PROTEOMICS PROFILING AND GENOMICS COMPUTATIONAL ANALYSIS OF ACUTE MYELOID LEUKEMIA

MD. MIRAZUL ISLAM

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF MEDICAL SCIENCE

FACULTY OF MEDICINE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: MD. MIRAZUL ISLAM (I.C/Passport No: [redacted]

Matric No: MGN140042

Name of Degree: Master of Medical Science

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

PROTEOMICS PROFILING AND GENOMICS COMPUTATIONAL ANALYSIS OF ACUTE MYELOID LEUKEMIA (AML)

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature [redacted] Date: 20/06/2017

Subscribed and solemnly declared before,

Witness's Signature [redacted] Date: 21/07/2017

Name: [redacted]

Designation

ABSTRACT

Abnormal genetic and/or cytogenetic changes in hematopoietic progenitors can lead to deregulation of protein expression. This study is aimed at examining differential expression patterns of plasma proteins in Malaysian acute myeloid leukemia (AML) patients and to analyze the molecular abnormalities using computational and experimental approaches. In a case-control study, 38 (case=19, control=19) plasma samples were analyzed for 32 cytokines and chemokines (C&Ckines), which were quantitatively measured in each sample. Genetic and cytogenetic aberrations were also detected by conventional methods and comparative analysis was performed with the Cancer Genome Atlas (TCGA) AML data. Chromosomal translocation is a common cytogenetic aberration in AML that causes gene fusion. To understand the functions and underlying oncogenic transformation mechanisms of a fusion protein (MLL-GAS7), differential expression patterns of growth arrest specific-7 (GAS7) were studied using neuro-2a (N2A) cell line. Levels of C&Ckines in AML plasma differed significantly ($p<0.005$) between AML patients and healthy controls, in which 13 were upregulated and 4 downregulated. In addition, circulating C&Ckines follow distinct expression patterns that have the potential to be used for subclassification of AML, complementing cytogenetic, genetic, and epigenetic information. Differential expression of ten novel C&Ckines in AML plasma has been demonstrated, which have not been previously reported in AML. Importantly, our results indicated that MPO, MIF, HGF, and Galactin-3 have diagnostic significance. We speculated that MIF and TRAIL are promoting survival of AML blasts using PI3K/Akt pathway through the anti-apoptotic mechanism. DNA methylation is a common epigenetic aberration in AML and we found almost 200 genes hypermethylated in 90% of TCGA cases. Distinct signatures have been found after combining methylation and mutation (three groups of patients and two groups of genes).

For patients with del (5q) positive, average survival was <1 year, whereas for T (15; 17) positive patients, there was significantly higher survival rate (>3 years) in Kaplan-Meier curve. MLL and GAS7 downstream targeted genes expression vary (particularly Hox genes) in AML patients due to fusion protein formation. Notable differences had been noticed in N2A cell line for attachment and plating efficiency as well as growth curve due to expression of GAS7 which is a partner protein of MLL, one of the reported causative factors in AML. In conclusion, C&Ckines expression profile in AML patients differs from normal and forms a distinct pattern of expression. Integrated analysis of genomic, proteomic, and transcriptomics data and regulation of fusion protein expression may reveal insight into the pathogenesis of AML.

ABSTRAK

Perubahan genetik dan / atau sitogenetik tidak normal dalam leluhur hematopoietic membawa kepada penyahkawalseliaan ekspresi protein. Kajian ini bertujuan untuk mengkaji corak ekspresi plasma protein pesakit leukemia miloid akut (AML) Malaysia dan menganalisa keabnormalan molekul menggunakan pendekatan komputasi dan eksperimental. Dalam satu kajian kes-kawalan, 38 (kes = 19, kawalan = 19) sampel plasma dianalisis bagi 32 sitokin dan kemokin (C & Ckines) telah diukur secara kuantitatif dalam setiap sampel. Penyimpangan genetik dan sitogenetik juga dikesan dengan kaedah konvensional dan analisis perbandingan dilakukan menggunakan data Cancer Genome Atlas (TCGA) AML. Translokasi kromosom merupakan penyimpangan sitogenetik biasa di dalam AML yang menyebabkan pelakuran gen. Untuk memahami fungsi dan asas mekanisme transformasi onkogenik daripada protein lakuran (MLL-GAS7), pembezaan corak ekspresi penghentian pertumbuhan tertentu-7 (growth arrest specific-7, GAS7) telah dikaji menggunakan warisan sel neuro-2a (N2A). Tahap C & Ckines dalam plasma AML berbeza dengan ketara ($p < 0.005$) daripada kawalan sihat di mana 13 dikawal naik dan 4 dikawal turun. Di samping itu, C&Ckines di dalam edaran menunjukkan corak ekspresi tersendiri yang mempunyai potensi untuk digunakan bagi subklasifikasi AML, bagi melengkapkan maklumat sitogenetik, genetik dan epigenetik. Sepuluh ekspresi pembezaan C & Ckines yang novel didalam plasma AML telah diperhatikan yang belum pernah dilaporkan bagi AML. Lebih penting lagi, keputusan kami menunjukkan bahawa MPO, MIF, HGF dan Galactin-3 mempunyai kepentingan diagnostik. Kami membuat spekulasi bahawa MIF dan TRAIL mempromosikan survival blas AML menggunakan laluan PI3K / Akt melalui mekanisme anti-apoptotic. Metilasi DNA merupakan suatu penyimpangan epigenetik yang biasa di dalam AML dan kami mendapati hampir 200 gen terhipermetilasi dalam 90% daripada kes-kes TCGA. 'Tandatangan' yang tersendiri telah ditemui selepas menggabungkan metilasi dan mutasi

