Identification of Interconnected Markers for T-cell Acute Lymphoblastic Leukemia

by

Emine Guven Maiorov

A Thesis Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science
in
Chemical and Biological Engineering

Koc University

August 2011

Koc University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Emine Guven Maiorov

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____

Ozlem Keskin, Ph.D. (Advisor)

_____

Attila Gursoy, Ph.D.

_____

Meltem Müftüoğlu, Ph.D.

Date: _____

ABSTRACT

T-cell acute lymphoblastic leukemia (T-ALL) is a very complex disease, resulting from proliferation of differentially arrested immature T-cells. The molecular mechanisms and the genes involved in the cause of T-ALL remain largely undefined. In this study, we found biomarkers to differentiate individuals with T-ALL from the non-leukemia/healthy ones, to discover markers that are not differential themselves but interconnect with highly differentially expressed genes, and to have a network-based view of T-ALL. Instead of applying only expression-based differential gene analysis and obtaining hundreds of candidate disease-causing genes, we integrated gene expression data of T-ALL and healthy samples with the human protein-protein interaction data in order to discover diagnostic biomarkers not as individual genes but as subnetworks. By using a network-based approach, we have identified 19 significant subnetworks, containing 102 genes (out of 409 genes). A given subnetwork contains up-to 12 genes. The classification/prediction accuracies of subnetworks are considerably high, as high as 98%. Some genes in the subnetworks were already known to be associated with T-ALL, but we found new ones to be involved in T-ALL development. The subnetworks were rich in transcription factors whose ectopic activation is known to be one of the reasons behind T-ALL. The Zinc-binding proteins are also abundant in subnetworks. Zinc levels are low in ALL-patients. Zinc supplement given to a T-ALL patient may increase the efficiency of chemotherapy. We recovered 6 tyrosine kinases which have important roles in T-cell survival, proliferation, and immune response. These important genes in our subnetworks may serve as an alternative to the traditional biomarkers used for the diagnosis of T-ALL. The aim of this study is also to help investigators to highlight potential disease gene candidates for further experimental validation.

We also applied a typical hierarchical clustering method to most differential 100 and 200 genes between T-ALL and healthy samples. As opposed to the presumption that most differential 100 or 200 genes would classify the diseased samples better, our subnetworks achieved the same or, in some cases, higher classification accuracies. Doing the same/better job with 10 genes in a subnetwork, instead of 100 or 200 genes might be regarded as an accomplishment. In short, network-based classification techniques help us to identify biologically more meaningful subnetworks than expression-based techniques which return thousands of differential genes.

ÖZET

Olgunlaşmamış T-hücrelerinin normalden daha fazla çoğalmasından kaynaklanan, T-hücresi Akut Lemfoblastik Lösemi (T-ALL) çok kompleks bir hastalıktır. Genel olarak T-ALL'ye sebep olan moleküler mekanizmalar ve genler bilinmiyor. Bu araştırma sonucunda, T-ALL hastaları ile normal/sağlıklı insanları birbirinden ayırt edebilen biyomarker genleri ve bu iki insan grubunda ekpresyonları farklılık göstermedigi halde, farklılık gösteren genleri birbirine bağlayan genleri bulduk. Ek olarak, T-ALL'ye sebep olabilecek yolakları belirledik. Bu sonuçları elde edebilmek ve biyomarker olarak tek başına genler değil, birbirine bağlı alt-ağlar bulmak için, T-ALL hastaları ve sağlıklı insanların gen ekspresyon verileri ile protein-protein etkileşim ağlarını entegre ettik. Ağa dayalı yaklaşımı kullanarak, 109 gen içeren 19 önemli alt-ağ bulduk (409 gen içerisinden). Elde edilen alt-ağların hastaları doğru sınıflandırma başarısı %98' e varıyor. Örneğin %98 sınıflandırma başarısı olan bir alt-ağ, hasta ile sağlıklı insan gruplarını %98 ölçüde birbirinden doğru ayırabiliyor demektir. Alt-ağlardaki bazı genlerin, önceden T-ALL ile ilişkileri bulunmuş, ama diğerlerinin literatürde bu hastalıkla ortaklıklarına rastlanmamıştır. Transkripsiyon faktörlerinin ektopik aktivasyonu, T-ALL'ye yol açan sebeplerden biridir. Bulunan alt-ağlar, transkripsiyon faktörleri açısından zengindir. Çinko elementine bağlanan proteinler de alt-ağlarda sıkça rastlanmıştır. Düşük çinko seviyesi Akut Lemfoblastik Lösemi hastalarının özelliklerinden biridir ve T-ALL hastalarına verilen çinko takviyesinin kemoterapinin etkinliğini artırdığı literatürde yer almaktadır. Alt-ağlarda bulduğumuz genlerin 6 tanesi Tirozin Kinaz enzimidir. T-hücerelerinin yaşamasında, üremesinde, bağışıklık sistemindeki görevlerini idame ettirmelerinde, bu enzimler çok önemli rol oynarlar. Alt-ağlarda bulunan bu önemli genler, şu an halihazırda T-ALL teşhisi için kullanılan genlere alternatif biyomarker olarak kullanılabilir. Bu çalışmanın bir diğer amacı

da bilim adamlarına, deneysel doğrulama için potansiyel hastalık genlerinin belirlenmesinde yardımcı olmaktır.

Bu araştırmada ayrıca, hasta ile sağlıklı bireyler arasında ekspresyonları en çok farklılık gösteren 100 ve 200 gene hiyerarşik kümeleme tekniği uyguladık. Ekspresyonları en farklı 100/200 gen hastaları ve sağlıklıları daha iyi ayırt edebilir varsayımının aksine, bizim alt-ağlarımız örnekleri sınıflandırmada aynı bazı durumlarda ise daha başarılı oldular. Aynı işi 100/200 gen yerine, alt-ağlarda bulunan10 gen ile yapmak bir başarı olarak değerlendirilebilir. Kısacası, ağa-bağlı sınıflandırma teknikleri, ekspresyonları farklılık gösteren binlerce geni bulan sadece-ekspresyona bağlı tekniklerden, biyolojik olarak daha anlamlı genler elde etmemizi sağlıyor.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude is to my advisors, Professor Ozlem Keskin and Professor Attila Gursoy for their sincerity, guidance, and support that I will never forget. This thesis would not have been possible without their invaluable recommendations and suggestions.

I am grateful to my officemates (Halil Peynirci, Ömer An, Serap Beldar, Büşra Topal, Pelin Atıcı, and Tuğçe Yıldızoğlu) for providing a very friendly and cheerful atmosphere in our office all the time and for the regular gossip sessions☺. I would like to express special thanks to Halil Peynirci who has always been helpful to me in dealing with algorithms.

My thanks and appreciations also go to my family members, especially my parents, for all their love and unlimited support.

Last but not least, words cannot express the gratitude I owe to my beloved husband, Alexander Maiorov, for his endless support and encouragement. He was the one who motivated me during the very difficult days that I faced and prepared a “Wonderland” for me.

Thank you…

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

T-ALL	T-cell Acute Lymphoblastic Leukemia
HSCs	Hemetopoietic Stem Cells HSCs
WBCs	White Blood Cells
NK	Natural Killer
APCs	Antigen Presenting Cells
RBCs	Red Blood Cells
CLP	Common Lymphoid Progenitor
TCR	T-cell receptor
BCR	B-cell receptor
V(D)J	Variable, Diversity and Joining
MHC	Major Histocompatibility Complex
T_H1	T-helper1
T_H2	T-helper2
T_H17	T-helper17
PPI	Protein-Protein Interaction
DE	Differentially Expressed
MILE	Microarray Innovations in Leukemia
GOTM	Gene Ontology Tree Machine
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
LSCs	Leukemic Stem Cells

Chapter 1

1. INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) is a very complex disease, resulting from proliferation of differentially arrested immature T-cells. Differentiation arrest may take place at almost all stages of thymocyte development [1]. The molecular mechanisms and the genes involved in the cause of T-ALL remain largely undefined. The analysis of genome-wide expression profiles is frequently used to discover new biomarkers [2]. In this study, we integrated gene expression data of T-ALL and healthy samples with the human protein-protein interaction data in order to discover diagnostic biomarkers not as individual genes but as subnetworks. Then, we calculated the classification accuracies and made the functional enrichment analysis of the resulting subnetworks. We found biomarkers to differentiate individuals with T-ALL from the non-leukemia/healthy ones, to discover markers that are not differential themselves but interconnect with highly differentially expressed genes, and to have a network-based view of T-ALL. We also applied a traditional hierarchical clustering method to 100 and 200most differential genes between T-ALL and healthy samples.

In Chapter 2, an extended literature review regarding Hematopoietic stem cell differentiation, overview of immunology and T-cell role in immune responses, T-ALL, importance of biomarkers, role of microarrays in biomarker discovery, and expression-based vs. network-based classification techniques is presented.

Chapter 3 includes the methods used in this study. Pinnacle-Z algorithm as a plug-in to Cytoscape is used to generate significant subnetworks by integrating gene expression (microarray) data with the protein protein interaction data. A traditional hierarchical clustering method implemented in Expander was also applied to 100 and 200 most differentially expressed genes. J48 and RBF-Network classifiers in WEKA are used to calculate the classification accuracies of the subnetworks that are found by Pinnacle-Z algorithm. Lastly, functional enrichments of the subnetworks are done by Gene Ontology Tree Machine (GOTM).

In Chapter 4, the results of Pinnacle-Z algorithm, hierarchical clustering, J48 and RBF-Network classifiers and functional enrichment analysis are given. We identified 19 significant subnetworks, containing 102 genes. The classification/prediction accuracies of subnetworks are significantly high, as high as 98%. Some genes in the subnetworks were already known to be associated with T-ALL, but we found new ones to be involved in T-ALL development.

In the last section, this dissertation is completed with a chapter providing the conclusion of this study, followed by a supplementary materials chapter and the bibliography.

Chapter 2

2. LITERATURE REVIEW

2.1 Hematopoietic stem cell differentiation

Hematopoiesis is the differentiation of pluripotent Hematopoietic Stem Cells (HSCs) in bone marrow, giving rise to all of the constituents of blood, including elements of immune system. The pluripotent HSCs first differentiate into three different progenitors, namely, “Common Lymphoid”, “ Common Myeloid”, and “Megakaryocyte/Erythrocyte” progenitors [3]. The progenitors further differentiate and produce all types of blood cells. As HSCs propagate downstream of the differentiation steps, they lose their renewing capability [4, 5]. Figure 2-1 shows the details of HSC differentiation. The self-renewing pluripotent stem cell resides in bone marrow and gives rise to “Common Lymphoid”, “Common Myeloid” progenitors and “Erythroblast”. These 3 lineages are also located in the bone marrow. The common lymphoid progenitor generates naive B-cells, T-cells and Natural Killer (NK) cells which migrate from bone marrow to blood first, then, to lymph nodes. Likewise, common myeloid progenitor produces Neutrophil, Eosinophil, Basophil, Monocyte, Mast cell, and Platelets. The progenitor is found in bone marrow, and their products/progenies are found in the blood. Only Mast cells and Macrophages migrate to the tissues. Lastly, erythroblasts in bone marrow yield Erythrocytes (RBCs) in the blood [6].

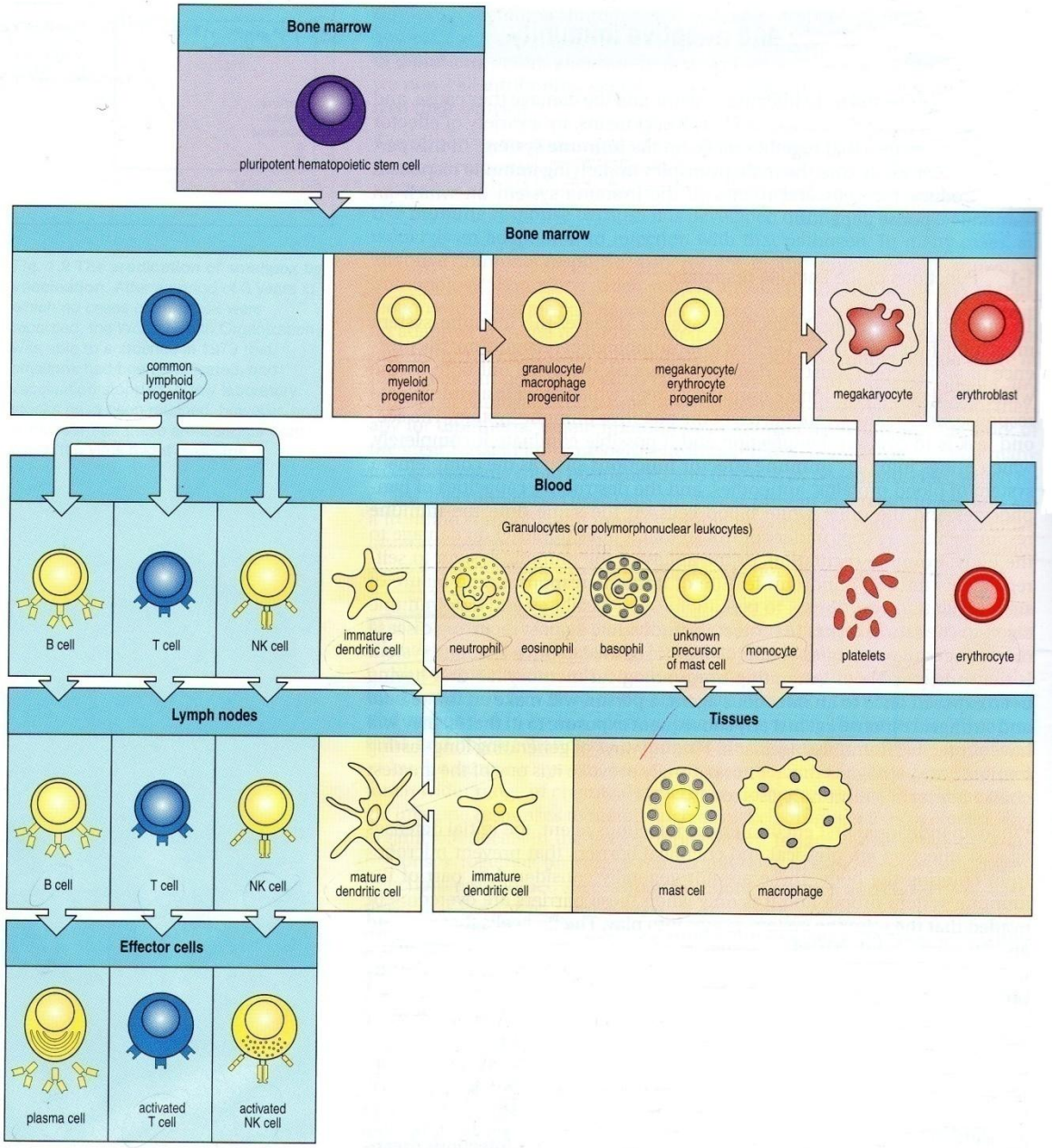

Figure 2-1: Hematopoietic stem cell differentiation, giving rise to different blood cell types [6].

The lymphoid progenitor yields lymphoid lineage which in turn produces leukocytes (White Blood Cells (WBCs)). The stem cells of lymphoid lineage are called "lymphoblast" (blast means stem/precursor cell). The "Natural Killer" (NK) cells, T and B lymphocytes compose the leukocytes, the main elements of adaptive immune system. T-lymphocytes mature in the Thymus, whereas B-lymphocytes mature in bone marrow [7]. T and B cells have antigen specificity, meaning that they can recognize parts of pathogens through their antigen receptors. Each lymphocyte bears a distinctive antigen receptor. Therefore, it can recognize only its partner antigen (small particles of the pathogen which is phagocytosed and processed by Antigen Presenting Cells (APCs)). But NK cells do not have antigen specificity [6].

The myeloid progenitor, or myeloblast produces myeloid lineage which comprises Erythrocytes (RBCs), Platelets, phagocytic cells such as Dendritic cells and Macrophages, and other cells of the blood [8]. Unlike leukocytes, some sorts of myeloid cells take part in innate immunity. For example, the macrophages and dendritic cells are capable of phagocytosis: they internalize and try to destroy the pathogens, then act as an APC to stimulate T or B lymphocytes [6].

2.2 T-cell Role in Immunology and Overview of Immune System

T-cells, so called T-lymphocytes, are major players of the adaptive immune system. They are derived from HSCs of bone marrow but they reach their maturation (Naive T-lymphocytes) in Thymus [4, 5]. There are different classes of T-cells with specific immunological functions, namely; "Cytotoxic T-cells", "Helper T-cells" and "Regulatory T-cells" [6]. Before going into details of T-cell mediated immunity, the key concepts of immune system will be introduced briefly.

Body has a defense mechanism against infection and pathogens. This protective mechanism is called “Immune System”. The body’s responses against the infection are divided into two: “Innate Immunity” (natural immunity), which is always available to fight with pathogens but not specific and long-lasting; and “Adaptive Immunity” (acquired immunity) which is built up upon infection with a particular pathogen.

Adaptive immune responses are highly specific; targeting only a type of pathogen. If innate immunity cannot wipe pathogens out from the body, adaptive immunity comes into play. It takes time to develop adaptive immune response, but once developed, it grants a “protective immunity” (also called as “immunological memory”) for a life span. Thanks to this memory, if an individual is re-infected with the same pathogen, the proper immune response is made immediately (much faster than developing an adaptive immune response after being exposed to a pathogen for the first time) [6]. Figure 2-2 shows the differences between innate and adaptive immunity. First, innate immunity combats with the pathogen. If it fails to do so, the adaptive immune system comes on the scene. When a pathogen enters the body, it is internalized by phagocytic cells such as Macrophages or Dentritic cells. Macrophages digest and eliminate the pathogen. The Toll-like Receptor (TLR) on the dentritic cell recognizes the components (such as lipid, protein or nucleic acid) of the pathogen: internalize and process it. Then, they produce a wide range of bioactive substances, like cytokines which induce inflammatory reaction or activate adaptive immune system and interferons which is an anti-viral cytokine. The dentritic cells may also serve as APCs by presenting the small particles of pathogen (so called, antigen) to T-cells through their MHC complex. Once the T-cell receptor matches with its antigen, T-cell differentiates into either T_H1 or T_H2 (helper T-cells). T_H1 excites cytotoxic T-cells which kills the infected cells (APCs). T_H2 stimulates B-cells and induces antibody production, which in turn tries to eliminate the pathogen [6, 9].

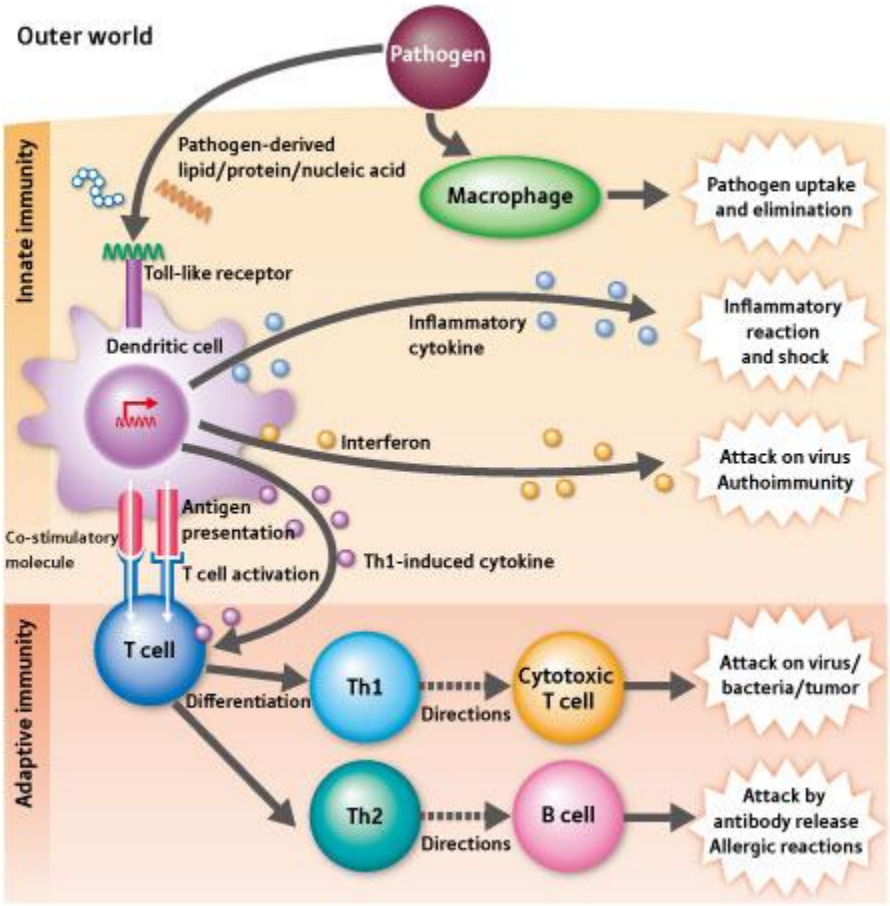

Figure 2-2: The main actors, responses and differences between adaptive and innate immunity [9].

The WBCs or Lymphocytes are the main players in adaptive immune system. These specialized lymphocytes recognize and target the infection or the origins of infection. T-cells, B-cells and NK cells are downstream progenies of common lymphoid progenitor (CLP). The lymphocytes that have not yet met their partner antigen and thus not activated are called “Naive lymphocytes”. However, the lymphocytes that met their partner and became activated are called “Effector lymphocytes”.

As noted in the previous part, T and B cells have antigen specificity. The receptor on the surface of B-lymphocytes and T-lymphocytes are called as “B-cell receptor (BCR)” and “T-cell receptor (TCR)”, respectively (see Figure 2-3). Each of them has receptors which specifically recognize only one type of antigen. BCR has 2 “light-chains” and 2 “heavy-chains”, while TCR has 1 “α-chain” and 1 “β-chain”. Both BCR and TCR have “constant” and “variable” regions. The variable regions are the sites where the antigen is specifically recognized and bound to. In addition, the variable region of a TCR (or BCR) on surface of a T-lymphocyte is not identical with variable regions of TCRs on other T-lymphocytes. In other words, each T-cell has a distinct TCR with the capacity to recognize different antigen.

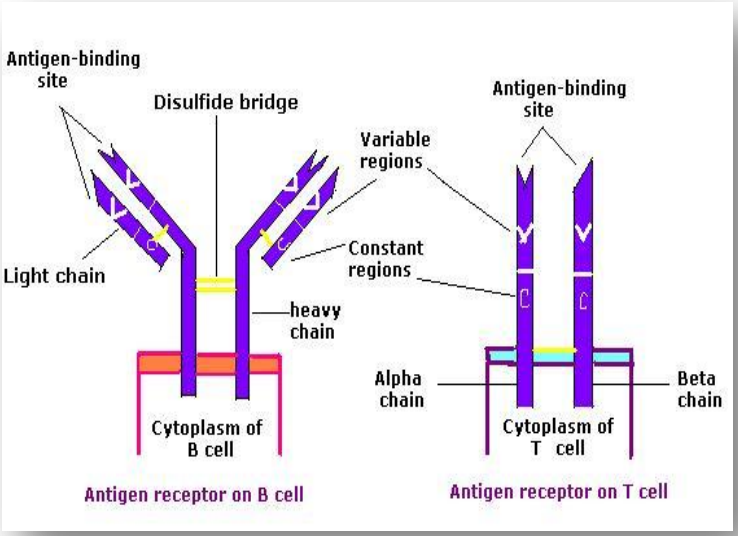

Figure 2-3: The differences between B-cell receptor and T-cell receptor [10]. BCR has 2 heavy and 2 light chains, but TCR has 1 α and 1 β chain. Both BCR and TCR have constant and variable regions.

Since there are various kinds of pathogens and their processed antigens, there should be a variety of lymphocyte receptors (BCR or TCR) to recognize them [6]. If all of the lymphocytes are originated from the same progenitor HSC and have the exact same genome, how do they achieve to produce different kinds of receptors or "variable regions" on their own receptors? Each developing lymphocyte expresses a unique receptor that is different than other cells by a mechanism, called "somatic recombination" in which a somatic cell re-arranges its own genome irreversibly [11]. This feature of lymphocytes is very extraordinary, because recombination event normally does not take place in somatic cells, but occur only during Meiosis-I in gonads ("chromosomal recombination" or "crossing-over" between homologous chromosomes). You can see the difference between somatic recombination and chromosomal recombination in Figure 2-4. Somatic recombination is an intra-chromosomal (in the same chromosome) event, generating diverse receptors on lymphocytes [12], while chromosomal recombination is inter-chromosomal, taking place between two different chromosomes.

There are several copies of "variable", "diversity", and "joining" segments on genomes of the lymphocytes, but only one copy is used to create a functional lymphocyte receptor (BCR or TCR). By re-arranging these copies of gene segments, different combinations can be gained, and thus different kinds of receptors can be produced. In this re-arrangement process, the excess copies of V, D, and J segments are excised out; only one copy for each segment remains intact. The somatic recombination ensures T or B cells to generate almost un-limited number of receptor types (since each lymphocyte can bear only one type of receptor, producing different receptors means producing different types of lymphocytes which can recognize different antigens). This somatic recombination event is irreversible and called as "V(D)J recombination" [13] which is shown on Figure 2-5.

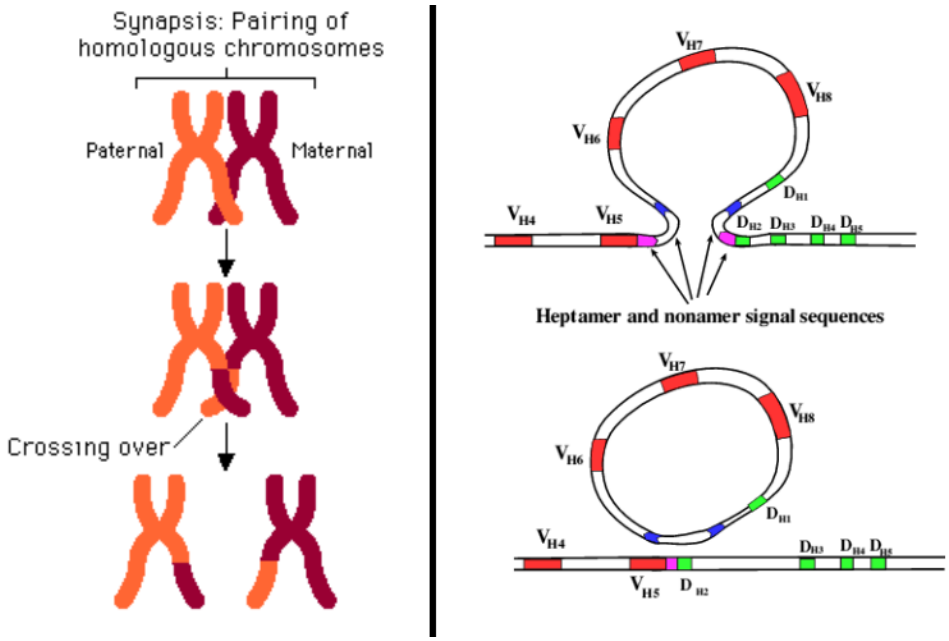

Figure 2-4: The difference between somatic and chromosomal recombination. On the right, somatic recombination [14] in a particular chromosome and on the left, chromosomal recombination between 2 homologous chromosomes during Meiosis-I are shown.

