. Dr. Osman Uğur Sezerman

Keywords: Epigenetics, CpG Island Methylation, Epilepsy related Pathways

Abstract

Epilepsy is a disorder which affects approximately 1% of the world's population. It is possible to explain the underlying mechanisms of epilepsy by not only alterations in genetic structure and environmental factors, but also epigenetic mechanisms. As a significant epigenetic modification, DNA methylation may play a crucial role for understanding biological pathways related to epilepsy.

Even though the relationship between CpG islands, DNA methylation and gene expression is quite complex, in this thesis we aimed to determine significant biological pathways related to epilepsy, using the difference of methylation levels at CpG loci in family trios. Dataset was gathered from 10 family trios (30 individuals) and it includes methylation level information for each CpG locus determined by Illumina HumanMethylation450 Bead Chip which is a microarray based analysis tool. Considering the difference of methylation levels at each CpG locus, we prepared lists of genes that we marked as significant. Applying several filters to the dataset allowed us to gather more reliable information from the dataset. We identified a number of significant pathways that may be related to epilepsy, especially immune system related pathways stand out as a significant and novel finding.

EPİLEPSİ İLE İLİŞKİLİ YOLAKLARIN
METİLOM ANALİZİ İLE SAPTANMASI

Ece Egemen

Bilgisayar Bilimi ve Mühendisliği, Yüksek Lisans Tezi, 2013

Tez Danışmanı: Prof. Dr. Osman Uğur Sezerman

Anahtar Kelimeler: Epigenetik, CpG Adası Metilasyonu, Epilepsi ile İlişkili Yolaklar

Özet

Epilepsi, dünya nüfusunun yaklaşık %1'ini etkileyen bir hastalıktır. Epilepsinin ardında bulunan sebepleri sadece genetik yapıda değişiklikler ve çevresel faktörler ile değil, aynı zamanda epigenetik mekanizmalarla açıklamak mümkündür. Önemli bir epigenetik değişikliği olan DNA metilasyonu, epilepsi ile ilgili biyolojik yolakları anlamak için önemli bir role sahip olma ihtimalini barındırmaktadır.

CpG adaları, DNA metilasyonu ve gen ifadesi arasındaki ilişki oldukça karmaşık olsa da, üç kişilik ailelere ait CpG bölgelerinin metilasyon seviyelerinin farkı incelenerek epilepsi ile ilgili önemli biyolojik yolakların belirlenmesi bu tezin hedeflerinden birisi olmuştur. Illumina HumanMethylation450 Bead Chip adlı mikrodizi tabanlı analiz aracı yardımıyla 10 farklı 3 kişilik aileden (30 kişi) toplanan, CpG bölgelerinin içindeki metilasyon düzeyleri içeriğine sahip veri kümesi analiz edilmiştir. Her CpG bölgesine ait metilasyon seviyelerindeki fark göz önüne alınarak, önemli görülen genlerin listesi oluşturulmuştur. Veri kümesine çeşitli filtreler uygulanarak daha güvenilir sonuçlar elde edilmiştir. Epilepsi ile ilgili olabilecek bir takım önemli yolaklar tespit edilmiş, bunların içinde özellikle bağışıklık sistemi ile ilgili olanların tespit edilmesi önemli ve yeni bir bulgu olarak dikkat çekmiştir.

to my beloved family

and my significant other…

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor Prof. O. Uğur Sezerman for his invaluable guidance and continuous support throughout my master studies and even before. He always believed in me and motivated me; I can't even imagine completing this study without his support and smiling face.

I am also thankful to my thesis defense committee members, Assoc. Prof. Berrin Yanıkoğlu, Assoc. Prof. Yücel Saygın, Assoc. Prof. Devrim Gözüaçık and Asst. Prof. Kemal Kılıç for their valuable comments and their presence.

I also want to express my gratitude to Prof. Rita Casadio for her support and guidance during my internship in Bologna University. I want to thank Castrense Savojardo and Piero Fariselli for their supervision and everyone in Casadio Lab for their warm and helpful approach.

I also want to express my gratitude to Prof. Dr. S. Ülker Öneş who has been an inspiration for me as a scientist role model. Her achievements in medical sciences and unmatched personality, also her support always motivated me to move forward.

I am also grateful to Dr. Burcu Bakır-Güngör for her guidance and support in the implementation process of PANOGA web-server, and for her ideas and direct support in the partial implementation of PANOGA Protocol.

Even though they are not directly involved in this thesis, I appreciate the support of my dear friends. They always motivated me and they are always confident about my success. Although some of them are far far away now, I know I can always count on them.

I don't even know how I can explain my gratitude to Emre İnal. As a colleague, he lent his help during the implementation of PANOGA web-server, and debugging of other software. As my dearest, he always supported me and guided me to solutions in my most frustrated moments. He was always there for me, we shared the happiest

moments and he never took the easy road when I feel down, instead he made me understand why I'm overrating the obstacles I had encountered, and I have the strength and ability to overcome them.

As the most important people in my life, the support of my family was invaluable for me. They always believed in me and supported my decisions. Since my childhood, my father always enjoyed the topics I'm working on, even though he is not aware of it, he always makes me feel I'm on the right track, doing something important. My mother always encouraged me on the steps I want to take, even if they mean I will be away from her. Her unconditional love and trust are immeasurably important for me. I love them so much that I don't ever want to let them down. I should also thank my dear grandmother, grandfather and aunt. Their love and support are precious and irreplaceable for me; I'll always be running back to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

BMIQ	Beta-Mixture Quantile Normalization
DNA	Deoxyribonucleic Acid
gDNA	Genomic DNA
GWAS	Genome-Wide Association Study
HTF	HpaII Tiny Fragments
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism

CHAPTER 1

Introduction

Epilepsy is a common neurological condition which affects approximately 1% of the world's population. It is defined by recurrent, unprovoked seizures that can cause motor, sensory, cognitive, psychic or autonomic disturbances. By convention, its diagnosis requires the patient whom has had at least two unprovoked seizures. Seizures may last between a few seconds and a few minutes (Steinlein, 2004). In most cases, epilepsy is controlled with medication, but not cured.

Instead of being considered as a single disorder, epilepsy can be differentiated into several clinical subtypes. Clinical classification of epilepsy defines the general categories as symptomatic, presumed symptomatic and idiopathic epilepsy. The cases where the patient develops recurrent unprovoked seizures for no obvious reason and without any other neurological abnormalities are called as idiopathic epilepsies, and they are mainly caused by genetic reasons. Most idiopathic epilepsies occur with a greater frequency in relatives of affected individuals (Kaneko & Wada, 1998), but they do not exhibit a single Mendelian inheritance pattern (Kaneko et al., 2002). Multiple genes are simultaneously involved in the process of epileptogenesis, which describes the process that leads to development of recurrent unprovoked seizures in a previously healthy brain.

Although all modes of inheritance are found in epilepsies, including autosomal, X-chromosomal, mitochondrial and complex inheritance (Steinlein, 2004), epigenetic modifications associated with transgenerational inheritance also play role in a number of circumstances related with epilepsy. Epigenetic modifications describe heritable

changes in gene function without any alterations in the nucleotide sequence. DNA methylation is considered as a significant epigenetic modification and it can be associated with epileptogenesis under some circumstances.

CpG islands (CpG rich regions in DNA) are usually hypomethylated, and they are associated with approximately 60% of human genes (Antequera & Bird, 1993), over 50% of them in their promoter regions (Larsen et al., 1992; Wang et al., 2004). DNA methylation in these regions may have effects in gene expression. Even though the role of DNA methylation in promoters and/or CpG islands has long been appreciated, the relationship between CpG islands, methylation and gene expression is quite complex, therefore making it an important and intriguing research area.

The aim of this study is to identify functionally important biological pathways related to epilepsy. Main focus of the thesis is to study the effects of the difference in methylation levels in upstream regions of the genes between disease and control groups. Methylation status of each CpG locus were determined and used to mark related genes and these related genes were used to determine significant pathways. In order to find common motifs in methylated and unmethylated CpG island sequences, a computational method that determines CpG island nucleotide composition is also applied.

Background information on related concepts can be found in Chapter 2, whereas details of the dataset and applied methods are presented in Chapter 3. Results are shown in Chapter 4 and discussion of the results and future projections can be found in Chapter 5. Several Appendices provide details about certain parts of analysis results.

CHAPTER 2

Background Information

DNA (Deoxyribonucleic Acid) is a molecule that contains heredity information in humans and almost all other organisms. Two polynucleotide chains, in other words strands, run in opposite directions around the common axis, hence having the name double helix structure. Each strand consists of smaller units called nucleotides, which are composed of nucleobases and a backbone made of sugars and phosphate groups. Those bases are Adenine (A), Guanine (G), Cytosine (C) and Thymine (T), and they are bonded to deoxyribose sugar with covalent bonds.

Sequences of nucleotides encode genetic information by specifying the sequence of amino acids within proteins. Segments of DNA that code for polypeptides and functional Ribonucleic Acid (RNA) chains are called genes. The information carried by genes is read in transcription process which transfers the sequence information to RNA to be used in protein synthesis.

There are some particular regions in DNA sequence which initiate transcription of genes. These regions are called promoters and they are located near genes. Promoters play a crucial role in gene expression since they are regulatory regions which control the transcription process of a gene.

2.1. Methylation

DNA Methylation is a significant epigenetic modification which plays a vital role in transcriptional regulation, chromatin remodeling (Wang et al., 2012), chromosomal stability, genomic imprinting and X-inactivation (Rakyan et al., 2008). Also the aberration of DNA methylation profile is associated with many human diseases including cancer (Wang et al., 2012).

DNA Methylation was first found to influence gene expression in 1975. It was accepted that both adenine and cytosine residues were capable of being methylated in bacteria, however only methylated cytosines were possessed by higher organisms (Harrison & Parle-McDermott, 2011). In eukaryotic genomes, DNA methylation appears as a cytosine methylated at the 5-carbon position (Antequera, 2007). In mammals, DNA methylation is primarily restricted to CpG dinucleotides (Smith and Meissner, 2013). Because of that, DNA methylation is often referred according to methylation or non-methylation of CpG dinucleotides. Since a methylated CpG is matched by another methylated CpG in the complementary strand, methylation on the DNA double helix is symmetrically organized (Antequera, 2007). Mammalian genomes are globally CpG-depleted, and approximately 60-80% of CpG in the human genome are methylated (Smith & Meissner, 2013). Huge number of possible methylation patterns may exist and the information that carried out by these patterns can be changed without any alteration to the nucleotide sequence. Moreover, during DNA replication part of the cell division, CpGs in the daughter strand are methylated only when they are complementary to methylated CpGs in the parental strand. This regulation provides a protection mechanism for the inheritance of DNA methylation (Antequera, 2007).

2.2. CpG Islands

One of the most significant features of DNA methylation appears in CpG rich regions, so called CpG islands, which are usually hypomethylated. Approximately 60% of human genes are associated with CpG islands (Antequera & Bird, 1993) and over 50% of human genes are associated with CpG islands in their promoter regions (Larsen

et al., 1992; Wang et al., 2004). However CpG islands can stay methylation-free when their associated gene is silent, because many CpG islands are located at genes that have a tissue-restricted expression pattern. On the other hand, a small but significant fraction of CpG islands become methylated during development, and their associated promoters remain silent. A significant amount of all human CpG islands are prone to progressive methylation in certain tissues during aging or in abnormal cells such as cancers and permanent cell lines (Bird, 2002).

Originally, CpG islands were identified by an epigenetic property, which is the absence of DNA methylation (Bock et al., 2007). When the genomic DNA was experimentally digested with restriction enzyme isoschizomers that differed only by their sensitivity to cytosine methylation (Singer et al., 1979), the genome were digested to completion by the methylation-insensitive MspI (5′-CCGG-3′) enzyme, whereas most of the DNA exposed to its methylation-sensitive isoschizomer HpaII remains high-molecular weight (Fazzari & Greally, 2004). It is concluded that a small but significant fraction of the genome is reproducibly unmethylated, and the hypomethylated part of the genome digested by HpaII were called HpaII tiny fragments (HTFs) (Bird et al., 1986). When HTFs were cloned and sequenced, it is discovered that they were highly (G+C) and CpG rich (Bird et al., 1986).

According to another research based on this information, CpG Island definition was made and its criteria known as Gardiner-Garden sequence criteria were set (Bock et al., 2007). According to this criteria, a genomic region must fulfill the following criteria to be classified as a CpG Island: (1) (G+C) content should be above 50%, (2) observed to expected CpG dinucleotide ratio should be 0.60 or greater, (3) the length of the sequence should be greater than 200 basepairs (bp) (Gardiner-Garden & Frommer, 1987).

The identification of CpG islands, still remains as a great model for today's epigenomic studies (Fazzari & Greally, 2004). After the discovery, CpG islands have been subject to extensive research (Bock et al., 2007). Today it is known that CpG islands are present in the promoter and exonic regions of approximately 50% of mammalian genes. On the other hand, other regions of the mammalian genome contain

few CpG dinucleotides and these are largely methylated (Larsen et al., 1992). Since methylated cytosines are mutational hotspots (Coulondre et al., 1978) which lead to CpG depletion during evolution, the decrease in the occurrence of CpGs can be explained by this fact (Takai & Jones, 2002). CpG islands are associated with more than 75% of all known transcriptional start sites (Bajic et al., 2006) and 88% of active promoters identified in primary fibroblasts (Kim et al., 2005). Moreover, CpG islands are hotspots of specific histone modifications (Bernstein et al., 2005), they frequently bind ubiquitous transcription factors such as SP1 (Cawley et al., 2004), and they exhibit particularly accessible chromatin structures (Crawford et al., 2006). For these reasons CpG Islands have become a significant element for genome analysis and annotation (Bock et al., 2007).

2.3. Epigenetics

There exist a few different definitions for the term "epigenetics", because the word itself has several different meanings with independent roots (Bird, 2007). One definition describes epigenetics as the study of epigenesis, which means how genotypes give rise to phenotypes during development (Waddington, 1957). Another definition explains the word as "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Riggs et al., 1996). Bird suggested a unifying definition for epigenetic events as the following: "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007).

DNA methylation and the Polycomb/Trithorax systems are considered classic epigenetic systems, because the alterations in these systems are inherited by subsequent generations of cells and sometimes organisms. Epimutations cause the epigenetic inheritance, without altering the nucleotide sequence. Transmission of DNA methylation at a locus from generation to generation is a significant type of epimutations. Even though examples of transgenereational epigenetics are well documented in plants, epigenetic inheritance in animals have been only detected by highly sensitive genetic assays (Bird, 2007).

2.4. Epilepsy

Epilepsy affects approximately 1% of the world's population, including one in 200 children (Cowan, 2002). Epilepsies are characterized by recurrent seizures which can cause motor, sensory, cognitive, psychic or autonomic disturbances. Even though its basic manifestations appear at the single neuron level, seizures are only possible because the brain is organized in a series of interconnected neuronal networks (Steinlein, 2004).

Epilepsy is not considered as a single disorder, but it can be differentiated into several clinical subtypes. According to clinical classification of epilepsy, the general categories are symptomatic, presumed symptomatic and idiopathic epilepsy. Over time, idiopathic epilepsy definition has become the epilepsy caused by an underlying genetic basis, and the majority of cases in this category are likely to be influenced by genetic susceptibility (Poduri, 2011).

Epileptogenesis is the definition of the process that leads to development of recurrent unprovoked seizures in a previously healthy brain. "Epileptogenesis is associated with complex temporal and spatial abnormalities of neural network structure and activity mediated by posttranslational modifications of proteins, activation of immediate early genes (IEGs), and other alterations in profiles of gene expression and function (e.g., GABAA receptor subunit, CREB, JAK-STAT, BDNF, and EGR3) that eventually lead to deregulated neural circuits with a predisposition for synchronous electrical activity" (Rakhade & Jensen, 2009).

Epilepsy is explained by not only alterations in genetic structure and environmental factors, but also epigenetic mechanisms. Recent studies show that, interconnected epigenetic regulatory networks play significant roles in modulation of cellular and organismal development, homeostasis and aging. Epigenetic mechanisms are also associated with transgenerational inheritance and "they are critical for promoting brain patterning, neural stem cell (NSC) maintenance and proliferation, neurogenesis and gliogenesis, cellular migration, and synaptic and neural network

connectivity and plasticity and are also implicated in orchestrating sophisticated cognitive functions including learning and memory" (Mehler, 2008).

Recent studies provide evidence for the fact that, DNA methylation, as a crucial epigenetic mechanism, results in a number of circumstances that are associated with epilepsy. DNA methyltransferase, an enzyme which catalyzes DNA methylation and methyl-CpG-binding domain proteins are regulated throughout a few number of processes (e.g., cell fate specification, maturation and survival and during activity-dependent synaptic plasticity) that can be associated with pathopsyiology of epilepsy (Sharma et al., 2008). Another study shows that epileptic syndromes that are associated with MeCP2 gene disruption are resulted by epigenetic dysregulations (Qureshi, 2010). High levels of DNA methylation in the promoter region of the critical neural factor, Reelin (RELN), are associated with the pathogenesis of temporal lobe epilepsy (Kobow et al., 2009).

2.5. Computational Methods

2.5.1. Prediction of Methylation Patterns

The relationship between genomic methylation patterns and underlying DNA sequence remains unclear. In order to provide an explanation for this relationship, different computational methods have been used to predict methylation patterns using DNA sequence information.

Mostly pattern recognition methods which rely on classifiers and clustering algorithms were used to predict methylation patterns. A classifier based on a Support Vector Machine (SVM) called MethCGI was developed. Classification features were nucleotide sequence contents as well as transcription factor binding sites (TFBSs) in this approach. The classifier achieved specificity of 84.65% and sensitivity of 84.32% on the brain data, additionally it two-third of the data from other tissues reported in MethDB database were correctly predicted by MethCGI (Fang et al., 2006).

Another method called Methylator also identifies support vector machine based classifier as better performing in comparison to classifiers built using alternative machine learning and statistical algorithms including artificial neural networks, Bayesian statistics, and decision trees (Bhasin et al., 2005). Datasets were generated by fragmenting the MethDB sequence patterns into overlapping fragments of fixed length (window size) and fragments with a methylated cytosine in the center were considered as positives, whereas fragments with non-methylated cytosine in the center were considered as negatives.

Support vector machine approach was reported as best performing classifier, in comparison to K means clustering, linear discriminant analysis and logistic regression, also in a study that presents another classifier called HDFINDER. It was reported that the classifier has a prediction accuracy of 86% on all 22 autosomal chromosomes (Das et al., 2006). For CpG islands they used 92 features including G+C content, di- and trinucleotide count, Alu coverage, and 10 hexamers and they applied recursive feature elimination which is a backward selection method and Principal Components Analysis (PCA) for feature subset reduction.

Another study on epigenetic alterations in human non-Hodgkin's lymphomas (NHL), attempted to discover methylation-expression relationships. A set of algorithms based on fuzzy sets theory, in particular Possibilistic c-Means (PCM) and cluster fuzzy density were used in the study. For each gene, these algorithms calculate measures of confidence of various methylation-expression relationships in each NHL subclass. The aim of the study was to provide a better understanding of the pathobiology of NHL and other types of cancer (Sjahputera et al., 2007).

In another analysis with a linear support vector machine predictor, three groups of DNA attributes, namely certain sequence patterns, specific DNA repeats, and predicted DNA structure, were found to be correlated with CpG island methylation with correlation coefficients of 0.64, 0.66, and 0.49, respectively (Bock et al., 2006).

2.5.2. Analysis of Methylation Data

Several modules were developed to be used in the analysis of Illumina Infinium methylation microarray data. These modules are mostly extensions of R package, which is a programming language package that is widely used in data mining for statistics and data analysis. Notable examples are IMA, which provides a pipeline that automates the tasks (Wang et al., 2012), lumi, which makes use of M-value instead of β value (Du et al., 2008) and HumMeth27QCReport which originally targets Illumina HumanMethylation27, but also is compatible with HumanMethylation450 Bead Chip (Mancuso et al., 2011).

CHAPTER 3

Materials and Methods

3.1. CpG Island Nucleotide Composition Determination

Analysis of CpG islands based on nucleotide sequences is a significant method for methylation analysis. It is frequently used for prediction of methylated CpG islands and determination of CpG islands that are prone to change in methylation levels.

In this study, methylated and unmethylated CpG island sequences were compared to determine if there exist any differences between them. In order to achieve this goal, methylated and unmethylated CpG island sequences are analyzed in different window sizes. Common motifs around methylated and unmethylated CpG islands are devised using the frequencies of motifs.

The data for this analysis is taken from MethDB database. Common motifs are found from the sequences using four different window sizes that are 14, 18, 22, and 30. The motifs are found in the form of dinucleotides and trinucleotides.

3.2. Microarray based Methylation Analysis

The foundations for high throughput analysis of single nucleotide polymorphisms and other genomic variants, was laid by the development of DNA microarray technology (Southern et al., 1999). DNA microarrays allow the whole

genome to be monitored on a single chip, so that the gene expression levels and the interaction between genes can be observed simultaneously (Roland, 2005; Brazma et al., 2000).

Three traditional methods of DNA methylation analysis, Immunoprecipitation, endonuclease digestion and sodium-bisulfite treatment, were all received a transformation with the development of DNA microarray technology. However the underlying principle regarding the microarray assays is same: Methylated and unmethylated fragments of the genome are separated and analyzed (Harrison & Parle-McDermott, 2011).

In this study, Infinium HumanMethylation450 BeadChip Kit by Illumina was used for methylation analysis. This platform is considered as a robust and reliable approach for DNA methylation analysis (Sandoval et al., 2011). This BeadChip is powered by the Infinium Methylation Assay which detects cytosine methylation at CpG islands based on highly multiplexed genotyping of bisulfite-converted genomic DNA (gDNA). Cytosine methylation can be detected because of the fact that, upon treatment with bisulfite, unmethylated cytosine bases are converted to uracil, while methylated cytosine bases remain unchanged (Illumina, 2012b).

Making use of two Methylation Assay chemistries (Illumina, 2012a), the BeadChip allows the interrogation of 485,000 methylation sites per sample at single-nucleotide resolution, thus it provides coverage of 99% of RefSeq genes, 96% of CpG islands, with multiple sites within islands and island shores. HumanMethylation450 BeadChip workflow does not require PCR, the kit includes all required reagents for methylation analysis, except for the bisulfite conversion kit, which is available separately (Illumina, 2012b).

3.3. GenomeStudio Software

GenomeStudio Data Analysis Software is used to visualize and analyze data generated by all platforms of Illumina. The GenomeStudio Software has seven application modules which can be added to the software. Application modules include;

DNA sequencing module, RNA sequencing module, ChIP sequencing module, Genotyping module, Gene expression module, Methylation module, Protein analysis module.

The software and its packages allow researchers to obtain a comprehensive view of the genome, gene expression and regulation. The data on the GenomeStudio software can be exported to a number of third-party software tools for further analysis. The GenomeStudio also allows the export of data to .csv format, which allows specific analysis of the data (Illumina, 2011).

Illumina HumanMethylation450 Bead Chip provides signal intensities of each CpG location from methylation assays. This information is provided in raw data format and β values calculated from raw signal intensities are provided in GenomeStudio project. Sample GenomeStudio data is shown in Table 3.1.

TargetID	Sample 1 AVG_Beta	Sample 1 Intensity	Sample 2 AVG_Beta	Sample 2 Intensity
cg00000029	0.4812995	1824.655	0.5044388	2351.673
cg00000108	0.9432696	5668.381	0.9389272	5423.673
cg00000109	0.8908905	2541.251	0.8681602	2197.432
cg00000165	0.1235855	1938.497	0.1880388	2583.251
cg00000236	0.7291482	2838.623	0.6898252	2644.511

Table 3.1 Sample GenomeStudio file with β values

3.3.1. CpG Loci Identification

Since CpG loci was lacking a formal method of nomenclature, unlike other genomic loci, Illumina developed a method to consistently designate CpG loci based on the actual or contextual sequence of each individual CpG locus.

Illumina's CpG loci database provides unique identification for each CpG locus and it can be accessed by HumanMethylation450 manifest file (Illumina, 2010).