(tiga kumpulan pesakit dan dua kumpulan gen). Bagi pesakit dengan del (5q) positif, purata survival adalah <1 tahun, manakala pesakit positif T (15; 17) menunjukkan kadar survival lebih tinggi (> 3 tahun) dengan signifikan pada keluk Kaplan-Meier. Ungkapan gen tersasar MLL dan GAS7 hiliran berbeza-beza (terutamanya gen Hox) di dalam pesakit AML disebabkan oleh pembentukan protein terlakur. Perbezaan yang ketara diperhatikan dalam sel warisan N2A untuk perlekatan dan kecekapan penyaduran serta keluk pertumbuhan kerana ungkapan GAS7, yang merupakan protein bersama MLL, yang merupakan salah satu faktor penyebab yang dilaporkan dalam AML. Kesimpulannya, profil ungkapan C&Ckines dalam AML berbeza daripada subjek kawalan dan membentuk corak ungkapan yang tersendiri. Analisis data secara berintegrasi antara genomik, proteomik, dan transkriptomik serta ungkapan protein terlakur dapat memberi pengertian yang mendalam mengenai patogenesis AML.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Prof. Datin Dr. Zahurin Mohamed and Dr. Elsa Haniffah Mejia Mohamed from the Department of Pharmacology, Faculty of Medicine, University of Malaya for giving me the opportunity to work on this project and their continuous guidance, encouragement, and support to finish my research project.

Next, I would like to acknowledge Dr. Yassen Assenov from Computational Epigenomics Group, German Cancer Research Center (DKFZ), Heidelberg, Germany, who had generously supported me to perform computational analysis. I would like to express gratitude to Dr. Sue Lin-Chao from the Institute of Molecular Biology, Academia Sinica, Taiwan, who allowed me to continue part of my project in her lab. I also would like to thank Prof. Dr. Paul Gaynon from Hematology Unit, Children Hospital Los Angeles (CHLA), University of Southern California and Dr. Altaf H. Sarker from Life Science Division, the Laurence Berkeley National Laboratory, University of California-Berkeley, California for their instruction regarding project direction and clinical data analysis.

My sincere appreciation goes to Hematology Unit, Institute for Medical Research, Malaysia for assisting me in the sample collection. I would also like to thank Dr. Shamsul Mohd Zain and to Ms Tan Hwa-Li for useful discussion and for sample collection.

Special thanks to Yayasan Khazanah for offering me a scholarship to study at the University of Malaya. Last but not least, a big thank you to my family and friends for supporting me continuously throughout the course of my study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AFP	:	Alpha-fetoprotein
AML	:	Acute myeloid leukemia
ANOVA	:	Analysis of variance
Akt	:	Protein kinase B
ATCC	:	American Type Culture Collection
ATRA	:	All-trans retinoic acid
CA	:	Cancer antigen
CEA	:	Carcinoembryonic antigen
CD74	:	Cluster of Differentiation 74
CGI	:	CpG Island
CN-AML	:	Cytogenetically normal-acute myeloid leukemia
C&Ckines	:	Cytokines and chemokines
CSF	:	Colony stimulating factor
CYFRA-21-1	:	Cytokine 19 fragment
DFS	:	Disease free survival
DMEM	:	Dulbecco's modified Eagle's medium
DNMT3A	:	DNA Methyltransferase-3A
ELISA	:	Enzyme-linked immunosorbent assay
FAB	:	French American British
FAP	:	Fibroblast activating protein
FGF2	:	Fibroblast growth factor beta
GAS7	:	Growth arrest specific-7
GO	:	Gene ontology