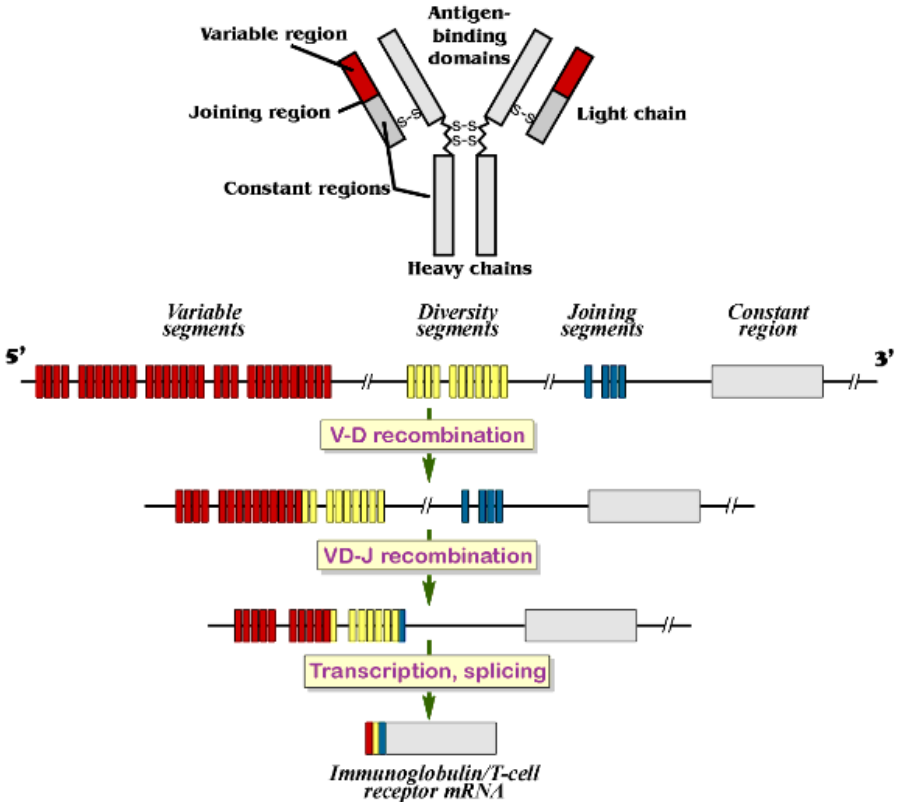

Figure 2-5: The somatic (V(D)J) recombination taking place during differentiation of both T and B lymphocytes [14]. There are many copies of each of the "Variable", "Diversity", and "Joining" segments. Different combinations can be created by taking different copies of these segments.

2.3 T-cell Types

HSCs reside in bone marrow and differentiate into different types of blood cells. T-precursor cells migrate to the Thymus and complete their maturation there. Once T-cells matured in Thymus, they enter bloodstream and re-circulate between blood and peripheral lymphoid organs/tissues. These mature T-lymphocytes that have not encountered with their antigen are called "naive T-cells". When they come across with their antigen and the proper "Major Histocompatibility Complex (MHC)" molecule, naive T-cells proliferate and further differentiate into 5 kinds of "Effector T-cells". These 5 kinds of effector T-cells and their key functions can be seen in Figure 2-6. Once the T-cell met its antigen on the APC, it starts to proliferate allowing uncommon antigen-specific T-cells to raise their number with the intention that they can successfully fight with the pathogen [6]. The first meeting of the T-cell with its antigen is called "Priming". Moreover, the proliferation of T-cells upon the priming event is called "Clonal Expansion".

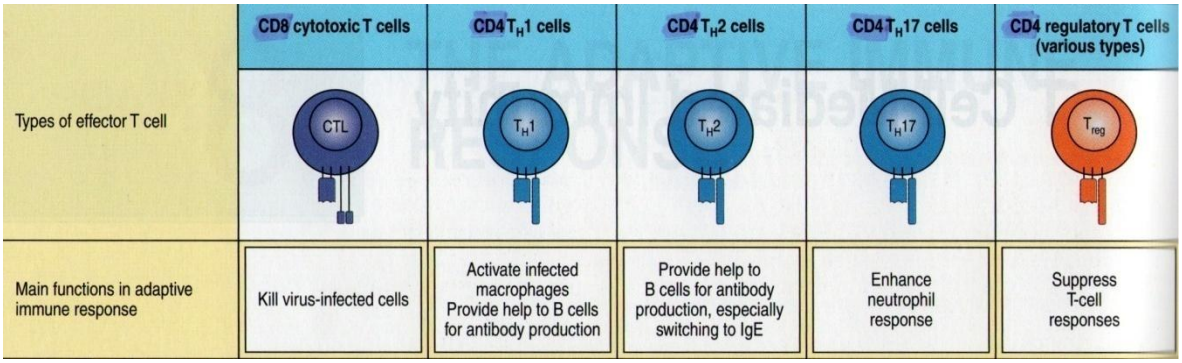

Figure 2-6: The 5 distinct types of "Effector T-lymphocytes" in adaptive immune system [6]. CD8 T-cells can only generate Cytotoxic T-cells, but CD4 T-cells can generate 4 different kinds of effector T-cells.

As Figure 2-6 illustrates, naive CD8 T-cells can give rise to only “Cytotoxic T-cells”. However, naive CD4 T-cells can lead to 4 kinds of effector T-cells. The function of these 5 effectors formed by either CD4 or CD8 naive T-cells are listed below:

1. Upon binding of CD8 naive T-cells to their antigen on APC, Cytotoxic T-cells are generated. Cytotoxic (Killer) T-cells kills the pathogen-infected APCs not the pathogen itself and try to eliminate all the APCs which internalized and processed the pathogen.
2. Unlike CD8 T-cells, upon binding of CD4 naive T-cells to their antigen on APC, 4 different types of effector T-cells are produced. The first type is T-helper1 (T_H1) which is responsible for tuning infected macrophages (a type of APC) on and stimulating B cells to secrete antibody.
3. The second type of CD4-derived effector T-cell is T-helper2 (T_H2) which is in charge of helping B cells to produce antibody.
4. The third effector type is T-helper17 (T_H17) which recruits Neutrophils to the site of infection and improves their response.
5. The last effector type is regulatory T-cells with the function of suppressing T-cell response.

Macrophages, Dendritic cells, and some B cells are antigen presenting cells (APCs). T-cells can recognize only infected APCs (host’s own cells) but B cell can directly recognize the pathogen. APCs present the antigens of pathogen with the help of a glycoprotein which is named as MHC. T-cells require MHC molecules to bind to the antigen. In addition, T-cells posses some cytosolic proteins which are needed to bind to the antigens on APCs. The cytosolic proteins are CD4 and CD8. A mature T-cell can bear only one of these 2 proteins. Actually, T-cells are divided into 2 main classes according to the

protein that they express. In other words, there are 2 groups of T cells, namely: CD4 T-cells and CD8 T-cells. Figure 2-7 shows the MHC & antigen on the APC and TCR & one of the glycoproteins (CD4 or CD8) on T-cell surface. This whole structure at the interface of APC and T-cell is called as "immunological synapse" [15]. "MHC class-I" molecules can associate with only CD8 T-cells and "MHC-class-II" molecules can connect to CD4 T-cells.

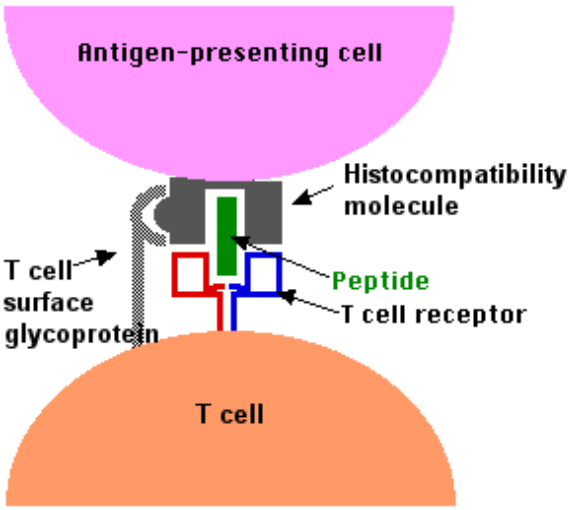

Figure 2-7: The immunological synapse or the overall view of T-cell and its interacting partner, APC [14]. On T-cell surface, there are TCR and one of the glycoproteins (CD4 or CD8) and on the APC, there are MHC molecules which are bound to the peptide (antigen).

Figure 2-8 demonstrates the action of effector T-cells when they are activated. When CD8 T-cells are activated by APCs with MHC class-I, they stimulate the virus infected cells (APCs) and induce apoptosis. T_H1 cell either causes the digestion of

internalized bacteria by digestive enzymes of macrophages or activates B-cells to secrete their membrane-bound immunoglobulins (antibodies), upon binding with its partner antigen on MHC class-II APCs. Activation of T_H2 cells also result in the secretion of antibodies. Activated T_H17 lymphocytes increase Neutrophil response which is phagocytosis of microorganism and destroy them by degradative enzymes. T_{reg} cells are regulatory T-cells which repress T-cell responses. They prevent auto-immunity, in other word immune responses against self antigens (body's own antigens) [6].

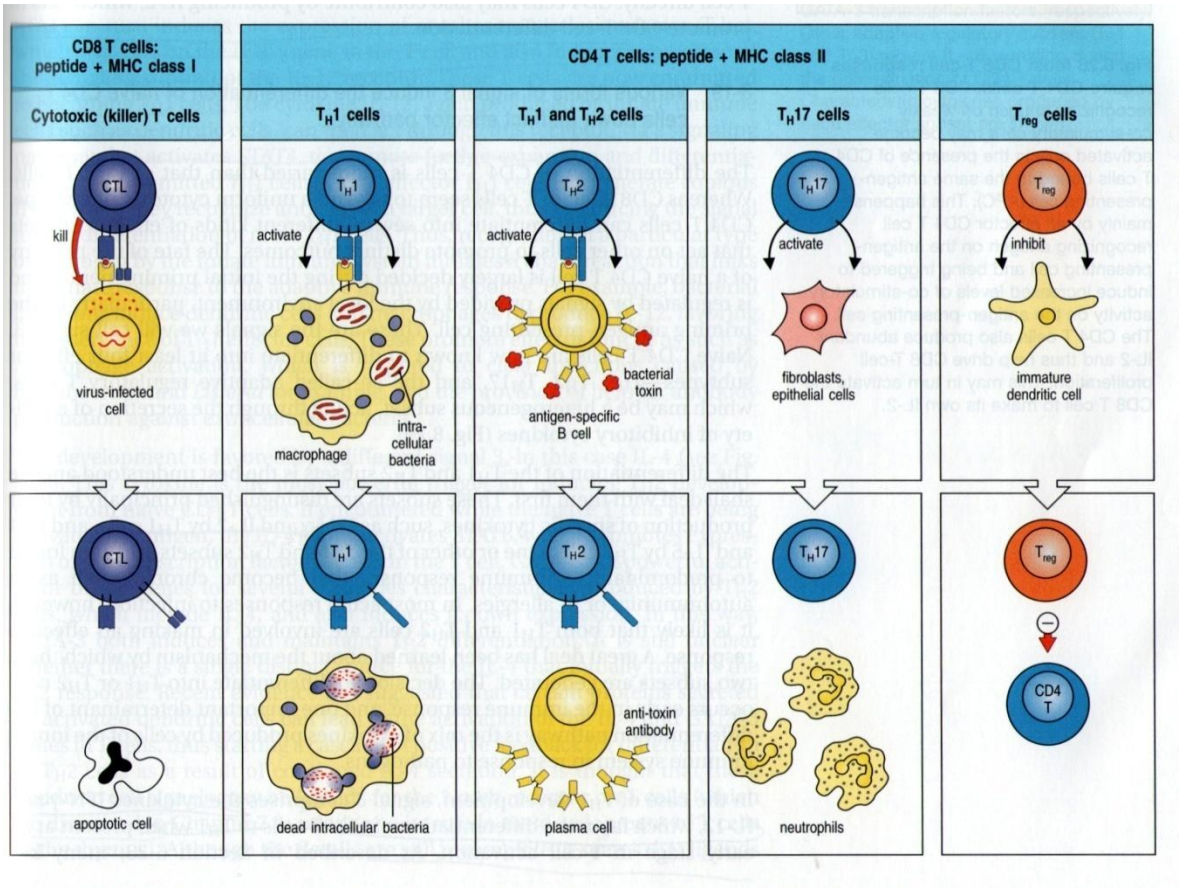

Figure 2-8: The effector T-cells and their responses when they are activated, upon meeting their antigen on APCs [6].

2.4 Leukemia

The mature blood cells get older in time and need to be replaced by the new ones on a continual basis. Therefore, hematopoiesis or the production of new blood cells occurs almost all the time. However, some hematopoietic stem cells accidentally become arrested at some stages of their differentiation, stop developing into functional blood cells, and stay as abnormal immature stem cells. The excess differentially arrested immature blood cells gives rise to "Leukemia". Normally, the course of proliferation and differentiation of precursor cells are firmly regulated. The overabundance of leukemic cells interrupts the production of normal blood cells and leads to deregulation of hematopoiesis [16]. Figure 2-9 shows the difference of blood cell precipitates of leukemic patient vs. normal person. The patient's blood has excess amount of white blood cell compared to normal person [16].

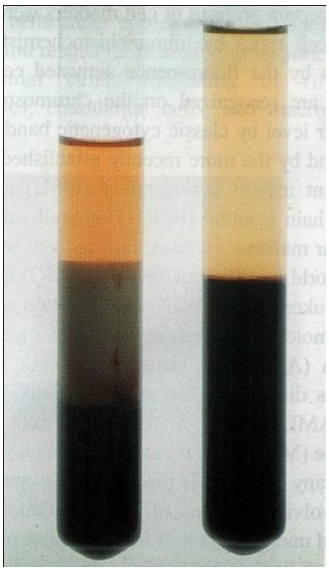

Figure 2-9: The blood of leukemic vs. normal control. On the left, leukemic patient has overabundance of white blood cells and on the right, white blood cells of a normal person is seen as very thin, almost invisible layer between plasma and red blood cells [16].

Leukemia is dived into 2 major groups: “Acute” vs “Chronic” leukemia. In chronic leukemia, the leukemic cells are derived mainly from mature cells. The mature cells accumulate and grow very slowly. In early phase of chronic leukemia, the leukemic cells still carry out the functions of normal white blood cells. But in acute leukemia, immature blast cells give rise to leukemic cells which proliferate more than their normal counterparts do. Furthermore, these 2 groups are further divided into “Lymphoid” vs “Myeloid” classes depending on at which lineage the transformed blast (stem) cell is arrested during its development [16]. Classification of leukemia is made by considering the key differentiation steps of hematopoietic precursor cells. For instance, if the precursor cell is arrested when it committed to B-lineage, then, the leukemia type is named as “B-lineage Leukemia”. Moreover, it is possible to name the leukemia in a more detailed manner, such as “mature B-cell leukemia” of “pro-B-cell leukemia”.

As stem cells propagate downstream differentiation steps, they lose their self-renewal capacity. Self-renewing capacity means ability to undergo lots of cycles of cell division and maintaining the undifferentiated (stem cell) state, at the same time [17]. Therefore, self renewing is not a synonym for only proliferation [18]. For the continuation of blood cells, self-renewing blast cells play very important role. Thanks to the self-renewing stem cells, the old blood cells are replaced by the new ones. Likewise, leukemic stem cells (LSCs) with self-renewing capacity are necessary for maintenance of leukemia [19]. Figure 2-10 demonstrates the normal blood cell and leukemic development from self-renewing normal stem cell and leukemic stem cell, respectively. In part-a of Figure 2-10, the initial transforming event (differentiation arrest or mutation giving rise to leukemia) takes place in early stages of development, the transformed cell can renew itself. Therefore, it is always ready to replace the old leukemic cells with the new ones. However, in part-c of Figure 2-10, the initial transforming event occurs at late stages of development and the non-

transformed cell has already lost its self renewing potential, so has the transformed cell. The leukemic cell cannot proliferate: it will get older and die [19].

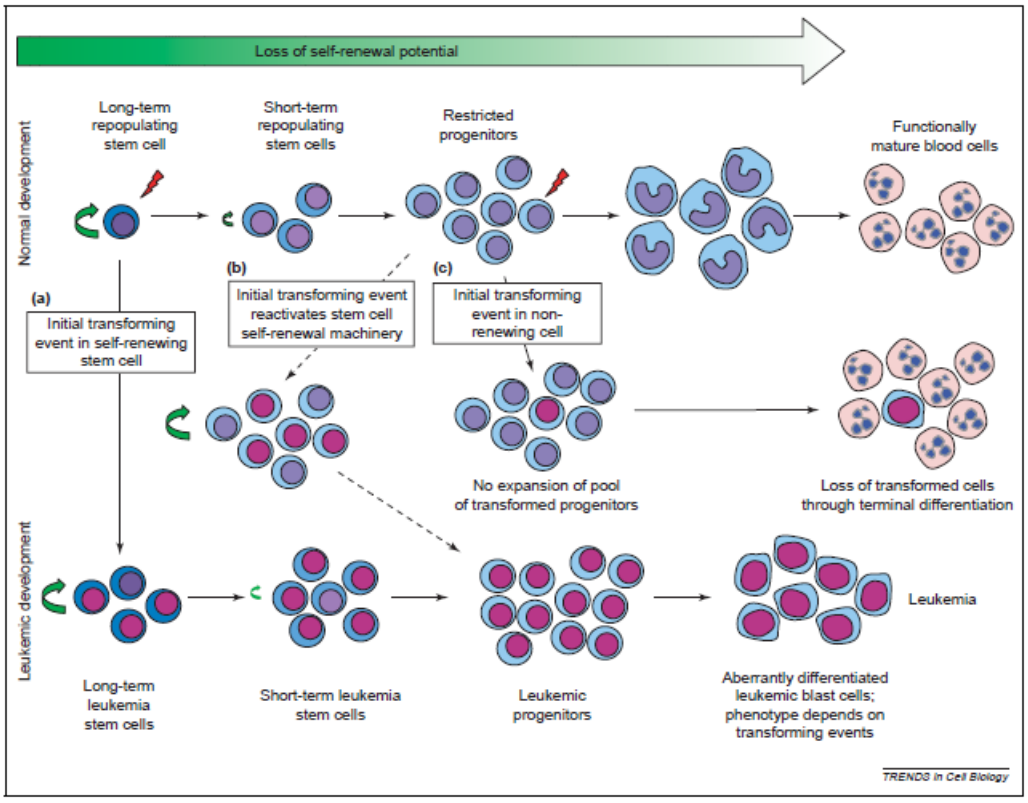

Figure 2-10: The difference between normal blood cell development and leukemic development. Self-renewing leukemic stem cells are vital for the maintenance of leukemia [19].

The cells in different lineages or at different differentiation steps express unique set of genes [3]. Thus, we can identify the development stage of a cell by just looking at its gene expression status (microarray data, qRT-PCR, etc.). Special signatures or expression patterns help us to determine the cell type.

2.5 T-cell Acute Lymphoblastic Leukemia

T-lineage acute lymphoblastic leukemia (T-ALL) is known to be resulted from the malignant transformation of hematopoietic precursor or blast cells at different maturation stages of T-cells, so called thymocytes [20]. Proliferation of developmentally arrested T-cells gives rise to T-ALL. Differentiation arrest may take place at almost all stages of thymocyte development [1]. T-cells undergo several differentiation steps until they reach their full maturation. Figure 2-11and Figure 2-12 display the differentiation stages that HSCs undergo in order to be CD4 or CD8 T-cells. Some of the T-cell development steps are “Antigen-independent” (very early stages, up to late double positive thymocyte), and some are “Antigen-dependent”. The HSCs differentiate and migrate from Bone Marrow to Thymus. In “early double-negative” stage (neither CD8 nor CD4 is expressed as surface markers), D-J segments and in “late double-negative” stage, V-DJ segments are re-arranged for β-chain of the TCR. Then the developing thymocyte enters the “early double-positive” period where α-chain of its receptor (TCR) re-arranges V-J segments. The double positive thymocyte expresses both CD4 and CD8, but a mature functional T-cell should bear either of them, not both. After “late double-positive” step, the further maturation of T-cell requires its partner antigen on which is presented by APCs. At this stage, the cell is called “naive T-cell” and it migrates from Thymus to Peripheral Lymphoid organs. If the TCR matches with its antigen, it undergoes additional differentiation steps and forms the “effector T-cell” [6, 20].

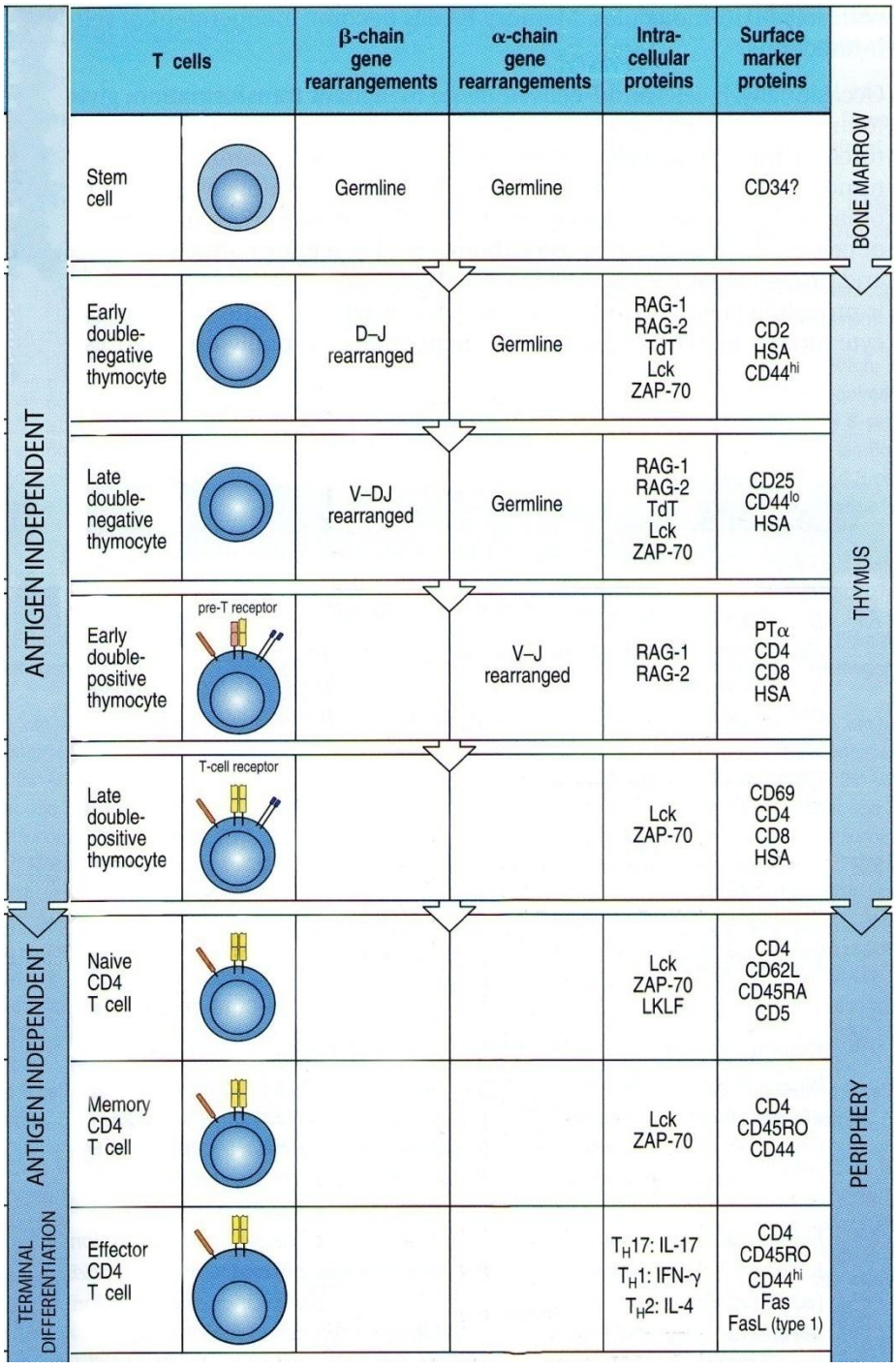

Figure 2-11: T-cell differentiation steps, giving rise to CD4 T-cell [6].

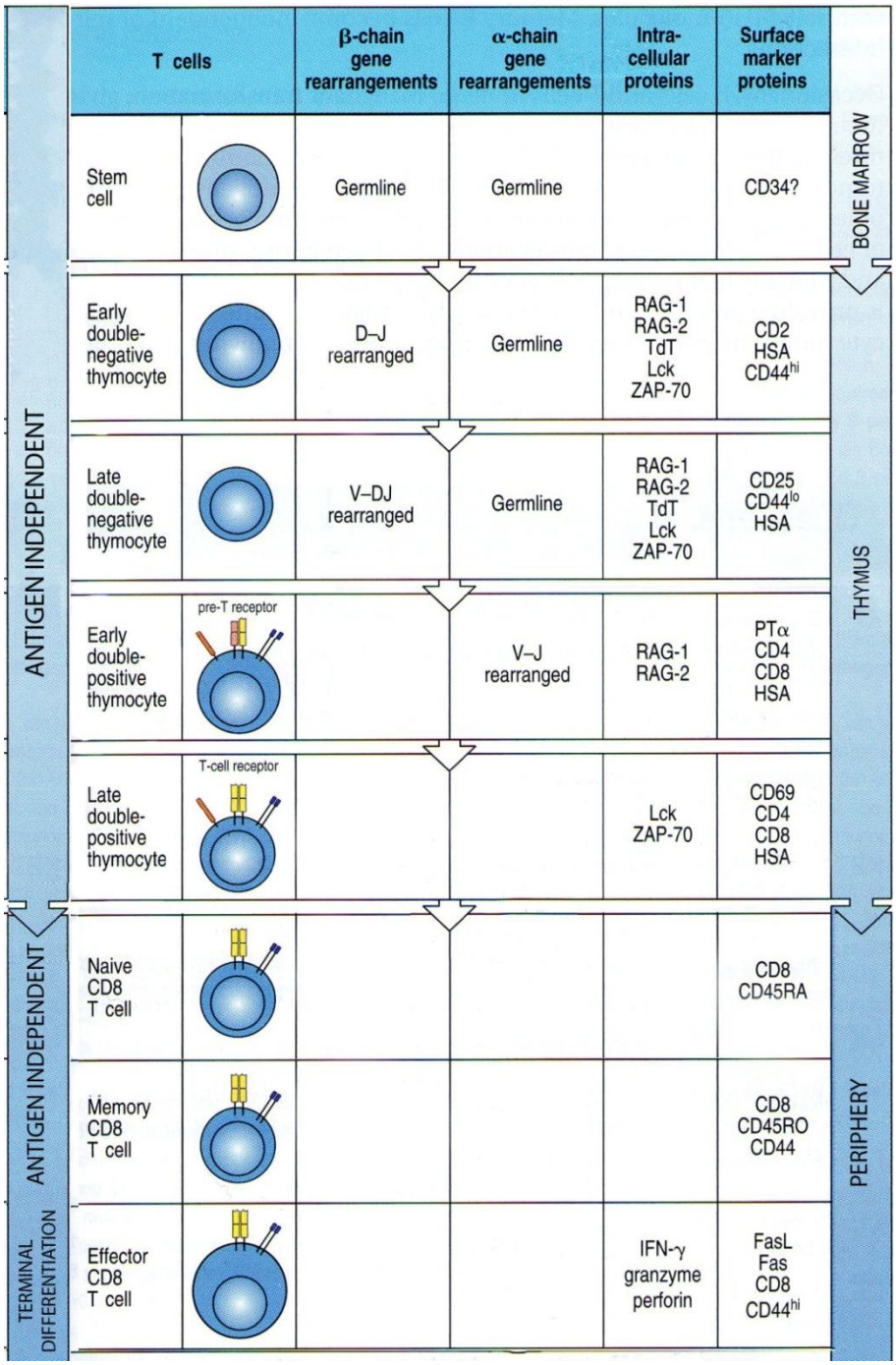

Figure 2-12: T-cell differentiation steps, giving rise to CD8 T-cell [6].