3.4. Data

The case-parent trio design is a method that widely used to detect genetic associations (Yu, 2011). In this design, child and one parent are affected and other parent is healthy, therefore healthy parent is considered as control parent. Use of case-parent trio design may play a crucial role to address population structure bias (Infante-Rivard et al., 2009). This design may play a crucial role in the researches regarding the analysis of transgenerational epigenetics, when it is used in a study on DNA methylation, which is an epigenetic system instead of a study on genetic inheritance.

The data used in this thesis were provided in the scope of IntenC project by Istanbul University Capa Faculty of Medicine. The data is gathered from 30 individuals from 10 family trios, which consists of an offspring who has the disorder, one parent who has the disorder and the other parent as control. Clinic information regarding the disorder, phenotypes and age information are provided for each individual.

DNA methylation analysis results from HumanMethylation450 BeadChip are provided. These results include β values and Intensity scores for each CpG locus. Methylation level on a CpG locus is expressed by β value. This data can be visualized in Illumina's GenomeStudio. Using export option in the software, data was converted to a comma-separated file (a file in .csv format). While rows represent each CpG locus in this file, columns indicate different individuals.

3.5. Implementation

3.5.1. Determination of Methylation Status

Average β values indicate the methylation level in a locus. In order to determine whether a loci is methylated or not, it was necessary to decide on particular range values for methylated status and non-methylated status. Since β value has a direct correspondence with methylation and its value can be interpreted as the percentage of methylation at a given loci, the necessary ranges were determined as the following: β

value between 0.0 and 0.2 indicates non-methylation and β value between 0.8 and 1.0 indicates methylation (Du et al., 2010).

3.5.2. Three Types of Analysis

For the analysis, each family trio considered separately. We decided on two different approaches which takes account methylation status information of each individual of a family trio. For each approach we created a CpG list according to the criteria in **Error! Reference source not found.**.

	First approach	**Second approach**
Offspring	Methylated	Non-methylated
Parent with disorder	Methylated	Non-methylated
Control parent	Non-methylated	Methylated

Table 3.2 Methylation statuses of a CpG locus for each individual in a family trio

3.5.2.1. First approach

CpG loci should be methylated in the offspring and the parent with disorder, and it should be non-methylated in the control parent.

3.5.2.2. Second approach

CpG loci should be non-methylated in the offspring and the parent with disorder, and it should be methylated in the control parent.

3.5.2.3. Third approach

This approach is combination of first and second approaches: if there is a difference in the methylation status between control parent and other two samples with disorder (child and parent with disorder) in a CpG locus, that CpG locus is marked.

For all three approaches, we marked each CpG loci that fits the criteria, and formed two lists that consist of CpG loci identification codes according to Illumina CpG loci database. Manifest file of HumanMethylation450 BeadChip also indicates the genes that are associated with each CpG loci. This information allowed us to generate a list of associated genes for each approach.

3.5.3. Filtering

After two lists (one with CpG identification codes, another with the list of associated genes) were generated according to both approaches that described in the previous section, we proceeded to filtering the dataset.

3.5.3.1. First filter

Probes with no β value, all rs and ch probes and all sex chromosome probes were removed from the dataset. Probes which are marked as rs indicates probes that measure SNPs, not probes containing SNPs and ch probes target non-CpG methylation.

Among the autosomal probes, some of them display significant methylation differences between males and females. According to the list which was generated by another study (Price et al., 2013) autosomal probes with sex-differences were also removed from the dataset.

After filtering the dataset based on previously described filtering criteria, we repeated the process that allows us to generate lists that include CpG identification codes and list of associated genes.

3.5.3.2. Second filter

In addition to the filtering previously described, we also decided to apply an extra filtering, and repeat the generation of lists. This filtering includes the removal of

cross-reactive probes and polymorphic CpGs. While polymorphic CpGs indicate the CpGs that overlap known SNPs, cross-reactive probes target repetitive sequences or they co-hybridize to alternate sequences that are highly homologous to the intended targets. Because of that, cross-reactive probes could generate spurious signals and they potentially lead to invalid conclusions (Chen et al., 2013). Dataset was filtered using the lists of cross-reactive probes and polymorphic CpGs that generated in a related study (Chen et al., 2013). After the extra filtering the lists of CpG identification codes and associated genes were generated again by repeating the process for both approaches.

3.5.3.3. Third filter

A third filter was developed with the aim of removal of age bias in dataset. It was necessary to find CpG loci in which the methylation level changes between individuals in correlation with their age. Since the CpG dataset is too large, Principal Components Analysis (PCA) was used to transform the data into principal components.

PCA was implemented in MatLab using the following function:

$$[coeff, score, latent] = princomp(x,' econ') \quad (3.1)$$

In this function *x* denotes our methylation level data, where rows denote each individual as observations and columns denote each CpG loci as variables. In the results, *coeff* denotes a matrix where the number of initial variables is the number of rows, and the number of principal components is the number of columns. It shows the coefficient of each variable in each principal component. Score denotes a matrix where number of observations is rows and number of principal components is columns. It shows the transformed version of the initial data. Latent explains the variance covered by each principal component. Total variance can be explained by the total number of variables.

In our analysis, 29 principal components were found. For each principal component, its correlation with age is determined by Pearson's linear correlation

coefficient. Therefore each of these principal component scores was analyzed against age data.

Only the 8th principal component shows low correlation with age with a significant p-value. Correlation is low and its latent value indicates that this principal component covers very low variance of the data; therefore it is safe to conclude that our data is not correlated with age.

3.5.4. Statistical Analysis

3.5.4.1. Wilcoxon rank-sum test

Wilcoxon Rank-Sum Test is a non-parametric test for two populations which tests null hypothesis that two populations are from continuous distributions with equal medians, against the alternative that they are not. Samples should be independent for the test. Wilcoxon Rank-Sum Test is equivalent to Mann-Whitney U Test.

The test was implemented in MatLab using ranksum(x,y) function.

$$[p, h] = ranksum(x, y) \tag{3.2}$$

In this function, p denotes the p-value of the two-sided Wilcoxon Rank-Sum test, and h denotes the logical value indicating the decision of the test. The result h = 0 indicates a failure to reject the null hypothesis at the 5% significance level and h = 1 indicates the rejection of the null hypothesis.

3.5.4.2. Student's t-test

Student's t-test is another statistical hypothesis test that can be used to compare two populations. It was introduced by William Sealy Gosset. Null hypothesis suggests that two populations come from distributions with equal means and equal variances. The alternative hypothesis is that the data from two populations have unequal means.

The test was implemented in MatLab using ttest2(x,y) function.

$$[h, p] = ttest2(x, y) \tag{3.3}$$

In this function, h denotes the test decision and p denotes the p-value of the test. The result h = 1 indicates that the null hypothesis is rejected with 5% significance level, and results h = 0 indicates otherwise.

3.5.5. Pathways

Biochemical pathways provide crucial information on the functions of individual genes and proteins. They highlight the systems and processes that contribute to normal physiology, as well as disease. Thus, the pathway-level analysis is a powerful approach to understand complex biological systems at multiple levels of biological organization (Bakir-Gungor et al., 2013).

In order to devise epilepsy-related pathways using the list of genes we generated from the methylation dataset, the following pathway analysis approaches were used:

3.5.5.1. DAVID

DAVID is a database which includes an integrated biological knowledgebase and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists (Huang et al., 2009). DAVID is able to visualize genes on pathway maps including KEGG, BioCarta, Reactome and Panther pathways. In order to use the web-tool, the list of genes is required to be uploaded and DAVID presents related results afterwards.

3.5.5.2. GOrilla

GOrilla is a web-based application that identifies enriched GO terms in ranked lists of genes. In order to use GOrilla, a ranked list of genes is required to be uploaded

to the server. GOrilla visualizes the output of the enrichment analysis as a hierarchical structures, therefore it provide a clear view of the relations between enriched GO terms (Eden et al., 2009).

3.5.5.3. KEGG

Kyoto Encyclopedia of Genes and Genomes (KEGG) is defined as a knowledge base for systematic analysis of gene functions, linking genomic information with higher order functional information (Kanehisa & Goto, 2000). The PATHWAY database stores the higher order functional information which contains graphical representations of cellular processes, such as metabolism, membrane transport, signal transduction and cell cycle. The PATHWAY database includes 2706 entries for pathway diagrams constructed from 143 manually drawn diagrams. According to Kanehisa and Goto (2000), the best organized part of the PATHWAY database is metabolism, which is represented by approximately 90 graphical diagrams.

3.5.6. PANOGA

3.5.6.1. Definition of the protocol

PANOGA (Pathway and Network Oriented GWAS Analysis) is a protocol that is designed to devise functionally important KEGG pathways through the identification of SNP targeted genes within these pathways. The original protocol begins with a list of SNPs that found to be associated with disease in Genome-Wide Association Studies (GWAS).

The protocol combines evidence from the following sources: i) genetic association information obtained through GWAS, ii) SNP functional information, iii) protein-protein interaction network, iv) linkage disequilibrium (LD), v) bio-chemical pathways. This multidimensional perspective strengthens the methodology (Bakir-Gungor & Sezerman, 2011).

The flow of the protocol is shown in
Figure 3.1.

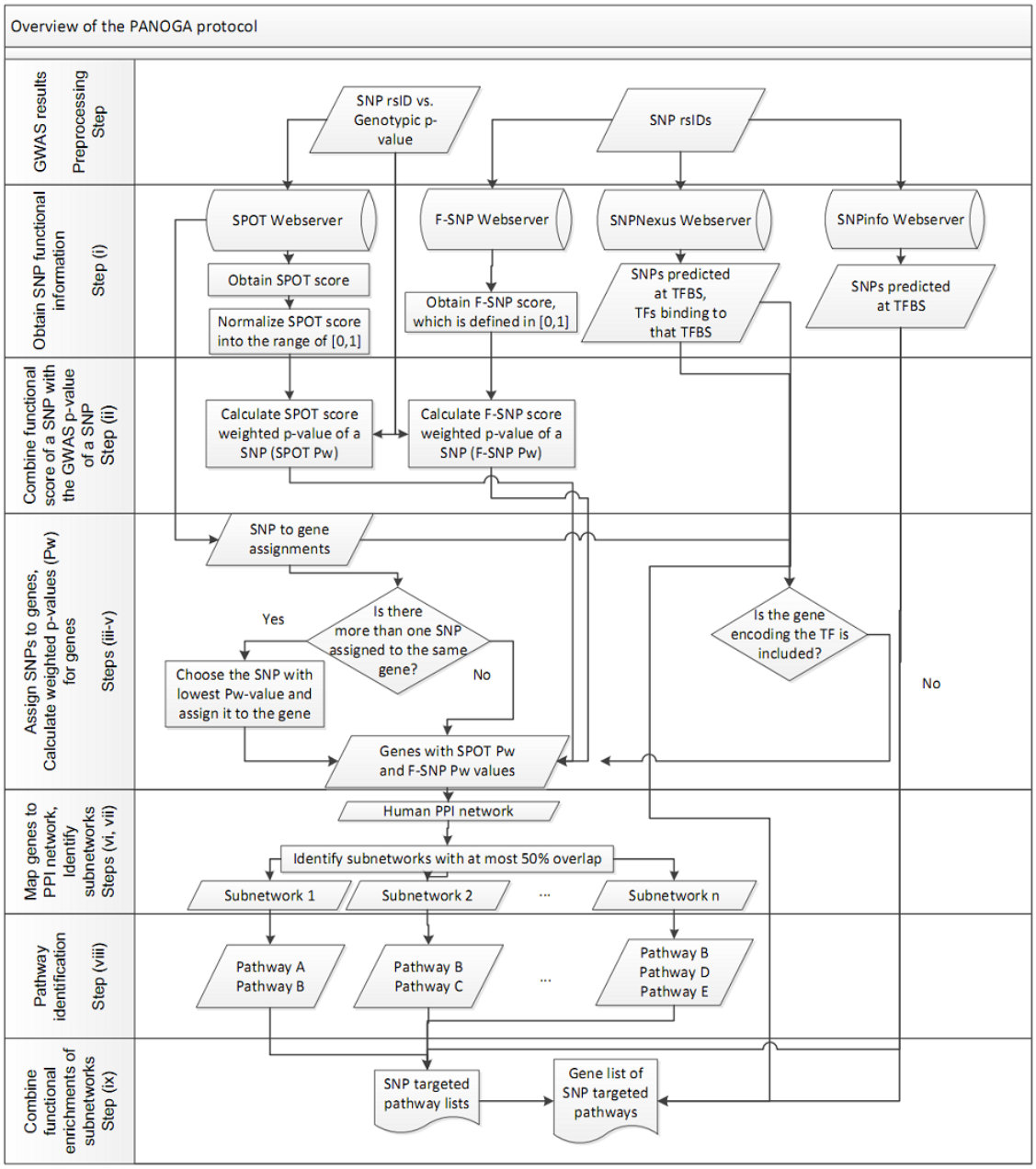

Figure 3.1 Overview of PANOGA Protocol (Bakir-Gungor et al., 2013)

3.5.6.2. Related tools and applications

3.5.6.2.1. F-SNP

F-SNP database provides integrated information about the functional effects of SNPs obtained from a number of different bioinformatics tools and servers (Lee & Shatkay, 2007).

3.5.6.2.2. SPOT

SPOT is a web-server for integrating biological databases into the prioritization of SNPs for further study after a genome-wide association study (GWAS) (Saccone et al., 2010).

3.5.6.2.3. SNPNexus

SNPNexus is a web-server that assesses the potential significance of known and novel SNPs on the major transcriptome, proteome, regulatory and structural variation models in order to identify the phenotypically important variants (Ullah et al., 2012).

3.5.6.2.4. SNPInfo

SNPInfo is another web-server that is used to get functional information of SNPs (Xu & Taylor, 2009).

3.5.6.2.5. Cytoscape

Cytoscape is an open source software platform for visualizing molecular interaction networks and biological pathways. It is also used for integrating these networks with annotations, gene expression profiles and other state data (Shannon et al., 2003). As an open source software platform, Cytoscape allows and encourages to

development of plugins that can be integrated into it. Following plugins were used in PANOGA.

3.5.6.2.5.1. jActive Modules

jActive Modules plugin is used to identify active sub-networks of interacting gene products that were also associated with the disease (Ideker et al., 2002).

3.5.6.2.5.2. ClueGO

ClueGO plugin is a functional enrichment tool that extracts the non-redundant biological information for groups of genes using GO, KEGG and BioCarta ontologies (Bindea et al., 2009). Among different ontologies, since KEGG database primarily categorizes genes into bona-fide biological pathways; and since biological interpretation of pathways is more straightforward compared to GO terms, functional enrichment results are only reported using KEGG pathways in PANOGA (Bakir-Gungor & Sezerman, 2011).

3.5.6.2.6. KEGG

As previously described, the PATHWAY database of KEGG stores functional information with graphical representations of cellular processes. In PANOGA protocol, most significant 10 pathways are chosen by their p-values, and the information about those pathways are uploaded to KEGG PATHWAY web-tool. Visualized results are provided and can be saved.

3.5.6.3. Complete implementation

PANOGA Protocol requires to access different websites in some of its steps, using previous step outputs as next step inputs. For this requirement, the protocol has

been automated. The automation both shortened the time that takes to apply the protocol and got rid of the errors caused by human interaction.

With the data gathered from listening and analyzing network protocols, POST and GET request methods were learned for most of the websites included in the protocol. Inputs were given with POST methods and outputs were gathered with GET methods.

A problem has arisen for the websites that had complicated request methods and used more complex systems that didn't allow for them to be sniffed. Web crawlers were used to overcome this problem as to simulate a real user browsing and using those websites.

For various steps of this protocol, software was developed with different working conditions and in different programming languages. In order to provide steps working flawlessly with each other, all software was re-implemented in Java with necessary upgrades.

The software supports working with multiple threads, so it's possible to have more than one active run theoretically. Yet, the final steps of the protocol require all available resources of the computer and also include the step that has the longest computation time amongst all the steps. Because of this bottleneck, the software uses only one thread at the moment.

The software also records its progress into separate log files. A simpler log file is used to notify the owner of the current run, and a detailed log file is also used to track the software from a developer’s point of view.

The protocol also required for local software to use the obtained data from websites, so the developed software was not only automated the flow between websites but also set the software up with the required inputs and parameters.

For a user to use this protocol from web, this software was put inside a webserver and a website to use this protocol has been generated in PHP. The website

takes user inputs to record them into a database and it also takes the user files to be uploaded into the server. The software periodically checks whether new entries are made and works through the protocol when it detects a new entry. The website also generates a unique id and unique link for the user to check their status by reading log files mentioned above. With this way, the database provides the record of all the protocol runs made by the website. The website is also used for contacting and giving brief information about the protocol itself.

A Linux based web host with terminal access was the first candidate for Panoga Protocol. First implementation and the tests for the website and the software were carried out in this environment. The developed software had no problems regarding Linux. But as Cytospace is a commercial product, although it supports working through command line with some alterations, the necessary add-ons used in the protocol did not; so web based hosting wasn't holding the requirements.

A Windows based virtual private server was the second candidate. With the experience from the first candidate, the system was transferred and started working quickly, and the results were satisfying. A complete run of the protocol has been initiated successfully, and the website was made online.

3.5.6.4. Partial use

Since PANOGA Protocol begins with a list of SNPs and makes use of their functional information, it wasn't a viable tool for devising functionally important pathways from a list of genes. On the other hand, we were able to skip the steps that are related to SNPs, and proceed with gene information only with small modifications in the code. It was also necessary to perform a calculation to determine significance score for each gene to provide ranking information for the network in jActive Modules. Methylation β values for the individuals from a family trio on the CpG locus that is associated with the chosen gene are used in the calculation of a score. Because of the fact that two different list of genes were generated using different information regarding methylation, it was also required to devise two different calculation methods.

Considering β value having a direct correspondence with methylation, the score calculations in Table 3.3 are applied.

First approach	((Offspring – 0.5) + (Parent – 0.5) + (0.5 – Control)) / 1.5
Second approach	((0.5 – Offspring) + (0.5 – Parent) + (Control – 0.5)) / 1.5

Table 3.3 Offspring: β value for the offspring on a locus, Parent: β value for the parent with disorder on a locus, Control: β value for the control parent on a locus

If more than one CpG loci are associated with a gene, only the highest score is used.

The flow of our approach begins with jActive Modules step, which identifies active sub-networks of interacting gene products, and continues with the protocol until the end where the important KEGG pathways are identified. The most significant 10 pathways are chosen according to their p-values, and their information are used for KEGG pathway identification.

3.5.7. Promoter Analysis

Since DNA methylation in promoter regions is considered crucial, its analysis is necessary to provide understanding about relationships between DNA methylation, gene expression and underlying pathways related to a disease. Therefore, we expected the results of promoter analysis to be more significant than a general analysis, and focused on promoter analysis while repeating each step for general analysis.

Illumina categorizes probes into six gene feature groups, namely TSS1500 (within 1500 bps of a transcription start site (TSS)), TSS200 (within 200 bps of a TSS), 5'UTR (5' untranslated region), first exon, body and 3'UTR (3' untranslated region). Manifest file of HumanMethylation450 Bead Chip provides information about the groups of CpG loci.

In order to accomplish a promoter based analysis, we determined CpG loci which belong to TSS1500 and TSS200 groups. We generated lists of genes only from

those CpG loci and applied all three filters separately and repeated the remaining parts of our analysis.

3.5.8. Normalization

Illumina Infinium HumanMethylation450 Bead Chip arrays include two different probe design types. Since some percentage of CpG loci are associated with type1 and the rest are associated with type2, it may result in biased conclusions in methylation analysis. In order to remove this bias, a normalization process may be required.

In this study, Beta-Mixture Quantile Normalization (BMIQ) is applied for normalization and its results are presented separately. BMIQ applies a normalization strategy that adjusts the β values of type2 design probes into a statistical distribution characteristic of type1 probes (Teschendorff et al., 2013). The method is presented in R language, and in this study, it was applied using R Software which is a free software environment for statistical computing.

3.5.9. Quality Checking

In the methylation dataset, three of the samples which were gathered from microarray analysis by HumanMethylation450 Bead Chip have non-satisfactory quality. The samples of the affected parent in family trios 4 and 7 have low signal intensities and the sample of the child in family trio 7 had difficulties in the bisulfite conversion step. Therefore, samples of family trio 4 and family trio 7 were removed from dataset and a few runs were repeated to check the effects of the removed data.

CHAPTER 4

Results

4.1. CpG Island Nucleotide Composition Determination

Common motifs around methylated and unmethylated CpG islands were determined. Most common motifs are chosen according to their frequencies and common motifs that have at least 0.1 difference in frequency are presented in Table 4.1.

	14				18				22				30			
	a		b		a		b		a		b		a		b	
	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2	3
before	tt	gca	cc cg	agc cag ccc ccg	gc tg	tgg	tc	cca ccc ccg ctc ggc tcc tgc	cc gc		ac tc	ccc ccg cgc gcc ggc	ac cc cg gc		tc	agc ccg cgc gcc ggc tcc
after	At Ca Gc		cg	ccc cct cgc cgg gag gcc ggc	ca cc gc ta	gct gtg		agc ccc ccg gcc ggc	ca cc gc			ccc ccg cgc cgg gcc ggc	ac at ca cc cg gc tc			ccg cgc cgg gag gcc ggc

Table 4.1 List of common motifs around methylated (a) and unmethylated (b) CpG islands for different window sizes

4.2. Application of First Filter

Firstly, results of the analysis without any filters are presented in APPENDIX A.

After applying the first filter which removes probes with no β value, all rs and ch probes, all sex chromosome probes and autosomal probes with sex differences; we gathered pathway analysis results also from partial implementation of PANOGA protocol as well as David and Gorilla.

4.2.1. General Results

4.2.1.1. DAVID results

Detailed results for the analysis with application of first filter on DAVID are shown in APPENDIX B. Most frequently devised pathways and the number of families

they are found are presented in Table 4.2 and Table 4.3 for KEGG, Table 4.4 and Table 4.5 for BioCarta, Table 4.6 and Table 4.7 for Panther, Table 4.8 and Table 4.9 for Reactome.

KEGG Pathways	Count
hsa04514:Cell adhesion molecules (CAMs)	5
hsa04612:Antigen processing and presentation	5
hsa05310:Asthma	4
hsa05332:Graft-versus-host disease	4
hsa05320:Autoimmune thyroid disease	4
hsa05322:Systemic lupus erythematosus	4
hsa05330:Allograft rejection	4
hsa04672:Intestinal immune network for IgA production	4
hsa04940:Type I diabetes mellitus	4
hsa05416:Viral myocarditis	4
hsa04530:Tight junction	3
hsa04640:Hematopoietic cell lineage	3
hsa04916:Melanogenesis	2
hsa04080:Neuroactive ligand-receptor interaction	2
hsa04310:Wnt signaling pathway	2
hsa04510:Focal adhesion	2
hsa04810:Regulation of actin cytoskeleton	2
hsa04270:Vascular smooth muscle contraction	2
hsa03440:Homologous recombination	2
hsa05200:Pathways in cancer	2
hsa04020:Calcium signaling pathway	2
hsa04650:Natural killer cell mediated cytotoxicity	2

Table 4.2 Number of KEGG pathways (first approach/first filter)

KEGG Pathways	Count
hsa04672:Intestinal immune network for IgA production	6
hsa05320:Autoimmune thyroid disease	5
hsa04940:Type I diabetes mellitus	5
hsa04514:Cell adhesion molecules (CAMs)	5
hsa05416:Viral myocarditis	5
hsa04612:Antigen processing and presentation	5
hsa05310:Asthma	5
hsa05322:Systemic lupus erythematosus	5
hsa05332:Graft-versus-host disease	5
hsa05330:Allograft rejection	5
hsa04740:Olfactory transduction	4
hsa05200:Pathways in cancer	3
hsa04530:Tight junction	3
hsa04310:Wnt signaling pathway	3
hsa04510:Focal adhesion	3
hsa04010:MAPK signaling pathway	3
hsa04070:Phosphatidylinositol signaling system	3
hsa04650:Natural killer cell mediated cytotoxicity	3
hsa04666:Fc gamma R-mediated phagocytosis	3
hsa04370:VEGF signaling pathway	2
hsa05214:Glioma	2
hsa05130:Pathogenic Escherichia coli infection	2
hsa04720:Long-term potentiation	2
hsa04640:Hematopoietic cell lineage	2
hsa04730:Long-term depression	2
hsa00512:O-Glycan biosynthesis	2
hsa04144:Endocytosis	2
hsa04270:Vascular smooth muscle contraction	2
hsa04912:GnRH signaling pathway	2
hsa05223:Non-small cell lung cancer	2
hsa04916:Melanogenesis	2
hsa04012:ErbB signaling pathway	2
hsa04540:Gap junction	2
hsa04020:Calcium signaling pathway	2
hsa04960:Aldosterone-regulated sodium reabsorption	2
hsa05110:Vibrio cholerae infection	2
hsa04670:Leukocyte transendothelial migration	2
hsa04664:Fc epsilon RI signaling pathway	2

Table 4.3 Number of KEGG pathways (second approach/first filter)

BioCarta Pathways	Count
h_il5Pathway:IL 5 Signaling Pathway	3
h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation	3
h_mhcPathway:Antigen Processing and Presentation	3
h_asbcellPathway:Antigen Dependent B Cell Activation	3
h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy	3
h_bbcellPathway:Bystander B Cell Activation	3
h_inflamPathway:Cytokines and Inflammatory Response	3
h_blymphocytePathway:B Lymphocyte Cell Surface Molecules	3
h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation	3
h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor	3
h_th1th2Pathway:Th1/Th2 Differentiation	3
h_metPathway:Signaling of Hepatocyte Growth Factor Receptor	2
h_integrinPathway:Integrin Signaling Pathway	2

Table 4.4 Number of BioCarta pathways (first approach/first filter)

BioCarta Pathways	Count
h_inflamPathway:Cytokines and Inflammatory Response	3
h_TPOPathway:TPO Signaling Pathway	3
h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway	3
h_ghPathway:Growth Hormone Signaling Pathway	3
h_myosinPathway:PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase	2
h_vegfPathway:VEGF, Hypoxia, and Angiogenesis	2
h_pyk2Pathway:Links between Pyk2 and Map Kinases	2
h_bcrPathway:BCR Signaling Pathway	2
h_ionPathway:Ion Channel and Phorbal Esters Signaling Pathway	2
h_arenrf2Pathway:Oxidative Stress Induced Gene Expression Via Nrf2	2
h_pkcPathway:Activation of PKC through G protein coupled receptor	2
h_blymphocytePathway:B Lymphocyte Cell Surface Molecules	2
h_tcrPathway:T Cell Receptor Signaling Pathway	2
h_calcineurinPathway:Effects of calcineurin in Keratinocyte Differentiation	2
h_il5Pathway:IL 5 Signaling Pathway	2
h_cardiacegfPathway:Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy	2
h_mef2dPathway:Role of MEF2D in T-cell Apoptosis	2
h_cblPathway:CBL mediated ligand-induced downregulation of EGF receptors	2
h_Par1Pathway:Thrombin signaling and protease-activated receptors	2
h_CCR3Pathway:CCR3 signaling in Eosinophils	2
h_plcPathway:Phospholipase C Signaling Pathway	2
h_Ccr5Pathway:Pertussis toxin-insensitive CCR5 Signaling in Macrophage	2
h_srcRPTPPathway:Activation of Src by Protein-tyrosine phosphatase alpha	2
h_cdMacPathway:Cadmium induces DNA synthesis and proliferation in macrophages	2
h_bbcellPathway:Bystander B Cell Activation	2
h_chemicalPathway:Apoptotic Signaling in Response to DNA Damage	2
h_agpcrPathway:Attenuation of GPCR Signaling	2
h_crebPathway:Transcription factor CREB and its extracellular signals	2
h_At1rPathway:Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling	2
h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor	2
h_keratinocytePathway:Keratinocyte Differentiation	2
h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation	2
h_mhcPathway:Antigen Processing and Presentation	2
h_cxcr4Pathway:CXCR4 Signaling Pathway	2
h_nos1Pathway:Nitric Oxide Signaling Pathway	2
h_edg1Pathway:Phospholipids as signalling intermediaries	2
h_pdgfPathway:PDGF Signaling Pathway	2
h_egfPathway:EGF Signaling Pathway	2
h_plcdPathway:Phospholipase C d1 in phospholipid associated cell signaling	2
h_eif4Pathway:Regulation of eIF4e and p70 S6 Kinase	2

h_pparaPathway:Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)	2
h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy	2
h_sppaPathway:Aspirin Blocks Signaling Pathway Involved in Platelet Activation	2
h_erbB4pathway:g-Secretase mediated ErbB4 Signaling Pathway	2
h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation	2
h_th1th2Pathway:Th1/Th2 Differentiation	2
h_telPathway:Telomeres, Telomerase, Cellular Aging, and Immortality	2
h_gpcrPathway:Signaling Pathway from G-Protein Families	2
h_asbcellPathway:Antigen Dependent B Cell Activation	2
h_trkaPathway:Trka Receptor Signaling Pathway	2

Table 4.5 Number of BioCarta pathways (second approach/first filter)

Panther Pathways	Count
P00057:Wnt signaling pathway	4
P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	4
P00031:Inflammation mediated by chemokine and cytokine signaling pathway	3
P00049:Parkinson disease	3
P00047:PDGF signaling pathway	3
P00010:B cell activation	3
P00012:Cadherin signaling pathway	3
P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	3
P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway	2
P00002:Alpha adrenergic receptor signaling pathway	2
P00019:Endothelin signaling pathway	2
P00053:T cell activation	2
P00041:Metabotropic glutamate receptor group I pathway	2
P00056:VEGF signaling pathway	2
P00018:EGF receptor signaling pathway	2
P04374:5HT2 type receptor mediated signaling pathway	2
P00021:FGF signaling pathway	2
P00006:Apoptosis signaling pathway	2
P00003:Alzheimer disease-amyloid secretase pathway	2
P04385:Histamine H1 receptor mediated signaling pathway	2
P00005:Angiogenesis	2
P04391:Oxytocin receptor mediated signaling pathway	2
P04394:Thyrotropin-releasing hormone receptor signaling pathway	2
P00029:Huntington disease	2
P00034:Integrin signalling pathway	2
P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin	2

Table 4.6 Number of Panther pathways (first approach/first filter)

Panther Pathways	Count
P00053:T cell activation	6
P00057:Wnt signaling pathway	4
P04394:Thyrotropin-releasing hormone receptor signaling pathway	3
P00003:Alzheimer disease-amyloid secretase pathway	3
P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway	3
P00005:Angiogenesis	3
P04385:Histamine H1 receptor mediated signaling pathway	3
P00006:Apoptosis signaling pathway	3
P00041:Metabotropic glutamate receptor group I pathway	3
P00010:B cell activation	3
P00047:PDGF signaling pathway	3
P00018:EGF receptor signaling pathway	3
P00056:VEGF signaling pathway	3
P00019:Endothelin signaling pathway	3
P04374:5HT2 type receptor mediated signaling pathway	3
P00021:FGF signaling pathway	3
P04391:Oxytocin receptor mediated signaling pathway	3
P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	3
P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin	3
P00031:Inflammation mediated by chemokine and cytokine signaling pathway	3
P00002:Alpha adrenergic receptor signaling pathway	3
P00029:Huntington disease	2

Table 4.7 Number of Panther pathways (second approach/first filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	3
REACT_6900:Signaling in Immune system	3
REACT_71:Gene Expression	2

Table 4.8 Number of Reactome pathways (first approach/first filter)

Reactome Pathways	Count
REACT_6900:Signaling in Immune system	5
REACT_604:Hemostasis	3
REACT_14797:Signaling by GPCR	3
REACT_15380:Diabetes pathways	3
REACT_13685:Synaptic Transmission	2
REACT_17015:Metabolism of proteins	2

Table 4.9 Number of Reactome pathways (second approach/first filter)

4.2.1.2. PANOGA results

Even though PANOGA partial implementation did not yield any results for family trios separately with first filter applied, when we combined the lists of genes of each family trio, the results in Table 4.10, Table 4.12 and Table 4.12 are obtained.

First Approach
Metabolic pathways
Pathways in cancer
Cell cycle
Neurotrophin signaling pathway
MAPK signaling pathway
Regulation of actin cytoskeleton
Insulin signaling pathway
Ribosome
ErbB signaling pathway
Apoptosis

Table 4.10 PANOGA results of general analysis with first filter (first approach)

Second Approach
Complement and coagulation cascades
Cholinergic synapse
GABAergic synapse
Retrograde endocannabinoid signaling
TGF-beta signaling pathway
Type II diabetes mellitus
Aldosterone-regulated sodium reabsorption
Bacterial invasion of epithelial cells
Glioma
Citrate cycle (TCA cycle)

Table 4.11 PANOGA results of general analysis with first filter (second approach)

Third Approach
Cell cycle
Chemokine signaling pathway
Ribosome
Glutamatergic synapse
Cholinergic synapse
TGF-beta signaling pathway
Retrograde endocannabinoid signaling
Insulin signaling pathway
Renal cell carcinoma
Long-term potentiation

Table 4.12 PANOGA results of general analysis with first filter (third approach)

4.2.2. Promoter Results

4.2.2.1. DAVID results

Detailed results for the analysis for promoter region with application of first filter on DAVID are shown in APPENDIX B. Most frequently devised pathways and the number of families they are found are presented in Table 4.13 and Table 4.14 for KEGG, Table 4.15 for Panther, Table 4.16 and Table 4.17 for Reactome. Most frequently found pathways for BioCarta are not reported for this analysis, because none of the families were found in more than one family.

KEGG Pathways	Count
hsa05322:Systemic lupus erythematosus	2
hsa05310:Asthma	2
hsa05332:Graft-versus-host disease	2
hsa04080:Neuroactive ligand-receptor interaction	2
hsa05320:Autoimmune thyroid disease	2
hsa04514:Cell adhesion molecules (CAMs)	2
hsa05330:Allograft rejection	2
hsa04612:Antigen processing and presentation	2
hsa05416:Viral myocarditis	2
hsa04940:Type I diabetes mellitus	2
hsa04672:Intestinal immune network for IgA production	2

Table 4.13 Number of KEGG pathways in promoter analysis (first approach/first filter)

KEGG Pathways	Count
hsa04740:Olfactory transduction	4

Table 4.14 Number of KEGG pathways in promoter analysis (second approach/first filter)

Panther Pathways	Count
P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	2
P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	2

Table 4.15 Number of Panther pathways in promoter analysis (first approach/first filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	3

Table 4.16 Number of Reactome pathways in promoter analysis (first approach/first filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	3

Table 4.17 Number of Reactome pathways in promoter analysis (second approach/first filter)

4.2.2.2. PANOGA results

Even though PANOGA partial implementation did not yield any results for family trios separately with first filter applied, when we combined the lists of genes of each family trio, the results in Table 4.18, Table 4.19 and Table 4.20 are obtained.

First Approach
Pathways in cancer
Spliceosome
Chemokine signaling pathway
Cell cycle
Neurotrophin signaling pathway
Ribosome
Focal adhesion
Proteasome
T cell receptor signaling pathway
Chronic myeloid leukemia

Table 4.18 PANOGA results of promoter analysis with first filter (first approach)

Second Approach
Pathways in cancer
Spliceosome
Chemokine signaling pathway
Cell cycle
Neurotrophin signaling pathway
Ribosome
Focal adhesion
Proteasome
T cell receptor signaling pathway
Chronic myeloid leukemia

Table 4.19 PANOGA results of promoter analysis with first filter (second approach)

Third Approach
Cell cycle
Nucleotide excision repair
Chemokine signaling pathway
Ribosome
Complement and coagulation cascades
Glutamatergic synapse
Cholinergic synapse
TGF-beta signaling pathway
Retrograde endocannabinoid signaling
Proteasome

Table 4.20 PANOGA results of promoter analysis with first filter (third approach)

4.3. Application of Second Filter

After the results of first filter are gathered, we applied the second filter which removes cross-reactive probes and polymorphic CpGs.

4.3.1. General Results

4.3.1.1. DAVID results

Detailed results for the analysis with application of second filter on DAVID are shown in APPENDIX B. Most frequently devised pathways and the number of families they are found are presented in Table 4.21 and Table 4.22 for KEGG, Table 4.23 and Table 4.24 for BioCarta, Table 4.25 and Table 4.26 for Panther, Table 4.27 and Table 4.28 for Reactome.

KEGG Pathways	Count
hsa05322:Systemic lupus erythematosus	3
hsa05310:Asthma	3
hsa05332:Graft-versus-host disease	3
hsa04514:Cell adhesion molecules (CAMs)	3
hsa05320:Autoimmune thyroid disease	3
hsa04612:Antigen processing and presentation	3
hsa05330:Allograft rejection	3
hsa04672:Intestinal immune network for IgA production	3
hsa05416:Viral myocarditis	3
hsa04940:Type I diabetes mellitus	3
hsa04530:Tight junction	2
hsa04640:Hematopoietic cell lineage	2

Table 4.21 Number of KEGG pathways (first approach/second filter)

KEGG Pathways	Count
hsa04740:Olfactory transduction	3
hsa04672:Intestinal immune network for IgA production	2

Table 4.22 Number of KEGG pathways (second approach/second filter)

BioCarta Pathways	Count
h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation	2
h_asbcellPathway:Antigen Dependent B Cell Activation	2
h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation	2
h_bbcellPathway:Bystander B Cell Activation	2
h_mhcPathway:Antigen Processing and Presentation	2
h_blymphocytePathway:B Lymphocyte Cell Surface Molecules	2
h_th1th2Pathway:Th1/Th2 Differentiation	2
h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor	2
h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy	2

Table 4.23 Number of BioCarta pathways (first approach/second filter)

BioCarta Pathways	Count
h_inflamPathway:Cytokines and Inflammatory Response	2

Table 4.24 Number of BioCarta pathways (second approach/second filter)

Panther Pathways	Count
P00057:Wnt signaling pathway	3
P00012:Cadherin signaling pathway	2

Table 4.25 Number of Panther pathways (first approach/second filter)

Panther Pathways	Count
P00057:Wnt signaling pathway	2

Table 4.26 Number of Panther pathways (second approach/second filter)

Reactome Pathways	Count
REACT_6900:Signaling in Immune system	3
REACT_14797:Signaling by GPCR	2

Table 4.27 Number of Reactome pathways (first approach/second filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	3

Table 4.28 Number of Reactome pathways (second approach/second filter)

4.3.1.2. PANOGA results

Even though PANOGA partial implementation did not yield any results for family trios separately with first filter applied, when we combined the lists of genes of each family trio, the results in Table 4.29, Table 4.30 and Table 4.31 are obtained.

First Approach
Pathways in cancer
Chronic myeloid leukemia
Spliceosome
Neurotrophin signaling pathway
Focal adhesion
Ribosome
ErbB signaling pathway
Proteasome
Cell cycle
T cell receptor signaling pathway

Table 4.29 PANOGA results of general analysis with second filter (first approach)

Second Approach
Proteasome
Ribosome
Circadian rhythm
Homologous recombination
Pathogenic Escherichia coli infection
Shigellosis
Complement and coagulation cascades
Bacterial invasion of epithelial cells
Pancreatic cancer
GABAergic synapse

Table 4.30 PANOGA results of general analysis with second filter (second approach)

Third Approach
Pathways in cancer
Cell cycle
Nucleotide excision repair
Neurotrophin signaling pathway
Chemokine signaling pathway
Ribosome
Chronic myeloid leukemia
Adherens junction
Complement and coagulation cascades
Prostate cancer

Table 4.31 PANOGA results of general analysis with second filter (third approach)

4.3.2. Promoter Results

4.3.2.1. DAVID results

Detailed results for analysis for promoter region with application of second filter on DAVID are shown in APPENDIX B. Most frequently devised pathways and the number of families they are found are presented in Table 4.32 for KEGG, Table 4.33 and Table 4.34 for Reactome. Rest of the results is not presented here, because none of the families were found in more than one family.

KEGG Pathways	Count
hsa04740:Olfactory transduction	2

Table 4.32 Number of KEGG pathways in promoter analysis (second approach/second filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	2

Table 4.33 Number of Reactome pathways in promoter analysis (first approach/second filter)

Reactome Pathways	Count
REACT_14797:Signaling by GPCR	2

Table 4.34 Number of Reactome pathways in promoter analysis (second approach/second filter)

4.3.2.2. PANOGA results

Even though PANOGA partial implementation did not yield any results for family trios separately with second filter applied, when we combined the lists of genes of each family trio, the results in Table 4.35, Table 4.36 and Table 4.37 are obtained.

First Approach
Pathways in cancer
Cell cycle
Neurotrophin signaling pathway
Focal adhesion
Chronic myeloid leukemia
Spliceosome
HTLV-I infection
Prostate cancer
Glioma
Long-term potentiation

Table 4.35 PANOGA results of promoter analysis with second filter (first approach)

Second Approach
Pathways in cancer
Cell cycle
Neurotrophin signaling pathway
Focal adhesion
Chronic myeloid leukemia
Spliceosome
HTLV-I infection
Prostate cancer
Glioma
Long-term potentiation

Table 4.36 PANOGA results of promoter analysis with second filter (second approach)

Third Approach
Pathways in cancer
Cell cycle
Nucleotide excision repair
Neurotrophin signaling pathway
Chemokine signaling pathway
Ribosome
Chronic myeloid leukemia
Adherens junction
Complement and coagulation cascades
Prostate cancer

Table 4.37 PANOGA results of promoter analysis with second filter (third approach)

4.4. Results of the Analysis with Normalization

Beta-mixture quantile normalization was applied to the data and pathway results were gathered from PANOGA.

4.4.1. General Results

Pathway results of the data pool which includes data from all families are gathered with PANOGA, and they are presented in Table 4.38. Most common pathways are presented in Table 4.39 and Table 4.40 with the number of families they are found.

First Filter	Second Filter
Spliceosome	Proteasome
Ribosome	TGF-beta signaling pathway
Pathways in cancer	Complement and coagulation cascades
Complement and coagulation cascades	ECM-receptor interaction
Cell cycle	Ribosome
Neurotrophin signaling pathway	Notch signaling pathway
Focal adhesion	Chronic myeloid leukemia
TGF-beta signaling pathway	Type II diabetes mellitus
T cell receptor signaling pathway	Protein digestion and absorption
ErbB signaling pathway	Homologous recombination

Table 4.38 Normalization results of general analysis (third approach)

Pathways	Count
Chronic myeloid leukemia	8
Neurotrophin signaling pathway	7
Pathways in cancer	7
T cell receptor signaling pathway	7
Ribosome	6
Focal adhesion	5
Cell cycle	5
ErbB signaling pathway	5
HTLV-I infection	5
Spliceosome	3
Glutamatergic synapse	3
GABAergic synapse	3
Retrograde endocannabinoid signaling	2
Cholinergic synapse	2
Non-small cell lung cancer	2
RNA polymerase	1
Prostate cancer	1
Type II diabetes mellitus	1
Endometrial cancer	1
Adherens junction	1
Acute myeloid leukemia	1
Cytosolic DNA-sensing pathway	1
MAPK signaling pathway	1
Propanoate metabolism	1
Measles	1
Proteasome	1
Citrate cycle (TCA cycle)	1
Complement and coagulation cascades	1
Endocytosis	1
Epithelial cell signaling in Helicobacter pylori infection	1
Notch signaling pathway	1
TGF-beta signaling pathway	1
Pancreatic cancer	1
Glioma	1
Colorectal cancer	1

Table 4.39 Most common pathways for general analysis with first filter (third approach)

Pathways	Count
Proteasome	3
RNA degradation	2
Notch signaling pathway	2
Type II diabetes mellitus	2
Circadian rhythm	2
Allograft rejection	2
Pentose phosphate pathway	2
Complement and coagulation cascades	2
TGF-beta signaling pathway	2
GABAergic synapse	2
Aldosterone-regulated sodium reabsorption	2
Type I diabetes mellitus	1
Staphylococcus aureus infection	1
Amphetamine addiction	1
Asthma	1
Homologous recombination	1
Folate biosynthesis	1
Intestinal immune network for IgA production	1
Autoimmune thyroid disease	1
Graft-versus-host disease	1

Table 4.40 Most common pathways for general analysis with second filter (third approach)

4.4.2. Promoter Results

Pathway results of the data pool which includes data from all families are gathered with PANOGA, and they are presented in Table 4.41. Most common pathways are presented in Table 4.42 and Table 4.43 with the number of families they are found.

First Filter	Second Filter
Complement and coagulation cascades	Complement and coagulation cascades
Neurotrophin signaling pathway	Notch signaling pathway
Pathways in cancer	Proteasome
Cell cycle	B cell receptor signaling pathway
Chronic myeloid leukemia	Acute myeloid leukemia
Prostate cancer	Malaria
T cell receptor signaling pathway	Endocrine and other factor-regulated calcium reabsorption
Ribosome	Aldosterone-regulated sodium reabsorption
Insulin signaling pathway	RNA polymerase
Glioma	Folate biosynthesis

Table 4.41 Normalization results of promoter analysis (third approach)

Pathways	Count
T cell receptor signaling pathway	3
Pancreatic cancer	3
Neurotrophin signaling pathway	3
Chronic myeloid leukemia	3
Pathways in cancer	3
Endocytosis	3
Focal adhesion	3
Cholinergic synapse	2
Ribosome	2
Complement and coagulation cascades	2
Adherens junction	1
Base excision repair	1
DNA replication	1
Mismatch repair	1
Prostate cancer	1
ErbB signaling pathway	1
Homologous recombination	1
Nucleotide excision repair	1
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1

Table 4.42 Most common pathways for promoter analysis with first filter (third approach)

Pathways	Count
Glutamatergic synapse	1
Renin-angiotensin system	1
Notch signaling pathway	1
Complement and coagulation cascades	1
Cholinergic synapse	1
DNA replication	1
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1
Dopaminergic synapse	1
Proteasome	1
ECM-receptor interaction	1
Retrograde endocannabinoid signaling	1
Folate biosynthesis	1
GABAergic synapse	1

Table 4.43 Most common pathways for promoter analysis with second filter (third approach)

4.5. Results of the Quality Checking

4.5.1. Results without Normalization

General and promoter analysis results from PANOGA with third approach are presented in Table 4.44. Second filter was applied in this analysis.

General Analysis	Promoter Analysis
Nucleotide excision repair	Hedgehog signaling pathway
Ribosome	
DNA replication	
Mismatch repair	
Base excision repair	
Non-homologous end-joining	
Proteasome	
Thyroid cancer	
Homologous recombination	
Circadian rhythm	

Table 4.44 General and promoter analysis results without normalization

4.5.2. Results with Normalization

General analysis results from PANOGA with third approach are presented in Table 4.45. Second filter was applied in this analysis. Promoter analysis did not yield any results for this case.

General Analysis
Neurotrophin signaling pathway
Pathways in cancer
T cell receptor signaling pathway
Endocytosis
Pancreatic cancer
Focal adhesion
Chronic myeloid leukemia
Adherens junction
ErbB signaling pathway
Prostate cancer

Table 4.45 General and promoter analysis results with normalization

4.6. Results of the Statistical Analysis

In order to determine whether there is a significant difference in methylation levels for individuals that have a different status in terms of being affected by the disorder, we applied a statistical analysis using Wilcoxon Rank-Sum Test and Student's t-test.

Initially we calculated the average, standard deviation and variance of methylation levels of each individual. Averages of those values are also calculated for each group (offspring, affected parent, control parent). Results for each group are shown in Table 4.46, Table 4.47 and Table 4.48 respectively.

	Variance	Std. Deviation	Average
Family trio 1	0,14	0,37	0,51
Family trio 2	0,13	0,36	0,50
Family trio 3	0,13	0,36	0,51
Family trio 4	0,14	0,37	0,50
Family trio 5	0,14	0,37	0,49
Family trio 6	0,13	0,37	0,51
Family trio 7	0,13	0,35	0,50
Family trio 8	0,13	0,36	0,50
Family trio 9	0,13	0,36	0,51
Family trio 10	0,14	0,37	0,52
Average	0,13	0,36	0,51

Table 4.46 Variance, standard deviation and average values for offsprings.