T-ALLs are a heterogeneous set of diseases, in terms of cytogenetics, molecular aberrations, and clinical characteristics [20]. Its pathogenesis and subtypes are usually undefined. T-ALL constitutes 15% of pediatric and 25% of adult ALL cases [1, 20-22].

In early stages of thymocyte differentiation, immature T-cells undergo V(D)J recombination [20]. During this time, many other genes, especially the TCR genes, are transcribed and in 'open chromatin' configuration, meaning that they are easily accessible to DNA binding proteins, like recombinases. An unusual recombinase action may lead to translocation of chromosomes [22-24]. Generally, the translocations involve abnormal juxtaposition of powerful enhancers or promoters of TCR with genes on other chromosomes, such as transcription factors or oncogenes [22, 25]. Translocations not only give rise to promoter exchange, but also fusion genes, encoding chimeric proteins [20]. The other molecular-genetic abnormalities in T-ALL involve deletions, amplifications, and point mutations which activate oncogenes or inhibit tumor suppressors, which in turn, cause differentiation arrest in thymocytes [1, 20, 22]. Deletions are the reason for loss of tumor suppressors. The correct diagnosis of acute leukemia requires wide-ranging diagnostic procedures together with cytochemistry, multiparamater flow cytometry, cytogenetics, fluorescence in situ hybridization and molecular-genetic methods [26]. Although chromosomal rearrangements are common to T-ALL, there is still a large fraction of incidents (50%) where normal karyotype is seen [20].

2.6 Importance of Biomarkers

The precise diagnosis of a tumor type is the most significant step in cancer treatments [27]. In order to apply the appropriate therapy, with maximum efficiency and minimum toxicity, the cancer should be diagnosed and classified correctly [28]. The

challenge is to identify new diagnostic biomarkers to differentiate diseased and healthy individuals properly. There are three types of biomarkers: diagnostic, prognostic or predictive. The diagnostic biomarkers mark the onset of a disease (e.g., leukemia); the prognostic biomarkers indicate the fate of the disease irrespective of the therapy; and lastly, the predictive biomarkers give information about response to therapy [29], in other words, they identify which patients are likely to benefit from a particular treatment. An optimum biomarker would be easily analyzed by a single test and measurable in body fluid (such as blood or urine). However, cancer, in our case T-ALL, is a complex disease and it is very difficult to find a single optimal biomarker at the molecular level [30-32]. The known cytogenetic and molecular changes in T-ALL are shown in Table 2-1, Table 2-2, Table 2-3, and Table 2-4.

Table 2-1: Translocations involving TCR genes in T-ALL [20].

Chromosomes involved	**Involved gene(s)**	**Protein(s)**	**Normal thymic expression**	**Function of fusion gene or expressed oncogene**
t(7;10)(q34;q24) and t(10;14)(q24;q11)	TLX1 (HOX11)	Class II homeodomain-containing	No	Transcription factor
t(5;14)(q35;q32) (cryptic)	TLX3 (HOX11L2)	Class II homeodomain-containing	No	Transcription factor
inv(7)(p15q34), t(7;7)	HOXA cluster	Class I homeodomain-containing	Some HOXA	Transcription factor
t(1;14)(p32;q11) and t(1;7)(p32;q34)	TAL1	bHLH type II	Very early stages	Transcription factor
t(7;9)(q34;q32)	TAL2	bHLH type II	No	Transcription factor
t(7;19)(q34;p13)	LYL1	bHLH type II	Very early stages	Transcription factor
t(14;21)(q11.2;q22)	BHLHB1	bHLH type II	No	Transcription factor
t(11;14)(p15;q11)	LMO1	LIM-only domain	Very early	Protein–protein

			stages	interaction
t(11;14)(p13;q11) and t(7;11)(q35;p13)	LMO2	LIM-only domain	Very early stages	Protein–protein interaction
t(1;7)(p34;q34)	LCK	SRC family of tyrosine kinase	Yes	Signal transduction
t(7;9)(q34;q34.3	NOTCH1	Notch receptor family	Yes	Fate determination, differentiation
t(7;12)(q34;p13) and t(12;14)(p13;q11)	CCND2	D-type cyclin		Cell cycle activator

Table 2-2: Known fusion genes in T-ALL [20].

Chromosomes involved	**Involved gene(s)**	**Protein(s)**	**Normal thymic expression**	**Function of fusion gene or expressed oncogene**
1p32 deletion (cryptic)	SIL/ TAL1	Proline-rich extensin/ bHLH type II	Very early stages	Transcription factor
t(10;11)(p13;q14) (often cryptic)	CALM/ AF10	ENTH motif containing/ Zinc fingers-leucine zipper containing	Ubiquitous	Transcription factor
t(11;19)(q23;p13)	MLL/ ENL	Mammalian homologue of Drosophila trithorax/ Nuclear targeting sequence containing		Transcriptional regulator during embryogenesis and hematopoiesis/Cfr supra
t(6;11)(q27;q23)	MLL/ AF6	Cfr supra/ GLGF motif containing		Cfr supra
t(10;11)(p13;q23)	MLL/ AF10	Cfr supra/ Zinc fingers-leucine zipper containing		Cfr supra
t(X;11)(q13;q23)	MLL/ AFX1	Cfr supra/ Forkhead family		Cfr supra
t(4;11)(q21;q23)	MLL/ AF4	Cfr supra/ Nuclear targeting sequence containing		Cfr supra
t(9;9)(q34;q34) (episomal or hsr)	NUP214/ ABL1	Nuclear pore complex component/	Yes	Signal transduction

		Intracellular tyrosine kinase		
t(9;14)(q34;q32) (cryptic	EML1/ ABL1	Echinoderm microtubule-associated/ Intracellular tyrosine kinase	Yes	Signal transduction
t(9;12)(q34;p13)	ETV6(TEL)/ ABL1	ETS DNA binding containing/ Intracellular tyrosine kinase	Yes	Signal transduction
t(9;12)(p24;p13)	ETV6(TEL)/ JAK2	ETS DNA binding containing/ Intracellular tyrosine kinase	Yes	Signal transduction
t(9;22)(q34;q11)	BCR/ ABL1	Serine-threonine kinase/ Intracellular tyrosine kinase	Yes	Signal transduction
t(4;11)(q21;p15)	NUP98/ RAP1GDS1	Nuclear pore complex component/ Cytoplasmic		?
t(10;11)(q25;p15)	NUP98/ ADD3	Nuclear pore complex component/ Cytoplasmic		?

Table 2-3: Mutations in T-ALL [20].

	Involved gene(s)	**Protein(s)**	**Normal thymic expression**	**Function of fusion gene or expressed oncogene**
NOTCH1	NOTCH1	Notch receptor family	Yes	Fate determination, differentiation
FLT3 ITD and m835	FLT3	Receptor tyrosine kinase	No	Development of hematopoietic stem cells
N-RAS	N-RAS	Signaling protein	Yes	Signal transduction

Table 2-4: (Cryptic) Deletions in T-ALL [20].

Chromosomes involved	Involved gene(s)	Protein(s)	Normal thymic expression	Function of fusion gene or expressed oncogene
9p21 (homozygous/hemizygous)	P16	INK4/ARF	Yes	G1/S cell cycle inhibitor (tumor suppressor gene)
Del (6q)	Unknown			Tumor supressor gene ?

2.7 Role of Microarrays in Biomarker Discovery

The analysis of genome-wide expression profiles is frequently used to discover new biomarkers [2]. Since microarrays give information about the expression of several genes in parallel [33], they are proposed to be a robust technology for the identification of signatures or expression patterns that vary significantly between diseased and healthy samples [30, 34, 35]. The current long-lasting diagnostic procedures for leukemia (e.g., cytomorphology, immunophenotyping, metaphase cytogenetics) might be replaced by the comprehensive microarray protocols which takes two or fewer days and allows the simultaneous detection of the expression of almost all genome in one experimental approach [27]. With the extensive usage of microarrays, an increasing number of methods have been developed to identify biomarkers [34, 36-40].

2.8 Limitations of Microarrays and Necessity of Network Information

Along with the advantages of microarrays, there are some limitations. For instance, some important genes of cancer are not differentially expressed at the level of transcription or the fate of cancer may not be controlled at the level of expression [30]. In addition, the transcriptional level does not always correlate with the translational level, in other words, the mRNA expression is not always equal to the protein expression [41]. The gene expression levels may vary even in the genetically identical cells with same histories of environmental exposure. These variations come from the random nature of biochemical reactions and they are called as "noise" [42]. Therefore, some differences in microarray data may result from the noise. Moreover, there is also an "experimental noise" other than the "expression noise" (biological variations), in which slight unintended differences in experimental setup may lead to huge differences in hybridization of probes. This kind of technical noise is also considered as one of the restrictions for successful use of this technology [43-45]. Another limitation of microarrays might be the lack of pathway knowledge. One way to overcome this problem is to integrate gene-expression profiles with protein-protein interaction (PPI) networks [46-49]. A biological function or a phenotype is not controlled by just one gene [50], rather pathways or cross-talks among proteins are responsible for the regulation of a function [51-53]. Thus, network information provides a functional insight when integrated with microarray data [30]. Identification of differential gene modules or subnetworks instead of individual differentially expressed genes may increase the reliability and robustness of biomarkers [30]. Chuang *et al.* recently established a network-based approach which integrated microarray data with PPIs and proved that their method predicts the breast cancer metastasis more accurately [2].

2.9 Expression-Based vs. Network-Based Classification Techniques

Essential patterns of gene expression which successfully distinguishes the diseased vs. healthy individuals can be extracted by numerous gene expression-based classification techniques. Some examples of these methods are Support Vector Machines (SVMs), Self Organizing Maps (SOMs), Hierarchical clustering [54], Predictive Analysis of Microarrays (PAM), Random Forest (RF), and k-top scoring pairs (kTSP) [55].

Traditional expression-based classification techniques identify only differentially expressed (DE) genes as signatures/markers, but a network-based approach returns subnetworks including both DE and non-DE genes [2]. The differential expression analysis of gene expression profiles returns massive numbers of genes which makes it difficult to conclude that the differential expression of a particular gene is resulted from the disease/abnormality. As noted in the previous part, even cells with identical genome and environmental exposure history show variations in their gene expression, which is referred as noise. Therefore, we cannot be so sure that the differential analysis identifies genes whose differential expression is only resulted from the disease: there may be another factor affecting the expression of that gene. Thus, just differential analysis alone is not reliable enough to conclude a gene as a biomarker of a disease.

However, network-based approaches do not depend on only expression data. They integrate microarray data with protein-protien interaction data and return subnetworks, rather than individual genes. The number of subnetworks is not as high as the number of individual genes identified by differential analysis. These subnetworks are differential as a whole, but individual genes in a subnetwork may not be differential. Irresponsive genes in subnetworks have very important role in interconnecting several DE-genes. The resulting modules represent models of the underlying molecular mechanisms. Each module corresponds to a different functional pathway or complex. Since genes function in

collaboration rather than alone, subnetworks are more rational than independent responsive genes, from a biological perspective [56].

Hierarchical clustering is one example of the expression-based classification techniques which groups together both genes and samples with similar expression pattern by using microarray data. It can also define subclasses of a disease, such as cancer [57]. Clustering algorithms try to organize genes or samples according to their similarity in expression. Genes with related pattern appear in the immediate vicinity of each other. In other words, it gathers co-expressed genes together. It also has a tendency to arrange genes with similar functions together since the functionally related genes are likely to be co-expressed [58]. Hierarchical clustering is a type of unsupervised clustering which does not use any prior knowledge regarding the sample classes.

2.10 Shortcomings of Network

The drawback of pathway-based classifiers is that most of the human genes have not been assigned to a definitive pathway yet [2]. As pathways become more complete, the classification performances of pathway-based approaches will increase [36].

Furthermore, "global" PPI networks are frequently used in bioinformatic analyses. However, the interacting partners may change in different tissues. For instance, protein-A may be interacting with protein-B in neural tissue, but not in muscle tissue. So, if a muscle disease is studied, the interaction of protein-A with protein-B should not be included. Some of the genes/proteins may localize and perform their function in specific tissues. In short, instead of global networks, tissue-specific networks should be constructed and used in studies with tissue-specific diseases. They are hypothesized to increase the efficiency of network-based algorithms [59].

Chapter 3

3. METHODS

3.1 Microarray Data (the MILE study)

We used a comprehensive group of Affymetrix microarray datasets to determine which genes or modules discriminate T-ALL samples from healthy individual samples. This study included bone marrow samples of 173 T-ALL patients at diagnosis (untreated patients) and 74 non-leukemia/healthy specimens (e.g., healthy, hemolysis, iron deficiency). The patient samples were heterogeneous; that is, there were samples from different stages of T-ALL. All microarray data were obtained from the Microarray Innovations in Leukemia (MILE) study, the National Center for Biotechnology Information's Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) under series accession number GSE13204. The MILE study is an international standardization program and was conducted by 11 laboratories across three continents [33]. In order to avoid experimental noise, the conductors of this study standardized the experiment across laboratories by supplying all laboratories with the same equipments and reagents. They proved both the "intra-laboratory" and "inter-laboratory" repeatability of the same microarray data [33]. There were two stages in this study containing microarray samples of 18 different classes of leukemia, myelodysplastic syndromes and non-leukemia. We used only the first stage in which the whole genome microarray platform (HG-U133

Plus 2.0; Affymetrix, Santa Clara, CA) was used [35]. In the GEO-series-matrix, the microarray data were already summarized and quantile normalized as described in [60].

In order to integrate microarray data with the human PPI network, it was necessary to convert Affymetrix probeset IDs to corresponding Entrez gene IDs. The annotation data was also provided under the same accession number.

It is important to note that values for multiple probes corresponding to the same Entrez ID were averaged so that a particular gene ID is seen only once throughout the microarray data.

3.2 Human Protein Interaction Network

The human PPI network is obtained from Human Protein Reference Database. There are 38788 PPIs whose interactions are experimentally verified and extracted from the literature [61].

3.3 Data Integration

In order to combine the human PPI data with microarray data, we used a previously described algorithm (Pinnacle-Z, plug-in to Cytoscape) [2]. Figure 3-1 demonstrates how Pinnacle-Z algorithm generates subnetworks by combining microarray data with PPI data. It is a type of supervised classification technique in which the user gives the phenotype information (diseased vs. healthy) as well as the normalized microarray data and PPI data. The algorithm superimposes gene expression values with corresponding network proteins and begins from every protein (seed) in the PPI network and greedily appends interactions to identify the subnetwork starting from each seed whose mean expression for each sample

best discriminates between the two sample types (in our case, the sample types are T-ALL and non-leukemia/healthy conditions). At first, the number of resulting subnetworks is equal to the number of proteins in the PPI network. Then, non-significant modules are filtered out by three types of permutation testing [50]. At last, subnetworks with high discriminative potential are obtained.

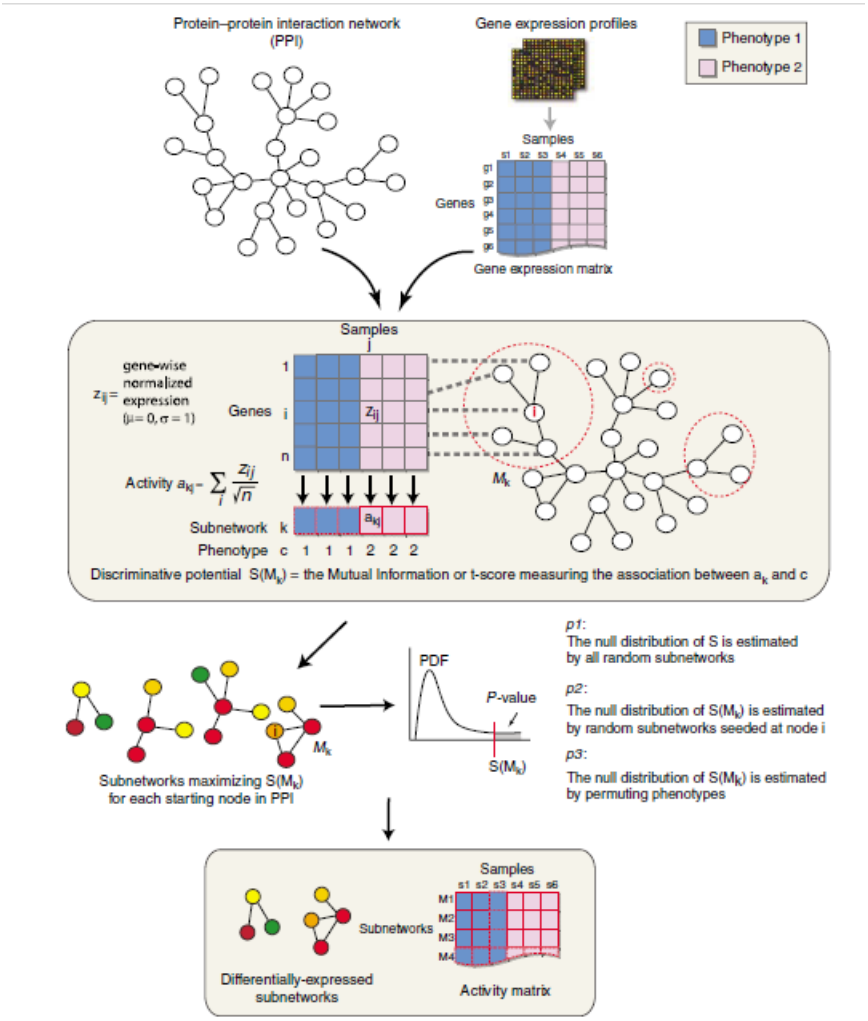

Figure 3-1: Integration of gene expression data with protein-protein interaction data [2].

To equate the number of healthy individuals (74) to the number of patients with T-ALL (173), we divided T-ALL patient samples into two groups randomly. If the number of diseased samples is much higher than number of healthy samples (or vice versa), the algorithm might not work properly. Thus, we gave the algorithm a microarray data which has almost equal numbers of healthy and diseased samples (87diseased+74healthy (combination1) and 86diseased+74healthy (combination2)). Then, we merged each half of that patient data with the healthy data, separately. For each merged data (87+74 and 86+74), we ran the algorithm 4 times. The parameters were chosen as described in (Chuang et al., 2007). Then, we re-sampled our patient samples 5 times and repeated the division of diseased samples, merging the two sample types (diseased and healthy) and algorithm-running steps. We obtained results of 40 runs in total.

3.4 Hierarchical Clustering

In order to identify groups of genes that have similar expression patterns and to show the difference between expression-based and network-based classification approaches, Hierarchical clustering method was applied to 173 T-ALL and 74 healthy samples together. Initially, there were 20148 probes. Then, these probes are filtered by T-test and determined 200 most differential and 100 most differential genes. Hierarchical clustering algorithm which is implemented in Expander was performed on both 200 and 100 most differentially expressed genes. Both samples and genes are clustered with complete linkage and Pearson correlation.

3.5 Classification Accuracy

The classification accuracy of a given subnetwork, or the overall rate of correct predictions of the patient and healthy datasets was estimated by 10-fold crossvalidation with two different classifiers (J48 and RBF Network Classifiers implemented in WEKA [62]). According to 10-fold cross-validation, the complete dataset was divided into 10 uniformly sized subgroups; the classifier was trained for nine subgroups and predictions were made for the remaining subgroup. High classification accuracy, like 90%, means that a given subnetwork differentiates/predicts the patient and healthy samples correctly, 90% of the time.

As pointed out, we randomly divided diseased samples into two groups: For each half we found subnetworks and determined their classification accuracy. We used the other half as a validation set to cross-test the prediction accuracy of a given subnetwork.

The classification accuracies of 100 and 200 most differentially expressed genes that are determined by Expander are also calculated by 10-fold cross validation with 2 classifiers (J48 and RBF Network).

3.6 Functional Enrichments

Functional enrichment analysis was achieved using Gene Ontology Tree Machine (GOTM) which searches for functional enrichments from GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes) categories for genes in the given subnetworks.

Chapter 4

4. RESULTS and DISCUSSION

4.1 Resulting Subnetworks and Their Characteristics

We applied a network-based approach, Pinnacle-Z algorithm [2] to distinguish T-ALL patients and healthy samples by integrating microarray data with the human PPI network. This approach enabled us to identify subnetworks/modules as markers which differentiate the patients from healthy individuals. The size of subnetworks or the number of genes in a subnetwork varies, ranging from 1 to 10 genes (the upper limit may change: a subnetwork may contain more genes but our subnetworks have up to 12 genes). As noted earlier, the individual genes may not be responsive, meaning that expression of a particular gene is different in patients and healthy samples, but an entire subnetwork is differential.

As explained in the methods section, we prepared different combinations (10 different combinations) of patient-healthy merged data and ran the algorithm 4 times with each combination. If the two classes of samples, namely, healthy and diseased samples contain similar number of samples, the algorithm might work more accurately. That's why the patient data set (173 samples) was divided into two randonmly and merged with the healthy data (74). Then, we focused on the most frequent subnetworks. We recovered most repeatedly observed 19 subnetworks, out of 183 subnetworks (see Figure 4-1, Figure 4-2,

Figure 4-3). Subnetworks cover 102 genes in total and the majority of them are differentially expressed (DE) (Please refer to Chapter-6, Supplementary Materials to get more detailed information regarding these genes and their functions). There were only 6 genes whose expression was not differential among T-ALL and healthy samples (see Table 4-2). The list of genes in the subnetworks and their functions can be found in Supplementary materials in Chapter-6 (Table 6-1 through Table 6-19) and the pathways they are involved in can be seen in Table 4-1. As Table 4-1 displays, each subnetwork is enriched and therefore represents a different pathway.

Table 4-1: KEGG and GO enrichments of the most frequent 19 subnetworks

subnetworks	**KEGG**	**GO:Function**
subnetwork1	No enrichments	Actin binding, Calmodulin binding
subnetwork2	Chemokine-signaling pathway, Chemokine-Cytokine receptor interaction, Endocytosis	Signal-transducer activity, chemokine activity, cytokine activity
subnetwork3	ECM-receptor interaction, Focal adhesion, Hematopoietic cell lineage	Growth-factor binding, Extracellular matrix structural constituent
subnetwork4	Pathways in cancer, Colorectal cancer, Pancreatic cancer, Chronic myeloid leukemia, Cell- cycle, Wnt signaling pathway, MAPK signaling pathway	Promoter binding, transcription activator activity, transcription factor activity
subnetwork5	Neuroactive ligand-receptor interaction	Signal transducer activity, G-Protein coupled receptor activity
subnetwork6	No enrichments	No enrichments
subnetwork7	No enrichments	No enrichments
subnetwork8	Pathways in cancer, Colorectal cancer, Wnt signaling pathway	No enrichments
subnetwork9	Leukocyte transendothelial migration, regulation of actin cytoskeleton, dilated	GTP binding, Actin binding

	cardiomyopathy, Fc gamma R-mediated phagocytosis, Hypertrophic cardiomyopathy (HCM)	
subnetwork10	Phagosome, Gap junction	GTP binding, GTPase activity, nucleotide binding, structural molecule activity
subnetwork11	No enrichments	Metal-ion binding, Calcium ion binding, SH3 domain binding, Phosholipid transporter activity
subnetwork12	No enrichments	No enrichments
subnetwork13	Complement and coagulation cascades	No enrichments
subnetwork14	Bladder cancer	Calcium ion binding, Matalloendopeptidase activity
subnetwork15	Toll-like receptor signaling pathway	Transmembrane receptor activity
subnetwork16	Pathways in cancer, Neurotrophin signaling pathway, ErbB signaling pathway, Non-homologous end-joining	ATP binding, protein C-terminus binding
subnetwork17	Toll-like receptor signaling pathway, Pathogeneic Escherichia Coli infection	Transmembrane receptor activity
subnetwork18	Tight junction, Endocytosis	No enrichments
subnetwork19	No enrichments	No enrichments

Table 4-2: GO enrichment of Non-Differentially Expressed Genes.

Gene Symbol	**Subnetwork**	**Gene title**	**GO: Function**	**GO: Process**
P2RX7	Sub-9	purinergic receptor P2X, ligand-gated ion channel, 7	ATP binding, ATP-gated cation channel activity, ion channel & receptor activity	ion transport, signal transduction
C9	Sub-13	complement component 9		caspase activation, complement activation, induction of apoptosis

CHGA	Sub-14	chromogranin A (parathyroid secretory protein 1)	calcium ion binding, protein binding	regulation of blood pressure
PLG	Sub-14	plasminogen	apolipoprotein binding, calcium ion binding, peptidase activity, plasmin activity	blood coagulation, induction of apoptosis, negative regulation of angiogenesis, cell proliferation, fibrinolysis, proteolysis, tissue remodeling
MEP1A	Sub-14	meprin A, alpha (PABA peptide hydrolase)	astacin activity, metal ion binding, metallopeptidase activity, zinc ion binding	digestion, proteolysis
HN1L	Sub-16	hematological and neurological expressed 1-like	No enrichments	No enrichments

Development of a cancer requires the accumulation of several mutations in several genes in different pathways [63]. Since the cell regulation is controlled by many pathways, a mutation in one pathway may be compensated by other pathways. But if there are many mutations in many pathways, the harmful impacts of these mutations cannot be compensated. For instance, if a tumor suppressor gene is inactivated, it is not enough to generate cancer; additional mutations are needed before cancer appears [16]. It is found that the subnetworks correspond to different pathways, in general (Table 4-1). This result supports the notion that multiple cell regulatory pathways are involved in production of leukomogenesis.

Among the genes in the subnetworks, some are known to be associated with T-ALL from previous studies: ABL1 from previous studies: ABL1 [20, 64, 65], CCL5 [66], CD99 [67], TP53 [68], and WT1 [69]. Table 4-3 exhibits in which subnetworks these T-ALL related genes and their behavior. Many other genes seen in subnetworks were not previously found to be linked to T-ALL (Some are related to cancer, but not specifically to T-ALL, such as VANGL1 [70], CEACAM5 [71], SSBP2 [72], LTB4R [73], TFAP2C [74], USP13 [75], CD36 [76], UBE2I [77], EWSR1 [78], SOX4 [79] and LASP1 [80] (see

Table 4-4, for corresponding subnetworks and the behavior of these genes).

Table 4-3: T-ALL related genes found in subnetworks and their behavior in our samples.

T-ALL related genes	Subnetwork	Behavior
ABL1	16	↑
CCL5	2	↓
CD99	12	↑
TP53	4	↑
WT1	4	↑

Table 4-4: Cancer related genes recovered in subnetworks and their behavior in our samples.

Cancer-related genes	Subnetwork	Behavior
VANGL1	8	↑
CEACAM5	7	↑
SSBP2	7	↑
LTB4R	5	↓
TFAP2C	4	↑
USP13	8	↑
CD36	3	↓
UBE2I	4	↑
EWSR1	6	↑
SOX4	4	↑
LASP1	9	↓

Subnetwork1 includes Calmodulin1 gene (CALM1) which encodes for a Calcium-binding protein and is very important for Ca2+ ion homeostasis [81]. CALM1 modulates

the control of several enzymes and other proteins by Ca2+ ion. It affects lots of protein kinases and phosphatases which have key roles in many pathways.

Chemokines are vital for recruiting the leukocytes to the sites of inflammation or infection. They also have an impact on the T-cell differentiation. In subnetwork2, chemokines CCR5 & CCL5 and PTK2 gene are found. These genes are important in regulating T-cell functions. CCR5 facilitates to the formation of the immunological synapse which is a special structure at the interface between APC and T-cell. CCR5 is chemotactic for T_H1 but not T_H2 [15]. CCL5 is required for T-cell proliferation. PTK2 (or FAK) is important for cell migration [82]. PTK2 activation is previously found to be related with generation transformed leukemic cells [83].

Subnetwork15 and subnetwork17 are dominated by genes of Toll-like Receptor (TLR) family: TLR1, TLR2, TLR10, TLR4, and TLR5. TLR family is very important in recognition of pathogen and production of cytokines, chemokines, interferons and immunoglobulins to build up the proper immune response [84]. Almost all tissues express one kind of TLR. Different TLRs exhibit different expression patterns [85]. When a pathogen is recognized by TLRs, complicated intracellular signaling pathways such as MAP kinase, NF-κB and interferon regulatory factors are activated. TLR1, TLR2, TLR4, and TLR5 recognize ligands on the pathogen membranes, such as peptidoglycans, LPS, flagellin and lipopeptides [84]. The biological meaning of a particular combination of expression pattern, ligand specificity, and signaling specificity for each TLR remains unknown [86].

Three genes appear more than once in our subnetworks (Table 4-5). A gene appearing only in one subnetwork means that that gene has only one function and is

involved in only one pathway, which is usually not the case in biological systems. Therefore, it is not a surprising result having a gene in more than one group.

Table 4-5: The genes that appear more than one subnetwork, which subnetworks they are seen and their functions.

Gene symbol	Gene title	Subnetworks	GO:Function	GO:Process
PRTFDC1	phosphoribosyl transferase domain containing 1	6, 7	hypoxanthine phosphoribosyltransferase activity	nucleoside metabolic process, purine ribonucleoside salvage
EWSR1	Ewing sarcoma breakpoint region 1	6, 7	RNA binding, calmodulin binding, metal ion binding, nucleic acid binding, sequence-specific DNA binding, transcription factor activity, zinc ion binding	regulation of transcription, DNA-dependent, transcription
SMAD4	SMAD family member 4	4, 8	SMAD binding, protein homodimerization activity, sequence-specific DNA binding, transcription activator activity, transcription factor activity	BMP signaling pathway, SMAD protein complex assembly, anterior/posterior pattern formation, negative regulation of cell growth, proliferation and transcription

Subnetworks are rich in transcription factors: there are 14 transcription factors involved in subnetworks (Table 4-6). This result is in accordance with the important assumption that abnormal or ectopic activation of specific transcription factor genes,

with/without chromosomal rearrangements, is the main event in transformation of immature T-cells [24].

Table 4-6: Transcription factors recovered in subnetworks

Gene Symbol	Gene title	Subnetwork	Behavior
TNNI2	troponin I type 2 (skeletal, fast)	sub1	Down
SMAD4	SMAD family member 4	sub4	Up
TP53	tumor protein p53	sub4	Up
ATF2	activating transcription factor 2	sub4	Up
WT1	Wilms tumor 1	sub4	Up
SOX4	SRY (sex determining region Y)-box 4	sub4	Up
TFAP2C	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	sub4	Up
ATXN3	ataxin 3	sub6	Up
SALL2	sal-like 2 (Drosophila)	sub6	Up
EWSR1	Ewing sarcoma breakpoint region 1	sub6, sub7	Up
SSBP2	single-stranded DNA binding protein 2	sub7	Up
GFI1B	growth factor independent 1B transcription repressor	sub11	Down
ABL1	c-abl oncogene 1, receptor tyrosine kinase	sub16	Up
XRCC6	X-ray repair complementing defective repair in Chinese hamster cells 6 (Ku autoantigen, 70kDa)	sub16	Up

There were 6 tyrosine kinases (Table 4-7) in our subnetworks. Tyrosine kinases have a critical role in TCR signaling, regulation of T-cell immune response, T-cell survival and proliferation. Expression of tyrosine kinases, like ABL1, affect pre-TCR and TCR signaling and give a proliferative and survival advantage [20]. ABL1 is an oncogene and is

often seen as ABL1-NUP214 fusion gene in T-ALL [20, 25]. As can be seen in subnetwork16 (Figure 4-3), ABL1 is up-regulated in T-ALL patients compared to healthy individuals. Although YWHAG was not found to be related with T-ALL in literature, it interacts with ABL1. The over expression of ABL1 might induce up-regulation of the YWHAG gene (Figure 4-1, subnetwork7). BTK is significantly down-regulated in T-lymphocytes [87], which is consistent with our results. PTK2, also known as FAK, Focal Adhesion Kinase, (Figure 4-1, subnetwork2) has a role in growth, differentiation, tumor metastasis, and wound healing [88-91]. It was not previously associated with T-ALL but its over-expression is associated with several types of cancer [92, 93]. But in a recent study, it has been shown that PTK2 protein is predominantly absent in both normal T-cells and T-lymphoblastic leukemia/lymphoma. Although it is negative in T-cell leukemia/ lymphoma, it is mostly positive in B-cell lymphomas [88]. Consistent with the literature, PTK2 gene is also down-regulated in T-ALL patients used in this study. Cell Adhesion molecules (CAMs) are necessary for interaction of hematopoietic cells with extracellular matrix with stromal and other cells [16]. Defects in adhesion were reported in other types of leukemia before, such as in Chronic Myeloid Leukemia (CML) [16, 94, 95]. The failure of HSCs to express the correct or fundamental adhesion molecules may contribute to transformation of a normal HSC to leukemic cell and to get arrested at a particular step of their differentiation. The adhesion deficiency may also help leukemic cells to escape from the recognition by immune system [16].

Table 4-7: Genes involved in tyrosine-kinase signaling pathway recovered in subnetworks

Gene Symbol	Gene title	Subnetwork	Behavior
BTK	Bruton agammaglobulinemia tyrosine kinase	sub17	down
ABL1	c-abl oncogene 1, receptor tyrosine kinase	sub16	up

YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	sub7	up
PTK2	PTK2 protein tyrosine kinase 2	sub2	down
COL1A2	collagen, type I, alpha 2	sub3	down
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	sub14	down

Interestingly, the subnetworks are also abundant in zinc-ion binding proteins (Table 4-8) which are generally enzymes, including those involved in DNA repair. Zinc finger motifs play key role in interaction of proteins with nucleic acids (DNA/RNA) [96]. They are essential for site-specific DNA recognition and transcriptional activation [97]. Zinc-ion (Zn2+) has both structural and regulatory roles in zinc-binding proteins; meaning that, Zn2+ maintains the three-dimensional structure of the proteins and it is required for the proper function of the proteins. For example, p53 needs Zn2+ to fold properly. Both excess and inadequate amounts of Zn2+ are the causes behind misfolding of p53 [97]. One of the molecular mechanisms in carcinogenesis is the deformation of zinc-finger domains in DNA repair proteins [96]. Zinc ion is also important for thymic immune responses [98]. Low levels of zinc are frequently reported in ALL cases. Normal lymphocytes contain more zinc than leukemic cells [99]. Treatment of T-ALL patients with zinc, in addition to chemotherapy was hypothesized to increase the overall ability to recover from T-ALL permanently and to endure toxic effects of chemotherapy [100]. Some of the Zn2+ ion binding proteins are up-regulated, but some are down-regulated (see Table 4-8). Although zinc ion binding proteins that we found do not behave similarly at the level of expression, it is evident that more attention should be paid on them.

Table 4-8: Zinc-ion binding proteins recovered in subnetworks

Gene Symbol	Gene title	Subnetwork	Behavior
TP53	tumor protein p53	sub4	up
ATF2	activating transcription factor 2	sub4	up
WT1	Wilms tumor 1	sub4	up
SALL2	sal-like 2 (Drosophila)	sub6	up
EWSR1	Ewing sarcoma breakpoint region 1	sub6, sub7	up
USP13	ubiquitin specific peptidase 13 (isopeptidase T-3)	sub8	up
CXXC4	CXXC finger 4	sub8	up
LASP1	LIM and SH3 protein 1	sub9	Down
GFI1B	growth factor independent 1B transcription repressor	sub11	Down
APP	amyloid beta (A4) precursor protein (peptidase nexin-II, Alzheimer disease)	sub13	Down
MEP1A	meprin A, alpha (PABA peptide hydrolase)	sub14	non-DE
MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	sub14	Down
RASA4	RAS p21 protein activator 4	sub16	Up
BTK	Bruton agammaglobulinemia tyrosine kinase	sub17	Down
USP20	ubiquitin specific peptidase 20	sub18	Up
PRKCI	protein kinase C, iota	sub18	Up

Two subnetworks are actually individual genes, rather than interconnected genes (subnetwork10, Figure 4-2 and subnetwork19, Figure 4-3). These genes are TUBB1 and HHIP. HHIP stands for Hedgehog Interacting Protein, which is negative regulator of Hedgehog signaling pathway. Over activity of Hedgehog signaling pathway is related with many cancer types [101]. HHIP is found to be associated with lung cancer [102] and brain tumor [103]. Although it is down-regulated in several tumor types, it is up-regulated in our T-ALL patients (subnetwork19, Figure 4-3) [104, 105]. The other individual gene that is found as a marker is TUBB1, Tubulin beta-1 (subnetwork10, Figure 4-2) which has a role in assembly of microtubules only in hematopoietic cells. Altered expression of beta-tubulin isotypes were observed in specific tumor types [106]. Tubulin mutations are involved in resistance to drugs that target microtubules in cancer patients [107].

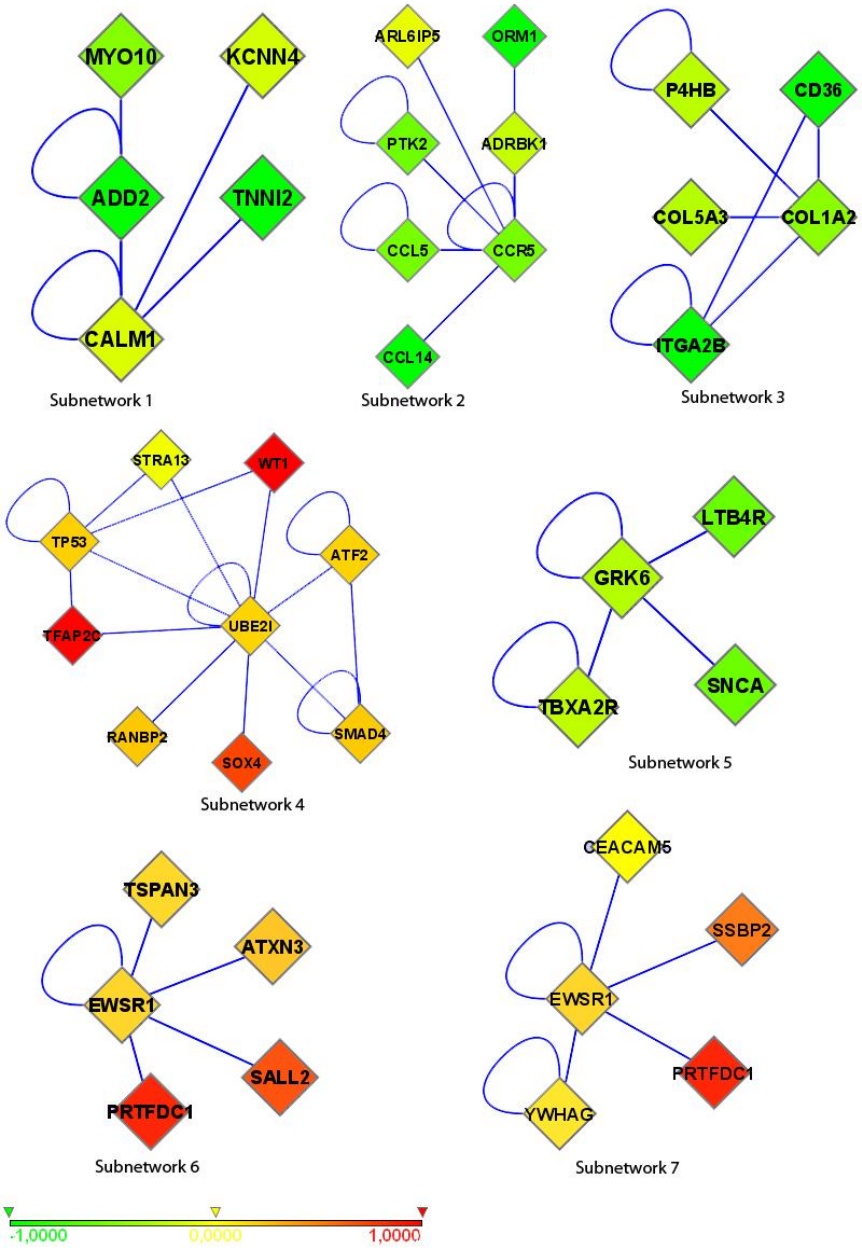

Figure 4-1: The most frequent subnetworks. Nodes represent proteins and edges represent interactions. The color of each node ranges in accordance with the change in expression of the corresponding gene for T-ALL versus healthy samples. The shape of each node shows whether its gene is significantly differentially expressed (diamond; $P<0.05$ from a two-tailed t-test) or not (circle).

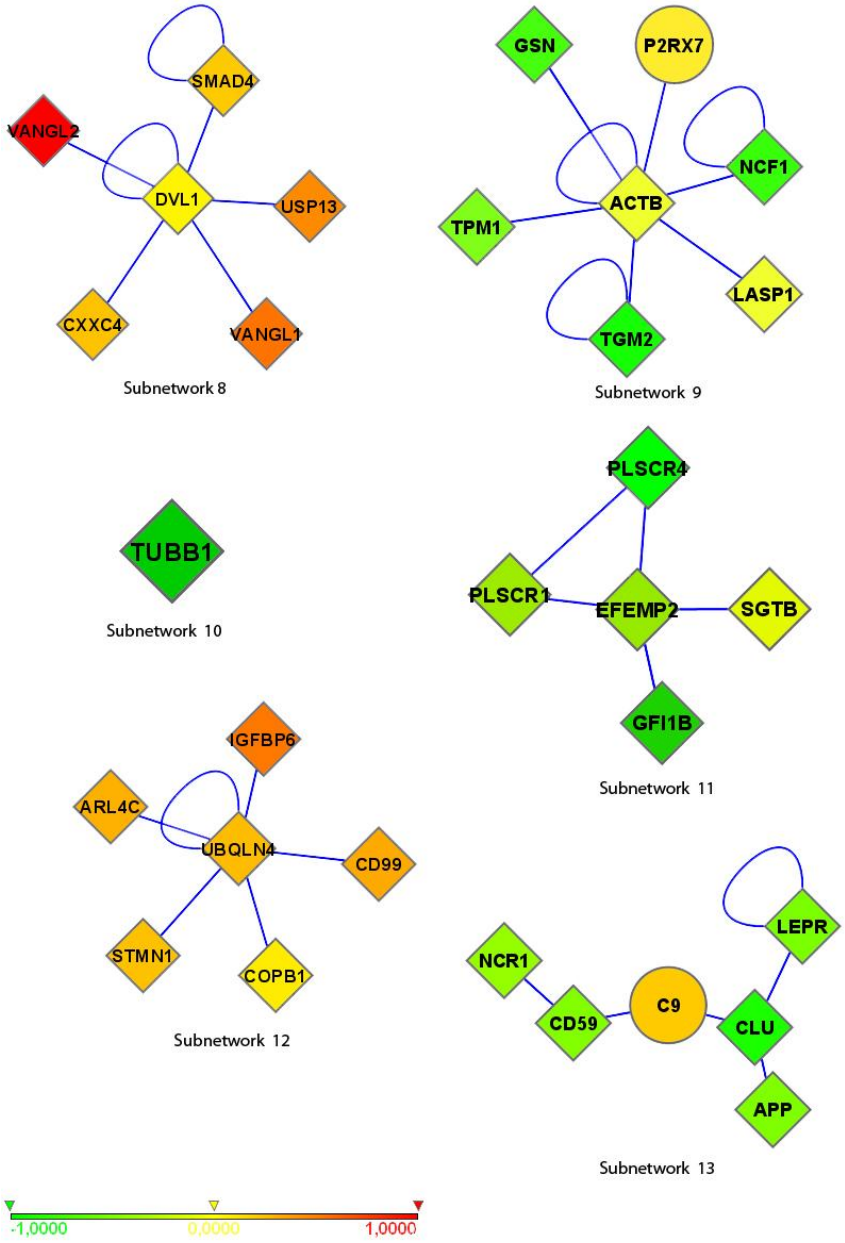

Figure 4-2: The most frequent subnetworks. Nodes represent proteins and edges represent interactions. The color of each node ranges in accordance with the change in expression of the corresponding gene for T-ALL versus healthy samples. The shape of each node shows whether its gene is significantly differentially expressed (diamond; $P<0.05$ from a two-tailed t-test) or not (circle).

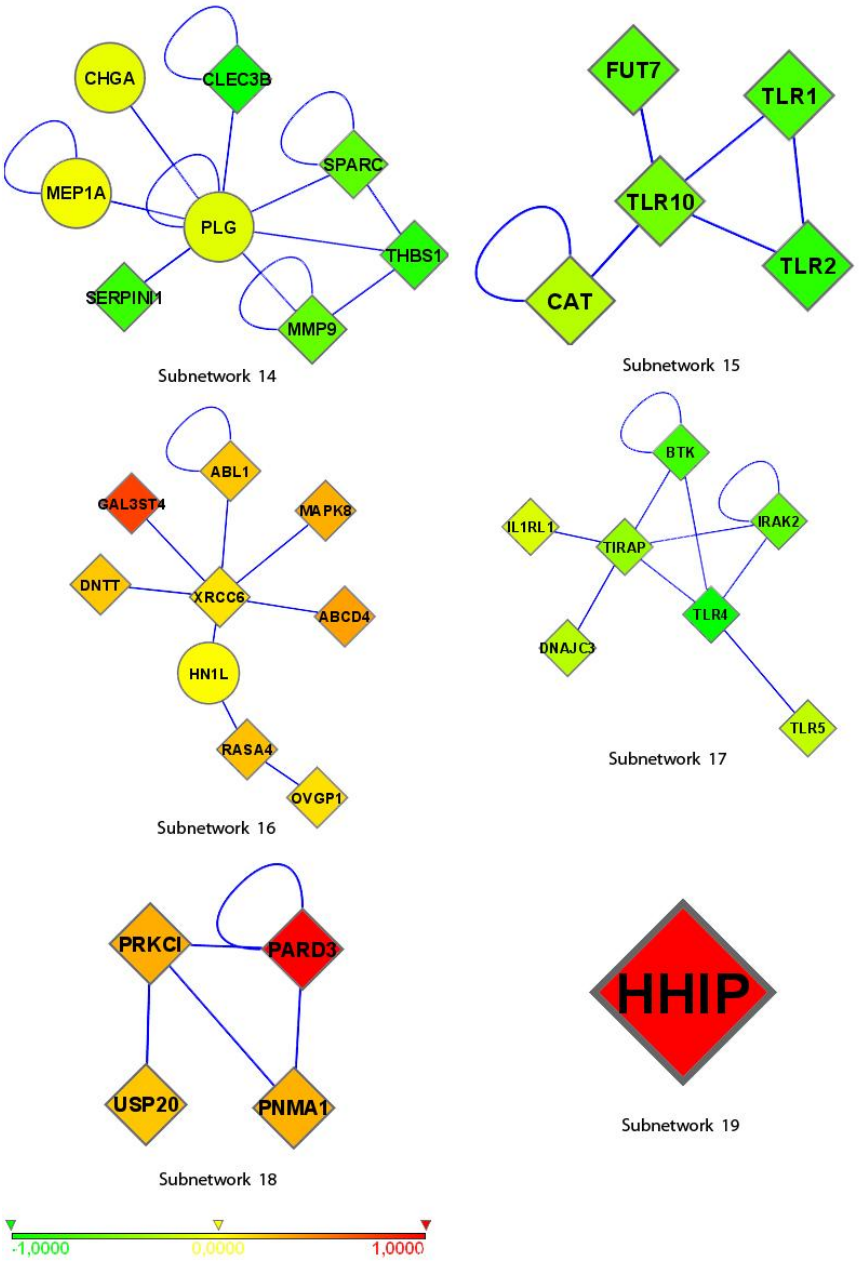

Figure 4-3: The most frequent subnetworks. Nodes represent proteins and edges represent interactions. The color of each node ranges in accordance with the change in expression of the corresponding gene for T-ALL versus healthy samples. The shape of each node shows whether its gene is significantly differentially expressed (diamond; $P<0.05$ from a two-tailed t-test) or not (circle).

4.2 Hiearchical Clustering Results

After filtering 20148 probes in microarray data with T-test and obtaining 100 and 200 most DE genes, hierarchical clustering was applied with complete linkage and Pearson correlation. The resulting clusters are shown in Figure 4-4 and Figure 4-5. Since there are too many samples and genes, the gene names are not visible on these figures. Please refer to Chapter-6, Supplementary Materials (Table 6-20, Table 6-21) to see the gene names.

In Figure 4-4 or cluster results of 100 most DE genes, 5 T-ALL samples and 1 healthy sample were misclassified. In other words, they are found in the wrong group. Moreover, there were 1 T-ALL and 1 healthy sample, which are misclassified in clusters of 200 most DE, Figure 4-5. This result is striking because one would expect that 100 most DE genes should classify the diseased and healthy samples more accurately than 200 most DE genes. But the number of genes and classification accuracy are not directly proportional.

A far more striking result is that 102 genes in our subnetworks and 100 most DE genes have only 2 genes in common (Table 4-9). Furthermore, subnetworks and 200 most DE genes have only 8 genes in common (Table 4-10). This result shows that 94 of our subnetwork genes are less differential than 200 most DE genes. There are also 6 non-DE genes in our subnetworks (in the remaining 94 genes). Therefore, it would not be wrong to expect much higher classification accuracies from 100 and 200 most DE genes than that of 102 subnetwork genes. But the classification accuracies of 100 and 200 most DE genes are not very different from those of subnetworks. Actually, 2 subnetworks (subnetwork 12 and 13) performed even higher classification accuracies than 100 and 200 most DE genes. Table 4-11 displays the classification accuracies of 100 and 200 most DE genes. Thus, we can safely conclude that our subnetworks (even they are less differential than 100 and 200 most DE genes and contain non-DE genes) can do the same or better job in distinguishing

diseased samples from healthy samples. Doing the same job with 10 genes, instead of 100 or 200 genes might be regarded as an accomplishment.

It is interesting that very well-known cancer genes such as TP53 and MAPK8 were not included in the 100/200 most DE genes but we were able to detect them by network-based approach.

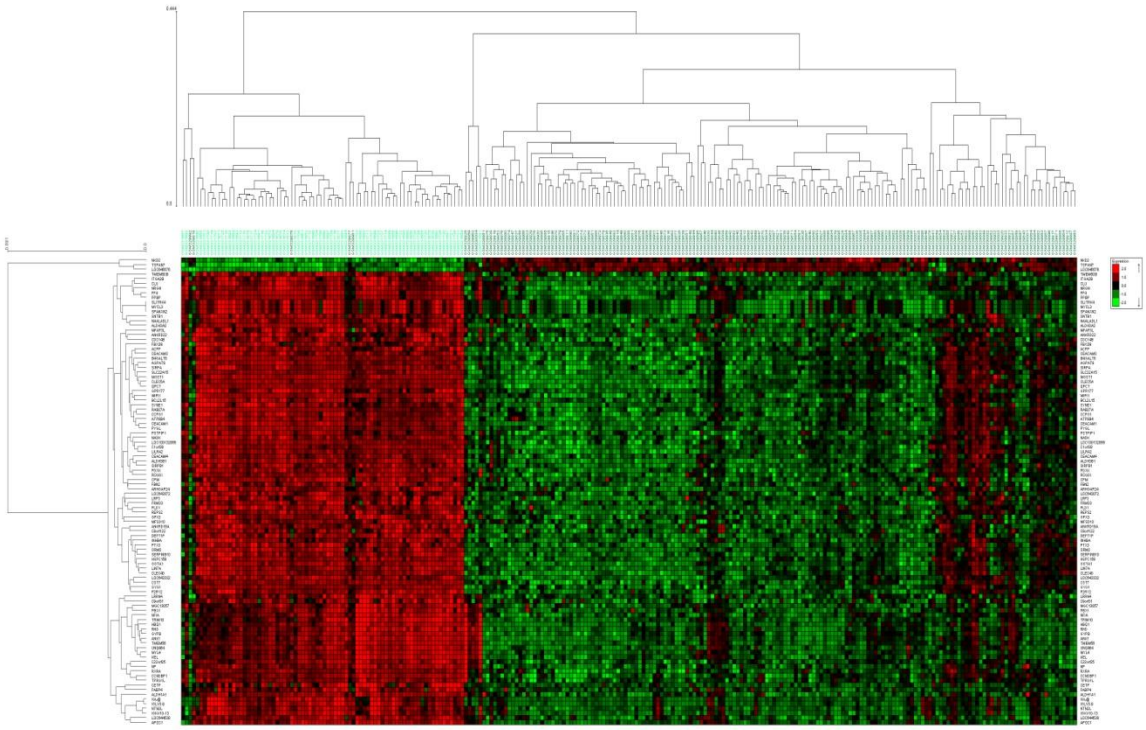

Figure 4-4: The heat-map shows the hierarchical clustering result of 100 most differentially expressed genes in T-ALL with respect to healthy individuals. Red and green spots represent up-regulated and down-regulated genes, respectively. Black spots denote equal expression. The columns labeled with light green belong to healthy individuals and the columns labeled with black are individuals with T-ALL.