	Variance	Std. Deviation	Average
Family trio 1	0,13	0,36	0,50
Family trio 2	0,13	0,36	0,50
Family trio 3	0,13	0,36	0,50
Family trio 4	0,09	0,30	0,51
Family trio 5	0,14	0,37	0,50
Family trio 6	0,13	0,36	0,50
Family trio 7	0,08	0,29	0,45
Family trio 8	0,13	0,36	0,50
Family trio 9	0,13	0,37	0,51
Family trio 10	0,14	0,37	0,51
Average	0,12	0,35	0,50

Table 4.47 Variance, standard deviation and average values for affected parents.

	Variance	Std. Deviation	Average
Family trio 1	0,13	0,36	0,51
Family trio 2	0,13	0,36	0,50
Family trio 3	0,13	0,37	0,51
Family trio 4	0,13	0,37	0,51
Family trio 5	0,13	0,36	0,50
Family trio 6	0,13	0,36	0,50
Family trio 7	0,13	0,36	0,51
Family trio 8	0,13	0,36	0,50
Family trio 9	0,13	0,36	0,51
Family trio 10	0,14	0,37	0,51
Average	0,13	0,36	0,50

Table 4.48 Variance, standard deviation and average values for control parents.

Then we applied Wilcoxon Rank-Sum Test and Student's t-test with the null hypothesis being "the difference between two samples has no significance". Comparisons were made between offspring versus control parent, affected parent versus control parent and all affected samples (offspring and affected parent) versus control parent. Results for both tests are shown in Table 4.49.

Test Results				
All families	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00592	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,20959	h = 0
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
Family No.1	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,01426	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
Family No.2	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
Family No.3	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,87041	h = 0
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,66145	h = 0
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00106	h = 1
Family No.4	All affected (child&parent) vs. Control	Wilcoxon	0,00313	h = 1
		Student's t-test	0,58478	h = 0
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00003	h = 1
	Affected Parent vs. Control	Wilcoxon	0,30672	h = 0
		Student's t-test	0,00019	h = 1
Family No.5	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00053	h = 1
Family No.6	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00103	h = 1

Family No.7	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00188	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
Family No.8	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Child vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00000	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,00567	h = 1
Family No.9	All affected (child&parent) vs. Control	Wilcoxon	0,35086	h = 0
		Student's t-test	0,91461	h = 0
	Affected Child vs. Control	Wilcoxon	0,26554	h = 0
		Student's t-test	0,47038	h = 0
	Affected Parent vs. Control	Wilcoxon	0,00125	h = 1
		Student's t-test	0,00480	h = 1
Family No.10	All affected (child&parent) vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,01864	h = 1
	Affected Child vs. Control	Wilcoxon	0,65677	h = 0
		Student's t-test	0,00028	h = 1
	Affected Parent vs. Control	Wilcoxon	0,00000	h = 1
		Student's t-test	0,88178	h = 0

Table 4.49 Results of the statistical analysis. Null hypothesis for both methods states “the difference between two samples has no importance”, therefore h=0 implies that the test has “failed to reject null hypothesis, difference has no importance”, and h=1 implies that the test “rejects null hypothesis, difference is important”

CHAPTER 5

Discussion and Conclusion

In this thesis, we aimed to devise functionally important pathways for epilepsy, using DNA methylation levels information which was provided by a microarray based analysis method, Illumina HumanMethylation450 Bead Chip. The dataset was constructed from 10 family trios. In each family trio, the child and one of the parents are affected by epilepsy, and one of the parents (control parent) is not. Therefore, the dataset allowed us to apply a trio analysis on the data set. Two approaches that we developed for trio analysis were originated from the requirement to determine the effects of both methylated status and non-methylated status at a CpG locus, which may be associated to a gene and therefore may lead us to a significant pathway.

At first the dataset was used without applying any filter. Although all pathway analysis methods which are described in this thesis were used in this stage, the results can't be considered as reliable as the results of analysis with filters. Even though the resulting pathway lists may include significant pathways that are associated with epilepsy, they may also include unrelated results, because a fraction of the CpG locus may yield biased DNA methylation level results according to sex. Moreover, a number of positions in the dataset yields DNA methylation level results from non-CpG methylation, or SNPs directly. Pathway results that are associated with those positions may compromise our assessment.

Along with application of two filters which were previously described in section 3.3 in this thesis, also additional analysis on promoter-associated CpGs were separately done and both filters were applied in promoter analysis. Pathway results for each family

trio by different analysis methods (i.e., DAVID, PANOGA) were presented in the previous chapter. Moreover, the list of all devised pathways by DAVID is presented in APPENDIX B. The numbers of occurrences of pathways in different family trio analyses were also provided in there. Those lists were used to compare our results with another study by Bakir-Gungor (2013), which aims to identify epilepsy-related pathways. In Table 5.1, the results of Bakir-Gungor are compared with the results we found in this thesis. They presented top 29 KEGG pathways, therefore we used KEGG results from DAVID and KEGG results from PANOGA for comparison. Matching pathways are marked in the table.

KEGG pathways	a	b	c	d	e	f	g	h
Complement and coagulation cascades							x	x
Cell cycle			x	x	x	x	x	x
Focal Adhesion			x		x	x		x
ECM—receptor interaction								
Jak-STAT signaling pathway			x	x				
MAPK signaling pathway			x	x			x	
Proteasome			x		x			x
Ribosome					x		x	x
Calcium signaling pathway	x	x	x	x				
Regulation of actin cytoskeleton			x				x	
Adherens junction			x	x				
Pathways in cancer			x	x	x	x	x	x
Gap junction			x					
Apoptosis							x	
Long-term depression			x					
Axon guidance			x	x				
Fc gamma R-mediated phagocytosis			x	x				
Tight junction	x	x	x	x				
ErbB signaling pathway			x				x	x
Wnt signaling pathway			x	x				
Chemokine signaling pathway			x	x	x			
GnRH signaling pathway								
Pentose phosphate pathway								
Long-term potentiation								
Neurotrophin signaling pathway					x	x	x	x
Glycolysis/Gluconeogenesis								
Notch signaling pathway			x					
Dilated cardiomyopathy			x	x				
TGF-beta signaling pathway			x	x			x	

Table 5.1 Comparison of top 29 KEGG pathways by Bakir-Gungor (2013) and findings of this thesis. a) Promoter analysis results with first filter from DAVID (17 pathways), b) Promoter analysis results with second filter from DAVID (5 pathways), c) General analysis results with first filter from DAVID (82 pathways), d) General analysis results with first second filter from DAVID (54 pathways), e) Promoter analysis results with first filter from PANOGA (10 pathways), f) Promoter analysis results with second filter from PANOGA (10 pathways), g) General analysis results with first filter from PANOGA (20 pathways), h) General analysis results with second filter from PANOGA (18 pathways).

General analysis with both filters show correspondence with the study (Bakir-Gungor et al., 2013), 20 out of 29 matching pathways were identified with the first filter and 13 out of 29 matching pathways were identified with the second filter. Since the second filter also contains the first filter, the decrease in the number of identified

pathways for the second filter is an expected outcome. Even though the promoter analysis results from DAVID do not show correspondence with the study (only 2 out of 29 pathways were identified), it is also an expected outcome, because the pathway lists for promoter analysis with DAVID contains relatively small number of pathways. The list with the first filter contains 17 pathways and the list with the second filter contains only 5 pathways. On the other hand, containing a number of 10 pathways, combined results (combination of the results of first and second approaches) of promoter analysis from PANOGA shows better results by identifying 7 (first filter) and 4 (second filter) matching pathways. Moreover, combined results of general analysis from PANOGA also provide better outcomes than DAVID by identifying 10 out of 20 pathways (first filter) and 8 out of 18 pathways (second filter).

Comparison of the results by Bakir-Gungor and PANOGA results of the analysis with third approach are presented in Table 5.2.

KEGG Pathways	a	b	c	d	e
Complement and coagulation cascades			x	x	x
Cell cycle		x	x	x	x
Focal Adhesion					
ECM—receptor interaction					
Jak-STAT signaling pathway					
MAPK signaling pathway					
Proteasome			x		
Ribosome	x	x	x	x	x
Calcium signaling pathway					
Regulation of actin cytoskeleton					
Adherens junction				x	x
Pathways in cancer	x			x	x
Gap junction					
Apoptosis	x				
Long-term depression					
Axon guidance					
Fc gamma R-mediated phagocytosis					
Tight junction					
ErbB signaling pathway					
Wnt signaling pathway					
Chemokine signaling pathway		x	x	x	x
GnRH signaling pathway					
Pentose phosphate pathway					
Long-term potentiation		x			
Neurotrophin signaling pathway	x			x	x
Glycolysis/Gluconeogenesis					
Notch signaling pathway					
Dilated cardiomyopathy					
TGF-beta signaling pathway		x	x		

Table 5.2 Comparison of top 29 KEGG pathways by Bakir-Gungor (2013) and findings of this thesis. a) Third approach results without filter from PANOGA (10 pathways), b) Third approach results of general analysis with first filter from PANOGA (10 pathways), c) Third approach results of general analysis with second filter from PANOGA (10 pathways), d) Third approach results of promoter analysis with first filter from PANOGA (10 pathways), e) Third approach results of promoter analysis with second filter from PANOGA (10 pathways)

Third approach results from PANOGA shows significant correspondence with the findings of Bakir-Gungor. Since the results of the analysis without any filters contains only 3 pathways out of 10, it can be concluded that filtered analysis from PANOGA provides better results. General analysis results contain 5 matching pathways with first filter and 6 matching pathways with second filter. Promoter analysis results

shows better correspondence with the findings off Bakir-Gungor and contains 7 matching pathways for both filters.

Comparison of the results of Bakir-Gungor and the third approach results with normalization from PANOGA are presented in Table 5.3.

KEGG Pathways	**a**	**b**	**c**	**d**
Complement and coagulation cascades	x	x	x	x
Cell cycle	x		x	
Focal Adhesion	x			
ECM—receptor interaction		x		
Jak-STAT signaling pathway				
MAPK signaling pathway				
Proteasome		x		x
Ribosome	x	x	x	
Calcium signaling pathway				
Regulation of actin cytoskeleton				
Adherens junction				
Pathways in cancer	x		x	
Gap junction				
Apoptosis				
Long-term depression				
Axon guidance				
Fc gamma R-mediated phagocytosis				
Tight junction				
ErbB signaling pathway	x			
Wnt signaling pathway				
Chemokine signaling pathway				
GnRH signaling pathway				
Pentose phosphate pathway				
Long-term potentiation				
Neurotrophin signaling pathway	x		x	
Glycolysis/Gluconeogenesis				
Notch signaling pathway		x		x
Dilated cardiomyopathy				
TGF-beta signaling pathway	x	x		

Table 5.3 Comparison of top 29 KEGG pathways by Bakir-Gungor (2013) and findings of this thesis. a) General analysis normalization results with first filter (10 pathways), b) General analysis normalization results with second filter (10 pathways), c) Promoter analysis normalization results with first filter (10 pathways), d) Promoter analysis normalization results with second filter (10 pathways)

General analysis results with normalization shows a significant correspondence with results of Bakir-Gungor by matching 8 pathways with first filter and 7 pathways with second filter. On the other hand, promoter analysis results with normalization shows less correspondence by matching 5 pathways with first filter and only three pathways with second filter. Even though promoter analysis with normalization seems to perform worse, showing less correspondence with previous findings may also indicate the possibility of newly found epilepsy-related pathways in the results of this thesis.

Comparison of the results of Bakir-Gungor and quality checking results with third approach are presented in Table 5.4.

KEGG Pathways	a	b
Complement and coagulation cascades		
Cell cycle		
Focal Adhesion		x
ECM—receptor interaction		
Jak-STAT signaling pathway		
MAPK signaling pathway		
Proteasome	x	
Ribosome	x	
Calcium signaling pathway		
Regulation of actin cytoskeleton		
Adherens junction		x
Pathways in cancer		x
Gap junction		
Apoptosis		
Long-term depression		
Axon guidance		
Fc gamma R-mediated phagocytosis		
Tight junction		
ErbB signaling pathway		x
Wnt signaling pathway		
Chemokine signaling pathway		
GnRH signaling pathway		
Pentose phosphate pathway		
Long-term potentiation		
Neurotrophin signaling pathway		x
Glycolysis/Gluconeogenesis		
Notch signaling pathway		
Dilated cardiomyopathy		
TGF-beta signaling pathway		

Table 5.4 Comparison of top 29 KEGG pathways by Bakir-Gungor (2013) and findings of this thesis. a) Quality checking without normalization results with second filter (general analysis), b) Quality checking with normalization results with second filter (general analysis)

Even though quality checking with normalization results shows moderate correspondence with the results of Bakir-Gungor, quality checking without normalization results shows less correlation. This may be resulted from the fact that family trio 7, which was removed in the process of quality checking, provides a high number of marked genes when compared to other family trios. Moreover, promoter analysis did not yield any results without normalization, and it only provides one pathway with normalization, and a possible reason behind this being the removal of

family trio 7 is highly likely. On the other hand, results of the quality checking includes some pathways that did not appear on the lists of previous studies and other analysis results of this thesis, but they may also indicate the possibility of newly found epilepsy-related pathways.

Other than those uniquely found pathways, the existence of many immune system related pathways is a remarkable outcome. This novel found may contribute to explain the relationship between immune system and epileptogenesis.

In this thesis, common motifs around methylated and unmethylated nucleotide sequences were also found. Since the sequence information of methylation data used in this thesis wasn't provided yet, MethDB database were used in this process. When the sequence information of methylation data is acquired, same analysis can be repeated to compare the common motifs found, and those common motifs can be searched around the marked CpG islands which are differentially methylated in family trios. If a common motif is found around a certain CpG island in a sample, and doesn't appear around the same CpG Island in other family members, this situation may lead to find a relationship between the mutations of certain motifs around a CpG island and methylation level of it.

The statistical analysis shows that there is a significant difference between DNA methylation levels affected samples and control samples. 29 out of 33 Wilcoxon rank-sum test results and 26 out of 33 Student's t-test results imply this conclusion. The work of this thesis can be continued by determining specific regions in the genome which are differentially methylated. Associated pathways can be identified by related genes in those differentially methylated regions. A comparison between trio analysis results and results from the new analysis may provide further validation for identified pathways.

REFERENCES

Antequera, F. (2007). CpG islands and DNA methylation. eLS. John Wiley & Sons Ltd, Chichester. doi: 10.1002/9780470015902.a0005027.pub2

Antequera, F., & Bird, A. (1993). Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences*, 90(24), 11995–11999. doi:10.1073/pnas.90.24.11995

Bajic, V. B., Tan, S. L., Christoffels, A., Schonbach, C., Lipovich, L., Yang, L., Hofmann, O., Kruger, A., Hide, W., Kai, C., Kawai, J., Hume, D. A., Carninci, P., & Hayashizaki, Y. (2006). Mice and men: Their promoter properties. *PLoS Genetics*, 2(4), e54. doi:10.1371/journal.pgen.0020054.

Bakir-Gungor, B., & Sezerman, O. U. (2011) A new methodology to associate SNPs with human diseases according to their pathway related context. *PLoS One*, 6. doi:10.1371/journal.pone.0026277

Bakir-Gungor, B., Baykan, B., Ugur Iseri, S., Tuncer, F. N., & Sezerman, O. U. (2013). Identifying SNP targeted pathways in partial epilepsies using genome-wide association study data. *Epilepsy Research*, 105(1-2), 92-102. doi: 10.1016/j.eplepsyres.2013.02.008.

Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J. 3rd, Gingeras, T. R., Schreiber, S. L., & Lander, E. S. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, 120(2), 169–181. doi: 10.1016/j.cell.2005.01.001.

Bhasin, M., Zhang, H., Reinherz, E., & Reche, P. (2005). Prediction of methylated CpGs in DNA sequences using a support vector machine. *FEBS Letters*, 579, 4302–8. doi: 10.1016/j.febslet.2005.07.002.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pages, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology annotation networks. *Bioinformatics*, 25(8), 1091-1093. doi:10.1093/bioinformatics/btp101

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, 16(1), 6–21. doi:10.1101/gad.947102

Bird, A. (2007). Perceptions of epigenetics. *Nature,* 447(7143), 396-398. doi:10.1038/nature05913

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature,* 321(6067), 209–213. doi:10.1038/321209a0

Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., & Walter, J. (2006). CpG ısland methylation in human lymphocytes ıs highly correlated with DNA sequence, repeats, and predicted DNA structure . *PLoS Genetics*. doi:10.1371/journal.pgen.0020026

Bock, C., Walter, J., Paulsen, M., & Lengauer T. (2007). CpG island mapping by epigenome prediction. *PLoS Computational Biology*, 3(6), e110. doi:10.1371/journal.pcbi.0030110

Brazma, A., Robinson, A., Cameron, G., & Ashburner, M. (2000). One-stop shop for microarray data. *Nature*, 403, 699-700. doi:10.1038/35001676

Cawley, S., Bekiranov, S., Ng, H. H., Kapranov, P., Sekinger, E. A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng J., Williams, A. J., Wheeler, R., Wong, B., Drenkow, J., Yamanaka, M., Patel, S., Brubaker, S., Tammana, H., Helt, G., Struhl, K., & Gingeras, T. R. (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell,* 116(4), 499–509. doi: 10.1016/S0092-8674(04)00127-8.

Chen, Y., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., Gallinger, S., Hudson, T. J., & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. doi:10.4161/epi.23470

Coulondre, C., Miller, J. H., Farabaugh, P. J., & Gilbert, W. (1978). Molecular basis of base substitution hotspots in Escherichia coli. *Nature (London),* 274(5673), 775–780. doi:10.1038/274775a0

Cowan, L. D. (2002). The epidemiology of the epilepsies in children. *Mental Retardation and Developmental Disabilities Research Reviews*, 8,171-181. doi:10.1002/mrdd.10035

Crawford, G. E., Holt, I. E., Whittle, J., Webb, B. D., Tai, D., Davis, S., Margulies, E. H., Chen, Y., Bernat, J. A., Ginsburg, D., Zhou, D., Luo, S., Vasicek, T. J., Daly, M. J., Wolfsberg, T. G., & Collins, F. S. (2006). Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Research,* 16(1), 123–131. doi: 10.1101/gr.4074106.

Das, R., Dimitrova, N., Xuan, Z., Rollins, R., Haghighi, F., Edwards, J., Ju, J., Bestor, T., & Zhang, M. (2006). Computational prediction of methylation status in human genomic sequences. *Proceedings of the National Academy of Sciences*, 22, 10713–10716. doi:10.1073/pnas.0602949103

Dayem Ullah, A. Z., Lemoine, N. R., & Chelala, C. (2012). SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic Acids Research*, 40, W65–70. doi: 10.1093/nar/gks364

Du, P., Kibbe, W. A., & Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, 24(13), 1547-1548. doi:10.1093/bioinformatics/btn224

Du, P., Zhang, X., Huang, C. C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11, 587. doi:10.1186/1471-2105-11-587

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. doi:10.1186/1471-2105-10-48.

Fang, F., Fan, S., Zhang, X., & Zhang, M. Q. (2006). Predicting methylation status of CpG islands in the human brain. *Bioinformatics*, 22, 2204–2209. doi: 10.1093/bioinformatics/btl377.

Fazzari, M. J., & Greally, J. M. (2004). Epigenomics: Beyond CpG islands. *Nature Reviews Genetics*, 5(6), 446–455. doi: 10.1038/nrg1349.

Gardiner-Garden, M., & Frommer, M. (1987). CpG islands in vertebrate genomes. *Journal of Molecular Biology*, 196(2), 261–282. doi: 10.1016/0022-2836(87)90689-9.

Grunau, C., Renault, E., Rosenthal, A., Roizes, G. (2001). MethDB–a public database for DNA methylation data. *Nucleic Acids Research*, 29(1), 270–274. doi: 10.1093/nar/29.1.270.

Harrison, A., & Parle-McDermott, A. (2011). DNA methylation: A timeline of methods and applications. *Frontiers in Genetics*, 2(74), 1-13. doi:10.3389/fgene.2011.00074

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols*, 4(1), 44-57. doi:10.1038/nprot.2008.211

Ideker, T., Ozier, O., Schwikowski, B., & Siegel, A. F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, 18(Suppl 1), 233–240. doi:10.1093/bioinformatics/18.suppl_1.S233

Illumina. (2010). CpG loci ıdentification technical note. Retrieved from http://res.illumina.com/documents/products/technotes/technote_cpg_loci_identification.pdf

Illumina. (2011). Data analysis software data sheet. Retrieved from http://res.illumina.com/documents/products/datasheets/datasheet_genomestudio_software.pdf

Illumina. (2012a). Human Methylation450 Bead Chip achieves breadth of coverage using two Infinium chemistries. Retrieved from http://res.illumina.com/documents/products/technotes/technote_hm450_data_analysis_optimization.pdf

Illumina. (2012b). Infinium HumanMethylation450 Bead Chip data sheet. Retrieved from http://res.illumina.com/documents/products/datasheets/datasheet_humanmethylation450.pdf

Infante-Rivard, C., Mirea, L., & Bull, S. B. (2009). Combining case-control and case-trio data from the same population in genetic association analyses: overview of approaches and illustration with a candidate gene study. *American Journal of Epidemiology*, 170(5), 657–664. doi: 10.1093/aje/kwp180.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28, 27-30. doi:10.1093/nar/28.1.27

Kaneko, S., & Wada, K. (1998). Molecular genetic studies of epilepsies. *Brain Nerve,* 50, 1071-1077.

Kaneko, S., Okada, M., Iwasa, H., Yamakawa, K., & Hirose, S. (2002). Genetics of epilepsy: currents status and perspectives. *Neuroscience Research*, 44, 11–30. doi:10.1016/S0168-0102(02)00065-2

Kim, T. H., Barrera, L.O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., & Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature,* 436(7052), 876–880. doi: 10.1038/nature03877.

Kobow, K., Jeske, I., Hildebrandt, M., Hauke, J., Hahnen, E., Buslei, R., Buchfelder, M., Weigel, D., Stefan, H., Kasper, B., Pauli, E., & Blümcke, I. (2009). Increased reelin promoter methylation is associated with granule cell dispersion in humantemporal lobe epilepsy. *Journal of Neuropathology and Experimental Neurology*, 68, 356–364. doi:10.1097/NEN.0b013e31819ba737

Larsen, F., Gundersen, G., Lopez, R., & Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics,* 13(4), 1095–1107. doi:10.1016/0888-7543(92)90024-M

Lee, P. H., & Shatkay, H. (2008). F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Research,* 13, 820–824. doi: 10.1093/nar/gkm904

Mancuso, F. M., Montfort, M., Carreras, A., Alibes, A., & Roma, G. (2011). HumMeth27QCReport: an R package for quality control and primary analysis of Illumina Infinium methylation data. *BMC Research Notes,* 4546 doi:10.1186/1756-0500-4-546

Mehler, M. F. (2008). Epigenetic principles and mechanisms underlying nervous system functions in health and disease. *Progress in Neurobiology,* 86, 305–341. doi:10.1016/j.pneurobio.2008.10.001

Poduri, A., & Lowenstein, D. (2011). Epilepsy genetics; past, present, and future. Current Opinion in *Genetics & Development*, 21, 325–332. doi:10.1016/j.gde.2011.01.005

Price, E. M., Cotton, A. M., Lam, L. L., Farre, P., Emberly, E., Brown, C. J., Robinson, W. P., & Kobor, M .S.(2013) Additional annotation enhances potential for biologically-relevant analysis of the illumina infinium Human Methylation 450 Bead Chip array. *Epigenetics & Chromatin*, 6(1), 4. doi:10.1186/1756-8935-6-4

Qureshi, I. A., & Mehler, M. F. (2010). Epigenetic mechanisms underlying human epileptic disorders and the process of epileptogenesis. *Neurobiology of Disease*, 39, 53–60. doi: 10.1016/j.nbd.2010.02.005.