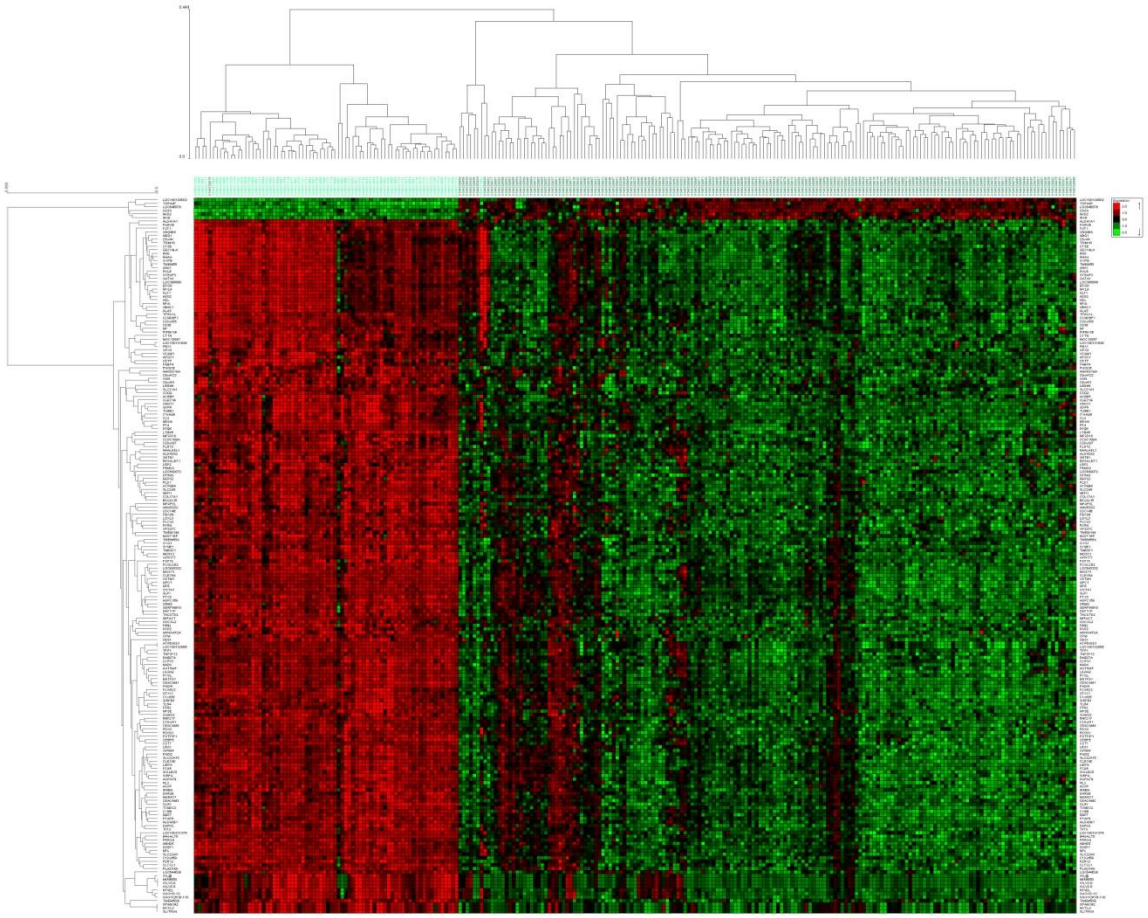

Figure 4-5: The heat-map shows the hierarchical clustering result of 200 most differentially expressed genes in T-ALL with respect to healthy individuals. Red and green spots represent up-regulated and down-regulated genes, respectively. Black spots denote equal expression. The columns labeled with light green belong to healthy individuals and the columns labeled with black are individuals with T-ALL.

Table 4-9: The genes common to subnetwork genes and 100 most differentially expressed genes.

Gene symbol	Gene title	GO:Function	GO:Process
CLU	clusterin	protein binding	anti-apoptosis, apoptosis, cell death, complement activation, classical pathway, endocrine pancreas development, innate immune response, lipid metabolic process, neurite morphogenesis, positive regulation of cell differentiation, positive regulation of cell proliferation, response to oxidative stress
ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	calcium ion binding, identical protein binding, protein binding, receptor activity	cell adhesion, cell adhesion, integrin-mediated signaling pathway

Table 4-10: The genes common to subnetwork genes and 200 most differentially expressed genes.

Gene symbol	Gene title	GO:Function	GO:Process
ADD2	adducin 2 (beta)	actin binding, calmodulin binding, metal ion binding	
CD36	CD36 molecule (thrombospondin receptor)	lipoprotein binding, low-density lipoprotein receptor activity, receptor activity	blood coagulation, cell adhesion, lipid metabolic process, lipoprotein transport, transport
CLU	clusterin	protein binding	anti-apoptosis, apoptosis, cell death, complement activation, classical pathway, endocrine pancreas development, innate immune response, lipid metabolic process, neurite morphogenesis, positive regulation of cell differentiation, positive regulation of cell proliferation, response to oxidative stress
GSN	gelsolin	actin binding, calcium ion	actin filament polymerization, actin

	(amyloidosis, Finnish type)	binding, protein binding	filament severing, barbed-end actin filament capping
ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	calcium ion binding, identical protein binding, protein binding, receptor activity	cell adhesion, cell adhesion, integrin-mediated signaling pathway
LTB4R	leukotriene B4 receptor	leukotriene B4 receptor activity, nucleotide binding, receptor activity	G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), cell motility, immune response, inflammatory response, muscle contraction, signal transduction
TLR4	toll-like receptor 4	lipopolysaccharide binding, protein binding, protein binding, transmembrane receptor activity, transmembrane receptor activity	I-kappaB kinase/NF-kappaB cascade, T-helper 1 type immune response, detection of fungus, inflammatory response, innate immune response, macrophage activation, negative regulation of osteoclast differentiation, positive regulation of interleukin-12 biosynthetic process, positive regulation of interleukin-12 biosynthetic process, positive regulation of interleukin-8 biosynthetic process, positive regulation of tumor necrosis factor biosynthetic process, signal transduction, signal transduction
TUBB1	tubulin, beta 1	GTP binding, GTPase activity, nucleotide binding, structural molecule activity	microtubule-based movement, protein polymerization

Table 4-11: The classification accuracies of 100 and 200 most differential genes between T-ALL and healthy samples.

Genes	J48	RBFNetwork
100mostDE	95	98
200mostDE	95	98

4.3 Classification accuracies of subnetworks

After finding subnetworks, we tested their classification accuracies with two different classifiers, namely, J48 and RBF-Network in WEKA. The prediction accuracy of each subnetwork was tested individually with 10-fold cross validation. As discussed, the patient samples were randomly divided into 2 groups, and for each sub-group, differential subnetworks were found. Classification accuracies of subnetworks were found by testing their original sub-group (the group for which the subnetworks were found), and by cross-testing the remaining sub-group (the other half of the patients). The cross-testing was applied to validate the prediction accuracies of modules also in different set of patient microarray data. All of the subnetworks achieved very high accuracies in prediction, higher than 90% (Table 4-12). There are some subnetworks that achieved 99% prediction accuracy (Figure 4-1, subnetwork4 and Figure 4-2, subnetwork12). These results also prove the success of the network-based classification approaches. Cross-comparisons between two independent halves of patient dataset revealed that subnetworks are good at distinguishing T-ALL patients from healthy individuals, regardless of the dataset in which they are found. In other words, they can classify patients in both independent datasets with similar accuracy.

Although integrating microarray data with network information is a promising way to identify functional biomarkers, the drawback of pathway-based classifiers is that most of the human genes have not been assigned to a definitive pathway yet [2]. As pathways become more complete, the classification performances of pathway-based approaches will increase [36].

Table 4-12: Classification accuracies of subnetworks found by two different classifiers in WEKA

subnetworks	J48	RBF Network	J48 (validation set)	RBF Network (validation set)
subnetwork1	93	97	95	96
subnetwork2	91	96	91	98
subnetwork3	93	97	96	98
subnetwork4	98.75	99	93	96
subnetwork5	94	97	92	98
subnetwork6	96	97	94	94
subnetwork7	96	97.5	96	96
subnetwork8	94	97	96	93.75
subnetwork9	93	98	93	94
subnetwork10	97	97	96	96.25
subnetwork11	91	98	93	97
subnetwork12	95	99	96	98
subnetwork13	98	98	95	99
subnetwork14	94	97.5	91	93
subnetwork15	93	97	92	93
subnetwork16	95	96	92.5	97
subnetwork17	89	97	90	94
subnetwork18	93	96	94	97
subnetwork19	96.27	96.89	96.25	97

Chapter 5

5. CONCLUSION

In conclusion, each subnetwork with high prediction accuracy provides a new suggestion for pathways and molecular mechanisms involved in the causes of T-ALL. All subnetworks serve as a biomarker which can be helpful in diagnosis, and in identifying potential drug targets for T-ALL, in the near feature. The accomplishment of network-based classification and subnetwork/pathway detection is in line with the idea that cancer is not result of solely one pathway, but instead, it is a "disease of pathways" [2, 63, 108]. Compared to only expression-based differential analysis which returns thousands of genes to be analyzed, network-based classification technique produces fewer genes with more meaningful and successful results. By network-based approach, we can identify the significant genes which are not differentially expressed but have a central role in interconnecting many differentially expressed genes/proteins. According to our results, we conclude that transcription factors, tyrosine kinases, and Zinc-ion binding proteins are the most important protein groups involved in production of T-ALL.

These important genes in our subnetworks may serve as an alternative to the traditional biomarkers used for the diagnosis of T-ALL. The aim of this study is also to help investigators to highlight potential disease gene candidates for further experimental validation. It is beyond the scope of this study to verify all genes that appear in subnetworks; experimental proof is vital. We recommend investigators with an interest in

a subnetwork/pathway to validate them with experimental techniques, like PCR techniques. However, the important point of this work is that a combination of bioinformatic methods, high throughput gene expression profiles and interactomics provide a promising way of identifying T-ALL specific modules and revealing transformation pathways involved in T-ALL.

Chapter 6

6. SUPPLEMENTARY MATERIALS

The genes in the subnetworks and their functions are given as tables 6-1 through 6-19 (for each individual subnetwork, there is one table).

Table 6-1: The genes in Subnetwork-1 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
MYO10	myosin X	ATP binding, actin binding, motor activity, nucleotide binding	signal transduction
TNNI2	troponin I type 2 (skeletal, fast)	Contributes to actin binding, protein binding, troponin T binding	positive regulation of transcription, skeletal muscle contraction
KCNN4	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4	calcium-activated potassium channel activity, calmodulin binding, ion channel activity, voltage-gated potassium channel activity	defense response, ion transport, positive regulation of protein secretion, potassium ion transport, saliva secretion
ADD2	adducin 2 (beta)	actin binding, calmodulin binding, metal ion binding	
CALM1	calmodulin 1 (phosphorylase kinase, delta)		

Table 6-2: The genes in Subnetwork-2 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
CCR5	chemokine (C-C motif) receptor 5	C-C chemokine receptor activity, C-C chemokine receptor activity, actin binding, coreceptor activity, phosphoinositide phospholipase C activity ,protein binding, receptor activity, rhodopsin-like receptor activity	G-protein coupled receptor protein signaling pathway, cell-cell signaling, cellular defense response, chemotaxis, elevation of cytosolic calcium ion concentration, immune response, inflammatory response, initiation of viral infection, signal transduction
CCL5	chemokine (C-C motif) ligand 5	chemoattractant activity, chemokine activity, signal transducer activity	cell adhesion, cell motility, cell-cell signaling, cellular calcium ion homeostasis, cellular defense response, chemotaxis, exocytosis, immune response, Inflammatory response, negative regulation of viral genome replication, response to oxidative stress, response to virus, signal transduction
PTK2	PTK2 protein tyrosine kinase 2	ATP binding, SH2 domain binding, non-membrane spanning protein tyrosine kinase activity, nucleotide binding, protein binding, protein tyrosine kinase activity, signal transducer activity, transferase activity	integrin-mediated signaling pathway, protein amino acid phosphorylation, signal complex assembly
ORM1	orosomucoid 1	protein binding	acute-phase response, inflammatory response
CCL14	chemokine (C-C motif) ligand 14	chemokine activity, protein binding, signal transducer activity	cellular calcium ion homeostasis, immune response, positive regulation of cell proliferation
ARL6IP5	ADP-ribosylation-like factor 6 interacting protein 5	protein binding	L-glutamate transport
ADRBK1	adrenergic	ATP binding, G-protein	cardiac muscle contraction, desensitization of

	, beta, receptor kinase 1	coupled receptor kinase activity, beta-adrenergic receptor kinase activity, nucleotide binding, protein kinase activity,protein serine/threonine kinase activity, receptor activity, signal transducer activity, transferase activity	G-protein coupled receptor protein signaling pathway, heart development, negative regulation of striated muscle contraction, negative regulation of the force of heart contraction by chemical signal, protein amino acid phosphorylation, regulation of the force of heart contraction, signal transduction, tachykinin signaling pathway

Table 6-3: The genes in Subnetwork-3 and their functions according to GO.

Gene Symbol	**Gene title**	**GO:Function**	**GO:Process**
CD36	CD36 molecule (thrombospondin receptor)	lipoprotein binding, low-density lipoprotein receptor activity, receptor activity	blood coagulation, cell adhesion, lipid metabolic process, lipoprotein transport, transport
ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	calcium ion binding, identical protein binding, protein binding, receptor activity	cell adhesion, integrin-mediated signaling pathway
COL5A3	collagen, type V, alpha 3	extracellular matrix structural constituent, heparin binding	cell adhesion ,cell-matrix adhesion, muscle development, organ morphogenesis, phosphate transport
P4HB	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide	isomerase activity, procollagen-proline 4-dioxygenase activity, protein binding, protein disulfide isomerase activity	cell redox homeostasis, peptidyl-proline hydroxylation to 4-hydroxy-L-proline
COL1A2	collagen, type I, alpha 2	extracellular matrix structural constituent, protein binding	phosphate transport, sensory perception of sound, skeletal development, transmembrane receptor protein tyrosine kinase signaling pathway

Table 6-4: The genes in Subnetwork-4 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
STRA13	stimulated by retinoic acid 13 homolog (mouse)	DNA binding	
SMAD4	SMAD family member 4	SMAD binding, protein homodimerization activity, sequence-specific DNA binding ,transcription activator activity, transcription factor activity	BMP signaling pathway, SMAD protein complex assembly, anterior/posterior pattern formation, kidney development, negative regulation of cell growth, negative regulation of cell proliferation, negative regulation of transcription, DNA-dependent, positive regulation of transcription from RNA polymerase II promoter, regulation of binding, regulation of transforming growth factor beta receptor signaling pathway, regulation of transforming growth factor-beta2 production, response to hypoxia, transcription, ureteric bud branching
UBE2I	ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)	HLH domain binding, ligase activity ,protein binding, small conjugating protein ligase activity	cell cycle, cell division, chromosome segregation, mitosis, regulation of protein metabolic process, ubiquitin cycle, ubiquitin-dependent protein catabolic process
TP53	tumor protein p53	ATP binding, DNA strand annealing activity, chromatin binding, copper ion binding, enzyme binding, metal ion binding, nuclease activity, protein N-terminus binding, protein heterodimerization activity, transcription factor activity, zinc ion binding	DNA damage response, apoptosis, signal transduction by p53 class mediator resulting in induction of apoptosis, ER overload response, RNA-protein covalent cross-linking, base-excision repair, caspase activation via cytochrome c, cell aging, cell cycle arrest, cell differentiation, cell proliferation, cellular response to glucose starvation, induction of apoptosis by intracellular signals, multicellular organismal development, negative regulation of cell cycle, negative regulation of cell growth, negative regulation of helicase activity, positive regulation of transcription from RNA polymerase II promoter, protein complex

			assembly, protein localization, regulation of mitochondrial membrane permeability, regulation of transcription, response to tumor cell
RANBP2	RAN binding protein 2	Ran GTPase binding, isomerase activity, metal ion binding, peptidyl-prolyl cis-trans isomerase activity, protein binding, zinc ion binding	intracellular protein transport across a membrane,I ntracellular transport, mRNA transport, protein folding, protein import into nucleus, ubiquitin cycle
ATF2	activating transcripti on factor 2	RNA polymerase II transcription factor activity, metal ion binding,p rotein dimerization activity, sequence-specific DNA binding, transcription coactivator activity, transcription factor activity, zinc ion binding	regulation of transcription
WT1	Wilms tumor 1	metal ion binding, transcription factor activity, zinc ion binding	negative regulation of cell cycle, negative regulation of transcription from RNA polymerase II promoter
SOX4	SRY (sex determinin g region Y)-box 4	transcription factor activity	regulation of transcription
TFAP2C	transcripti on factor AP-2 gamma (activating enhancer binding protein 2 gamma)	protein dimerization activity, transcription factor activity	cell differentiation, cell-cell signaling, regulation of transcription from RNA polymerase II promoter, transcription

Table 6-5: The genes in Subnetwork-5 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
GRK6	G protein-coupled receptor kinase 6	ATP binding, nucleotide binding, signal transducer activity, transferase activity	amino acid phosphorylation, regulation of G-protein coupled receptor protein signaling pathway, signal transduction
LTB4R	leukotriene B4 receptor	nucleotide binding, receptor activity	G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), cell motility, immune response, inflammatory response, muscle contraction, signal transduction
SNCA	synuclein, alpha (non A4 component of amyloid precursor)	identical protein binding	anti-apoptosis, central nervous system development, dopamine biosynthetic process, phospholipid metabolic process, positive regulation of neurotransmitter secretion, regulation of locomotion, regulation of long-term neuronal synaptic plasticity, response to drug, synaptic transmission, dopaminergic, synaptic vesicle transport
TBXA2R	thromboxane A2 receptor	rhodopsin-like receptor activity	G-protein coupled receptor protein signaling pathway, signal transduction

Table 6-6: The genes in Subnetwork-6 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
ATXN3	ataxin 3	hydrolase activity ,protein binding	cell death, nervous system development, nucleotide-excision repair, regulation of transcription, DNA-dependent, synaptic transmission, transcription
SALL2	sal-like 2 (Drosophila)	metal ion binding, transcription factor activity, zinc ion binding	regulation of transcription
TSPAN3	tetraspanin 3		
PRTFDC1	phosphoribosyl	hypoxanthine	nucleoside metabolic process,

	transferase domain containing 1	phosphoribosyltransferase activity	purine ribonucleoside salvage
EWSR1	Ewing sarcoma breakpoint region 1	RNA binding, calmodulin binding, metal ion binding, nucleotide binding, sequence-specific DNA binding, transcription factor activity, zinc ion binding	regulation of transcription, DNA-dependent, transcription

Table 6-7: The genes in Subnetwork-7 and their functions according to GO.

Gene Symbol	**Gene title**	**GO:Function**	**GO:Process**
PRTFDC1	phosphoribosyl transferase domain containing 1	hypoxanthine phosphoribosyltransferase activity	nucleoside metabolic process, purine ribonucleoside salvage
EWSR1	Ewing sarcoma breakpoint region 1	RNA binding, calmodulin binding, metal ion binding, nucleotide binding, sequence-specific DNA binding, transcription factor activity, zinc ion binding	regulation of transcription, DNA-dependent, transcription
SSBP2	single-stranded DNA binding protein 2	single-stranded DNA binding, transcription regulator activity	regulation of transcription
CEACAM5	carcinoembryonic antigen-related cell adhesion molecule 5		
YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	insulin-like growth factor receptor binding, protein domain specific binding, protein kinase C binding, protein kinase C inhibitor activity	Negative regulation of protein kinase activity, protein targeting, regulation of neuron differentiation, regulation of signal transduction, regulation of synaptic plasticity

Table 6-8: The genes in Subnetwork-8 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
DVL1	dishevelled, dsh homolog 1 (Drosophila)	protein binding, signal transducer activity	Wnt receptor signaling pathway, anatomical structure morphogenesis, heart development, intracellular signaling cascade, multicellular organismal development
VANGL1	vang-like 1 (van gogh, Drosophila)	protein binding	multicellular organismal development
SMAD4	SMAD family member 4	SMAD binding, protein homodimerization activity, sequence-specific DNA binding, transcription activator activity, transcription factor activity	BMP signaling pathway, SMAD protein complex assembly, anterior/posterior pattern formation, kidney development, negative regulation of cell growth, negative regulation of cell proliferation, negative regulation of transcription, DNA-dependent, positive regulation of transcription from RNA polymerase II promoter, regulation of binding, regulation of transforming growth factor beta receptor signaling pathway, regulation of transforming growth factor-beta2 production, response to hypoxia, transcription, ureteric bud branching
USP13	ubiquitin specific peptidase 13 (isopeptidase T-3)	cysteine-type endopeptidase activity, metal ion binding, ubiquitin thiolesterase activity, zinc ion binding	ubiquitin cycle, ubiquitin-dependent protein catabolic process
VANGL2	vang-like 2 (van gogh, Drosophila)	protein binding	apical protein localization, establishment of planar polarity, heart looping, multicellular organismal development, neural tube closure, sensory cilium biogenesis
CXXC4	CXXC finger 4	DNA binding, PDZ domain binding, metal ion binding, zinc ion binding	negative regulation of Wnt receptor signaling pathway

Table 6-9: The genes in Subnetwork-9 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
GSN	gelsolin (amyloidosis, Finnish type)	actin binding, calcium ion binding, protein binding	actin filament polymerization, actin filament severing, barbed-end actin filament capping
TPM1	tropomyosin 1 (alpha)	actin binding, structural constituent of cytoskeleton, structural constituent of muscle	cell motility, regulation of heart contraction, regulation of muscle contraction
LASP1	LIM and SH3 protein1	SH3/SH2 adaptor activity, actin binding,ion transmembrane transporter activity, metal ion binding, zinc ion binding	cortical cytoskeleton organization and biogenesis, ion transport
TGM2	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)	GTP binding, acyltransferase activity, calcium ion binding, protein domain specific binding, protein-glutamine gamma-glutamyltransferase activity, transferase activity	induction of apoptosis, blood vessel remodeling, elevation of cytosolic calcium ion concentration during G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), isopeptide cross-linking via N6-(L-isoglutamyl)-L-lysine, peptide cross-linking, positive regulation of I-kappaB kinase/NF-kappaB cascade, positive regulation of cell adhesion, positive regulation of inflammatory response, positive regulation of smooth muscle cell proliferation, protein homooligomerization
NCF1	neutrophil cytosolic factor 1, (chronic granulomatous disease, autosomal1)	GTP binding, GTPase activity, electron carrier activity, phosphoinositide binding, protein binding	cell communication, cellular defense response
ACTB	actin, beta	ATP binding, nucleotide binding, protein binding,	cell motility, sensory perception of sound

		structural constituent of cytoskeleton, structural molecule activity	
P2RX7	purinergic receptor P2X, ligand-gated ion channel, 7	ATP binding, ATP-gated cation channel activity, ion channel activity	ion transport, signal transduction

Table 6-10: The genes in Subnetwork-10 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
TUBB1	tubulin, beta 1	GTP binding, GTPase activity, nucleotide binding, structural molecule activity	microtubule-based movement, protein polymerization

Table 6-11: The genes in Subnetwork-11 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
PLSCR4	phospholipid scramblase 4	calcium ion binding, phospholipid scramblase activity	blood coagulation,phospholipid scrambling
SGTB	small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta	binding	
PLSCR1	phospholipid scramblase 1	calcium ion binding, phospholipid scramblase activity, protein binding	phospholipid scrambling,platelet activation,response to virus
GFI1B	growth factor independent 1B transcription repressor	DNA binding, metal ion binding, specific RNA polymerase II transcription factor activity, zinc ion binding	G1-specific transcription in mitotic cell cycle, cell proliferation, chromatin modification, multicellular organismal development, negative regulation of transcription from RNA polymerase II promoter, regulation of transcription

EFEMP2	EGF-containing fibulin-like extracellular matrix protein 2	calcium ion binding, extracellular matrix structural constituent, protein binding, transmembrane receptor activity	blood coagulation

Table 6-12: The genes in Subnetwork-12 and their functions according to GO.

Gene Symbol	**Gene title**	**GO:Function**	**GO:Process**
ARL4C	ADP-ribosylation factor-like 4C	GTP binding, GTPase activity, nucleotide binding	small GTPase mediated signal transduction
UBQLN4	ubiquilin 4	identical protein binding	biological_process,protein modification process
STMN1	stathmin 1/oncoprotein 18	protein binding, signal transducer activity, tubulin binding	cell differentiation, intracellular signaling cascade, microtubule depolymerization, mitotic spindle organization and biogenesis, multicellular organismal development, nervous system development
IGFBP6	insulin-like growth factor binding protein 6	insulin-like growth factor binding	negative regulation of cell proliferation, regulation of cell growth, signal transduction
COPB1	coatomer protein complex, subunit beta 1	protein binding, structural molecule activity	COPI coating of Golgi vesicle, intra-Golgi vesicle-mediated transport, intracellular protein transport, membrane organization and biogenesis, retrograde vesicle-mediated transport, Golgi to ER, vesicle-mediated transport
CD99	CD99 molecule	protein binding	cell adhesion

Table 6-13: The genes in Subnetwork-13 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
NCR1	natural cytotoxicity triggering receptor 1	protein binding,receptor activity	cellular defense response, natural killer cell activation ,regulation of natural killer cell mediated cytotoxicity, signal transduction
C9	complement component 9		caspase activation, complement activation-alternative pathway ,complement activation-classical pathway, cytolysis, hemolysis by symbiont of host red blood cells, induction of apoptosis
CD59	CD59 molecule, complement regulatory protein	protein binding	blood coagulation, cell surface receptor linked signal transduction
CLU	clusterin	protein binding	anti-apoptosis, apoptosis, cell death, complement activation, classical pathway, endocrine pancreas development ,innate immune response, lipid metabolic process, neurite morphogenesis, positive regulation of cell differentiation, positive regulation of cell proliferation, response to oxidative stress
LEPR	leptin receptor	hematopoietin/interferon-class (D200-domain) cytokine receptor activity, protein binding	cell surface receptor linked signal transduction, energy reserve metabolic process, multicellular organismal development
APP	amyloid beta (A4) precursor protein (peptidase nexin-II, Alzheimer disease)	acetylcholine receptor binding, copper ion binding, heparin binding, iron ion binding, serine-type endopeptidase inhibitor activity, zinc ion binding	Notch signaling pathway, apoptosis, cell adhesion, cellular copper ion homeostasis, endocytosis, neuromuscular process

Table 6-14: The genes in Subnetwork-14 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
MEP1A	meprin A, alpha (PABA peptide hydrolase)	astacin activity, metal ion binding, metallopeptidase activity, zinc ion binding	digestion, proteolysis
SERPINI1	serpin peptidase inhibitor, clade I (neuroserpin), member 1	serine-type endopeptidase inhibitor activity	central nervous system development, peripheral nervous system development
THBS1	thrombospondin 1	calcium ion binding, endopeptidase inhibitor activity, heparin binding, protein binding, signal transducer activity, structural molecule activity	cell adhesion, multicellular organismal development
MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	calcium ion binding, collagen binding, collagenase activity, gelatinase B activity, zinc ion binding	collagen catabolic process, extracellular matrix organization and biogenesis, macrophage differentiation, metabolic process, positive regulation of apoptosis, proteolysis, skeletal development
CHGA	chromogranin A (parathyroid secretory protein 1)	calcium ion binding, protein binding	regulation of blood pressure
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	calcium ion binding, collagen binding, copper ion binding	ossification, transmembrane receptor protein tyrosine kinase signaling pathway
CLEC3B	C-type lectin domain family 3, member B	sugar binding	skeletal development
PLG	plasminogen	apolipoprotein binding, calcium ion binding, peptidase activity, plasmin activity	blood coagulation, induction of apoptosis, negative regulation of angiogenesis, negative regulation of blood vessel endothelial cell migration, negative regulation of cell proliferation, negative regulation of fibrinolysis, proteolysis, tissue remodeling

Table 6-15: The genes in Subnetwork-15 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
CAT	catalase	NADP binding, catalase activity, heme binding, iron ion binding, protein homodimerization activity	UV protection, hydrogen peroxide catabolic process, negative regulation of apoptosis, oxidation reduction, response to reactive oxygen species
TLR10	toll-like receptor 10	protein binding,transmembrane receptor activity	I-kappaB kinase/NF-kappaB cascade,inflammatory response,innate immune response,signal transduction
TLR1	toll-like receptor 1	protein binding, transmembrane receptor activity	inflammatory response, innate immune response, macrophage activation, positive regulation of interleukin-6 biosynthetic process, positive regulation of tumor necrosis factor biosynthetic process, signal transduction
TLR2	toll-like receptor 2	Gram-positive bacterial binding, lipopolysaccharide receptor activity, peptidoglycan binding, protein binding, transmembrane receptor activity	induction of apoptosis, I-kappaB kinase/NF-kappaB cascade, MyD88-dependent toll-like receptor signaling pathway, cell surface pattern recognition receptor signaling pathway, chloramphenicol transport, inflammatory response, innate immune response, positive regulation of tumor necrosis factor production, response to molecule of fungal origin, signal transduction
FUT7	fucosyltransferase 7 (alpha (1,3) fucosyltransferase)	transferase activity, transferring glycosyl groups	L-fucose catabolic process, protein amino acid glycosylation

Table 6-16: The genes in Subnetwork-16 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
ABL1	c-abl oncogene 1, receptor tyrosine kinase	ATP binding,bDNA binding,bkinase activity,bmagnesium ion bindingb,manganese ion binding, membrane spanning protein tyrosine kinase activity, nucleotide binding, protein C-terminus binding, transferase activity	DNA damage response, signal transduction resulting in induction of apoptosis, S-phase-specific transcription in mitotic cell cycle, actin cytoskeleton organization and biogenesis, cell adhesion, mismatch repair, peptidyl-tyrosine phosphorylation, positive regulation of oxidoreductase activity, protein modification process, regulation of transcription
XRCC6	X-ray repair complementing defective repair in Chinese hamster cells 6 (Ku autoantigen, 70kDa)	ATP-dependent DNA helicase activity, DNA binding, double-stranded DNA binding, protein binding	DNA ligation, DNA repair, double-strand break repair via nonhomologous end joining, initiation of viral infection, positive regulation of transcription, DNA-dependent, provirus integration
GAL3ST 4	galactose-3-O-sulfotransferase 4	3'-phosphoadenosine 5'-phosphosulfate binding ,proteoglycan sulfotransferase activity, transferase activity	biosynthetic process, cell-cell signaling, glycoprotein metabolic process, oligosaccharide metabolic process, proteoglycan biosynthetic process, sulfur metabolic process
OVGP1	oviductal glycoprotein 1, 120kDa (mucin 9, oviductin)	catalytic activity, cation binding, chitinase activity	carbohydrate metabolic process, chitin catabolic process, female pregnancy, single fertilization
RASA4	RAS p21 protein activator 4	GTPase activator activity, metal ion binding, zinc ion binding	intracellular signaling cascade, regulation of small GTPase mediated signal transduction
HN1L	hematological and neurological expressed 1-like	0	0
DNTT	deoxynucleotidyltransferase, terminal	DNA binding, DNA-directed DNA polymerase activity, catalytic activity, magnesium ion binding,	DNA modification, DNA replication

		transferase activity	
ABCD4	ATP-binding cassette, sub-family D (ALD), member 4	ATP binding, ATPase activity, coupled to transmembrane movement of substances, nucleotide binding,transporter activity	transport
MAPK8	mitogen-activated protein kinase 8	ATP binding, JUN kinase activity, MAP kinase activity, nucleotide binding, protein binding, protein serine/threonine kinase activity, transferase activity	activation of pro-apoptotic gene products, JUN phosphorylation, cell motility, negative regulation of apoptosis, protein amino acid phosphorylation, response to UV, response to stress, signal transduction

Table 6-17: The genes in Subnetwork-17 and their functions according to GO.

Gene Symbol	**Gene title**	**GO:Function**	**GO:Process**
DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	heat shock protein binding, protein kinase inhibitor activity	response to unfolded protein, response to virus
IRAK2	interleukin-1 receptor-associated kinase 2	ATP binding, NF-kappaB-inducing kinase activity, identical protein binding ,protein kinase activity	I-kappaB kinase/NF-kappaB cascade, inflammatory response, amino acid phosphorylation, signal transduction
BTK	Bruton agammaglobulinemia tyrosine kinase	ATP binding, kinase activity, metal ion binding, non-membrane spanning protein tyrosine kinase activity, nucleotide binding, protein tyrosine kinase activity, transferase activity, zinc ion binding	calcium-mediated signaling, induction of apoptosis by extracellular signals, intracellular signaling cascade, mesoderm development, amino acid phosphorylation
TIRAP	toll-interleukin 1 receptor (TIR) domain containing adaptor protein	protein binding, transmembrane receptor activity	I-kappaB kinase/NF-kappaB cascade, inflammatory response, Innate immune response, signal transduction
IL1RL1	interleukin 1	interleukin-1 receptor activity,	innate immune response, signal

	receptor-like 1	transmembrane receptor activity	transduction
TLR5	toll-like receptor 5	transmembrane receptor activity	inflammatory response, innate immune response, signal transduction
TLR4	toll-like receptor 4	lipopolysaccharide binding,transmembrane receptor activity	I-kappaB kinase/NF-kappaB cascade, T-helper 1 type immune response, detection of fungus, inflammatory response, innate immune response ,macrophage activation, negative regulation of osteoclast differentiation, positive regulation of interleukin-12 biosynthetic process, positive regulation of tumor necrosis factor biosynthetic process, signal transduction

Table 6-18: The genes in Subnetwork-18 and their functions according to GO.

Gene Symbol	**Gene title**	**GO:Function**	**GO:Process**
USP20	ubiquitin specific peptidase 20	cysteine-type peptidase activity, metal ion binding, ubiquitin thiolesterase activity, zinc ion binding	protein deubiquitination, ubiquitin-dependent protein catabolic process
PRKCI	protein kinase C, iota	ATP binding, atypical protein kinase C activity, diacylglycerol binding, metal ion binding, nucleotide binding, phospholipid binding, transferase activity, zinc ion binding	Golgi vesicle budding ,actin filament organization, cell-cell junction assembly and maintenance, cytoskeleton organization and biogenesis, establishment and/or maintenance of epithelial cell polarity ,establishment of apical/basal cell polarity ,eye photoreceptor cell development, intracellular signaling cascade, membrane organization and biogenesis, amino acid phosphorylation, protein targeting to membrane, secretion, vesicle-mediated transport
PARD3	par-3	protein binding	activation of protein kinase C activity,

	partitioning defective 3 homolog (C. elegans)		asymmetric cell division, axonogenesis, cell cycle, establishment and/or maintenance of cell polarity, protein complex assembly
PNMA1	paraneoplastic antigen MA1	protein binding	central nervous system development, spermatogenesis

Table 6-19: The genes in Subnetwork-19 and their functions according to GO.

Gene Symbol	Gene title	GO:Function	GO:Process
HHIP	hedgehog interacting protein	catalytic activity, protein binding	dorsal/ventral pattern formation, lung development, negative regulation of smoothened signaling pathway, neuroblast proliferation, organ morphogenesis, regulation of fibroblast growth factor receptor signaling pathway, signal transduction

The following tables (Table 6-20, Table 6-21) represent the functions and titles of 100 and 200 most differentially expressed genes.

Table 6-20: The titles and functions of 100 most differentially expressed genes.

Gene symbol	Gene title	GO:Function	GO:Process
ACPP	acid phosphatase, prostate	acid phosphatase activity, hydrolase activity	
AGPAT9	1-acylglycerol-3-phosphate O-acyltransferase 9	1-acylglycerol-3-phosphate O-acyltransferase activity, acyltransferase activity, transferase activity	metabolic process, phospholipid biosynthetic process
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	Ras GTPase activator activity, aldehyde dehydrogenase (NAD) activity, aldehyde	aldehyde metabolic process, metabolic process, oxidation reduction

		dehydrogenase (NAD) activity, androgen binding, oxidoreductase activity, retinal dehydrogenase activity	
ALDH3A2	aldehyde dehydrogenase 3 family, member A2	3-chloroallyl aldehyde dehydrogenase activity, aldehyde dehydrogenase (NAD) activity, aldehyde dehydrogenase [NAD(P)+] activity, oxidoreductase activity	aldehyde metabolic process, central nervous system development, epidermis development, lipid metabolic process, oxidation reduction, peripheral nervous system development
ALDH3B1	aldehyde dehydrogenase 3 family, member B1	3-chloroallyl aldehyde dehydrogenase activity, aldehyde dehydrogenase [NAD(P)+] activity, oxidoreductase activity	alcohol metabolic process, aldehyde metabolic process, lipid metabolic process, metabolic process, oxidation reduction
ANK1	ankyrin 1, erythrocytic	cytoskeletal adaptor activity, enzyme binding, spectrin binding, structural constituent of cytoskeleton	cytoskeleton organization and biogenesis, exocytosis, maintenance of epithelial cell polarity, signal transduction
ANKRD18A	ankyrin repeat domain 18A		
ANKRD22	ankyrin repeat domain 22		
APOC1	apolipoprotein C-I	lipid transporter activity	cholesterol efflux, lipid metabolic process, lipid transport, lipoprotein metabolic process, phospholipid efflux
ARHGAP24	Rho GTPase activating protein 24	GTPase activator activity, protein binding	angiogenesis, cell differentiation, multicellular organismal development, signal transduction
ATP8B4	ATPase, class I, type 8B, member 4	ATP binding, ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism, hydrolase activity, magnesium ion binding, nucleotide binding, phospholipid-	phospholipid transport, transport

		translocating ATPase activity	
B4GALT5	UDP-Gal:betaGlcNAc beta 1, 4-galactosyltransferase, polypeptide 5	galactosyltransferase activity, manganese ion binding, metal ion binding, transferase activity, transferring glycosyl groups	carbohydrate metabolic process
BCL2L15	BCL2-like 15		apoptosis
C1orf38	chromosome 1 open reading frame 38		cell adhesion
C22orf25	chromosome 22 open reading frame 25		
C9orf122	chromosome 9 open reading frame 122		
CCNDBP1	cyclin D-type binding-protein 1	protein binding	cell cycle
CCPG1	cell cycle progression 1		cell cycle
CDC14B	CDC14 cell division cycle 14 homolog B (S. cerevisiae)	hydrolase activity, protein serine/threonine phosphatase activity, protein tyrosine phosphatase activity, protein tyrosine/serine/threonine phosphatase activity	cell division, protein amino acid dephosphorylation
CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	molecular_function	angiogenesis, biological_process, cell migration, homophilic cell adhesion, integrin-mediated signaling pathway
CEACAM3	carcinoembryonic antigen-related cell adhesion molecule 3		
CEACAM4	carcinoembryonic antigen-related cell adhesion molecule 4		
CETP	cholesteryl ester transfer protein, plasma	cholesterol binding, cholesterol transporter activity, phosphatidylcholine binding, phospholipid transporter activity, triglyceride binding	acylglycerol homeostasis, cholesterol homeostasis, cholesterol metabolic process, high-density lipoprotein particle remodeling, lipid metabolic process, low-density lipoprotein particle remodeling, phosphatidylcholine metabolic process, phospholipid homeostasis, phospholipid

			transport, reverse cholesterol transport, steroid metabolic process, triacylglycerol metabolic process, very-low-density lipoprotein particle remodeling
CLEC4D	C-type lectin domain family 4, member D	binding, sugar binding	immune response
CLEC5A	C-type lectin domain family 5, member A	binding, binding, sugar binding	cellular defense response, immune response, signal transduction
CLU	clusterin	protein binding	anti-apoptosis, apoptosis, cell death, complement activation, classical pathway, endocrine pancreas development, innate immune response, lipid metabolic process, neurite morphogenesis, positive regulation of cell differentiation, positive regulation of cell proliferation, response to oxidative stress
CPM	carboxypeptidase M	carboxypeptidase A activity, metal ion binding, metallopeptidase activity, zinc ion binding	anatomical structure morphogenesis, proteolysis
CST7	cystatin F (leukocystatin)	cysteine protease inhibitor activity	immune response
DEFT1P	defensin, theta 1 pseudogene		
FABP4	fatty acid binding protein 4, adipocyte	fatty acid binding, protein binding, transcription repressor activity, transporter activity	cholesterol homeostasis, cytokine production, negative regulation of protein kinase activity, negative regulation of transcription, positive regulation of inflammatory response, transport
FBN2	fibrillin 2 (congenital contractural arachnodactyly)	binding, calcium ion binding, extracellular matrix structural constituent	anatomical structure morphogenesis, embryonic limb morphogenesis
FBXO9	F-box protein 9	ubiquitin-protein ligase activity	protein ubiquitination

FRMD3	FERM domain containing 3	binding, cytoskeletal protein binding	
GGTA1	glycoprotein, alpha-galactosyltransferase 1	transferase activity, transferring hexosyl groups	carbohydrate metabolic process
GPR177	G protein-coupled receptor 177	receptor activity, signal transducer activity	positive regulation of I-kappaB kinase/NF-kappaB cascade
GPX3	glutathione peroxidase 3 (plasma)	glutathione binding, glutathione peroxidase activity, oxidoreductase activity, selenium binding, transcription factor binding	glutathione metabolic process, hydrogen peroxide catabolic process, oxidation reduction, protein homotetramerization, response to lipid hydroperoxide, response to oxidative stress
GYG1	glycogenin 1	glycogenin glucosyltransferase activity, glycogenin glucosyltransferase activity, protein binding, transferase activity, transferase activity, transferring hexosyl groups	carbohydrate biosynthetic process, glycogen biosynthetic process
GYPB	glycophorin B (MNS blood group)		
HBQ1	hemoglobin, theta 1	heme binding, iron ion binding, metal ion binding, oxygen binding, oxygen transporter activity	oxygen transport, transport
HSPC159	galectin-related protein	sugar binding	
IGKV1D-13	immunoglobulin kappa variable 1D-13		
INHBA	inhibin, beta A	activin inhibitor activity, cytokine activity, follistatin binding, growth factor activity, hormone activity, identical protein binding	G1/S transition of mitotic cell cycle, cell cycle arrest, cell differentiation, cell surface receptor linked signal transduction, cell-cell signaling, defense response, growth, hemoglobin biosynthetic process, induction of apoptosis, negative regulation of B cell differentiation,

			negative regulation of cell cycle, negative regulation of cell growth, negative regulation of follicle-stimulating hormone secretion, negative regulation of interferon-gamma biosynthetic process, negative regulation of macrophage differentiation, negative regulation of phosphorylation, nervous system development, ovarian follicle development, positive regulation of erythrocyte differentiation, positive regulation of follicle-stimulating hormone secretion, positive regulation of transcription from RNA polymerase II promoter, regulation of activin receptor signaling pathway, response to external stimulus, skeletal development
ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	calcium ion binding, identical protein binding, protein binding, receptor activity	cell adhesion, cell adhesion, integrin-mediated signaling pathway
KEL	Kell blood group, metallo-endopeptidase	endopeptidase activity, endothelin-converting enzyme activity, metal ion binding, metallopeptidase activity, neprilysin activity, protein binding, zinc ion binding	protein amino acid N-linked glycosylation, proteolysis, vasoconstriction
LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2	antigen binding, receptor activity, receptor activity	defense response, immune response, signal transduction
LIN7A	lin-7 homolog A (C. elegans)	protein binding	exocytosis, neurotransmitter secretion, protein complex

			assembly, protein transport, synaptic vesicle transport
LOC643072	hypothetical LOC643072		
LOC643332	similar to Nonsecretory ribonuclease precursor (Ribonuclease US) (Eosinophil-derived neurotoxin) (RNase UpI-2) (Ribonuclease 2) (RNase 2)	nucleic acid binding, pancreatic ribonuclease activity	
LOC644538	hypothetical protein LOC644538		
LOC646576	hypothetical LOC646576		
LRP3	low density lipoprotein receptor-related protein 3	receptor activity	receptor-mediated endocytosis
LRRN4	leucine rich repeat neuronal 4	protein binding	
MFAP3L	microfibrillar-associated protein 3-like		
MFSD10	major facilitator superfamily domain containing 10	tetracycline transporter activity, transporter activity	apoptosis, transport
MGC13057	hypothetical protein MGC13057		
MGST1	microsomal glutathione S-transferase 1	glutathione transferase activity, transferase activity	glutathione metabolic process
MYL4	myosin, light chain 4, alkali; atrial, embryonic	actin filament binding, actin monomer binding, calcium ion binding, motor activity, myosin II heavy chain binding, structural constituent of muscle	cardiac muscle contraction, muscle development, positive regulation of ATPase activity, regulation of the force of heart contraction, striated muscle contraction
NAALADL1	N-acetylated alpha-linked acidic dipeptidase-like 1	carboxypeptidase activity, dipeptidase activity, glutamate carboxypeptidase II activity, metal ion binding,	proteolysis

		metallopeptidase activity, zinc ion binding	
NADK	NAD kinase	NAD+ kinase activity, NAD+ kinase activity, NAD+ kinase activity, metal ion binding, protein binding, transferase activity	ATP metabolic process, NAD metabolic process, metabolic process, phosphorylation
NFIA	nuclear factor I/A	transcription factor activity, transcription factor activity, transcription factor binding	DNA replication, regulation of transcription, DNA-dependent, regulation of transcription, DNA-dependent, transcription, viral genome replication
NKD2	naked cuticle homolog 2 (Drosophila)	calcium ion binding, protein binding	Wnt receptor signaling pathway, exocytosis, transport
NP	nucleoside phosphorylase	purine-nucleoside phosphorylase activity, purine-nucleoside phosphorylase activity, purine-nucleoside phosphorylase activity	DNA modification, inosine catabolic process, nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, positive regulation of T cell proliferation, positive regulation of alpha-beta T cell differentiation, urate biosynthetic process
NRGN	neurogranin (protein kinase C substrate, RC3)	calmodulin binding	nervous system development, signal transduction
NTN2L	netrin 2-like (chicken)	molecular_function, receptor binding	axon guidance
ORM2	orosomucoid 2	binding	acute-phase response
P2RY2	purinergic receptor P2Y, G-protein coupled, 2	purinergic nucleotide receptor activity, G-protein coupled, receptor activity, rhodopsin-like receptor activity	G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), cellular ion homeostasis, signal transduction
PBX1	pre-B-cell leukemia homeobox 1	protein binding, sequence-specific DNA	C21-steroid hormone biosynthetic process, cell differentiation,

		binding, transcription factor activity	regulation of transcription, DNA-dependent, sex determination
PDXK	pyridoxal (pyridoxine, vitamin B6) kinase	ATP binding, lithium ion binding, magnesium ion binding, nucleotide binding, potassium ion binding, protein homodimerization activity, pyridoxal kinase activity, pyridoxal phosphate binding, sodium ion binding, transferase activity, zinc ion binding	cell proliferation, pyridoxal phosphate biosynthetic process, pyridoxine biosynthetic process, vitamin B6 metabolic process
PF4	platelet factor 4 (chemokine (C-X-C motif) ligand 4)	chemokine activity, heparin binding	cytokine and chemokine mediated signaling pathway, immune response, leukocyte chemotaxis, negative regulation of angiogenesis, negative regulation of megakaryocyte differentiation, platelet activation
PLD1	phospholipase D1, phosphatidylcholine-specific	hydrolase activity, phosphoinositide binding, phospholipase D activity, protein binding	Ras protein signal transduction, cell communication, chemotaxis, lipid catabolic process, metabolic process
PPBP	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	chemokine activity, glucose transmembrane transporter activity, growth factor activity	cell proliferation, chemotaxis, defense response to bacterium, glucose transport, immune response
PSTPIP1	proline-serine-threonine phosphatase interacting protein 1	catalytic activity, protein binding	cell adhesion, signal transduction
PTX3	pentraxin-related gene, rapidly induced by IL-1 beta	zymosan binding	inflammatory response, opsonization, positive regulation of nitric oxide biosynthetic process, positive regulation of phagocytosis, response to yeast
PYGL	phosphorylase, glycogen; liver (Hers	AMP binding, ATP binding, bile acid binding,	carbohydrate metabolic process, glucose homeostasis, glycogen

	disease, glycogen storage disease type VI)	drug binding, glucose binding, glycogen phosphorylase activity, nucleotide binding, phosphorylase activity, protein homodimerization activity, purine binding, pyridoxal phosphate binding, transferase activity, transferring glycosyl groups	metabolic process
QPCT	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	acyltransferase activity, glutaminyl-peptide cyclotransferase activity, metal ion binding, peptidase activity, transferase activity, zinc ion binding	protein modification process, proteolysis
RAB27A	RAB27A, member RAS oncogene family	GTP binding, GTPase activity, nucleotide binding	protein transport, small GTPase mediated signal transduction
REPS2	RALBP1 associated Eps domain containing 2	calcium ion binding, calcium ion binding, protein binding	epidermal growth factor receptor signaling pathway, protein complex assembly
RHD	Rh blood group, D antigen	molecular_function	biological_process
ROGDI	rogdi homolog (Drosophila)		
RXRA	retinoid X receptor, alpha	ligand-regulated transcription factor activity, metal ion binding, molecular_function, protein binding, retinoid-X receptor activity, sequence-specific DNA binding, steroid binding, steroid hormone receptor activity, tRNA-pseudouridine synthase	biological_process, positive regulation of transcription from RNA polymerase II promoter, regulation of transcription, DNA-dependent, signal transduction, transcription, vitamin metabolic process

		activity, transcription coactivator activity, transcription factor activity, zinc ion binding	
SERPINB10	serpin peptidase inhibitor, clade B (ovalbumin), member 10	protein binding, serine-type endopeptidase inhibitor activity	
SIRPA	signal-regulatory protein alpha		cell adhesion
SIRPB1	signal-regulatory protein beta 1		cell surface receptor linked signal transduction
SLC22A15	solute carrier family 22, member 15	transporter activity	transport
SNTB1	syntrophin, beta 1 (dystrophin-associated protein A1, 59kDa, basic component 1)	actin binding, calcium ion binding, calmodulin binding	muscle contraction
SYNE1	spectrin repeat containing, nuclear envelope 1	actin binding, actin binding, lamin binding, structural molecule activity	Golgi organization and biogenesis, muscle cell differentiation, nuclear organization and biogenesis
TMEM50B	transmembrane protein 50B	molecular_function	biological_process
TMEM56	transmembrane protein 56		
TPRG1L	tumor protein p63 regulated 1-like		
TRIM10	tripartite motif-containing 10	metal ion binding, protein binding, zinc ion binding, zinc ion binding	hemopoiesis
TSPAN7	tetraspanin 7		protein amino acid N-linked glycosylation
UNQ464	GLPG464		
WIPI1	WD repeat domain, phosphoinositide interacting 1	androgen receptor binding, estrogen receptor binding, phosphatidylinositol 3-phosphate binding, receptor binding	autophagy, vesicle targeting, trans-Golgi to endosome

Table 6-21: The titles and functions of 200 most differentially expressed genes.