Rakyan, V. K., Down, T. A., Thorne, N. P., Flicek, P., Kulesha, E., Gräf, S., Tomazou, E. M., Bäckdahl, L., Johnson, N., Herberth, M., Howe, K. L., Jackson, D. K., Miretti, M. M., Fiegler, H., Marioni, J. C., Birney, E., Hubbard, T. J., Carter, N. P., Tavaré, S., & Beck, S. (2008) An integrated resource for genome-wide identification and analysis of human tissue-specific differential methylated regions (tDMRs). *Genome Research,* 18(9), 1518–1529. doi:10.1101/gr.077479.108

Russo, V. E. A., Martienssen, R. A., & Riggs, A. D. (eds) (1996). Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, Woodbury.

Saccone, S. F., Bolze, R., Thomas, P., Quan, J. X., Mehta, G., Deelman, E., Tischfield, J. A., & Rice, J. P. (2010) SPOT: a web-based tool for using biological databases to prioritize SNPs after a genome-wide association study. *Nucleic Acids Research,* 38, W201–W209. doi:10.1093/nar/gkq513

Sandoval, J., Heyn, H. A., Moran, S., Serra-Musach, J., Pujana, M. A., Bibikova, M., & Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*, 6, 692–702. doi:10.4161/epi.6.6.16196

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13, 2498–504. doi:10.1101/gr.1239303

Sharma, R. P., Tun, N., & Grayson, D. R. (2008). Depolarization induces downregulation of DNMT1 and DNMT3a in primary cortical cultures. *Epigenetics*, 3, 74–80. doi:10.4161/epi.3.2.6103.

Singer, J., Roberts-Ems, J., & Riggs, A. D. (1979). Methylation of mouse liver DNA studied by means of the restriction enzymes msp I and hpa II. *Science*, 203(4384), 1019–1021. doi:10.1126/science.424726

Sjahputera, O., Keller, J. M., Davis, J. W., Taylor, K. H., Rahmatpanah, F., Shi, H., Anderson, D. T., Blisard, S. N., Luke, R. H., Popescu, M., Arthur, G. C., & Caldwell, C. W. (2007). Relational analysis of CpG islands methylation and gene expression in human lymphomas using possibilistic C-means clustering and modified cluster fuzzy density. IEEE/*ACM Transactions on Computational Biology and Bioinformatics*, 4, 176–189. doi:10.1109/TCBB.2007.070205

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, 14(3): 204-220. doi:10.1038/nrg3354.

Southern, E., Mir, K., & Schchepinov, M. (1999). Molecular interactions on microarrays. *Nature Genetics*, 21, 5–9. doi: 10.1038/4429

Steinlein, O. K. (2004). Genetic mechanisms that underlie epilepsy. *Nature Reviews Neuroscience*. doi:10.1038/nrn1388.

Stoughton, R. B. (2005). Applications of DNA microarrays in biology. *Annual Review of Biochemistry*, 74, 53–82. doi: 10.1146/annurev.biochem.74.082803.133212.

Takai, D., & Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences USA*, 99(6), 3740–3745. doi:10.1073/pnas.052410099

Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., Beck, S. (2013). A Beta-Mixture Quantile Normalisation method for correcting probe design bias in illumina infinium 450k DNA methylation data. *Bioinformatics*, 29(2), 189–196. doi: 10.1093/bioinformatics/bts680

Waddington, C. H. (1957). The strategy of the genes. London: George Allen & Unwin.

Wang, D., Yan, L., Hu, Q., Sucheston, L. E., Higgins, M. J., Ambrosone, C. B., Johnson, C.S., Smiraglia, D. J., & Liu, S. (2012). IMA: an R package for high-throughput analysis of Illumina's 450 K Infinium methylation data. *Bioinformatics*, 28(5), 729–730. doi: 10.1093/bioinformatics/bts013

Wang, Y., & Leung, F. C. C. (2004). An evaluation of new criteria for CpG islands in the human genome as gene markers, *Bioinformatics*, 20(7), 1170–1177. doi:10.1093/bioinformatics/bth059

Xu, Z., & Taylor, J. A. (2009). SNPinfo: integrating GWAS and candidate gene ınformation into functional snp selection for genetic association studies. *Nucleic Acids Research*. doi:10.1093/nar/gkp290

Yu, Z. (2011). Testing gene-gene interactions in the case-parents design. *Human Heredity*, 71, 171-179. doi:10.1159/000327355

APPENDIX A

Results without Filtering

Ap denotes the approach used in the analysis and F denotes the number of the family.

Method	Ap	Pathway	F	Data
DAVID	1st	KEGG	1	hsa04070:Phosphatidylinositol signaling system
DAVID	1st	KEGG	1	hsa00562:Inositol phosphate metabolism
DAVID	1st	KEGG	6	hsa04940:Type I diabetes mellitus
DAVID	1st	KEGG	6	hsa05330:Allograft rejection
DAVID	1st	KEGG	6	hsa05332:Graft-versus-host disease
DAVID	1st	KEGG	6	hsa05320:Autoimmune thyroid disease
DAVID	1st	KEGG	6	hsa05416:Viral myocarditis
DAVID	1st	KEGG	6	hsa04612:Antigen processing and presentation
DAVID	1st	KEGG	6	hsa04514:Cell adhesion molecules (CAMs)
DAVID	1st	KEGG	6	hsa05310:Asthma
DAVID	1st	KEGG	6	hsa04672:Intestinal immune network for IgA production
DAVID	1st	KEGG	7	hsa04310:Wnt signaling pathway
DAVID	1st	KEGG	7	hsa00100:Steroid biosynthesis
DAVID	1st	KEGG	7	hsa05200:Pathways in cancer
DAVID	1st	KEGG	9	hsa04514:Cell adhesion molecules (CAMs)
DAVID	2nd	KEGG	7	hsa04144:Endocytosis
DAVID	2nd	KEGG	7	hsa04062:Chemokine signaling pathway
DAVID	2nd	KEGG	7	hsa04110:Cell cycle
DAVID	2nd	KEGG	8	hsa05330:Allograft rejection
DAVID	2nd	KEGG	8	hsa05332:Graft-versus-host disease
DAVID	2nd	KEGG	8	hsa04940:Type I diabetes mellitus
DAVID	2nd	KEGG	8	hsa05320:Autoimmune thyroid disease
DAVID	1st	BioCarta	7	h_p38mapkPathway:p38 MAPK Signaling Pathway
DAVID	1st	Panther	7	P00057:Wnt signaling pathway
DAVID	1st	Panther	3	P00047:PDGF signaling pathway
DAVID	1st	Panther	3	P00041:Metabotropic glutamate receptor group I pathway
DAVID	1st	Panther	3	P00019:Endothelin signaling pathway
DAVID	2nd	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
DAVID	2nd	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
DAVID	2nd	Panther	7	P00052:TGF-beta signaling pathway
DAVID	1st	Reactome	3	REACT_9480:Gap junction trafficking and regulation
DAVID	1st	Reactome	9	REACT_604:Hemostasis
GOrilla	1st	-	3	GO:0009605:response to external stimulus
GOrilla	1st	-	7	GO:0002253:activation of immune response
GOrilla	1st	-	7	GO:0002429:immune response-activating cell surface receptor signaling pathway

GOrilla	1st	-	7	GO:0002757:immune response-activating signal transduction
GOrilla	1st	-	7	GO:0002768:immune response-regulating cell surface receptor signaling pathway
GOrilla	1st	-	7	GO:0002764:immune response-regulating signaling pathway
GOrilla	1st	-	7	GO:0002696:positive regulation of leukocyte activation
GOrilla	1st	-	7	GO:0050778:positive regulation of immune response
GOrilla	1st	-	7	GO:0050870:positive regulation of T cell activation
GOrilla	1st	-	7	GO:0050852:T cell receptor signaling pathway
GOrilla	1st	-	7	GO:0050851:antigen receptor-mediated signaling pathway
GOrilla	1st	-	7	GO:0050863:regulation of T cell activation
GOrilla	1st	-	7	GO:0050867:positive regulation of cell activation
GOrilla	1st	-	7	GO:0051251:positive regulation of lymphocyte activation
GOrilla	1st	-	7	GO:0051249:regulation of lymphocyte activation
PANOGA	1st	KEGG	1	Metabolic pathways
PANOGA	1st	KEGG	1	Pathways in cancer
PANOGA	1st	KEGG	1	Neurotrophin signaling pathway
PANOGA	1st	KEGG	1	Focal adhesion
PANOGA	1st	KEGG	1	Cell cycle
PANOGA	1st	KEGG	1	HTLV-I infection
PANOGA	1st	KEGG	1	Small cell lung cancer
PANOGA	1st	KEGG	1	Ribosome
PANOGA	1st	KEGG	1	ErbB signaling pathway
PANOGA	1st	KEGG	1	MAPK signaling pathway
PANOGA	2nd	KEGG	2	Metabolic pathways
PANOGA	2nd	KEGG	2	Pathways in cancer
PANOGA	2nd	KEGG	2	Cell cycle
PANOGA	2nd	KEGG	2	Neurotrophin signaling pathway
PANOGA	2nd	KEGG	2	Focal adhesion
PANOGA	2nd	KEGG	2	HTLV-I infection
PANOGA	2nd	KEGG	2	Prostate cancer
PANOGA	2nd	KEGG	2	Small cell lung cancer
PANOGA	2nd	KEGG	2	Chronic myeloid leukemia
PANOGA	2nd	KEGG	2	Pancreatic cancer
PANOGA	1st	KEGG	5	ErbB signaling pathway
PANOGA	1st	KEGG	5	Adherens junction
PANOGA	1st	KEGG	5	Bacterial invasion of epithelial cells
PANOGA	1st	KEGG	5	Dorso-ventral axis formation

PANOGA	1st	KEGG	5	Thyroid cancer
PANOGA	2nd	KEGG	5	Non-homologous end-joining
PANOGA	1st	KEGG	6	Focal adhesion
PANOGA	1st	KEGG	6	Pathways in cancer
PANOGA	1st	KEGG	6	Ribosome
PANOGA	1st	KEGG	6	Regulation of actin cytoskeleton
PANOGA	1st	KEGG	6	Chemokine signaling pathway
PANOGA	1st	KEGG	6	ErbB signaling pathway
PANOGA	1st	KEGG	6	Cell cycle
PANOGA	1st	KEGG	6	Nucleotide excision repair
PANOGA	1st	KEGG	6	Metabolic pathways
PANOGA	1st	KEGG	6	Neurotrophin signaling pathway
PANOGA	1st	KEGG	7	Chronic myeloid leukemia
PANOGA	1st	KEGG	7	Long-term potentiation
PANOGA	1st	KEGG	7	NOD-like receptor signaling pathway
PANOGA	1st	KEGG	7	Proteasome
PANOGA	1st	KEGG	7	Acute myeloid leukemia
PANOGA	1st	KEGG	7	Thyroid cancer
PANOGA	1st	KEGG	7	Prion diseases
PANOGA	2nd	KEGG	8	Glyoxylate and dicarboxylate metabolism
PANOGA	1st	KEGG	9	RNA polymerase
PANOGA	1st	KEGG	9	Cytosolic DNA-sensing pathway
PANOGA	1st	KEGG	9	Prion diseases
PANOGA	1st	KEGG	10	Colorectal cancer
PANOGA	1st	KEGG	10	Non-homologous end-joining
PANOGA	1st	KEGG	10	Thyroid cancer
PANOGA	2nd	KEGG	10	Citrate cycle (TCA cycle)
PANOGA	2nd	KEGG	10	Valine
PANOGA	2nd	KEGG	10	Glyoxylate and dicarboxylate metabolism

APPENDIX B

Results with Filtering

Fi denotes the number of the filter used in the analysis, P shows whether the analysis is based on promoter regions or not (i.e. 1 means it is based on promoter regions, 0 means it is not) and finally F shows number of the family.

Fi	P	Ap	Pathway	F	Data
1st	0	1st	BioCarta	1	h_rhodopsinPathway:Visual Signal Transduction
1st	0	1st	BioCarta	2	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	0	1st	BioCarta	2	h_bbcellPathway:Bystander B Cell Activation
1st	0	1st	BioCarta	2	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	0	1st	BioCarta	2	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	0	1st	BioCarta	2	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	0	1st	BioCarta	2	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	0	1st	BioCarta	2	h_il5Pathway:IL 5 Signaling Pathway
1st	0	1st	BioCarta	2	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	1st	BioCarta	2	h_metPathway:Signaling of Hepatocyte Growth Factor Receptor
1st	0	1st	BioCarta	2	h_mhcPathway:Antigen Processing and Presentation
1st	0	1st	BioCarta	2	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
1st	0	1st	BioCarta	2	h_th1th2Pathway:Th1/Th2 Differentiation
1st	0	1st	BioCarta	3	h_actinYPathway:Y branching of actin filaments
1st	0	1st	BioCarta	3	h_alkPathway:ALK in cardiac myocytes
1st	0	1st	BioCarta	3	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	0	1st	BioCarta	3	h_bbcellPathway:Bystander B Cell Activation
1st	0	1st	BioCarta	3	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	0	1st	BioCarta	3	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	0	1st	BioCarta	3	h_ctcfPathway:CTCF: First Multivalent Nuclear Factor
1st	0	1st	BioCarta	3	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	0	1st	BioCarta	3	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	0	1st	BioCarta	3	h_il5Pathway:IL 5 Signaling Pathway
1st	0	1st	BioCarta	3	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	1st	BioCarta	3	h_mhcPathway:Antigen Processing and Presentation
1st	0	1st	BioCarta	3	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
1st	0	1st	BioCarta	3	h_th1th2Pathway:Th1/Th2 Differentiation
1st	0	1st	BioCarta	3	h_tob1Pathway:Role of Tob in T-cell activation
1st	0	1st	BioCarta	5	h_integrinPathway:Integrin Signaling Pathway

1st	0	1st	BioCarta	5	h_metPathway:Signaling of Hepatocyte Growth Factor Receptor
1st	0	1st	BioCarta	6	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	0	1st	BioCarta	6	h_bbcellPathway:Bystander B Cell Activation
1st	0	1st	BioCarta	6	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	0	1st	BioCarta	6	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	0	1st	BioCarta	6	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	0	1st	BioCarta	6	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	0	1st	BioCarta	6	h_il5Pathway:IL 5 Signaling Pathway
1st	0	1st	BioCarta	6	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	1st	BioCarta	6	h_integrinPathway:Integrin Signaling Pathway
1st	0	1st	BioCarta	6	h_longevityPathway:The IGF-1 Receptor and Longevity
1st	0	1st	BioCarta	6	h_mhcPathway:Antigen Processing and Presentation
1st	0	1st	BioCarta	6	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
1st	0	1st	BioCarta	6	h_th1th2Pathway:Th1/Th2 Differentiation
1st	0	1st	BioCarta	8	h_reckPathway:Inhibition of Matrix Metalloproteinases
1st	0	1st	BioCarta	9	h_agpcrPathway:Attenuation of GPCR Signaling
1st	0	1st	BioCarta	9	h_arenrf2Pathway:Oxidative Stress Induced Gene Expression Via Nrf2
1st	0	1st	BioCarta	9	h_At1rPathway:Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling
1st	0	1st	BioCarta	9	h_bcrPathway:BCR Signaling Pathway
1st	0	1st	BioCarta	9	h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway
1st	0	1st	BioCarta	9	h_calcineurinPathway:Effects of calcineurin in Keratinocyte Differentiation
1st	0	1st	BioCarta	9	h_cardiacegfPathway:Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy
1st	0	1st	BioCarta	9	h_cblPathway:CBL mediated ligand-induced downregulation of EGF receptors
1st	0	1st	BioCarta	9	h_CCR3Pathway:CCR3 signaling in Eosinophils
1st	0	1st	BioCarta	9	h_Ccr5Pathway:Pertussis toxin-insensitive CCR5 Signaling in Macrophage
1st	0	1st	BioCarta	9	h_cdMacPathway:Cadmium induces DNA synthesis and proliferation in macrophages
1st	0	1st	BioCarta	9	h_chemicalPathway:Apoptotic Signaling in Response to DNA Damage
1st	0	1st	BioCarta	9	h_cxcr4Pathway:CXCR4 Signaling Pathway

1st	0	1st	BioCarta	9	h_edg1Pathway:Phospholipids as signalling intermediaries
1st	0	1st	BioCarta	9	h_egfPathway:EGF Signaling Pathway
1st	0	1st	BioCarta	9	h_eif4Pathway:Regulation of eIF4e and p70 S6 Kinase
1st	0	1st	BioCarta	9	h_erbB4pathway:g-Secretase mediated ErbB4 Signaling Pathway
1st	0	1st	BioCarta	9	h_ghPathway:Growth Hormone Signaling Pathway
1st	0	1st	BioCarta	9	h_gpcrPathway:Signaling Pathway from G-Protein Families
1st	0	1st	BioCarta	9	h_ionPathway:Ion Channel and Phorbal Esters Signaling Pathway
1st	0	1st	BioCarta	9	h_keratinocytePathway:Keratinocyte Differentiation
1st	0	1st	BioCarta	9	h_mef2dPathway:Role of MEF2D in T-cell Apoptosis
1st	0	1st	BioCarta	9	h_myosinPathway:PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase
1st	0	1st	BioCarta	9	h_nos1Pathway:Nitric Oxide Signaling Pathway
1st	0	1st	BioCarta	9	h_Par1Pathway:Thrombin signaling and protease-activated receptors
1st	0	1st	BioCarta	9	h_pdgfPathway:PDGF Signaling Pathway
1st	0	1st	BioCarta	9	h_pkcPathway:Activation of PKC through G protein coupled receptor
1st	0	1st	BioCarta	9	h_plcdPathway:Phospholipase C d1 in phospholipid associated cell signaling
1st	0	1st	BioCarta	9	h_plcPathway:Phospholipase C Signaling Pathway
1st	0	1st	BioCarta	9	h_pparaPathway:Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)
1st	0	1st	BioCarta	9	h_pyk2Pathway:Links between Pyk2 and Map Kinases
1st	0	1st	BioCarta	9	h_sppaPathway:Aspirin Blocks Signaling Pathway Involved in Platelet Activation
1st	0	1st	BioCarta	9	h_srcRPTPPathway:Activation of Src by Protein-tyrosine phosphatase alpha
1st	0	1st	BioCarta	9	h_tcrPathway:T Cell Receptor Signaling Pathway
1st	0	1st	BioCarta	9	h_telPathway:Telomeres, Telomerase, Cellular Aging, and Immortality,h_TPOPathway:TPO Signaling Pathway
1st	0	1st	BioCarta	9	h_trkaPathway:Trka Receptor Signaling Pathway
1st	0	1st	BioCarta	9	h_vegfPathway:VEGF, Hypoxia, and Angiogenesis
2nd	0	1st	BioCarta	2	h_asbcellPathway:Antigen Dependent B Cell Activation
2nd	0	1st	BioCarta	2	h_bbcellPathway:Bystander B Cell Activation
2nd	0	1st	BioCarta	2	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
2nd	0	1st	BioCarta	2	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor

2nd	0	1st	BioCarta	2	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
2nd	0	1st	BioCarta	2	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
2nd	0	1st	BioCarta	2	h_il5Pathway:IL 5 Signaling Pathway,h_inflamPathway:Cytokines and Inflammatory Response
2nd	0	1st	BioCarta	2	h_mhcPathway:Antigen Processing and Presentation
2nd	0	1st	BioCarta	2	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
2nd	0	1st	BioCarta	2	h_th1th2Pathway:Th1/Th2 Differentiation
2nd	0	1st	BioCarta	3	h_asbcellPathway:Antigen Dependent B Cell Activation
2nd	0	1st	BioCarta	3	h_bbcellPathway:Bystander B Cell Activation
2nd	0	1st	BioCarta	3	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
2nd	0	1st	BioCarta	3	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
2nd	0	1st	BioCarta	3	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
2nd	0	1st	BioCarta	3	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
2nd	0	1st	BioCarta	3	h_il5Pathway:IL 5 Signaling Pathway
2nd	0	1st	BioCarta	3	h_inflamPathway:Cytokines and Inflammatory Response
2nd	0	1st	BioCarta	3	h_mhcPathway:Antigen Processing and Presentation
2nd	0	1st	BioCarta	3	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
2nd	0	1st	BioCarta	3	h_th1th2Pathway:Th1/Th2 Differentiation
1st	1	1st	BioCarta	3	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	1	1st	BioCarta	3	h_bbcellPathway:Bystander B Cell Activation
1st	1	1st	BioCarta	3	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	1	1st	BioCarta	3	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	1	1st	BioCarta	3	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	1	1st	BioCarta	3	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	1	1st	BioCarta	3	h_il5Pathway:IL 5 Signaling Pathway
1st	1	1st	BioCarta	3	h_inflamPathway:Cytokines and Inflammatory Response
1st	1	1st	BioCarta	3	h_mhcPathway:Antigen Processing and Presentation
1st	1	1st	BioCarta	3	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation

1st	1	1st	BioCarta	3	h_th1th2Pathway:Th1/Th2 Differentiation
1st	1	1st	BioCarta	6	h_longevityPathway:The IGF-1 Receptor and Longevity
1st	0	2nd	BioCarta	3	h_agpcrPathway:Attenuation of GPCR Signaling
1st	0	2nd	BioCarta	3	h_arenrf2Pathway:Oxidative Stress Induced Gene Expression Via Nrf2
1st	0	2nd	BioCarta	3	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	0	2nd	BioCarta	3	h_At1rPathway:Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling
1st	0	2nd	BioCarta	3	h_bbcellPathway:Bystander B Cell Activation
1st	0	2nd	BioCarta	3	h_bcrPathway:BCR Signaling Pathway
1st	0	2nd	BioCarta	3	h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway
1st	0	2nd	BioCarta	3	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	0	2nd	BioCarta	3	h_calcineurinPathway:Effects of calcineurin in Keratinocyte Differentiation
1st	0	2nd	BioCarta	3	h_cardiacegfPathway:Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy
1st	0	2nd	BioCarta	3	h_cblPathway:CBL mediated ligand-induced downregulation of EGF receptors
1st	0	2nd	BioCarta	3	h_CCR3Pathway:CCR3 signaling in Eosinophils
1st	0	2nd	BioCarta	3	h_Ccr5Pathway:Pertussis toxin-insensitive CCR5 Signaling in Macrophage
1st	0	2nd	BioCarta	3	h_cdMacPathway:Cadmium induces DNA synthesis and proliferation in macrophages
1st	0	2nd	BioCarta	3	h_chemicalPathway:Apoptotic Signaling in Response to DNA Damage
1st	0	2nd	BioCarta	3	h_crebPathway:Transcription factor CREB and its extracellular signals
1st	0	2nd	BioCarta	3	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	0	2nd	BioCarta	3	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	0	2nd	BioCarta	3	h_cxcr4Pathway:CXCR4 Signaling Pathway
1st	0	2nd	BioCarta	3	h_edg1Pathway:Phospholipids as signalling intermediaries
1st	0	2nd	BioCarta	3	h_egfPathway:EGF Signaling Pathway
1st	0	2nd	BioCarta	3	h_eif4Pathway:Regulation of eIF4e and p70 S6 Kinase
1st	0	2nd	BioCarta	3	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	0	2nd	BioCarta	3	h_erbB4pathway:g-Secretase mediated ErbB4 Signaling Pathway
1st	0	2nd	BioCarta	3	h_ghPathway:Growth Hormone Signaling Pathway
1st	0	2nd	BioCarta	3	h_il5Pathway:IL 5 Signaling Pathway