Gene symbol	Gene title	GO:Function	GO:Process
ABHD5	abhydrolase domain containing 5	aminopeptidase activity, hydrolase activity, protein binding	lipid metabolic process, proteolysis, sensory perception of sound
ACPP	acid phosphatase, prostate	acid phosphatase activity, hydrolase activity	
ACRBP	acrosin binding protein	molecular_function	biological_process
ADD2	adducin 2 (beta)	actin binding, calmodulin binding, metal ion binding	
AGPAT9	1-acylglycerol-3-phosphate O-acyltransferase 9	1-acylglycerol-3-phosphate O-acyltransferase activity, acyltransferase activity, transferase activity	metabolic process, phospholipid biosynthetic process
AGTRAP	angiotensin II receptor-associated protein		
ALAD	aminolevulinate, delta-, dehydratase	lyase activity, metal ion binding, porphobilinogen synthase activity, porphobilinogen synthase activity, zinc ion binding	heme biosynthetic process, metabolic process, porphyrin biosynthetic process
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	Ras GTPase activator activity, aldehyde dehydrogenase (NAD) activity, aldehyde dehydrogenase (NAD) activity, androgen binding, oxidoreductase activity, retinal dehydrogenase activity	aldehyde metabolic process, metabolic process, oxidation reduction
ALDH3A2	aldehyde dehydrogenase 3 family, member A2	3-chloroallyl aldehyde dehydrogenase activity, aldehyde dehydrogenase (NAD) activity, aldehyde dehydrogenase [NAD(P)+] activity, oxidoreductase activity	aldehyde metabolic process, central nervous system development, epidermis development, lipid metabolic process, oxidation reduction, peripheral nervous system development
ALDH3B1	aldehyde	3-chloroallyl aldehyde	alcohol metabolic process,

	dehydrogenase 3 family, member B1	dehydrogenase activity, aldehyde dehydrogenase [NAD(P)+] activity, oxidoreductase activity	aldehyde metabolic process, lipid metabolic process, metabolic process, oxidation reduction
ANK1	ankyrin 1, erythrocytic	cytoskeletal adaptor activity, enzyme binding, spectrin binding, structural constituent of cytoskeleton	cytoskeleton organization and biogenesis, exocytosis, maintenance of epithelial cell polarity, signal transduction
ANKRD18A	ankyrin repeat domain 18A		
ANKRD22	ankyrin repeat domain 22		
APOC1	apolipoprotein C-I	lipid transporter activity	cholesterol efflux, lipid metabolic process, lipid transport, lipoprotein metabolic process, phospholipid efflux
ARHGAP24	Rho GTPase activating protein 24	GTPase activator activity, protein binding	angiogenesis, cell differentiation, multicellular organismal development, signal transduction
ATP8B4	ATPase, class I, type 8B, member 4	ATP binding, ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism, hydrolase activity, magnesium ion binding, nucleotide binding, phospholipid-translocating ATPase activity	phospholipid transport, transport
B3GALNT1	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	UDP-galactose:beta-N-acetylglucosamine beta-1,3-galactosyltransferase activity, galactosylgalactosylglucosylceramide beta-D-acetylgalactosaminyltransferase activity, galactosyltransferase	oligosaccharide biosynthetic process, protein amino acid glycosylation

		activity, magnesium ion binding, transferase activity, transferring glycosyl groups	
B4GALT5	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5	galactosyltransferase activity, manganese ion binding, metal ion binding, transferase activity, transferring glycosyl groups	carbohydrate metabolic process
BCL2L15	BCL2-like 15		apoptosis
C10orf11	chromosome 10 open reading frame 11	protein binding	
C12orf59	chromosome 12 open reading frame 59		
C1orf38	chromosome 1 open reading frame 38		cell adhesion
C20orf27	chromosome 20 open reading frame 27		
C22orf25	chromosome 22 open reading frame 25		
C5orf4	chromosome 5 open reading frame 4	catalytic activity	metabolic process
C9orf122	chromosome 9 open reading frame 122		
CCDC109A	coiled-coil domain containing 109A		
CCNDBP1	cyclin D-type binding-protein 1	protein binding	cell cycle
CCPG1	cell cycle progression 1		cell cycle
CD36	CD36 molecule (thrombospondin receptor)	lipoprotein binding, low-density lipoprotein receptor activity, receptor activity	blood coagulation, cell adhesion, lipid metabolic process, lipoprotein transport, transport
CD3G	CD3g molecule, gamma (CD3-TCR complex)	T cell receptor binding, protein heterodimerization activity, receptor signaling complex scaffold activity, transmembrane receptor activity	T cell activation, cell surface receptor linked signal transduction, establishment and/or maintenance of cell polarity, protein complex assembly, protein transport, regulation of apoptosis
CDC14B	CDC14 cell division cycle 14 homolog B (S. cerevisiae)	hydrolase activity, protein serine/threonine phosphatase activity,	cell division, protein amino acid dephosphorylation

		protein tyrosine phosphatase activity, protein tyrosine/serine/threonine phosphatase activity	
CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	molecular_function	angiogenesis, biological_process, cell migration, homophilic cell adhesion, integrin-mediated signaling pathway
CEACAM3	carcinoembryonic antigen-related cell adhesion molecule 3		
CEACAM4	carcinoembryonic antigen-related cell adhesion molecule 4		
CEBPE	CCAAT/enhancer binding protein (C/EBP), epsilon	protein dimerization activity, sequence-specific DNA binding, transcription factor activity	cytokine biosynthetic process, defense response to bacterium, macrophage differentiation, phagocytosis, regulation of transcription, DNA-dependent, transcription
CES1	carboxylesterase 1 (monocyte/macrophage serine esterase 1)	carboxylesterase activity, carboxylesterase activity, hydrolase activity	metabolic process, response to toxin
CETP	cholesteryl ester transfer protein, plasma	cholesterol binding, cholesterol transporter activity, phosphatidylcholine binding, phospholipid transporter activity, triglyceride binding	acylglycerol homeostasis, cholesterol homeostasis, cholesterol metabolic process, high-density lipoprotein particle remodeling, lipid metabolic process, low-density lipoprotein particle remodeling, phosphatidylcholine metabolic process, phospholipid homeostasis, phospholipid transport, reverse cholesterol transport, steroid metabolic process, triacylglycerol metabolic process, very-low-density lipoprotein particle remodeling
CLEC1B	C-type lectin domain	binding, sugar binding,	cell surface receptor linked

	family 1, member B	transmembrane receptor activity	signal transduction, defense response
CLEC4D	C-type lectin domain family 4, member D	binding, sugar binding	immune response
CLEC5A	C-type lectin domain family 5, member A	binding, binding, sugar binding	cellular defense response, immune response, signal transduction
CLTCL1	clathrin, heavy chain-like 1	protein binding, signal transducer activity, structural molecule activity	anatomical structure morphogenesis, intracellular protein transport, receptor-mediated endocytosis, vesicle-mediated transport
CLU	clusterin	protein binding	anti-apoptosis, apoptosis, cell death, complement activation, classical pathway, endocrine pancreas development, innate immune response, lipid metabolic process, neurite morphogenesis, positive regulation of cell differentiation, positive regulation of cell proliferation, response to oxidative stress
COL17A1	collagen, type XVII, alpha 1	structural molecule activity	cell-matrix adhesion, epidermis development, phosphate transport
COQ2	coenzyme Q2 homolog, prenyltransferase (yeast)	prenyltransferase activity, transferase activity	biosynthetic process, glycerol metabolic process, isoprenoid biosynthetic process, ubiquinone biosynthetic process
CPM	carboxypeptidase M	carboxypeptidase A activity, metal ion binding, metallopeptidase activity, zinc ion binding	anatomical structure morphogenesis, proteolysis
CST7	cystatin F (leukocystatin)	cysteine protease inhibitor activity	immune response
CTSE	cathepsin E	cathepsin E activity, pepsin A activity, peptidase activity	antigen processing and presentation of exogenous peptide antigen via MHC class

			II, digestion, proteolysis
CTTN	cortactin	protein binding	
CYBB	cytochrome b-245, beta polypeptide (chronic granulomatous disease)	FAD binding, electron carrier activity, iron ion binding, metal ion binding, voltage-gated ion channel activity	inflammatory response, ion transport, oxidation reduction
DAPK2	death-associated protein kinase 2	ATP binding, ATP binding, calmodulin binding, calmodulin-dependent protein kinase activity, identical protein binding, nucleotide binding, protein kinase activity, transferase activity	apoptosis, induction of apoptosis, protein amino acid phosphorylation, protein amino acid phosphorylation, protein kinase cascade
DDEF1	development and differentiation enhancing factor 1	ARF GTPase activator activity, metal ion binding, protein binding, zinc ion binding	regulation of ARF GTPase activity
DEFT1P	defensin, theta 1 pseudogene		
DHRS9	dehydrogenase/reductase (SDR family) member 9	3-alpha(17-beta)-hydroxysteroid dehydrogenase (NAD+) activity, 3-alpha(17-beta)-hydroxysteroid dehydrogenase (NAD+) activity, alcohol dehydrogenase activity, alcohol dehydrogenase activity, binding, oxidoreductase activity, racemase and epimerase activity, retinol dehydrogenase activity, retinol dehydrogenase activity	9-cis-retinoic acid biosynthetic process, 9-cis-retinoic acid biosynthetic process, androgen metabolic process, androgen metabolic process, epithelial cell differentiation, oxidation reduction, progesterone metabolic process, retinol metabolic process
ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1	calcium ion binding, hydrolase activity, magnesium ion binding	blood coagulation, cell adhesion

EPOR	erythropoietin receptor	erythropoietin receptor activity, guanyl-nucleotide exchange factor activity, protein binding	brain development, elevation of cytosolic calcium ion concentration, heart development, signal transduction, signal transduction, small GTPase mediated signal transduction
FABP4	fatty acid binding protein 4, adipocyte	fatty acid binding, protein binding, transcription repressor activity, transporter activity	cholesterol homeostasis, cytokine production, negative regulation of protein kinase activity, negative regulation of transcription, positive regulation of inflammatory response, transport
FBN2	fibrillin 2 (congenital contractural arachnodactyly)	binding, calcium ion binding, extracellular matrix structural constituent	anatomical structure morphogenesis, embryonic limb morphogenesis
FBXO9	F-box protein 9	ubiquitin-protein ligase activity	protein ubiquitination
FCAR	Fc fragment of IgA, receptor for	IgA binding, receptor activity	immune response
FCGR2C	Fc fragment of IgG, low affinity IIc, receptor for (CD32)	IgG binding, protein binding, receptor activity, transmembrane receptor activity	immune response
FGF13	fibroblast growth factor 13	growth factor activity, protein binding, protein kinase activator activity	MAPKKK cascade, cell-cell signaling, nervous system development, signal transduction
FLOT2	flotillin 2	protein binding	cell adhesion, epidermis development
FPR2	formyl peptide receptor 2	N-formyl peptide receptor activity, receptor activity, rhodopsin-like receptor activity	G-protein coupled receptor protein signaling pathway, cell adhesion, cell motility, chemotaxis, inflammatory response, signal transduction
FRMD3	FERM domain containing 3	binding, cytoskeletal protein binding	
FUT1	fucosyltransferase 1	fucosyltransferase activity,	L-fucose catabolic process,

	(galactoside 2-alpha-L-fucosyltransferase, H blood group)	galactoside 2-alpha-L-fucosyltransferase activity, transferase activity, transferring glycosyl groups	carbohydrate metabolic process, protein amino acid glycosylation
GATA1	GATA binding protein 1 (globin transcription factor 1)	metal ion binding, sequence-specific DNA binding, transcription factor activity, zinc ion binding	multicellular organismal development, regulation of transcription from RNA polymerase II promoter, transcription
GGTA1	glycoprotein, alpha-galactosyltransferase 1	transferase activity, transferring hexosyl groups	carbohydrate metabolic process
GNG11	guanine nucleotide binding protein (G protein), gamma 11	GTPase activity, signal transducer activity	G-protein coupled receptor protein signaling pathway, hormone-mediated signaling, signal transduction
GPR177	G protein-coupled receptor 177	receptor activity, signal transducer activity	positive regulation of I-kappaB kinase/NF-kappaB cascade
GPR84	G protein-coupled receptor 84	receptor activity, rhodopsin-like receptor activity	G-protein coupled receptor protein signaling pathway, biological_process, signal transduction
GPX3	glutathione peroxidase 3 (plasma)	glutathione binding, glutathione peroxidase activity, oxidoreductase activity, selenium binding, transcription factor binding	glutathione metabolic process, hydrogen peroxide catabolic process, oxidation reduction, protein homotetramerization, response to lipid hydroperoxide, response to oxidative stress
GSN	gelsolin (amyloidosis, Finnish type)	actin binding, calcium ion binding, protein binding	actin filament polymerization, actin filament severing, barbed-end actin filament capping
GYG1	glycogenin 1	glycogenin glucosyltransferase activity, glycogenin glucosyltransferase activity, protein binding,	carbohydrate biosynthetic process, glycogen biosynthetic process

		transferase activity, transferase activity, transferring hexosyl groups	
GYPB	glycophorin B (MNS blood group)		
HBQ1	hemoglobin, theta 1	heme binding, iron ion binding, metal ion binding, oxygen binding, oxygen transporter activity	oxygen transport, transport
HLX	H2.0-like homeobox	sequence-specific DNA binding, sequence-specific DNA binding, transcription factor activity	cell differentiation, multicellular organismal development, regulation of transcription, DNA-dependent
HPR	haptoglobin-related protein	catalytic activity, hemoglobin binding, serine-type endopeptidase activity	proteolysis
HPSE	heparanase	beta-glucuronidase activity, calcium ion binding, hydrolase activity, magnesium ion binding	proteoglycan metabolic process
HSPC159	galectin-related protein	sugar binding	
HTRA3	HtrA serine peptidase 3	insulin-like growth factor binding, peptidase activity, protein binding, serine-type endopeptidase activity	proteolysis, regulation of cell growth
IGKV1D-13	immunoglobulin kappa variable 1D-13		
IGKV1OR15-118	immunoglobulin kappa variable 1/OR15-118		
IMPACT	Impact homolog (mouse)	molecular_function	biological_process, regulation of translation
INHBA	inhibin, beta A	activin inhibitor activity, cytokine activity, follistatin binding, growth factor activity, hormone activity, identical protein binding	G1/S transition of mitotic cell cycle, cell cycle arrest, cell differentiation, cell surface receptor linked signal transduction, cell-cell signaling, defense response,

			growth, hemoglobin biosynthetic process, induction of apoptosis, negative regulation of B cell differentiation, negative regulation of cell cycle, negative regulation of cell growth, negative regulation of follicle-stimulating hormone secretion, negative regulation of interferon-gamma biosynthetic process, negative regulation of macrophage differentiation, negative regulation of phosphorylation, nervous system development, ovarian follicle development, positive regulation of erythrocyte differentiation, positive regulation of follicle-stimulating hormone secretion, positive regulation of transcription from RNA polymerase II promoter, regulation of activin receptor signaling pathway, response to external stimulus, skeletal development
IRX5	iroquois homeobox 5	sequence-specific DNA binding, transcription factor activity	cell development, regulation of transcription, DNA-dependent
ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	calcium ion binding, identical protein binding, protein binding, receptor activity	cell adhesion, cell adhesion, integrin-mediated signaling pathway
KEL	Kell blood group, metallo-endopeptidase	endopeptidase activity, endothelin-converting enzyme activity, metal ion binding, metallopeptidase	protein amino acid N-linked glycosylation, proteolysis, vasoconstriction

		activity, neprilysin activity, protein binding, zinc ion binding	
KLF1	Kruppel-like factor 1 (erythroid)	metal ion binding, transcription factor activity, zinc ion binding	chromatin remodeling, erythrocyte maturation, regulation of transcription, DNA-dependent, transcription
LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2	antigen binding, receptor activity, receptor activity	defense response, immune response, signal transduction
LIN7A	lin-7 homolog A (C. elegans)	protein binding	exocytosis, neurotransmitter secretion, protein complex assembly, protein transport, synaptic vesicle transport
LOC388588	hypothetical LOC388588		
LOC643072	hypothetical LOC643072		
LOC643332	similar to Nonsecretory ribonuclease precursor (Ribonuclease US) (Eosinophil-derived neurotoxin) (RNase UpI-2) (Ribonuclease 2) (RNase 2)	nucleic acid binding, pancreatic ribonuclease activity	
LOC644538	hypothetical protein LOC644538		
LOC646576	hypothetical LOC646576		
LOXL3	lysyl oxidase-like 3	copper ion binding, copper ion binding, metal ion binding, oxidoreductase activity, protein-lysine 6-oxidase activity, protein-lysine 6-oxidase activity, scavenger receptor activity	biological_process, oxidation reduction
LRG1	leucine-rich alpha-2-glycoprotein 1	molecular_function, protein binding	biological_process

LRP3	low density lipoprotein receptor-related protein 3	receptor activity	receptor-mediated endocytosis
LRRN4	leucine rich repeat neuronal 4	protein binding	
LTB4R	leukotriene B4 receptor	leukotriene B4 receptor activity, nucleotide binding, receptor activity	G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), cell motility, immune response, inflammatory response, muscle contraction, signal transduction
MANSC1	MANSC domain containing 1	chitinase activity	
MFAP3L	microfibrillar-associated protein 3-like		
MFAP3L	microfibrillar-associated protein 3-like		
MFSD10	major facilitator superfamily domain containing 10	tetracycline transporter activity, transporter activity	apoptosis, transport
MGC13057	hypothetical protein MGC13057		
MGST1	microsomal glutathione S-transferase 1	glutathione transferase activity, transferase activity	glutathione metabolic process
MOSC1	MOCO sulphurase C-terminal domain containing 1	molybdenum ion binding, oxidoreductase activity, pyridoxal phosphate binding	oxidation reduction
MYL4	myosin, light chain 4, alkali; atrial, embryonic	actin filament binding, actin monomer binding, calcium ion binding, motor activity, myosin II heavy chain binding, structural constituent of muscle	cardiac muscle contraction, muscle development, positive regulation of ATPase activity, regulation of the force of heart contraction, striated muscle contraction
NAALADL1	N-acetylated alpha-linked acidic dipeptidase-like 1	carboxypeptidase activity, dipeptidase activity, glutamate carboxypeptidase II activity, metal ion binding,	proteolysis

		metallopeptidase activity, zinc ion binding	
NADK	NAD kinase	NAD+ kinase activity, NAD+ kinase activity, NAD+ kinase activity, metal ion binding, protein binding, transferase activity	ATP metabolic process, NAD metabolic process, metabolic process, phosphorylation
NFIA	nuclear factor I/A	transcription factor activity, transcription factor activity, transcription factor binding	DNA replication, regulation of transcription, DNA-dependent, regulation of transcription, DNA-dependent, transcription, viral genome replication
NKD2	naked cuticle homolog 2 (Drosophila)	calcium ion binding, protein binding	Wnt receptor signaling pathway, exocytosis, transport
NP	nucleoside phosphorylase	purine-nucleoside phosphorylase activity, purine-nucleoside phosphorylase activity, purine-nucleoside phosphorylase activity	DNA modification, inosine catabolic process, nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, positive regulation of T cell proliferation, positive regulation of alpha-beta T cell differentiation, urate biosynthetic process
NPL	N-acetylneuraminate pyruvate lyase (dihydrodipicolinate synthase)	N-acetylneuraminate lyase activity, lyase activity	carbohydrate metabolic process, metabolic process
NRGN	neurogranin (protein kinase C substrate, RC3)	calmodulin binding	nervous system development, signal transduction
NTN2L	netrin 2-like (chicken)	molecular_function, receptor binding	axon guidance
NUDT16P	nudix (nucleoside diphosphate linked moiety X)-type motif 16 pseudogene		

OLR1	oxidized low density lipoprotein (lectin-like) receptor 1	low-density lipoprotein receptor activity, protein binding, receptor activity, sugar binding	blood circulation, cell adhesion, immune response, inflammatory response, proteolysis
ORM2	orosomucoid 2	binding	acute-phase response
P2RY2	purinergic receptor P2Y, G-protein coupled, 2	purinergic nucleotide receptor activity, G-protein coupled, receptor activity, rhodopsin-like receptor activity	G-protein signaling, coupled to IP3 second messenger (phospholipase C activating), cellular ion homeostasis, signal transduction
PADI2	peptidyl arginine deiminase, type II	calcium ion binding, hydrolase activity, protein-arginine deiminase activity	peptidyl-citrulline biosynthetic process from peptidyl-arginine, protein modification process
PADI4	peptidyl arginine deiminase, type IV	calcium ion binding, hydrolase activity, protein-arginine deiminase activity, protein-arginine deiminase activity	chromatin modification, peptidyl-citrulline biosynthetic process from peptidyl-arginine, protein modification process, protein modification process, regulation of transcription, DNA-dependent, transcription
PARVB	parvin, beta	actin binding, protein binding	cell adhesion
PBX1	pre-B-cell leukemia homeobox 1	protein binding, sequence-specific DNA binding, transcription factor activity	C21-steroid hormone biosynthetic process, cell differentiation, regulation of transcription, DNA-dependent, sex determination
PCOLCE2	procollagen C-endopeptidase enhancer 2	heparin binding, protein binding	
PDXK	pyridoxal (pyridoxine, vitamin B6) kinase	ATP binding, lithium ion binding, magnesium ion binding, nucleotide binding, potassium ion binding, protein homodimerization activity, pyridoxal kinase activity, pyridoxal	cell proliferation, pyridoxal phosphate biosynthetic process, pyridoxine biosynthetic process, vitamin B6 metabolic process

		phosphate binding, sodium ion binding, transferase activity, zinc ion binding	
PF4	platelet factor 4 (chemokine (C-X-C motif) ligand 4)	chemokine activity, heparin binding	cytokine and chemokine mediated signaling pathway, immune response, leukocyte chemotaxis, negative regulation of angiogenesis, negative regulation of megakaryocyte differentiation, platelet activation
PIK3CB	phosphoinositide-3-kinase, catalytic, beta polypeptide	1-phosphatidylinositol-3-kinase activity, binding, inositol or phosphatidylinositol kinase activity, insulin receptor substrate binding, phosphatidylinositol-4,5-bisphosphate 3-kinase activity, phosphotransferase activity, alcohol group as acceptor, transferase activity	G-protein coupled receptor protein signaling pathway, activation of MAPK activity, cellular calcium ion homeostasis, chemotaxis, embryonic cleavage, homophilic cell adhesion, phosphoinositide phosphorylation, phosphoinositide-mediated signaling, platelet activation, regulation of cell-matrix adhesion, signal transduction
PIP5K1B	phosphatidylinositol-4-phosphate 5-kinase, type I, beta	1-phosphatidylinositol-4-phosphate 5-kinase activity, kinase activity, protein binding, transferase activity	phosphatidylinositol metabolic process, phosphorylation
PKLR	pyruvate kinase, liver and RBC	magnesium ion binding, potassium ion binding, pyruvate kinase activity, pyruvate kinase activity, transferase activity	glycolysis, glycolysis, response to other organism
PKP2	plakophilin 2	binding, protein binding, structural molecule activity	cell-cell adhesion

PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	calcium ion binding, hydrolase activity, lysophospholipase activity, phospholipase A2 activity	icosanoid metabolic process, lipid catabolic process, metabolic process, phospholipid catabolic process, platelet activating factor biosynthetic process
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	calcium ion binding, hydrolase activity, phosphoinositide phospholipase C activity, protein binding, signal transducer activity	intracellular signaling cascade, lipid catabolic process
PLD1	phospholipase D1, phosphatidylcholine-specific	hydrolase activity, phosphoinositide binding, phospholipase D activity, protein binding	Ras protein signal transduction, cell communication, chemotaxis, lipid catabolic process, metabolic process
PPBP	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	chemokine activity, glucose transmembrane transporter activity, growth factor activity	cell proliferation, chemotaxis, defense response to bacterium, glucose transport, immune response
PRRG4	proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane)	calcium ion binding, molecular_function	biological_process
PSTPIP1	proline-serine-threonine phosphatase interacting protein 1	catalytic activity, protein binding	cell adhesion, signal transduction
PTAFR	platelet-activating factor receptor	phospholipid binding, platelet activating factor receptor activity, receptor activity	G-protein coupled receptor protein signaling pathway, chemotaxis, cytokine production, immune response, inflammatory response, phosphoinositide-mediated signaling, signal transduction
PTX3	pentraxin-related gene, rapidly induced by IL-1 beta	zymosan binding	inflammatory response, opsonization, positive regulation of nitric oxide

			biosynthetic process, positive regulation of phagocytosis, response to yeast
PYGL	phosphorylase, glycogen; liver (Hers disease, glycogen storage disease type VI)	AMP binding, ATP binding, bile acid binding, drug binding, glucose binding, glycogen phosphorylase activity, nucleotide binding, phosphorylase activity, protein homodimerization activity, purine binding, pyridoxal phosphate binding, transferase activity, transferring glycosyl groups	carbohydrate metabolic process, glucose homeostasis, glycogen metabolic process
QPCT	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	acyltransferase activity, glutaminyl-peptide cyclotransferase activity, metal ion binding, peptidase activity, transferase activity, zinc ion binding	protein modification process, proteolysis
RAB27A	RAB27A, member RAS oncogene family	GTP binding, GTPase activity, nucleotide binding	protein transport, small GTPase mediated signal transduction
RBP7	retinol binding protein 7, cellular	lipid binding, retinal binding, retinol binding, transporter activity	transport
REPS2	RALBP1 associated Eps domain containing 2	calcium ion binding, calcium ion binding, protein binding	epidermal growth factor receptor signaling pathway, protein complex assembly
RHAG	Rh-associated glycoprotein	ammonium transmembrane transporter activity, ammonium transmembrane transporter activity, ankyrin binding, molecular_function,	ammonium transport, ammonium transport, blood circulation, cellular ion homeostasis, transport

		phosphopantetheine binding	
RHD	Rh blood group, D antigen	molecular_function	biological_process
RNF217	ring finger protein 217	ligase activity, metal ion binding, protein binding, ubiquitin-protein ligase activity, zinc ion binding	protein ubiquitination during ubiquitin-dependent protein catabolic process, ubiquitin cycle
ROGDI	rogdi homolog (Drosophila)		
RXRA	retinoid X receptor, alpha	ligand-regulated transcription factor activity, metal ion binding, molecular_function, protein binding, retinoid-X receptor activity, sequence-specific DNA binding, steroid binding, steroid hormone receptor activity, tRNA-pseudouridine synthase activity, transcription coactivator activity, transcription factor activity, zinc ion binding	biological_process, positive regulation of transcription from RNA polymerase II promoter, regulation of transcription, DNA-dependent, signal transduction, transcription, vitamin metabolic process
SDPR	serum deprivation response (phosphatidylserine binding protein)	phosphatidylserine binding, protein binding	
SEC14L4	SEC14-like 4 (S. cerevisiae)	lipid binding, transporter activity	transport
SERPINB10	serpin peptidase inhibitor, clade B (ovalbumin), member 10	protein binding, serine-type endopeptidase inhibitor activity	
SGMS2	sphingomyelin synthase 2	ceramide cholinephosphotransferase activity, kinase activity, sphingomyelin synthase activity, transferase activity	lipid metabolic process, sphingolipid metabolic process, sphingomyelin biosynthetic process

SIGLEC5	sialic acid binding Ig-like lectin 5	protein binding, sugar binding	cell adhesion
SIRPA	signal-regulatory protein alpha		cell adhesion
SIRPB1	signal-regulatory protein beta 1		cell surface receptor linked signal transduction
SLC22A15	solute carrier family 22, member 15	transporter activity	transport
SLC22A4	solute carrier family 22 (organic cation/ergothioneine transporter), member 4	ATP binding, DNA binding, carnitine transporter activity, ion transmembrane transporter activity, nucleotide binding, secondary active organic cation transmembrane transporter activity, sodium ion binding, symporter activity, transporter activity	body fluid secretion, carnitine transport, ion transport, nucleosome assembly, organic cation transport, sodium ion transport
SLC2A5	solute carrier family 2 (facilitated glucose/fructose transporter), member 5	fructose transmembrane transporter activity, fructose transmembrane transporter activity, glucose transmembrane transporter activity, substrate-specific transmembrane transporter activity, sugar:hydrogen symporter activity, transporter activity	carbohydrate metabolic process, carbohydrate transport, fructose transport, glucose transport, transmembrane transport
SLC31A1	solute carrier family 31 (copper transporters), member 1	copper ion binding, copper ion transmembrane transporter activity, copper ion transmembrane transporter activity	copper ion transport, copper ion transport, ion transport
SLPI	secretory leukocyte peptidase inhibitor	serine-type endopeptidase inhibitor activity	

SNTB1	syntrophin, beta 1 (dystrophin-associated protein A1, 59kDa, basic component 1)	actin binding, calcium ion binding, calmodulin binding	muscle contraction
STEAP3	STEAP family member 3	FAD binding, binding, copper ion binding, electron carrier activity, iron ion binding, metal ion binding	apoptosis, cell cycle, ion transport, iron ion transport, oxidation reduction
STX11	syntaxin 11	SNAP receptor activity	intracellular protein transport, membrane fusion, vesicle-mediated transport
SYNE1	spectrin repeat containing, nuclear envelope 1	actin binding, actin binding, lamin binding, structural molecule activity	Golgi organization and biogenesis, muscle cell differentiation, nuclear organization and biogenesis
TACSTD2	tumor-associated calcium signal transducer 2	receptor activity	cell proliferation, cell surface receptor linked signal transduction, response to stimulus, visual perception
TFF3	trefoil factor 3 (intestinal)		defense response, digestion
TLR4	toll-like receptor 4	lipopolysaccharide binding, protein binding, protein binding, transmembrane receptor activity, transmembrane receptor activity	I-kappaB kinase/NF-kappaB cascade, T-helper 1 type immune response, detection of fungus, inflammatory response, innate immune response, macrophage activation, negative regulation of osteoclast differentiation, positive regulation of interleukin-12 biosynthetic process, positive regulation of interleukin-12 biosynthetic process, positive regulation of interleukin-8 biosynthetic process, positive regulation of tumor necrosis factor biosynthetic process, signal transduction, signal

			transduction
TM6SF1	transmembrane 6 superfamily member 1	molecular_function	biological_process
TMEM144	transmembrane protein 144		
TMEM50B	transmembrane protein 50B	molecular_function	biological_process
TMEM55A	transmembrane protein 55A	hydrolase activity	
TMEM56	transmembrane protein 56		
TNFSF13	tumor necrosis factor (ligand) superfamily, member 13	cytokine activity, tumor necrosis factor receptor binding	immune response, positive regulation of cell proliferation, positive regulation of isotype switching to IgA isotypes, signal transduction
TPP1	tripeptidyl peptidase I	protein binding, subtilase activity, tripeptidyl-peptidase I activity, tripeptidyl-peptidase I activity, tripeptidyl-peptidase I activity	bone resorption, lipid metabolic process, nervous system development, peptide catabolic process, protein catabolic process, proteolysis
TPRG1L	tumor protein p63 regulated 1-like		
TRIM10	tripartite motif-containing 10	metal ion binding, protein binding, zinc ion binding, zinc ion binding	hemopoiesis
TSPAN7	tetraspanin 7		protein amino acid N-linked glycosylation
TUBB1	tubulin, beta 1	GTP binding, GTPase activity, nucleotide binding, structural molecule activity	microtubule-based movement, protein polymerization
TXNDC3	thioredoxin domain containing 3 (spermatozoa)	ATP binding, nucleoside diphosphate kinase activity	CTP biosynthetic process, GTP biosynthetic process, UTP biosynthetic process, cell differentiation, cell redox homeostasis, multicellular organismal development, spermatogenesis
UBAC1	UBA domain containing 1	protein binding	ubiquitin cycle

UGCGL2	UDP-glucose ceramide glucosyltransferase-like 2	UDP-glucose:glycoprotein glucosyltransferase activity, UDP-glucose:glycoprotein glucosyltransferase activity, protein binding, transferase activity, transferring glycosyl groups	protein amino acid glycosylation
UNQ464	GLPG464		
VCAM1	vascular cell adhesion molecule 1	protein binding	cell-cell adhesion
VPS37C	vacuolar protein sorting 37 homolog C (S. cerevisiae)		protein transport
VSTM1	V-set and transmembrane domain containing 1		
WIPI1	WD repeat domain, phosphoinositide interacting 1	androgen receptor binding, estrogen receptor binding, phosphatidylinositol 3-phosphate binding, receptor binding	autophagy, vesicle targeting, trans-Golgi to endosome

Chapter 7

7. BIBLIOGRAPHY

1. Van Vlierberghe P, Pieters R, Beverloo HB, Meijerink JPP: **Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia**. *British Journal of Haematology* 2008, **143**(2):153-168.