1st	0	2nd	BioCarta	3	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	2nd	BioCarta	3	h_ionPathway:Ion Channel and Phorbal Esters Signaling Pathway
1st	0	2nd	BioCarta	3	h_keratinocytePathway:Keratinocyte Differentiation
1st	0	2nd	BioCarta	3	h_mef2dPathway:Role of MEF2D in T-cell Apoptosis
1st	0	2nd	BioCarta	3	h_mhcPathway:Antigen Processing and Presentation
1st	0	2nd	BioCarta	3	h_myosinPathway:PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase
1st	0	2nd	BioCarta	3	h_nos1Pathway:Nitric Oxide Signaling Pathway
1st	0	2nd	BioCarta	3	h_Par1Pathway:Thrombin signaling and protease-activated receptors
1st	0	2nd	BioCarta	3	h_pdgfPathway:PDGF Signaling Pathway
1st	0	2nd	BioCarta	3	h_pkcPathway:Activation of PKC through G protein coupled receptor
1st	0	2nd	BioCarta	3	h_plcdPathway:Phospholipase C d1 in phospholipid associated cell signaling
1st	0	2nd	BioCarta	3	h_plcPathway:Phospholipase C Signaling Pathway
1st	0	2nd	BioCarta	3	h_pparaPathway:Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)
1st	0	2nd	BioCarta	3	h_pyk2Pathway:Links between Pyk2 and Map Kinases
1st	0	2nd	BioCarta	3	h_sppaPathway:Aspirin Blocks Signaling Pathway Involved in Platelet Activation
1st	0	2nd	BioCarta	3	h_srcRPTPPathway:Activation of Src by Protein-tyrosine phosphatase alpha
1st	0	2nd	BioCarta	3	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
1st	0	2nd	BioCarta	3	h_tcrPathway:T Cell Receptor Signaling Pathway
1st	0	2nd	BioCarta	3	h_telPathway:Telomeres, Telomerase, Cellular Aging, and Immortality
1st	0	2nd	BioCarta	3	h_th1th2Pathway:Th1/Th2 Differentiation
1st	0	2nd	BioCarta	3	h_TPOPathway:TPO Signaling Pathway
1st	0	2nd	BioCarta	3	h_trkaPathway:Trka Receptor Signaling Pathway
1st	0	2nd	BioCarta	3	h_vegfPathway:VEGF, Hypoxia, and Angiogenesis
1st	0	2nd	BioCarta	5	h_PDZsPathway:Synaptic Proteins at the Synaptic Junction,
1st	0	2nd	BioCarta	6	h_agpcrPathway:Attenuation of GPCR Signaling
1st	0	2nd	BioCarta	6	h_arenrf2Pathway:Oxidative Stress Induced Gene Expression Via Nrf2
1st	0	2nd	BioCarta	6	h_asbcellPathway:Antigen Dependent B Cell Activation
1st	0	2nd	BioCarta	6	h_At1rPathway:Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling
1st	0	2nd	BioCarta	6	h_bbcellPathway:Bystander B Cell Activation
1st	0	2nd	BioCarta	6	h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway

1st	0	2nd	BioCarta	6	h_bcrPathway:BCR Signaling Pathway
1st	0	2nd	BioCarta	6	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
1st	0	2nd	BioCarta	6	h_calcineurinPathway:Effects of calcineurin in Keratinocyte Differentiation
1st	0	2nd	BioCarta	6	h_cardiacegfPathway:Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy
1st	0	2nd	BioCarta	6	h_cblPathway:CBL mediated ligand-induced downregulation of EGF receptors
1st	0	2nd	BioCarta	6	h_CCR3Pathway:CCR3 signaling in Eosinophils
1st	0	2nd	BioCarta	6	h_Ccr5Pathway:Pertussis toxin-insensitive CCR5 Signaling in Macrophage
1st	0	2nd	BioCarta	6	h_cdMacPathway:Cadmium induces DNA synthesis and proliferation in macrophages
1st	0	2nd	BioCarta	6	h_chemicalPathway:Apoptotic Signaling in Response to DNA Damage
1st	0	2nd	BioCarta	6	h_crebPathway:Transcription factor CREB and its extracellular signals
1st	0	2nd	BioCarta	6	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
1st	0	2nd	BioCarta	6	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
1st	0	2nd	BioCarta	6	h_cxcr4Pathway:CXCR4 Signaling Pathway
1st	0	2nd	BioCarta	6	h_edg1Pathway:Phospholipids as signalling intermediaries
1st	0	2nd	BioCarta	6	h_egfPathway:EGF Signaling Pathway
1st	0	2nd	BioCarta	6	h_eif4Pathway:Regulation of eIF4e and p70 S6 Kinase
1st	0	2nd	BioCarta	6	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
1st	0	2nd	BioCarta	6	h_erbB4pathway:g-Secretase mediated ErbB4 Signaling Pathway
1st	0	2nd	BioCarta	6	h_ghPathway:Growth Hormone Signaling Pathway
1st	0	2nd	BioCarta	6	h_gpcrPathway:Signaling Pathway from G-Protein Families
1st	0	2nd	BioCarta	6	h_il5Pathway:IL 5 Signaling Pathway
1st	0	2nd	BioCarta	6	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	2nd	BioCarta	6	h_ionPathway:Ion Channel and Phorbal Esters Signaling Pathway
1st	0	2nd	BioCarta	6	h_keratinocytePathway:Keratinocyte Differentiation
1st	0	2nd	BioCarta	6	h_mef2dPathway:Role of MEF2D in T-cell Apoptosis
1st	0	2nd	BioCarta	6	h_mhcPathway:Antigen Processing and Presentation
1st	0	2nd	BioCarta	6	h_myosinPathway:PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase
1st	0	2nd	BioCarta	6	h_nos1Pathway:Nitric Oxide Signaling Pathway

1st	0	2nd	BioCarta	6	h_Par1Pathway:Thrombin signaling and protease-activated receptors
1st	0	2nd	BioCarta	6	h_pdgfPathway:PDGF Signaling Pathway
1st	0	2nd	BioCarta	6	h_pkcPathway:Activation of PKC through G protein coupled receptor
1st	0	2nd	BioCarta	6	h_plcdPathway:Phospholipase C d1 in phospholipid associated cell signaling
1st	0	2nd	BioCarta	6	h_plcPathway:Phospholipase C Signaling Pathway
1st	0	2nd	BioCarta	6	h_pparaPathway:Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)
1st	0	2nd	BioCarta	6	h_pyk2Pathway:Links between Pyk2 and Map Kinases
1st	0	2nd	BioCarta	6	h_sppaPathway:Aspirin Blocks Signaling Pathway Involved in Platelet Activation
1st	0	2nd	BioCarta	6	h_srcRPTPPathway:Activation of Src by Protein-tyrosine phosphatase alpha
1st	0	2nd	BioCarta	6	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
1st	0	2nd	BioCarta	6	h_tcrPathway:T Cell Receptor Signaling Pathway
1st	0	2nd	BioCarta	6	h_telPathway:Telomeres, Telomerase, Cellular Aging, and Immortality
1st	0	2nd	BioCarta	6	h_th1th2Pathway:Th1/Th2 Differentiation
1st	0	2nd	BioCarta	6	h_TPOPathway:TPO Signaling Pathway
1st	0	2nd	BioCarta	6	h_trkaPathway:Trka Receptor Signaling Pathway
1st	0	2nd	BioCarta	6	h_vegfPathway:VEGF, Hypoxia, and Angiogenesis
1st	0	2nd	BioCarta	7	h_aktPathway:AKT Signaling Pathway
1st	0	2nd	BioCarta	7	h_alkPathway:ALK in cardiac myocytes
1st	0	2nd	BioCarta	7	h_aMANpathway:Steps in the Glycosylation of Mammalian N-linked Oligosaccarides
1st	0	2nd	BioCarta	7	h_arapPathway:ADP-Ribosylation Factor
1st	0	2nd	BioCarta	7	h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway
1st	0	2nd	BioCarta	7	h_ctcfPathway:CTCF: First Multivalent Nuclear Factor
1st	0	2nd	BioCarta	7	h_eponfkbPathway:Erythropoietin mediated neuroprotection through NF-kB
1st	0	2nd	BioCarta	7	h_epoPathway:EPO Signaling Pathway
1st	0	2nd	BioCarta	7	h_eradPathway:ER???associated degradation (ERAD) Pathway
1st	0	2nd	BioCarta	7	h_erythPathway:Erythrocyte Differentiation Pathway
1st	0	2nd	BioCarta	7	h_g1Pathway:Cell Cycle: G1/S Check Point
1st	0	2nd	BioCarta	7	h_ghPathway:Growth Hormone Signaling Pathway
1st	0	2nd	BioCarta	7	h_gleevecpathway:Inhibition of Cellular Proliferation by Gleevec
1st	0	2nd	BioCarta	7	h_IL12Pathway:IL12 and Stat4 Dependent Signaling Pathway in Th1 Development
1st	0	2nd	BioCarta	7	h_ifngPathway:IFN gamma signaling pathway
1st	0	2nd	BioCarta	7	h_il1rPathway:Signal transduction through IL1R

1st	0	2nd	BioCarta	7	h_il3Pathway:IL 3 signaling pathway
1st	0	2nd	BioCarta	7	h_inflamPathway:Cytokines and Inflammatory Response
1st	0	2nd	BioCarta	7	h_mapkPathway:MAPKinase Signaling Pathway
1st	0	2nd	BioCarta	7	h_nktPathway:Selective expression of chemokine receptors during T-cell polarization
1st	0	2nd	BioCarta	7	h_no1Pathway:Actions of Nitric Oxide in the Heart
1st	0	2nd	BioCarta	7	h_no2il12Pathway:NO2-dependent IL 12 Pathway in NK cells
1st	0	2nd	BioCarta	7	h_p38mapkPathway:p38 MAPK Signaling Pathway
1st	0	2nd	BioCarta	7	h_ptdinsPathway:Phosphoinositides and their downstream targets.
1st	0	2nd	BioCarta	7	h_ranbp2Pathway:Sumoylation by RanBP2 Regulates Transcriptional Repression
1st	0	2nd	BioCarta	7	h_rhodopsinPathway:Visual Signal Transduction
1st	0	2nd	BioCarta	7	h_slrp2Pathway:Function of SLRP in Bone: An Integrated View
1st	0	2nd	BioCarta	7	h_tgfbPathway:TGF beta signaling pathway
1st	0	2nd	BioCarta	7	h_tidPathway:Chaperones modulate interferon Signaling Pathway
1st	0	2nd	BioCarta	7	h_tob1Pathway:Role of Tob in T-cell activation
1st	0	2nd	BioCarta	7	h_TPOPathway:TPO Signaling Pathway
1st	0	2nd	BioCarta	7	h_vitCBPathway:Vitamin C in the Brain
1st	0	2nd	BioCarta	10	h_metPathway:Signaling of Hepatocyte Growth Factor Receptor
2nd	0	2nd	BioCarta	6	h_asbcellPathway:Antigen Dependent B Cell Activation
2nd	0	2nd	BioCarta	6	h_bbcellPathway:Bystander B Cell Activation
2nd	0	2nd	BioCarta	6	h_blymphocytePathway:B Lymphocyte Cell Surface Molecules
2nd	0	2nd	BioCarta	6	h_CSKPathway:Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
2nd	0	2nd	BioCarta	6	h_ctla4Pathway:The Co-Stimulatory Signal During T-cell Activation
2nd	0	2nd	BioCarta	6	h_eosinophilsPathway:The Role of Eosinophils in the Chemokine Network of Allergy
2nd	0	2nd	BioCarta	6	h_il5Pathway:IL 5 Signaling Pathway
2nd	0	2nd	BioCarta	6	h_inflamPathway:Cytokines and Inflammatory Response
2nd	0	2nd	BioCarta	6	h_mhcPathway:Antigen Processing and Presentation
2nd	0	2nd	BioCarta	6	h_tcraPathway:Lck and Fyn tyrosine kinases in initiation of TCR Activation
2nd	0	2nd	BioCarta	6	h_th1th2Pathway:Th1/Th2 Differentiation
2nd	0	2nd	BioCarta	7	h_aktPathway:AKT Signaling Pathway
2nd	0	2nd	BioCarta	7	h_alkPathway:ALK in cardiac myocytes
2nd	0	2nd	BioCarta	7	h_aMANpathway:Steps in the Glycosylation of Mammalian N-linked Oligosaccarides

2nd	0	2nd	BioCarta	7	h_arapPathway:ADP-Ribosylation Factor
2nd	0	2nd	BioCarta	7	h_biopeptidesPathway:Bioactive Peptide Induced Signaling Pathway
2nd	0	2nd	BioCarta	7	h_ctcfPathway:CTCF: First Multivalent Nuclear Factor
2nd	0	2nd	BioCarta	7	h_eponfkbPathway:Erythropoietin mediated neuroprotection through NF-kB
2nd	0	2nd	BioCarta	7	h_epoPathway:EPO Signaling Pathway
2nd	0	2nd	BioCarta	7	h_eradPathway:ER???associated degradation (ERAD) Pathway
2nd	0	2nd	BioCarta	7	h_erythPathway:Erythrocyte Differentiation Pathway
2nd	0	2nd	BioCarta	7	h_g1Pathway:Cell Cycle: G1/S Check Point
2nd	0	2nd	BioCarta	7	h_ghPathway:Growth Hormone Signaling Pathway
2nd	0	2nd	BioCarta	7	h_gleevecpathway:Inhibition of Cellular Proliferation by Gleevec
2nd	0	2nd	BioCarta	7	h_IL12Pathway:IL12 and Stat4 Dependent Signaling Pathway in Th1 Development
2nd	0	2nd	BioCarta	7	h_ifngPathway:IFN gamma signaling pathway
2nd	0	2nd	BioCarta	7	h_il1rPathway:Signal transduction through IL1R
2nd	0	2nd	BioCarta	7	h_il3Pathway:IL 3 signaling pathway
2nd	0	2nd	BioCarta	7	h_inflamPathway:Cytokines and Inflammatory Response
2nd	0	2nd	BioCarta	7	h_mapkPathway:MAPKinase Signaling Pathway
2nd	0	2nd	BioCarta	7	h_nktPathway:Selective expression of chemokine receptors during T-cell polarization
2nd	0	2nd	BioCarta	7	h_no1Pathway:Actions of Nitric Oxide in the Heart
2nd	0	2nd	BioCarta	7	h_no2il12Pathway:NO2-dependent IL 12 Pathway in NK cells
2nd	0	2nd	BioCarta	7	h_p38mapkPathway:p38 MAPK Signaling Pathway
2nd	0	2nd	BioCarta	7	h_ptdinsPathway:Phosphoinositides and their downstream targets.
2nd	0	2nd	BioCarta	7	h_ranbp2Pathway:Sumoylation by RanBP2 Regulates Transcriptional Repression
2nd	0	2nd	BioCarta	7	h_rhodopsinPathway:Visual Signal Transduction
2nd	0	2nd	BioCarta	7	h_slrp2Pathway:Function of SLRP in Bone: An Integrated View
2nd	0	2nd	BioCarta	7	h_tgfbPathway:TGF beta signaling pathway
2nd	0	2nd	BioCarta	7	h_tidPathway:Chaperones modulate interferon Signaling Pathway
2nd	0	2nd	BioCarta	7	h_tob1Pathway:Role of Tob in T-cell activation
2nd	0	2nd	BioCarta	7	h_TPOPathway:TPO Signaling Pathway
2nd	0	2nd	BioCarta	7	h_vitCBPathway:Vitamin C in the Brain
1st	1	2nd	BioCarta	5	h_PDZsPathway:Synaptic Proteins at the Synaptic Junction
1st	1	2nd	BioCarta	7	h_no1Pathway:Actions of Nitric Oxide in the Heart
2nd	1	2nd	BioCarta	7	h_no1Pathway:Actions of Nitric Oxide in the Heart
1st	0	1st	KEGG	2	hsa04510:Focal adhesion

1st	0	1st	KEGG	2	hsa04514:Cell adhesion molecules (CAMs)
1st	0	1st	KEGG	2	hsa04612:Antigen processing and presentation
1st	0	1st	KEGG	2	hsa04640:Hematopoietic cell lineage
1st	0	1st	KEGG	2	hsa04672:Intestinal immune network for IgA production
1st	0	1st	KEGG	2	hsa04810:Regulation of actin cytoskeleton
1st	0	1st	KEGG	2	hsa04940:Type I diabetes mellitus
1st	0	1st	KEGG	2	hsa05310:Asthma
1st	0	1st	KEGG	2	hsa05320:Autoimmune thyroid disease
1st	0	1st	KEGG	2	hsa05322:Systemic lupus erythematosus
1st	0	1st	KEGG	2	hsa05330:Allograft rejection
1st	0	1st	KEGG	2	hsa05332:Graft-versus-host disease
1st	0	1st	KEGG	2	hsa05416:Viral myocarditis
1st	0	1st	KEGG	3	hsa03050:Proteasome
1st	0	1st	KEGG	3	hsa04514:Cell adhesion molecules (CAMs)
1st	0	1st	KEGG	3	hsa04612:Antigen processing and presentation
1st	0	1st	KEGG	3	hsa04640:Hematopoietic cell lineage
1st	0	1st	KEGG	3	hsa04672:Intestinal immune network for IgA production
1st	0	1st	KEGG	3	hsa04940:Type I diabetes mellitus
1st	0	1st	KEGG	3	hsa05310:Asthma
1st	0	1st	KEGG	3	hsa05320:Autoimmune thyroid disease
1st	0	1st	KEGG	3	hsa05322:Systemic lupus erythematosus
1st	0	1st	KEGG	3	hsa05330:Allograft rejection
1st	0	1st	KEGG	3	hsa05332:Graft-versus-host disease
1st	0	1st	KEGG	3	hsa05416:Viral myocarditis
1st	0	1st	KEGG	5	hsa03440:Homologous recombination
1st	0	1st	KEGG	5	hsa04330:Notch signaling pathway
1st	0	1st	KEGG	5	hsa04360:Axon guidance
1st	0	1st	KEGG	5	hsa04510:Focal adhesion
1st	0	1st	KEGG	5	hsa04530:Tight junction
1st	0	1st	KEGG	5	hsa04810:Regulation of actin cytoskeleton
1st	0	1st	KEGG	6	hsa04080:Neuroactive ligand-receptor interaction
1st	0	1st	KEGG	6	hsa04144:Endocytosis
1st	0	1st	KEGG	6	hsa04514:Cell adhesion molecules (CAMs)
1st	0	1st	KEGG	6	hsa04612:Antigen processing and presentation
1st	0	1st	KEGG	6	hsa04640:Hematopoietic cell lineage
1st	0	1st	KEGG	6	hsa04650:Natural killer cell mediated cytotoxicity
1st	0	1st	KEGG	6	hsa04672:Intestinal immune network for IgA production
1st	0	1st	KEGG	6	hsa04940:Type I diabetes mellitus
1st	0	1st	KEGG	6	hsa05310:Asthma
1st	0	1st	KEGG	6	hsa05320:Autoimmune thyroid disease
1st	0	1st	KEGG	6	hsa05322:Systemic lupus erythematosus
1st	0	1st	KEGG	6	hsa05330:Allograft rejection

1st	0	1st	KEGG	6	hsa05332:Graft-versus-host disease
1st	0	1st	KEGG	6	hsa05416:Viral myocarditis
1st	0	1st	KEGG	7	hsa03040:Spliceosome
1st	0	1st	KEGG	7	hsa04020:Calcium signaling pathway
1st	0	1st	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
1st	0	1st	KEGG	7	hsa04270:Vascular smooth muscle contraction
1st	0	1st	KEGG	7	hsa04310:Wnt signaling pathway
1st	0	1st	KEGG	7	hsa04514:Cell adhesion molecules (CAMs)
1st	0	1st	KEGG	7	hsa04520:Adherens junction
1st	0	1st	KEGG	7	hsa04530:Tight junction
1st	0	1st	KEGG	7	hsa04612:Antigen processing and presentation
1st	0	1st	KEGG	7	hsa04672:Intestinal immune network for IgA production
1st	0	1st	KEGG	7	hsa04916:Melanogenesis
1st	0	1st	KEGG	7	hsa04940:Type I diabetes mellitus
1st	0	1st	KEGG	7	hsa05200:Pathways in cancer
1st	0	1st	KEGG	7	hsa05210:Colorectal cancer
1st	0	1st	KEGG	7	hsa05211:Renal cell carcinoma
1st	0	1st	KEGG	7	hsa05213:Endometrial cancer
1st	0	1st	KEGG	7	hsa05215:Prostate cancer
1st	0	1st	KEGG	7	hsa05216:Thyroid cancer
1st	0	1st	KEGG	7	hsa05217:Basal cell carcinoma
1st	0	1st	KEGG	7	hsa05221:Acute myeloid leukemia
1st	0	1st	KEGG	7	hsa05310:Asthma
1st	0	1st	KEGG	7	hsa05320:Autoimmune thyroid disease
1st	0	1st	KEGG	7	hsa05322:Systemic lupus erythematosus
1st	0	1st	KEGG	7	hsa05330:Allograft rejection
1st	0	1st	KEGG	7	hsa05332:Graft-versus-host disease
1st	0	1st	KEGG	7	hsa05412:Arrhythmogenic right ventricular cardiomyopathy (ARVC)
1st	0	1st	KEGG	7	hsa05416:Viral myocarditis
1st	0	1st	KEGG	8	hsa00590:Arachidonic acid metabolism
1st	0	1st	KEGG	8	hsa02010:ABC transporters
1st	0	1st	KEGG	9	hsa04010:MAPK signaling pathway
1st	0	1st	KEGG	9	hsa04012:ErbB signaling pathway
1st	0	1st	KEGG	9	hsa04020:Calcium signaling pathway
1st	0	1st	KEGG	9	hsa04070:Phosphatidylinositol signaling system
1st	0	1st	KEGG	9	hsa04270:Vascular smooth muscle contraction
1st	0	1st	KEGG	9	hsa04310:Wnt signaling pathway
1st	0	1st	KEGG	9	hsa04370:VEGF signaling pathway
1st	0	1st	KEGG	9	hsa04510:Focal adhesion
1st	0	1st	KEGG	9	hsa04514:Cell adhesion molecules (CAMs)
1st	0	1st	KEGG	9	hsa04530:Tight junction
1st	0	1st	KEGG	9	hsa04540:Gap junction
1st	0	1st	KEGG	9	hsa04612:Antigen processing and presentation

1st	0	1st	KEGG	9	hsa04622:RIG-I-like receptor signaling pathway
1st	0	1st	KEGG	9	hsa04650:Natural killer cell mediated cytotoxicity
1st	0	1st	KEGG	9	hsa04664:Fc epsilon RI signaling pathway
1st	0	1st	KEGG	9	hsa04666:Fc gamma R-mediated phagocytosis
1st	0	1st	KEGG	9	hsa04670:Leukocyte transendothelial migration
1st	0	1st	KEGG	9	hsa04720:Long-term potentiation
1st	0	1st	KEGG	9	hsa04730:Long-term depression
1st	0	1st	KEGG	9	hsa04912:GnRH signaling pathway
1st	0	1st	KEGG	9	hsa04916:Melanogenesis
1st	0	1st	KEGG	9	hsa04960:Aldosterone-regulated sodium reabsorption
1st	0	1st	KEGG	9	hsa05110:Vibrio cholerae infection
1st	0	1st	KEGG	9	hsa05130:Pathogenic Escherichia coli infection
1st	0	1st	KEGG	9	hsa05200:Pathways in cancer
1st	0	1st	KEGG	9	hsa05214:Glioma
1st	0	1st	KEGG	9	hsa05223:Non-small cell lung cancer
1st	0	1st	KEGG	9	hsa05340:Primary immunodeficiency
1st	0	1st	KEGG	10	hsa03440:Homologous recombination
1st	0	1st	KEGG	10	hsa04340:Hedgehog signaling pathway
1st	0	1st	KEGG	10	hsa04740:Olfactory transduction
2nd	0	1st	KEGG	2	hsa04514:Cell adhesion molecules (CAMs)
2nd	0	1st	KEGG	2	hsa04612:Antigen processing and presentation
2nd	0	1st	KEGG	2	hsa04640:Hematopoietic cell lineage
2nd	0	1st	KEGG	2	hsa04672:Intestinal immune network for IgA production
2nd	0	1st	KEGG	2	hsa04940:Type I diabetes mellitus
2nd	0	1st	KEGG	2	hsa05310:Asthma
2nd	0	1st	KEGG	2	hsa05320:Autoimmune thyroid disease
2nd	0	1st	KEGG	2	hsa05322:Systemic lupus erythematosus
2nd	0	1st	KEGG	2	hsa05330:Allograft rejection
2nd	0	1st	KEGG	2	hsa05332:Graft-versus-host disease
2nd	0	1st	KEGG	2	hsa05416:Viral myocarditis
2nd	0	1st	KEGG	3	hsa04514:Cell adhesion molecules (CAMs)
2nd	0	1st	KEGG	3	hsa04612:Antigen processing and presentation
2nd	0	1st	KEGG	3	hsa04640:Hematopoietic cell lineage
2nd	0	1st	KEGG	3	hsa04672:Intestinal immune network for IgA production
2nd	0	1st	KEGG	3	hsa04940:Type I diabetes mellitus
2nd	0	1st	KEGG	3	hsa05310:Asthma
2nd	0	1st	KEGG	3	hsa05320:Autoimmune thyroid disease
2nd	0	1st	KEGG	3	hsa05322:Systemic lupus erythematosus
2nd	0	1st	KEGG	3	hsa05330:Allograft rejection
2nd	0	1st	KEGG	3	hsa05332:Graft-versus-host disease
2nd	0	1st	KEGG	3	hsa05416:Viral myocarditis
2nd	0	1st	KEGG	4	hsa04530:Tight junction