2. Chuang H-Y, Lee E, Liu Y-T, Lee D, Ideker T: **Network-based classification of breast cancer metastasis**. *Molecular Systems Biology* 2007, **3**.

3. Akashi K: **Lineage promiscuity and plasticity in hematopoietic development**. *Ann N Y Acad Sci* 2005, **1044**:125-131.

4. Yang Q, Jeremiah Bell J, Bhandoola A: **T-cell lineage determination**. *Immunol Rev* 2010, **238**(1):12-22.

5. Chi AW, Bell JJ, Zlotoff DA, Bhandoola A: **Untangling the T branch of the hematopoiesis tree**. *Curr Opin Immunol* 2009, **21**(2):121-126.

6. Murphy KP, Travers P, Walport M, Janeway C: **Janeway's immunobiology**, 7th edn. New York: Garland Science; 2008.

7. Wu L: **T lineage progenitors: the earliest steps en route to T lymphocytes**. *Curr Opin Immunol* 2006, **18**(2):121-126.

8. Kawamoto H, Katsura Y: **A new paradigm for hematopoietic cell lineages: revision of the classical concept of the myeloid-lymphoid dichotomy**. *Trends Immunol* 2009, **30**(5):193-200.

9. **Elucidating the mechanism behind immunity using dentritic cells** [http://www.rikenresearch.riken.jp/eng/frontline/5028]

10. **Structural Biochemistry** [http://www.thefullwiki.org/Structural_Biochemistry/Protein_function/Antigen]

11. Kinoshita K, Honjo T: **Linking class-switch recombination with somatic hypermutation**. *Nat Rev Mol Cell Biol* 2001, **2**(7):493-503.

12. Ramsden DA, Weed BD, Reddy YV: **V(D)J recombination: Born to be wild**. *Semin Cancer Biol* 2010, **20**(4):254-260.

13. Soulas-Sprauel P, Rivera-Munoz P, Malivert L, Le Guyader G, Abramowski V, Revy P, de Villartay JP: **V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining**. *Oncogene* 2007, **26**(56):7780-7791.

14. [http://www.cartage.org.lb/en/themes/sciences/LifeScience/GeneralBiology/Immunology/Recognition/AntigenRecognition/Antibodydiversity/loop.gif]

15. Wong MM, Fish EN: **Chemokines: attractive mediators of the immune response**. *Seminars in Immunology* 2003, **15**:5-14.

16. Wiernik PH, American Cancer Society.: **Adult leukemias**. Hamilton, Ont. ; Lewiston, NY: BC Decker; 2001.

17. Shi Y, Sun G, Zhao C, Stewart R: **Neural stem cell self-renewal**. *Crit Rev Oncol Hematol* 2008, **65**(1):43-53.

18. Al-Hajj M, Clarke MF: **Self-renewal and solid tumor stem cells**. *Oncogene* 2004, **23**(43):7274-7282.

19. Wang JC, Dick JE: **Cancer stem cells: lessons from leukemia**. *Trends Cell Biol* 2005, **15**(9):494-501.

20. C Graux JC, L Michaux, P Vandenberghe, A Hagemeijer: **Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast.** *Leukemia* 2006, **20**:1496–1510.

21. Staal FJT, Langerak AW: **Signaling pathways involved in the development of T-cell acute lymphoblastic leukemia**. *Haematologica* 2008, **93**(4):493-497.

22. Aifantis I, Raetz E, Buonamici S: **Molecular pathogenesis of T-cell leukaemia and lymphoma**. *Nature Reviews Immunology* 2008, **8**(5):380-390.

23. Lluis Espinosa SC, Teresa D'Altri, Thomas Trimarchi, Alexander Statnikov, Jordi Guiu,, Veronica Rodilla JIs-E, Josep Nomdedeu, Beatriz Bellosillo, Carles Besses, Omar Abdel-Wahab,, Nicole Kucine S-CS, Guangchan Song, Charles C. Mullighan, Ross L. Levine, Klaus Rajewsky,, Iannis Aifantis AB: **The Notch/Hes1 Pathway Sustains NF-kB Activation through CYLD Repression in T Cell Leukemia**. *Cancer Cell* 2010, **18**(September 14):268-281.

24. Adolfo A. Ferrando, Donna S. Neuberg, Jane Staunton, Mignon L. Loh, Christine Huard, Susana C. Raimondi, Fred G. Behm, Ching-Hon Pui, James R. Downing, D. Gary Gilliland *et al*: **Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia**. *Cancer Cell* 2002, **1**.

25. Trang Hoang, Hoang T: **The T-ALL paradox in cancer**. *Nature Medicine* 2010, **16**(11):1185-1186.

26. Haferlach T, Kern W, Schnittger S, Schoch C: **Modern diagnostics in acute leukemias.** *Critical Reviews in Oncology/Hematology* 2005, **56**(2):223-234.

27. Haferlach T: **Global approach to the diagnosis of leukemia using gene expression profiling**. *Blood* 2005, **106**(4):1189-1198.

28. T. R. Golub DKS, P. Tamayo, C. Huard,, M. Gaasenbeek JPM, H. Coller, M. L. Loh,, J. R. Downing MAC, C. D. Bloomfield,, Lander ES: **Molecular Classifcation of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring.** *SCIENCE* 1999, **286**:531-537.

29. Samuel W. Beenken WEG, D. Ralph Crowe, Michael G. Conner,, Heidi L. Weiss MTS, Helen Krontiras, * Marshall M. Urist, Kirby I. Bland: **Molecular Biomarkers for Breast Cancer Prognosis: Coexpression of c-erbB-2 and p53**. *ANNALS OF SURGERY* 2001, **233**(5):630-638.

30. Nibbe RK, Chance MR: **Approaches to biomarkers in human colorectal cancer: looking back, to go forward**. *Biomarkers in Medicine* 2009, **3**(4):385-396.

31. Kelloff GJ, Boone CW, Crowell JA, Nayfield SG, Hawk E, Malone WF, Steele VE, Lubet RA, Sigman CC: **Risk biomarkers and current strategies for cancer chemoprevention**. *J Cell Biochem Suppl* 1996, **25**:1-14.

32. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y: **Phases of biomarker development for early detection of cancer**. *J Natl Cancer Inst* 2001, **93**(14):1054-1061.

33. Alexander Kohlmann TJK, Laura Z. Rassenti JRD, Sheila A. Shurtleff KIM, Amanda F. Gilkes W-K, Hofmann GB, Marta, Campo Dell'Orto RF, Sabina, Chiaretti JDV, Sonja Rauhut,, Peter R. Papenhausen JM, Herna´ndez EL, Allen E., Yeoh ESK, Rachel Li, *et al*: **An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase**. *british Journal of Haematology* 2008, **142** 802–807.

34. Bas J. Wouters, Bob Lo¨wenberg, Delwel R: **A decade of genome-wide gene expression profiling in acute myeloid leukemia: flashback and prospects**. *Blood* 2009, **113**(2):291-298.

35. Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Bene MC, De Vos J, Hernandez JM, Hofmann WK, Mills KI *et al*: **Clinical Utility of Microarray-Based Gene Expression Profiling in the Diagnosis and Subclassification of Leukemia: Report From the International Microarray Innovations in Leukemia Study Group**. *Journal of Clinical Oncology* 2010, **28**(15):2529-2537.

36. Junjie Su B-JY, Edward R. Dougherty: **Accurate and Reliable Cancer Classification Based on Probabilistic Inference of Pathway Activity.** *PLoS ONE* 2009, **4**(12).

37. Ash A. Alizadeh, Michael B. Eisen, R. Eric Davis, Chi Ma, Izidore S. Lossos, Andreas Rosenwald, Jennifer C. Boldrick, Hajeer Sabet, Truc Tran, Xin Yu *et al*: **Distinct types of diffuse large B-cell lymphoma identified by gene expresion profiling**. *Nature* 2000, **403**(6769):503-511.

38. Amir Ben-Dor, Laurakay Bruhn, Nir Friedman, Iftach Nachman, Michel Schummer, Yakhini Z: **Tissue Classification with Gene Expression Profiles**. *Journal of Computational Biology* 2000, **7**(3/4):559-583.

39. Ramaswamy S, Ross KN, Lander ES, Golub TR: **A molecular signature of metastasis in primary solid tumors**. *Nature Genetics* 2002, **33**(1):49-54.

40. TREY IDEKER, VESTEINN THORSSON, ANDREW F. SIEGEL, HOOD LE: **Testing for Differentially-Expressed Genes by Maximum-Likelihood Analysis of Microarray Data**. *Journal of Computational Biology* 2000, **7**(6):805-817.

41. Jansen R: **Relating Whole-Genome Expression Data with Protein-Protein Interactions**. *Genome Research* 2002, **12**(1):37-46.

42. Raser JM, O'Shea EK: **Noise in gene expression: origins, consequences, and control**. *SCIENCE* 2005, **309**(5743):2010-2013.

43. Klebanov L, Yakovlev A: **How high is the level of technical noise in microarray data?** *Biol Direct* 2007, **2**:9.

44. Tu Y, Stolovitzky G, Klein U: **Quantitative noise analysis for gene expression microarray experiments**. *Proc Natl Acad Sci U S A* 2002, **99**(22):14031-14036.

45. Aris VM, Cody MJ, Cheng J, Dermody JJ, Soteropoulos P, Recce M, Tolias PP: **Noise filtering and nonparametric analysis of microarray data underscores discriminating markers of oral, prostate, lung, ovarian and breast cancer**. *BMC Bioinformatics* 2004, **5**:185.

46. Trey Ideker, Owen Ozier, Benno Schwikowski, SIEGEL AF: **Discovering regulatory and signalling circuits in molecular interaction networks**. *Bioinformatics* 2002, **18**(1):233-240.

47. Segal E: **Discovering molecular pathways from protein interaction and gene expression data**. *Bioinformatics* 2003, **19**(90001):264i-272.

48. Chen J: **Detecting functional modules in the yeast protein-protein interaction network**. *Bioinformatics* 2006, **22**(18):2283-2290.

49. Gersten M, Alirezaei M, Marcondes MCG, Flynn C, Ravasi T, Ideker T, Fox HS: **An Integrated Systems Analysis Implicates EGR1 Downregulation in Simian Immunodeficiency Virus Encephalitis-Induced Neural Dysfunction**. *Journal of Neuroscience* 2009, **29**(40):12467-12476.

50. Yu-Qing Qiu SZ, Xiang-Sun Zhang, Luonan Chen: **Detecting disease associated modules and prioritizing active genes based on high throughput data.** *BMC Bioinformatics* 2010, **11**(26).

51. Vidal M: **A Biological Atlas of Functional Maps**. *Cell* 2001, **104**:333-339.

52. Gu J, Chen Y, Li S, Li Y: **Identification of responsive gene modules by network-based gene clustering and extending: application to inflammation and angiogenesis**. *BMC Systems Biology* 2010, **4**(1):47.

53. Aderem A: **Systems Biology: Its Practice and Challenges**. *Cell* 2005, **121**(4):511-513.

54. Wiltgen M, Tilz GP: **DNA microarray analysis: principles and clinical impact**. *Hematology* 2007, **12**(4):271-287.

55. Buness A, Ruschhaupt M, Kuner R, Tresch A: **Classification across gene expression microarray studies**. *BMC Bioinformatics* 2009, **10**:453.

56. Nacu S, Critchley-Thorne R, Lee P, Holmes S: **Gene expression network analysis and applications to immunology**. *Bioinformatics* 2007, **23**(7):850-858.

57. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M *et al*: **Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses**. *Proc Natl Acad Sci U S A* 2001, **98**(24):13790-13795.

58. MICHAEL B. EISEN PTS, PATRICK O. BROWN, DAVID BOTSTEIN: **Cluster analysis and display of genome-wide expression patterns** *Genetics* 1998, **95**:14863–14868.

59. Jiang B, Wang J, Wang Y, Xiao J: **Gene Prioritization for Type 2 Diabetes in Tissue-specific Protein Interaction Networks**. *The Third International Symposium on Optimization and Systems Biology (OSB'09)* 2009.

60. Liu W-m, Li R, Sun JZ, Wang J, Tsai J, Wen W, Kohlmann A, Mickey Williams P: **PQN and DQN: Algorithms for expression microarrays**. *Journal of Theoretical Biology* 2006, **243**(2):273-278.

61. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A *et al*: **Human Protein Reference Database--2009 update**. *Nucleic Acids Research* 2009, **37**(Database):D767-D772.

62. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH: **The WEKA Data Mining Software: An Update**. *SIGKDD Explorations* 2009, **11**(1).

63. Douglas Hanahan RAW: **The Hallmarks of Cancer**. *Cell* 2000, **100**: 57–70.

64. Graux C, Cools J, Melotte C, Quentmeier H, Ferrando A, Levine R, Vermeesch JR, Stul M, Dutta B, Boeckx N *et al*: **Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia**. *nature genetics* 2004, **36**(10):1084-1089.

65. De Keersmaecker K: **Fusion of EML1 to ABL1 in T-cell acute lymphoblastic leukemia with cryptic t(9;14)(q34;q32)**. *Blood* 2005, **105**(12):4849-4852.

66. Mori N, Krensky AM, Ohshima K, Tomita M, Matsuda T, Ohta T, Yamada Y, Tomonaga M, Ikeda S, Yamamoto N: **Elevated expression of CCL5/RANTES in adult T-cell leukemia cells: Possible transactivation of theCCL5 gene by human T-cell leukemia virus type I tax**. *International Journal of Cancer* 2004, **111**(4):548-557.

67. Dworzak MN, Fröschl G, Printz D, Zen LD, Gaipa G, Ratei R, Basso G, Biondi A, Ludwig WD, Gadner H: **CD99 expression in T-lineage ALL: implications for flow cytometric detection of minimal residual disease**. *Leukemia* 2004, **18**(4):703-708.

68. Kim De Keersmaecker, Peter Marynen, Cools J: **Genetics insights in the pathogenesis of T-cell acute lymphoblastic leukemia**. *Haematologica* 2005, **90**(8).

69. Xu Bing, Song Xiaoyan, Yip Nga Chi, Xiao Pingnan, Zhang Yanyan, Wang Weiguang, Shuyun Z: **Simultaneous detection of MDR1 and WT1 gene expression to predict the prognosis of adult acute lymphoblastic leukemia**. *Hematology* 2010, **15**(7):74-80.

70. Yagyu R, Hamamoto R, Furukawa Y, Okabe H, Yamamura T, Nakamura Y: **Isolation and characterization of a novel human gene, VANGL1, as a therapeutic target for hepatocellular carcinoma**. *Int J Oncol* 2002, **20**(6):1173-1178.

71. Blumenthal RD, Leon E, Hansen HJ, Goldenberg DM: **Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers**. *BMC Cancer* 2007, **7**:2.

72. Y. Wang, S. Klumpp, H. M. Amin, H. Liang, J. Li, Z. Estrov, P. Zweidler-McKay, S. J. Brandt, A. Agulnick, Nagarajan L: **SSBP2 is an in vivo tumor suppressor and regulator of LDB1 stability**. *Oncogene* 2010:3044-3053.

73. Oehler VG, Yeung KY, Choi YE, Bumgarner RE, Raftery AE, Radich JP: **The derivation of diagnostic markers of chronic myeloid leukemia progression from microarray data**. *Blood* 2009, **114**(15):3292-3298.

74. Soulier J, Clappier E, Cayuela JM, Regnault A, Garcia-Peydro M, Dombret H, Baruchel A, Toribio ML, Sigaux F: **HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL)**. *Blood* 2005, **106**(1):274-286.

75. Ovaa H: **Activity-based ubiquitin-specific protease (USP) profiling of virus-infected and malignant human cells**. *Proceedings of the National Academy of Sciences* 2004, **101**(8):2253-2258.

76. Ge Yimin, Terek EM: **CD36: A Multiligand Molecule**. *Hematology* 2005, **11**(1):31-37.

77. Mo YY, Moschos SJ: **Targeting Ubc9 for cancer therapy**. *Expert Opin Ther Targets* 2005, **9**(6):1203-1216.

78. Alessandra Martini, Roberta La Starza, Hilde Janssen, Chryste`le Bilhou-Nabera, Anniek Corveleyn, Riet Somers, Ana Aventin, Robin Foa`, Anne Hagemeijer, Christina Mecucci *et al*: **Recurrent Rearrangement of the Ewing's Sarcoma Gene, *EWSR1,* or Its Homologue, *TAF15,* with the Transcription Factor *CIZ*/*NMP4* in Acute Leukemia**. *Cancer Research* 2002, **62**:5408-5412.

79. Moreno CS: **The Sex-Determining Region Y-Box 4 and Homeobox C6 Transcriptional Networks in Prostate Cancer ProgressionCrosstalk with the Wnt, Notch, and PI3K Pathways**. *The American Journal of Pathology* 2010, **176**(2):518-527.

80. Strehl S, Borkhardt A, Slany R, Fuchs UE, Konig M, Haas OA: **The human LASP1 gene is fused to MLL in an acute myeloid leukemia with t(11;17)(q23;q21)**. *Oncogene* 2003, **22**(1):157-160.

81. Dagher R, Briere C, Feve M, Zeniou M, Pigault C, Mazars C, Chneiweiss H, Ranjeva R, Kilhoffer MC, Haiech J: **Calcium fingerprints induced by calmodulin interactors in eukaryotic cells**. *Biochim Biophys Acta* 2009, **1793**(6):1068-1077.

82. Sieg DJ, Hauck CR, Schlaepfer DD: **Required role of focal adhesion kinase (FAK) for integrin-stimulated cell migration**. *J Cell Sci* 1999, **112 (Pt 16)**:2677-2691.

83. Abdel-Ghany M, Cheng HC, Elble RC, Pauli BU: **Focal adhesion kinase activated by beta(4) integrin ligation to mCLCA1 mediates early metastatic growth**. *J Biol Chem* 2002, **277**(37):34391-34400.

84. Ishii KJ, Coban C, Akira S: **Manifold mechanisms of Toll-like receptor-ligand recognition**. *J Clin Immunol* 2005, **25**(6):511-521.

85. Zarember KA, Godowski PJ: **Tissue expression of human Toll-like receptors and differential regulation of Toll-like receptor mRNAs in leukocytes in response to microbes, their products, and cytokines**. *J Immunol* 2002, **168**(2):554-561.

86. Iwasaki A, Medzhitov R: **Toll-like receptor control of the adaptive immune responses**. *Nat Immunol* 2004, **5**(10):987-995.

87. Smith CI, Baskin B, Humire-Greiff P, Zhou JN, Olsson PG, Maniar HS, Kjellen P, Lambris JD, Christensson B, Hammarstrom L *et al*: **Expression of Bruton's agammaglobulinemia tyrosine kinase gene, BTK, is selectively down-regulated in T lymphocytes and plasma cells**. *J Immunol* 1994, **152**(2):557-565.

88. Ozkal S, Paterson JC, Tedoldi S, Hansmann ML, Kargi A, Manek S, Mason DY, Marafioti T: **Focal adhesion kinase (FAK) expression in normal and neoplastic lymphoid tissues**. *Pathol Res Pract* 2009, **205**(11):781-788.

89. Reddig PJ, Juliano RL: **Clinging to life: cell to matrix adhesion and cell survival**. *Cancer Metastasis Rev* 2005, **24**(3):425-439.

90. Jarvis LJ, Maguire JE, LeBien TW: **Contact between human bone marrow stromal cells and B lymphocytes enhances very late antigen-4/vascular cell adhesion molecule-1-independent tyrosine phosphorylation of focal adhesion kinase, paxillin, and ERK2 in stromal cells**. *Blood* 1997, **90**(4):1626-1635.

91. Gilmore AP, Romer LH: **Inhibition of focal adhesion kinase (FAK) signaling in focal adhesions decreases cell motility and proliferation**. *Mol Biol Cell* 1996, **7**(8):1209-1224.

92. McLean GW, Carragher NO, Avizienyte E, Evans J, Brunton VG, Frame MC: **The role of focal-adhesion kinase in cancer - a new therapeutic opportunity**. *Nat Rev Cancer* 2005, **5**(7):505-515.

93. Gabarra-Niecko V, Schaller MD, Dunty JM: **FAK regulates biological processes important for the pathogenesis of cancer**. *Cancer Metastasis Rev* 2003, **22**(4):359-374.

94. Gordon MY, Dowding CR, Riley GP, Goldman JM, Greaves MF: **Altered adhesive interactions with marrow stroma of haematopoietic progenitor cells in chronic myeloid leukaemia**. *Nature* 1987, **328**(6128):342-344.

95. Takahashi M, Keating A, Singer JW: **A functional defect in irradiated adherent layers from chronic myelogenous leukemia long-term marrow cultures**. *Exp Hematol* 1985, **13**(9):926-931.

96. Witkiewicz-Kucharczyk A, Bal W: **Damage of zinc fingers in DNA repair proteins, a novel molecular mechanism in carcinogenesis**. *Toxicol Lett* 2006, **162**(1):29-42.

97. Loh SN: **The missing zinc: p53 misfolding and cancer**. *Metallomics* 2010, **2**(7):442-449.

98. Mocchegiani E, Costarelli L, Giacconi R, Cipriano C, Muti E, Malavolta M: **Zinc-binding proteins (metallothionein and alpha-2 macroglobulin) and immunosenescence**. *Exp Gerontol* 2006, **41**(11):1094-1107.

99. Sahin G, Ertem U, Duru F, Birgen D, Yuksek N: **High prevelance of chronic magnesium deficiency in T cell lymphoblastic leukemia and chronic zinc deficiency in children with acute lymphoblastic leukemia and malignant lymphoma**. *Leuk Lymphoma* 2000, **39**(5-6):555-562.

100. Eby GA: **Treatment of acute lymphocytic leukemia using zinc adjuvant with chemotherapy and radiation--a case history and hypothesis**. *Med Hypotheses* 2005, **64**(6):1124-1126.

101. Martin ST, Sato N, Dhara S, Chang R, Hustinx SR, Abe T, Maitra A, Goggins M: **Aberrant methylation of the Human Hedgehog interacting protein (HHIP) gene in pancreatic neoplasms**. *Cancer Biol Ther* 2005, **4**(7):728-733.

102. Young RP, Whittington CF, Hopkins RJ, Hay BA, Epton MJ, Black PN, Gamble GD: **Chromosome 4q31 locus in COPD is also associated with lung cancer**. *Eur Respir J* 2010, **36**(6):1375-1382.

103. Shahi MH, Afzal M, Sinha S, Eberhart CG, Rey JA, Fan X, Castresana JS: **Human hedgehog interacting protein expression and promoter methylation in medulloblastoma cell lines and primary tumor samples**. *J Neurooncol* 2010.

104. Olsen CL, Hsu PP, Glienke J, Rubanyi GM, Brooks AR: **Hedgehog-interacting protein is highly expressed in endothelial cells but down-regulated during angiogenesis and in several human tumors**. *BMC Cancer* 2004, **4**:43.

105. Tada M, Kanai F, Tanaka Y, Tateishi K, Ohta M, Asaoka Y, Seto M, Muroyama R, Fukai K, Imazeki F *et al*: **Down-regulation of hedgehog-interacting protein through genetic and epigenetic alterations in human hepatocellular carcinoma**. *Clin Cancer Res* 2008, **14**(12):3768-3776.

106. Leandro-Garcia LJ, Leskela S, Landa I, Montero-Conde C, Lopez-Jimenez E, Leton R, Cascon A, Robledo M, Rodriguez-Antona C: **Tumoral and tissue-specific expression of the major human beta-tubulin isotypes**. *Cytoskeleton (Hoboken)* 2010, **67**(4):214-223.

107. Yin S, Bhattacharya R, Cabral F: **Human mutations that confer paclitaxel resistance**. *Mol Cancer Ther* 2010, **9**(2):327-335.

108. Petricoin EF, 3rd, Bichsel VE, Calvert VS, Espina V, Winters M, Young L, Belluco C, Trock BJ, Lippman M, Fishman DA *et al*: **Mapping molecular networks using proteomics: a vision for patient-tailored combination therapy**. *J Clin Oncol* 2005, **23**(15):3614-3621.