2nd	0	1st	KEGG	5	hsa04530:Tight junction
2nd	0	1st	KEGG	6	hsa04514:Cell adhesion molecules (CAMs)
2nd	0	1st	KEGG	6	hsa04612:Antigen processing and presentation
2nd	0	1st	KEGG	6	hsa04672:Intestinal immune network for IgA production
2nd	0	1st	KEGG	6	hsa04940:Type I diabetes mellitus
2nd	0	1st	KEGG	6	hsa05310:Asthma
2nd	0	1st	KEGG	6	hsa05320:Autoimmune thyroid disease
2nd	0	1st	KEGG	6	hsa05322:Systemic lupus erythematosus
2nd	0	1st	KEGG	6	hsa05330:Allograft rejection
2nd	0	1st	KEGG	6	hsa05332:Graft-versus-host disease
2nd	0	1st	KEGG	6	hsa05416:Viral myocarditis
2nd	0	1st	KEGG	7	hsa03040:Spliceosome
2nd	0	1st	KEGG	7	hsa04020:Calcium signaling pathway
2nd	0	1st	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
2nd	0	1st	KEGG	7	hsa04310:Wnt signaling pathway
2nd	0	1st	KEGG	7	hsa04520:Adherens junction
2nd	0	1st	KEGG	7	hsa04916:Melanogenesis
2nd	0	1st	KEGG	7	hsa05200:Pathways in cancer
2nd	0	1st	KEGG	7	hsa05210:Colorectal cancer
2nd	0	1st	KEGG	7	hsa05211:Renal cell carcinoma
2nd	0	1st	KEGG	7	hsa05213:Endometrial cancer
2nd	0	1st	KEGG	7	hsa05215:Prostate cancer
2nd	0	1st	KEGG	7	hsa05216:Thyroid cancer
2nd	0	1st	KEGG	7	hsa05217:Basal cell carcinoma
2nd	0	1st	KEGG	7	hsa05221:Acute myeloid leukemia
2nd	0	1st	KEGG	7	hsa05412:Arrhythmogenic right ventricular cardiomyopathy (ARVC)
2nd	0	1st	KEGG	10	hsa04340:Hedgehog signaling pathway
2nd	0	1st	KEGG	10	hsa04740:Olfactory transduction
1st	1	1st	KEGG	3	hsa04514:Cell adhesion molecules (CAMs)
1st	1	1st	KEGG	3	hsa04612:Antigen processing and presentation
1st	1	1st	KEGG	3	hsa04640:Hematopoietic cell lineage
1st	1	1st	KEGG	3	hsa04672:Intestinal immune network for IgA production
1st	1	1st	KEGG	3	hsa04940:Type I diabetes mellitus
1st	1	1st	KEGG	3	hsa05310:Asthma
1st	1	1st	KEGG	3	hsa05320:Autoimmune thyroid disease
1st	1	1st	KEGG	3	hsa05322:Systemic lupus erythematosus
1st	1	1st	KEGG	3	hsa05330:Allograft rejection
1st	1	1st	KEGG	3	hsa05332:Graft-versus-host disease
1st	1	1st	KEGG	3	hsa05416:Viral myocarditis
1st	1	1st	KEGG	4	hsa04530:Tight junction
1st	1	1st	KEGG	6	hsa04080:Neuroactive ligand-receptor interaction
1st	1	1st	KEGG	7	hsa04020:Calcium signaling pathway

1st	1	1st	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
1st	1	1st	KEGG	7	hsa04514:Cell adhesion molecules (CAMs)
1st	1	1st	KEGG	7	hsa04612:Antigen processing and presentation
1st	1	1st	KEGG	7	hsa04672:Intestinal immune network for IgA production
1st	1	1st	KEGG	7	hsa04940:Type I diabetes mellitus
1st	1	1st	KEGG	7	hsa05310:Asthma
1st	1	1st	KEGG	7	hsa05320:Autoimmune thyroid disease
1st	1	1st	KEGG	7	hsa05322:Systemic lupus erythematosus
1st	1	1st	KEGG	7	hsa05330:Allograft rejection
1st	1	1st	KEGG	7	hsa05332:Graft-versus-host disease
1st	1	1st	KEGG	7	hsa05416:Viral myocarditis
1st	1	1st	KEGG	8	hsa02010:ABC transporters
1st	1	1st	KEGG	10	hsa04740:Olfactory transduction
2nd	1	1st	KEGG	4	hsa04530:Tight junction
2nd	1	1st	KEGG	7	hsa04020:Calcium signaling pathway
2nd	1	1st	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
2nd	1	1st	KEGG	10	hsa04740:Olfactory transduction
1st	0	2nd	KEGG	1	hsa04740:Olfactory transduction
1st	0	2nd	KEGG	3	hsa04010:MAPK signaling pathway
1st	0	2nd	KEGG	3	hsa04012:ErbB signaling pathway
1st	0	2nd	KEGG	3	hsa04020:Calcium signaling pathway
1st	0	2nd	KEGG	3	hsa04070:Phosphatidylinositol signaling system
1st	0	2nd	KEGG	3	hsa04270:Vascular smooth muscle contraction
1st	0	2nd	KEGG	3	hsa04310:Wnt signaling pathway
1st	0	2nd	KEGG	3	hsa04370:VEGF signaling pathway
1st	0	2nd	KEGG	3	hsa04510:Focal adhesion
1st	0	2nd	KEGG	3	hsa04514:Cell adhesion molecules (CAMs)
1st	0	2nd	KEGG	3	hsa04530:Tight junction
1st	0	2nd	KEGG	3	hsa04540:Gap junction
1st	0	2nd	KEGG	3	hsa04612:Antigen processing and presentation
1st	0	2nd	KEGG	3	hsa04640:Hematopoietic cell lineage
1st	0	2nd	KEGG	3	hsa04650:Natural killer cell mediated cytotoxicity
1st	0	2nd	KEGG	3	hsa04664:Fc epsilon RI signaling pathway
1st	0	2nd	KEGG	3	hsa04666:Fc gamma R-mediated phagocytosis
1st	0	2nd	KEGG	3	hsa04670:Leukocyte transendothelial migration
1st	0	2nd	KEGG	3	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	3	hsa04720:Long-term potentiation
1st	0	2nd	KEGG	3	hsa04730:Long-term depression
1st	0	2nd	KEGG	3	hsa04912:GnRH signaling pathway
1st	0	2nd	KEGG	3	hsa04916:Melanogenesis
1st	0	2nd	KEGG	3	hsa04940:Type I diabetes mellitus
1st	0	2nd	KEGG	3	hsa04960:Aldosterone-regulated sodium reabsorption

1st	0	2nd	KEGG	3	hsa05110:Vibrio cholerae infection
1st	0	2nd	KEGG	3	hsa05130:Pathogenic Escherichia coli infection
1st	0	2nd	KEGG	3	hsa05200:Pathways in cancer
1st	0	2nd	KEGG	3	hsa05214:Glioma
1st	0	2nd	KEGG	3	hsa05223:Non-small cell lung cancer
1st	0	2nd	KEGG	3	hsa05310:Asthma
1st	0	2nd	KEGG	3	hsa05320:Autoimmune thyroid disease
1st	0	2nd	KEGG	3	hsa05322:Systemic lupus erythematosus
1st	0	2nd	KEGG	3	hsa05330:Allograft rejection
1st	0	2nd	KEGG	3	hsa05332:Graft-versus-host disease
1st	0	2nd	KEGG	3	hsa05416:Viral myocarditis
1st	0	2nd	KEGG	4	hsa04340:Hedgehog signaling pathway
1st	0	2nd	KEGG	4	hsa04740:Olfactory transduction
1st	0	2nd	KEGG	6	hsa04010:MAPK signaling pathway
1st	0	2nd	KEGG	6	hsa04012:ErbB signaling pathway
1st	0	2nd	KEGG	6	hsa04020:Calcium signaling pathway
1st	0	2nd	KEGG	6	hsa04070:Phosphatidylinositol signaling system
1st	0	2nd	KEGG	6	hsa04270:Vascular smooth muscle contraction
1st	0	2nd	KEGG	6	hsa04310:Wnt signaling pathway
1st	0	2nd	KEGG	6	hsa04370:VEGF signaling pathway
1st	0	2nd	KEGG	6	hsa04510:Focal adhesion
1st	0	2nd	KEGG	6	hsa04514:Cell adhesion molecules (CAMs)
1st	0	2nd	KEGG	6	hsa04530:Tight junction
1st	0	2nd	KEGG	6	hsa04540:Gap junction
1st	0	2nd	KEGG	6	hsa04612:Antigen processing and presentation
1st	0	2nd	KEGG	6	hsa04640:Hematopoietic cell lineage
1st	0	2nd	KEGG	6	hsa04650:Natural killer cell mediated cytotoxicity
1st	0	2nd	KEGG	6	hsa04664:Fc epsilon RI signaling pathway
1st	0	2nd	KEGG	6	hsa04666:Fc gamma R-mediated phagocytosis
1st	0	2nd	KEGG	6	hsa04670:Leukocyte transendothelial migration
1st	0	2nd	KEGG	6	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	6	hsa04720:Long-term potentiation
1st	0	2nd	KEGG	6	hsa04730:Long-term depression
1st	0	2nd	KEGG	6	hsa04912:GnRH signaling pathway
1st	0	2nd	KEGG	6	hsa04916:Melanogenesis
1st	0	2nd	KEGG	6	hsa04940:Type I diabetes mellitus
1st	0	2nd	KEGG	6	hsa04960:Aldosterone-regulated sodiumreabsorption
1st	0	2nd	KEGG	6	hsa05110:Vibrio cholerae infection
1st	0	2nd	KEGG	6	hsa05130:Pathogenic Escherichia coli infection
1st	0	2nd	KEGG	6	hsa05200:Pathways in cancer
1st	0	2nd	KEGG	6	hsa05214:Glioma
1st	0	2nd	KEGG	6	hsa05223:Non-small cell lung cancer
1st	0	2nd	KEGG	6	hsa05310:Asthma
1st	0	2nd	KEGG	6	hsa05320:Autoimmune thyroid disease

1st	0	2nd	KEGG	6	hsa05322:Systemic lupus erythematosus
1st	0	2nd	KEGG	6	hsa05330:Allograft rejection
1st	0	2nd	KEGG	6	hsa05332:Graft-versus-host disease
1st	0	2nd	KEGG	6	hsa05416:Viral myocarditis
1st	0	2nd	KEGG	7	hsa00510:N-Glycan biosynthesis
1st	0	2nd	KEGG	7	hsa00512:O-Glycan biosynthesis
1st	0	2nd	KEGG	7	hsa00531:Glycosaminoglycan degradation
1st	0	2nd	KEGG	7	hsa00533:Keratan sulfate biosynthesis
1st	0	2nd	KEGG	7	hsa00562:Inositol phosphate metabolism
1st	0	2nd	KEGG	7	hsa00770:Pantothenate and CoA biosynthesis
1st	0	2nd	KEGG	7	hsa03040:Spliceosome
1st	0	2nd	KEGG	7	hsa03060:Protein export
1st	0	2nd	KEGG	7	hsa04010:MAPK signaling pathway
1st	0	2nd	KEGG	7	hsa04060:Cytokine-cytokine receptor interaction
1st	0	2nd	KEGG	7	hsa04062:Chemokine signaling pathway
1st	0	2nd	KEGG	7	hsa04070:Phosphatidylinositol signaling system
1st	0	2nd	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
1st	0	2nd	KEGG	7	hsa04110:Cell cycle
1st	0	2nd	KEGG	7	hsa04142:Lysosome
1st	0	2nd	KEGG	7	hsa04144:Endocytosis
1st	0	2nd	KEGG	7	hsa04310:Wnt signaling pathway
1st	0	2nd	KEGG	7	hsa04320:Dorso-ventral axis formation
1st	0	2nd	KEGG	7	hsa04350:TGF-beta signaling pathway
1st	0	2nd	KEGG	7	hsa04360:Axon guidance
1st	0	2nd	KEGG	7	hsa04530:Tight junction
1st	0	2nd	KEGG	7	hsa04630:Jak-STAT signaling pathway
1st	0	2nd	KEGG	7	hsa04666:Fc gamma R-mediated phagocytosis
1st	0	2nd	KEGG	7	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	7	hsa04740:Olfactory transduction
1st	0	2nd	KEGG	7	hsa04910:Insulin signaling pathway
1st	0	2nd	KEGG	7	hsa04920:Adipocytokine signaling pathway
1st	0	2nd	KEGG	7	hsa04930:Type II diabetes mellitus
1st	0	2nd	KEGG	7	hsa05200:Pathways in cancer
1st	0	2nd	KEGG	7	hsa05210:Colorectal cancer
1st	0	2nd	KEGG	7	hsa05211:Renal cell carcinoma
1st	0	2nd	KEGG	7	hsa05212:Pancreatic cancer
1st	0	2nd	KEGG	7	hsa05215:Prostate cancer
1st	0	2nd	KEGG	7	hsa05220:Chronic myeloid leukemia
1st	0	2nd	KEGG	7	hsa05410:Hypertrophic cardiomyopathy (HCM)
1st	0	2nd	KEGG	7	hsa05414:Dilated cardiomyopathy
1st	0	2nd	KEGG	8	hsa04144:Endocytosis
1st	0	2nd	KEGG	8	hsa04514:Cell adhesion molecules (CAMs)
1st	0	2nd	KEGG	8	hsa04612:Antigen processing and presentation
1st	0	2nd	KEGG	8	hsa04650:Natural killer cell mediated cytotoxicity

1st	0	2nd	KEGG	8	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	8	hsa04740:Olfactory transduction
1st	0	2nd	KEGG	8	hsa04940:Type I diabetes mellitus
1st	0	2nd	KEGG	8	hsa05016:Huntington's disease
1st	0	2nd	KEGG	8	hsa05310:Asthma
1st	0	2nd	KEGG	8	hsa05320:Autoimmune thyroid disease
1st	0	2nd	KEGG	8	hsa05322:Systemic lupus erythematosus
1st	0	2nd	KEGG	8	hsa05330:Allograft rejection
1st	0	2nd	KEGG	8	hsa05332:Graft-versus-host disease
1st	0	2nd	KEGG	8	hsa05416:Viral myocarditis
1st	0	2nd	KEGG	9	hsa04514:Cell adhesion molecules (CAMs)
1st	0	2nd	KEGG	9	hsa04612:Antigen processing and presentation
1st	0	2nd	KEGG	9	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	9	hsa04940:Type I diabetes mellitus
1st	0	2nd	KEGG	9	hsa05310:Asthma
1st	0	2nd	KEGG	9	hsa05320:Autoimmune thyroid disease
1st	0	2nd	KEGG	9	hsa05322:Systemic lupus erythematosus
1st	0	2nd	KEGG	9	hsa05330:Allograft rejection
1st	0	2nd	KEGG	9	hsa05332:Graft-versus-host disease
1st	0	2nd	KEGG	9	hsa05416:Viral myocarditis
1st	0	2nd	KEGG	10	hsa00512:O-Glycan biosynthesis
1st	0	2nd	KEGG	10	hsa04510:Focal adhesion
1st	0	2nd	KEGG	10	hsa04514:Cell adhesion molecules (CAMs)
1st	0	2nd	KEGG	10	hsa04612:Antigen processing and presentation
1st	0	2nd	KEGG	10	hsa04672:Intestinal immune network for IgA production
1st	0	2nd	KEGG	10	hsa04810:Regulation of actin cytoskeleton
1st	0	2nd	KEGG	10	hsa04940:Type I diabetes mellitus
1st	0	2nd	KEGG	10	hsa05310:Asthma
1st	0	2nd	KEGG	10	hsa05320:Autoimmune thyroid disease
1st	0	2nd	KEGG	10	hsa05322:Systemic lupus erythematosus
1st	0	2nd	KEGG	10	hsa05330:Allograft rejection
1st	0	2nd	KEGG	10	hsa05332:Graft-versus-host disease
1st	0	2nd	KEGG	10	hsa05416:Viral myocarditis
2nd	0	2nd	KEGG	1	hsa04740:Olfactory transduction
2nd	0	2nd	KEGG	4	hsa04340:Hedgehog signaling pathway
2nd	0	2nd	KEGG	4	hsa04740:Olfactory transduction
2nd	0	2nd	KEGG	6	hsa04514:Cell adhesion molecules (CAMs)
2nd	0	2nd	KEGG	6	hsa04612:Antigen processing and presentation
2nd	0	2nd	KEGG	6	hsa04640:Hematopoietic cell lineage
2nd	0	2nd	KEGG	6	hsa04672:Intestinal immune network for IgA production
2nd	0	2nd	KEGG	6	hsa04940:Type I diabetes mellitus

2nd	0	2nd	KEGG	6	hsa05310:Asthma
2nd	0	2nd	KEGG	6	hsa05320:Autoimmune thyroid disease
2nd	0	2nd	KEGG	6	hsa05322:Systemic lupus erythematosus
2nd	0	2nd	KEGG	6	hsa05330:Allograft rejection
2nd	0	2nd	KEGG	6	hsa05332:Graft-versus-host disease
2nd	0	2nd	KEGG	6	hsa05416:Viral myocarditis
2nd	0	2nd	KEGG	7	hsa00510:N-Glycan biosynthesis
2nd	0	2nd	KEGG	7	hsa00531:Glycosaminoglycan degradation
2nd	0	2nd	KEGG	7	hsa00533:Keratan sulfate biosynthesis
2nd	0	2nd	KEGG	7	hsa00562:Inositol phosphate metabolism
2nd	0	2nd	KEGG	7	hsa00770:Pantothenate and CoA biosynthesis
2nd	0	2nd	KEGG	7	hsa03060:Protein export
2nd	0	2nd	KEGG	7	hsa04010:MAPK signaling pathway
2nd	0	2nd	KEGG	7	hsa04060:Cytokine-cytokine receptor interaction
2nd	0	2nd	KEGG	7	hsa04062:Chemokine signaling pathway
2nd	0	2nd	KEGG	7	hsa04070:Phosphatidylinositol signaling system
2nd	0	2nd	KEGG	7	hsa04080:Neuroactive ligand-receptor interaction
2nd	0	2nd	KEGG	7	hsa04110:Cell cycle
2nd	0	2nd	KEGG	7	hsa04142:Lysosome
2nd	0	2nd	KEGG	7	hsa04144:Endocytosis
2nd	0	2nd	KEGG	7	hsa04310:Wnt signaling pathway
2nd	0	2nd	KEGG	7	hsa04320:Dorso-ventral axis formation
2nd	0	2nd	KEGG	7	hsa04350:TGF-beta signaling pathway
2nd	0	2nd	KEGG	7	hsa04360:Axon guidance
2nd	0	2nd	KEGG	7	hsa04530:Tight junction
2nd	0	2nd	KEGG	7	hsa04630:Jak-STAT signaling pathway
2nd	0	2nd	KEGG	7	hsa04666:Fc gamma R-mediated phagocytosis
2nd	0	2nd	KEGG	7	hsa04672:Intestinal immune network for IgA production
2nd	0	2nd	KEGG	7	hsa04740:Olfactory transduction
2nd	0	2nd	KEGG	7	hsa04910:Insulin signaling pathway
2nd	0	2nd	KEGG	7	hsa04920:Adipocytokine signaling pathway
2nd	0	2nd	KEGG	7	hsa04930:Type II diabetes mellitus
2nd	0	2nd	KEGG	7	hsa05200:Pathways in cancer
2nd	0	2nd	KEGG	7	hsa05210:Colorectal cancer
2nd	0	2nd	KEGG	7	hsa05211:Renal cell carcinoma
2nd	0	2nd	KEGG	7	hsa05212:Pancreatic cancer
2nd	0	2nd	KEGG	7	hsa05215:Prostate cancer
2nd	0	2nd	KEGG	7	hsa05220:Chronic myeloid leukemia
2nd	0	2nd	KEGG	7	hsa05410:Hypertrophic cardiomyopathy (HCM)
2nd	0	2nd	KEGG	7	hsa05414:Dilated cardiomyopathy
1st	1	2nd	KEGG	1	hsa04740:Olfactory transduction
1st	1	2nd	KEGG	4	hsa04740:Olfactory transduction
1st	1	2nd	KEGG	7	hsa00533:Keratan sulfate biosynthesis
1st	1	2nd	KEGG	7	hsa04740:Olfactory transduction

1st	1	2nd	KEGG	8	hsa04740:Olfactory transduction
2nd	1	2nd	KEGG	1	hsa04740:Olfactory transduction
2nd	1	2nd	KEGG	4	hsa04740:Olfactory transduction
2nd	1	2nd	KEGG	7	hsa00533:Keratan sulfate biosynthesis
1st	0	1st	Panther	1	P00010:B cell activation
1st	0	1st	Panther	1	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	0	1st	Panther	1	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	1st	Panther	1	P00028:Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction
1st	0	1st	Panther	1	P00043:Muscarinic acetylcholine receptor 2 and 4 signaling pathway
1st	0	1st	Panther	1	P00049:Parkinson disease
1st	0	1st	Panther	2	P00012:Cadherin signaling pathway
1st	0	1st	Panther	2	P00034:Integrin signalling pathway
1st	0	1st	Panther	2	P00057:Wnt signaling pathway
1st	0	1st	Panther	3	P00047:PDGF signaling pathway
1st	0	1st	Panther	3	P00049:Parkinson disease
1st	0	1st	Panther	5	P00007:Axon guidance mediated by semaphorins
1st	0	1st	Panther	5	P00029:Huntington disease
1st	0	1st	Panther	5	P00034:Integrin signalling pathway
1st	0	1st	Panther	6	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	0	1st	Panther	6	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	1st	Panther	7	P00002:Alpha adrenergic receptor signaling pathway
1st	0	1st	Panther	7	P00003:Alzheimer disease-amyloid secretase pathway
1st	0	1st	Panther	7	P00004:Alzheimer disease-presenilin pathway
1st	0	1st	Panther	7	P00005:Angiogenesis
1st	0	1st	Panther	7	P00006:Apoptosis signaling pathway
1st	0	1st	Panther	7	P00010:B cell activation
1st	0	1st	Panther	7	P00012:Cadherin signaling pathway
1st	0	1st	Panther	7	P00018:EGF receptor signaling pathway
1st	0	1st	Panther	7	P00019:Endothelin signaling pathway
1st	0	1st	Panther	7	P00021:FGF signaling pathway
1st	0	1st	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	0	1st	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	1st	Panther	7	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	1st	Panther	7	P00041:Metabotropic glutamate receptor group I pathway
1st	0	1st	Panther	7	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway

1st	0	1st	Panther	7	P00045:Notch signaling pathway
1st	0	1st	Panther	7	P00047:PDGF signaling pathway
1st	0	1st	Panther	7	P00053:T cell activation
1st	0	1st	Panther	7	P00056:VEGF signaling pathway
1st	0	1st	Panther	7	P00057:Wnt signaling pathway
1st	0	1st	Panther	7	P04374:5HT2 type receptor mediated signaling pathway
1st	0	1st	Panther	7	P04385:Histamine H1 receptor mediated signaling pathway
1st	0	1st	Panther	7	P04391:Oxytocin receptor mediated signaling pathway
1st	0	1st	Panther	7	P04394:Thyrotropin-releasing hormone receptor signaling pathway
1st	0	1st	Panther	7	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	1st	Panther	7	P05912:Dopamine receptor mediated signaling pathway
1st	0	1st	Panther	8	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	1st	Panther	9	P00002:Alpha adrenergic receptor signaling pathway
1st	0	1st	Panther	9	P00003:Alzheimer disease-amyloid secretase pathway
1st	0	1st	Panther	9	P00005:Angiogenesis
1st	0	1st	Panther	9	P00006:Apoptosis signaling pathway
1st	0	1st	Panther	9	P00010:B cell activation
1st	0	1st	Panther	9	P00012:Cadherin signaling pathway
1st	0	1st	Panther	9	P00018:EGF receptor signaling pathway
1st	0	1st	Panther	9	P00019:Endothelin signaling pathway
1st	0	1st	Panther	9	P00021:FGF signaling pathway
1st	0	1st	Panther	9	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	1st	Panther	9	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	1st	Panther	9	P00037:Ionotropic glutamate receptor pathway
1st	0	1st	Panther	9	P00039:Metabotropic glutamate receptor group III pathway
1st	0	1st	Panther	9	P00041:Metabotropic glutamate receptor group I pathway
1st	0	1st	Panther	9	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway
1st	0	1st	Panther	9	P00047:PDGF signaling pathway
1st	0	1st	Panther	9	P00056:VEGF signaling pathway
1st	0	1st	Panther	9	P00057:Wnt signaling pathway
1st	0	1st	Panther	9	P04374:5HT2 type receptor mediated signaling pathway
1st	0	1st	Panther	9	P04385:Histamine H1 receptor mediated signaling pathway

1st	0	1st	Panther	9	P00053:T cell activation
1st	0	1st	Panther	9	P04394:Thyrotropin-releasing hormone receptor signaling pathway
1st	0	1st	Panther	9	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	1st	Panther	10	P00015:Circadian clock system
1st	0	1st	Panther	10	P00025:Hedgehog signaling pathway
1st	0	1st	Panther	10	P00029:Huntington disease
1st	0	1st	Panther	10	P00049:Parkinson disease
1st	0	1st	Panther	10	P00057:Wnt signaling pathway
2nd	0	1st	Panther	2	P00012:Cadherin signaling pathway
2nd	0	1st	Panther	2	P00057:Wnt signaling pathway
2nd	0	1st	Panther	5	P00029:Huntington disease
2nd	0	1st	Panther	7	P00003:Alzheimer disease-amyloid secretase pathway
2nd	0	1st	Panther	7	P00004:Alzheimer disease-presenilin pathway
2nd	0	1st	Panther	7	P00005:Angiogenesis
2nd	0	1st	Panther	7	P00012:Cadherin signaling pathway
2nd	0	1st	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
2nd	0	1st	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
2nd	0	1st	Panther	7	P00057:Wnt signaling pathway
2nd	0	1st	Panther	7	P05912:Dopamine receptor mediated signaling pathway
2nd	0	1st	Panther	10	P00015:Circadian clock system
2nd	0	1st	Panther	10	P00025:Hedgehog signaling pathway
2nd	0	1st	Panther	10	P00049:Parkinson disease
2nd	0	1st	Panther	10	P00057:Wnt signaling pathway
1st	1	1st	Panther	6	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	1	1st	Panther	6	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	1	1st	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	1	1st	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	1	1st	Panther	7	P00045:Notch signaling pathway
1st	1	1st	Panther	7	P00057:Wnt signaling pathway
1st	1	1st	Panther	7	P05912:Dopamine receptor mediated signaling pathway
1st	1	1st	Panther	9	P00037:Ionotropic glutamate receptor pathway
1st	1	1st	Panther	9	P00039:Metabotropic glutamate receptor group III pathway
1st	1	1st	Panther	9	P00047:PDGF signaling pathway
2nd	1	1st	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway

2nd	1	1st	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
2nd	1	1st	Panther	7	P00057:Wnt signaling pathway
2nd	1	1st	Panther	7	P05912:Dopamine receptor mediated signaling pathway
1st	0	2nd	Panther	3	P00002:Alpha adrenergic receptor signaling pathway
1st	0	2nd	Panther	3	P00003:Alzheimer disease-amyloid secretase pathway
1st	0	2nd	Panther	3	P00005:Angiogenesis
1st	0	2nd	Panther	3	P00006:Apoptosis signaling pathway
1st	0	2nd	Panther	3	P00010:B cell activation
1st	0	2nd	Panther	3	P00018:EGF receptor signaling pathway
1st	0	2nd	Panther	3	P00019:Endothelin signaling pathway
1st	0	2nd	Panther	3	P00021:FGF signaling pathway
1st	0	2nd	Panther	3	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	2nd	Panther	3	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	2nd	Panther	3	P00041:Metabotropic glutamate receptor group I pathway
1st	0	2nd	Panther	3	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway
1st	0	2nd	Panther	3	P00047:PDGF signaling pathway
1st	0	2nd	Panther	3	P00053:T cell activation
1st	0	2nd	Panther	3	P00056:VEGF signaling pathway
1st	0	2nd	Panther	3	P00057:Wnt signaling pathway
1st	0	2nd	Panther	3	P04374:5HT2 type receptor mediated signaling pathway
1st	0	2nd	Panther	3	P04385:Histamine H1 receptor mediated signaling pathway
1st	0	2nd	Panther	3	P04391:Oxytocin receptor mediated signaling pathway
1st	0	2nd	Panther	3	P04394:Thyrotropin-releasing hormone receptor signaling pathway
1st	0	2nd	Panther	3	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	2nd	Panther	4	P00015:Circadian clock system
1st	0	2nd	Panther	4	P00025:Hedgehog signaling pathway
1st	0	2nd	Panther	4	P00049:Parkinson disease
1st	0	2nd	Panther	4	P00057:Wnt signaling pathway
1st	0	2nd	Panther	6	P00002:Alpha adrenergic receptor signaling pathway
1st	0	2nd	Panther	6	P00003:Alzheimer disease-amyloid secretase pathway
1st	0	2nd	Panther	6	P00005:Angiogenesis
1st	0	2nd	Panther	6	P00006:Apoptosis signaling pathway
1st	0	2nd	Panther	6	P00010:B cell activation
1st	0	2nd	Panther	6	P00018:EGF receptor signaling pathway

1st	0	2nd	Panther	6	P00019:Endothelin signaling pathway
1st	0	2nd	Panther	6	P00021:FGF signaling pathway
1st	0	2nd	Panther	6	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	2nd	Panther	6	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	2nd	Panther	6	P00041:Metabotropic glutamate receptor group I pathway
1st	0	2nd	Panther	6	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway
1st	0	2nd	Panther	6	P00047:PDGF signaling pathway
1st	0	2nd	Panther	6	P00053:T cell activation
1st	0	2nd	Panther	6	P00056:VEGF signaling pathway
1st	0	2nd	Panther	6	P00057:Wnt signaling pathway
1st	0	2nd	Panther	6	P04374:5HT2 type receptor mediated signaling pathway
1st	0	2nd	Panther	6	P04385:Histamine H1 receptor mediated signaling pathway
1st	0	2nd	Panther	6	P04391:Oxytocin receptor mediated signaling pathway
1st	0	2nd	Panther	6	P04394:Thyrotropin-releasing hormone receptor signaling pathway
1st	0	2nd	Panther	6	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	2nd	Panther	7	P00002:Alpha adrenergic receptor signaling pathway
1st	0	2nd	Panther	7	P00003:Alzheimer disease-amyloid secretase pathway
1st	0	2nd	Panther	7	P00005:Angiogenesis
1st	0	2nd	Panther	7	P00006:Apoptosis signaling pathway
1st	0	2nd	Panther	7	P00007:Axon guidance mediated by semaphorins
1st	0	2nd	Panther	7	P00010:B cell activation
1st	0	2nd	Panther	7	P00018:EGF receptor signaling pathway
1st	0	2nd	Panther	7	P00019:Endothelin signaling pathway
1st	0	2nd	Panther	7	P00021:FGF signaling pathway
1st	0	2nd	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
1st	0	2nd	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
1st	0	2nd	Panther	7	P00028:Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction
1st	0	2nd	Panther	7	P00029:Huntington disease
1st	0	2nd	Panther	7	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
1st	0	2nd	Panther	7	P00033:Insulin/IGF pathway-protein kinase B signaling cascade
1st	0	2nd	Panther	7	P00035:Interferon-gamma signaling pathway
1st	0	2nd	Panther	7	P00036:Interleukin signaling pathway

1st	0	2nd	Panther	7	P00037:Ionotropic glutamate receptor pathway
1st	0	2nd	Panther	7	P00038:JAK/STAT signaling pathway
1st	0	2nd	Panther	7	P00039:Metabotropic glutamate receptor group III pathway
1st	0	2nd	Panther	7	P00041:Metabotropic glutamate receptor group I pathway
1st	0	2nd	Panther	7	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway
1st	0	2nd	Panther	7	P00044:Nicotinic acetylcholine receptor signaling pathway
1st	0	2nd	Panther	7	P00047:PDGF signaling pathway
1st	0	2nd	Panther	7	P00048:PI3 kinase pathway
1st	0	2nd	Panther	7	P00052:TGF-beta signaling pathway
1st	0	2nd	Panther	7	P00053:T cell activation
1st	0	2nd	Panther	7	P00056:VEGF signaling pathway
1st	0	2nd	Panther	7	P00057:Wnt signaling pathway
1st	0	2nd	Panther	7	P02736:Coenzyme A biosynthesis
1st	0	2nd	Panther	7	P04374:5HT2 type receptor mediated signaling pathway
1st	0	2nd	Panther	7	P04385:Histamine H1 receptor mediated signaling pathway
1st	0	2nd	Panther	7	P04391:Oxytocin receptor mediated signaling pathway
1st	0	2nd	Panther	7	P04394:Thyrotropin-releasing hormone receptor signaling pathway
1st	0	2nd	Panther	7	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	2nd	Panther	7	P05913:Enkephalin release
1st	0	2nd	Panther	7	P05915:Opioid proenkephalin pathway
1st	0	2nd	Panther	7	P05917:Opioid proopiomelanocortin pathway
1st	0	2nd	Panther	8	P00029:Huntington disease
1st	0	2nd	Panther	8	P00053:T cell activation
1st	0	2nd	Panther	9	P00053:T cell activation
1st	0	2nd	Panther	10	P00034:Integrin signalling pathway
1st	0	2nd	Panther	10	P00053:T cell activation
2nd	0	2nd	Panther	4	P00015:Circadian clock system
2nd	0	2nd	Panther	4	P00025:Hedgehog signaling pathway
2nd	0	2nd	Panther	4	P00049:Parkinson disease
2nd	0	2nd	Panther	4	P00057:Wnt signaling pathway
2nd	0	2nd	Panther	7	P00002:Alpha adrenergic receptor signaling pathway
2nd	0	2nd	Panther	7	P00003:Alzheimer disease-amyloid secretase pathway
2nd	0	2nd	Panther	7	P00005:Angiogenesis
2nd	0	2nd	Panther	7	P00006:Apoptosis signaling pathway
2nd	0	2nd	Panther	7	P00007:Axon guidance mediated by semaphorins
2nd	0	2nd	Panther	7	P00010:B cell activation
2nd	0	2nd	Panther	7	P00018:EGF receptor signaling pathway

2nd	0	2nd	Panther	7	P00019:Endothelin signaling pathway
2nd	0	2nd	Panther	7	P00021:FGF signaling pathway
2nd	0	2nd	Panther	7	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway
2nd	0	2nd	Panther	7	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway
2nd	0	2nd	Panther	7	P00028:Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction
2nd	0	2nd	Panther	7	P00029:Huntington disease
2nd	0	2nd	Panther	7	P00031:Inflammation mediated by chemokine and cytokine signaling pathway
2nd	0	2nd	Panther	7	P00033:Insulin/IGF pathway-protein kinase B signaling cascade
2nd	0	2nd	Panther	7	P00035:Interferon-gamma signaling pathway
2nd	0	2nd	Panther	7	P00036:Interleukin signaling pathway
2nd	0	2nd	Panther	7	P00037:Ionotropic glutamate receptor pathway
2nd	0	2nd	Panther	7	P00038:JAK/STAT signaling pathway
2nd	0	2nd	Panther	7	P00039:Metabotropic glutamate receptor group III pathway
2nd	0	2nd	Panther	7	P00041:Metabotropic glutamate receptor group I pathway
2nd	0	2nd	Panther	7	P00042:Muscarinic acetylcholine receptor 1 and 3 signaling pathway
2nd	0	2nd	Panther	7	P00047:PDGF signaling pathway
2nd	0	2nd	Panther	7	P00048:PI3 kinase pathway
2nd	0	2nd	Panther	7	P00052:TGF-beta signaling pathway
2nd	0	2nd	Panther	7	P00053:T cell activation
2nd	0	2nd	Panther	7	P00056:VEGF signaling pathway
2nd	0	2nd	Panther	7	P00057:Wnt signaling pathway
2nd	0	2nd	Panther	7	P02736:Coenzyme A biosynthesis
2nd	0	2nd	Panther	7	P04374:5HT2 type receptor mediated signaling pathway
2nd	0	2nd	Panther	7	P04385:Histamine H1 receptor mediated signaling pathway
2nd	0	2nd	Panther	7	P04391:Oxytocin receptor mediated signaling pathway
2nd	0	2nd	Panther	7	P04394:Thyrotropin-releasing hormone receptor signaling pathway
2nd	0	2nd	Panther	7	P05911:Angiotensin II-stimulated signaling through G proteins and beta-arrestin
1st	0	1st	Reactome	2	REACT_6900:Signaling in Immune system,
1st	0	1st	Reactome	3	REACT_11045:Signaling by Wnt
1st	0	1st	Reactome	3	REACT_13:Metabolism of amino acids
1st	0	1st	Reactome	3	REACT_13635:Regulation of activated PAK-2p34 by proteasome mediated degradation
1st	0	1st	Reactome	3	REACT_152:Cell Cycle, Mitotic
1st	0	1st	Reactome	3	REACT_1538:Cell Cycle Checkpoints

1st	0	1st	Reactome	3	REACT_383:DNA Replication
1st	0	1st	Reactome	3	REACT_578:Apoptosis
1st	0	1st	Reactome	3	REACT_6185:HIV Infection
1st	0	1st	Reactome	3	REACT_6850:Cdc20:Phospho-APC/C mediated degradation of Cyclin A
1st	0	1st	Reactome	3	REACT_6900:Signaling in Immune system
1st	0	1st	Reactome	3	REACT_9035:APC/C:Cdh1-mediated degradation of Skp2
1st	0	1st	Reactome	3	REACT_9480:Gap junction trafficking and regulation
1st	0	1st	Reactome	5	REACT_71:Gene Expression
1st	0	1st	Reactome	6	REACT_14797:Signaling by GPCR
1st	0	1st	Reactome	6	REACT_6900:Signaling in Immune system
1st	0	1st	Reactome	7	REACT_14797:Signaling by GPCR
1st	0	1st	Reactome	7	REACT_6900:Signaling in Immune system
1st	0	1st	Reactome	8	REACT_15518:Transmembrane transport of small molecules
1st	0	1st	Reactome	9	REACT_13685:Synaptic Transmission
1st	0	1st	Reactome	9	REACT_15380:Diabetes pathways
1st	0	1st	Reactome	9	REACT_604:Hemostasis
1st	0	1st	Reactome	9	REACT_71:Gene Expression
1st	0	1st	Reactome	10	REACT_14797:Signaling by GPCR
2nd	0	1st	Reactome	2	REACT_6900:Signaling in Immune system
2nd	0	1st	Reactome	3	REACT_6900:Signaling in Immune system
2nd	0	1st	Reactome	6	REACT_6900:Signaling in Immune system
2nd	0	1st	Reactome	7	REACT_14797:Signaling by GPCR
2nd	0	1st	Reactome	10	REACT_14797:Signaling by GPCR
1st	1	1st	Reactome	5	REACT_71:Gene Expression,
1st	1	1st	Reactome	6	REACT_14797:Signaling by GPCR
1st	1	1st	Reactome	7	REACT_14797:Signaling by GPCR
1st	1	1st	Reactome	7	REACT_6900:Signaling in Immune system
1st	1	1st	Reactome	8	REACT_15518:Transmembrane transport of small molecules
1st	1	1st	Reactome	9	REACT_13685:Synaptic Transmission
1st	1	1st	Reactome	9	REACT_71:Gene Expression
1st	1	1st	Reactome	10	REACT_14797:Signaling by GPCR
2nd	1	1st	Reactome	7	REACT_14797:Signaling by GPCR
2nd	1	1st	Reactome	10	REACT_14797:Signaling by GPCR
1st	0	2nd	Reactome	1	REACT_14797:Signaling by GPCR
1st	0	2nd	Reactome	2	REACT_17015:Metabolism of proteins
1st	0	2nd	Reactome	3	REACT_13685:Synaptic Transmission
1st	0	2nd	Reactome	3	REACT_15380:Diabetes pathways
1st	0	2nd	Reactome	3	REACT_604:Hemostasis
1st	0	2nd	Reactome	3	REACT_6900:Signaling in Immune system
1st	0	2nd	Reactome	4	REACT_14797:Signaling by GPCR
1st	0	2nd	Reactome	6	REACT_13685:Synaptic Transmission

1st	0	2nd	Reactome	6	REACT_15380:Diabetes pathways
1st	0	2nd	Reactome	6	REACT_604:Hemostasis
1st	0	2nd	Reactome	6	REACT_6900:Signaling in Immune system
1st	0	2nd	Reactome	7	REACT_11061:Signalling by NGF
1st	0	2nd	Reactome	7	REACT_11193:Metabolism of vitamins and cofactors
1st	0	2nd	Reactome	7	REACT_13:Metabolism of amino acids
1st	0	2nd	Reactome	7	REACT_13698:Regulation of beta-cell development
1st	0	2nd	Reactome	7	REACT_14797:Signaling by GPCR
1st	0	2nd	Reactome	7	REACT_15295:Opioid Signalling
1st	0	2nd	Reactome	7	REACT_15380:Diabetes pathways
1st	0	2nd	Reactome	7	REACT_15518:Transmembrane transport of small molecules
1st	0	2nd	Reactome	7	REACT_17015:Metabolism of proteins
1st	0	2nd	Reactome	7	REACT_1762:3' -UTR-mediated translational regulation
1st	0	2nd	Reactome	7	REACT_18266:Axon guidance
1st	0	2nd	Reactome	7	REACT_216:DNA Repair
1st	0	2nd	Reactome	7	REACT_604:Hemostasis
1st	0	2nd	Reactome	7	REACT_6167:Influenza Infection
1st	0	2nd	Reactome	7	REACT_6844:Signaling by TGF beta
1st	0	2nd	Reactome	7	REACT_71:Gene Expression
1st	0	2nd	Reactome	8	REACT_6900:Signaling in Immune system
1st	0	2nd	Reactome	9	REACT_6900:Signaling in Immune system
1st	0	2nd	Reactome	10	REACT_6900:Signaling in Immune system
2nd	0	2nd	Reactome	1	REACT_14797:Signaling by GPCR
2nd	0	2nd	Reactome	4	REACT_14797:Signaling by GPCR
2nd	0	2nd	Reactome	6	REACT_6900:Signaling in Immune system
2nd	0	2nd	Reactome	7	REACT_11061:Signalling by NGF
2nd	0	2nd	Reactome	7	REACT_11193:Metabolism of vitamins and cofactors
2nd	0	2nd	Reactome	7	REACT_13:Metabolism of amino acids
2nd	0	2nd	Reactome	7	REACT_13698:Regulation of beta-cell development
2nd	0	2nd	Reactome	7	REACT_14797:Signaling by GPCR
2nd	0	2nd	Reactome	7	REACT_15380:Diabetes pathways
2nd	0	2nd	Reactome	7	REACT_15518:Transmembrane transport of small molecules
2nd	0	2nd	Reactome	7	REACT_17015:Metabolism of proteins
2nd	0	2nd	Reactome	7	REACT_1762:3' -UTR-mediated translational regulation
2nd	0	2nd	Reactome	7	REACT_216:DNA Repair
2nd	0	2nd	Reactome	7	REACT_604:Hemostasis
2nd	0	2nd	Reactome	7	REACT_6167:Influenza Infection
2nd	0	2nd	Reactome	7	REACT_6844:Signaling by TGF beta
2nd	0	2nd	Reactome	7	REACT_71:Gene Expression
1st	1	2nd	Reactome	1	REACT_14797:Signaling by GPCR

1st	1	2nd	Reactome	4	REACT_14797:Signaling by GPCR
1st	1	2nd	Reactome	7	REACT_13:Metabolism of amino acids
1st	1	2nd	Reactome	7	REACT_14797:Signaling by GPCR
1st	1	2nd	Reactome	7	REACT_15518:Transmembrane transport of small molecules
2nd	1	2nd	Reactome	1	REACT_14797:Signaling by GPCR
2nd	1	2nd	Reactome	4	REACT_14797:Signaling by GPCR
2nd	1	2nd	Reactome	7	REACT_13:Metabolism of amino acids

APPENDIX C

Family Results for the Analysis with Normalization

List of most significant pathways for each family are presented in this section. Pathways are ranked according to p-values and most significant 10 pathways are included for each family. If there are less than 10 pathways, all pathways are presented for that family. If the analysis did not yield any results for a family, a table is not provided.

All results in this section are obtained by third approach analysis with normalization using second filter.

KEGG Pathway	P-values
Allograft rejection	1,98E+08

Table C.1 General analysis results for family trio 3

KEGG Pathway	P-values
Circadian rhythm	7,20E+03
Amphetamine addiction	3,05E+05
Proteasome	5,86E+07

Table C.2 General analysis results for family trio 4

KEGG Pathway	P-values
Asthma	5,60E+01
Allograft rejection	2,01E+04
Graft-versus-host disease	2,78E+04
Type I diabetes mellitus	3,76E+02
Intestinal immune network for IgA production	9,80E+03
Autoimmune thyroid disease	1,59E+04
Staphylococcus aureus infection	2,23E+04

Table C.3 General analysis results for family trio 6

KEGG Pathway	P-values
Complement and coagulation cascades	4,80E-14
TGF-beta signaling pathway	7,41E-06
GABAergic synapse	2,52E+04
Proteasome	5,92E+04
RNA degradation	8,40E+04
Notch signaling pathway	3,90E+06
Type II diabetes mellitus	2,69E+06
Aldosterone-regulated sodium reabsorption	7,50E+05
Pentose phosphate pathway	1,37E+09
Homologous recombination	3,22E+07

Table C.4 General analysis results for family trio 7

KEGG Pathway	P-values
Complement and coagulation cascades	4,80E-14
TGF-beta signaling pathway	7,41E-06
GABAergic synapse	2,52E+04
Proteasome	5,92E+04
RNA degradation	8,40E+04
Notch signaling pathway	3,90E+06
Circadian rhythm	1,77E+07
Type II diabetes mellitus	2,69E+06
Aldosterone-regulated sodium reabsorption	7,50E+05
Pentose phosphate pathway	1,37E+09

Table C.5 General analysis results for family trio 8

KEGG Pathway	P-values
Folate biosynthesis	2,83E+11

Table C.6 General analysis results for family trio 9

KEGG Pathway	P-values
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1,15E+10
Renin-angiotensin system	4,23E+11

Table C.7 Promoter analysis results for family trio 6

KEGG Pathway	P-values
DNA replication	6,19E-11
Complement and coagulation cascades	9,95E-05
Cholinergic synapse	4,31E-03
Glutamatergic synapse	3,49E-02
GABAergic synapse	1,45E+00
Retrograde endocannabinoid signaling	7,79E-01
Dopaminergic synapse	1,88E+01
Notch signaling pathway	1,22E+03
Proteasome	7,16E+03
ECM-receptor interaction	7,44E+03

Table C.8Promoter analysis results for family trio 7

KEGG Pathway	P-values
Folate biosynthesis	1,42E+11

Table C.9 Promoter analysis results for family trio 9