GWAS	:	Genome wide association studies
HCG	:	Human chorionic gonadotropin
HDAC	:	Histone deacetylase
HE4	:	Human epididymis protein 4
HGF	:	Hepatocyte growth factor
HSP90	:	Heat shock protein 90
IFN	:	Interferon
IGFBP3	:	Insulin-like growth factor binding protein 3
IL	:	Interleukin
IPA	:	Ingenuity pathway analysis
KAT	:	Lysine acetyltransferase
lncRNAs	:	Long non-coding RNAs
LPD	:	Lymphoproliferative disorders
MDS	:	Multidimensional scaling
miRNA	:	MicroRNA
miRES	:	miRNA encoding sites
miRDB	:	miRNA database
MLL	:	Mixed lineage leukemia
MOI	:	Multiplicity of infection
MPO	:	Myeloperoxidase
MIA	:	Melanoma inhibitory activity protein
MIF	:	Macrophage migration inhibitory factor
N2A	:	Neuro-2A
OOR	:	Out-of-range
OPN	:	Osteopontin
OS	:	Overall survival

PCA	:	Principal component analysis
PI3K	:	Phosphoinositide 3-kinase
PSA	:	Prostate specific antigen
PVDF	:	Polyvinylidene difluoride
RnBeads	:	Software
SNP	:	Single nucleotide polymorphism
STP	:	Signal transduction pathway
sFasL	:	Soluble Fas Ligand
SCF	:	Stem cell factor
SHBG	:	Sex Hormone Binding Globulin
SDS-PAGE	:	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TCGA	:	The cancer genome atlas
TNF	:	Tumor nacrosis factor
TRAIL	:	TNF-related apoptosis-inducing ligand
VEGF	:	Vascular endothelial growth factor
WHO	:	World Health Organization

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

Acute myeloid leukemia (AML) is a hematological cancer characterized by the excessive proliferation of immature myeloid cells as well as disruption of differentiation (Tenen, 2003) in the bone marrow that infiltrate liver, spleen, lymph node, and circulating blood (Mehdipour, Santoro, & Minucci, 2015). This cancer type progresses rapidly and is relatively fatal due to acquired genetic and/or cytogenetic aberrations. In 2015, AML was the most common type of leukemia diagnosed and formed 1.75% of expected total cancer death in 2016 in the United States (Siegel, Miller, & Jemal, 2016). Overall 5-year survival rate varies drastically based on the cytogenetic risk classification: 55%, 24%, and 5% overall survival for favorable, intermediate, and unfavorable risk patients groups, respectively (Byrd et al., 2002). Relapse is the major reason for poor survival rate as it could occur in up to 80% of AML patients (Christopeit & Bartholdy, 2014). Almost 50% of all AML cases belonging to the intermediate risk group have been reported to lack any cytogenetic abnormalities (Vardiman et al., 2009). Furthermore, a significant proportion of the patients carry no reported genetic mutations in any known AML associated driver genes (Patel et al., 2012; Shen et al., 2011). These findings clearly indicate that there are other elements predisposing to and driving the disease especially in the case of cytogenetically normal AML (CN-AML).

Molecular abnormalities (including mutation of AML associated driver genes, SNPs in protein coding sequences and miRNA binding sites of gene transcripts, and copy number variation of the genes) eventually cause over/under expression of proteins. Deregulation of epigenetic modifications have been reported in AML (Conway O'Brien, Prideaux, & Chevassut, 2014) thereby causing changes to normal gene expression patterns, which then leads to abnormal proteins expressions. Cytogenetic aberrations (including chromosomal deletions, insertions, duplications, substitutions, and

translocations) are also directly or indirectly affecting the translated proteins levels in AML patients. So, genetic, epigenetic, and cytogenetic abnormalities are ultimately cause deregulated proteins expressions which are associated with the abnormal myeloid development and poor differentiation.

While deregulated expression of circulating C&Ckines in AML is well documented, considerable baseline variability of inter- and intra-assay plasma C&Ckines levels are reported in AML. Also, plasma baselines for some C&Ckines that are biomarkers for other cancers have not been previously reported in AML. It would also be interesting to see whether those deregulated C&Ckines expressions follow any patterns that may help to figure out novel biomarkers for early diagnosis and prognosis. Protein expression may also vary between the different geographic and ethnic groups. A comprehensive profiling of plasma C&Ckines might give us a greater insight of baselines variability and expression signatures, potentially leading to the identification of novel diagnostic biomarkers and therapeutic targets in AML.

It has been reported that particular epigenetic modifications are involved in the regulation of hematopoietic development (Bock et al., 2012; Ji et al., 2010). Epigenetic modifications, such as DNA methylation and histone acetylation, are well known and contribute to cellular identity, development, and maturation through regulation of genes expression in myeloid progenitor lineages. Any unexpected change in epigenetic modifiers (like a mutation in DNMT3A) could affect downstream gene expressions and may cause diseases like cancer. So, differential and exploratory analysis of epigenomic data from a large cohort (TCGA) may reveal insight into AML pathogenesis.

Chromosomal translocation is one of the common cytogenetic abnormalities in AML that is often linked to the generation of oncogenic fusion proteins. One of the frequently fused genes is mixed-lineage leukemia (MLL), which form a fusion protein with GAS7

through the translocation at t(11;17)(q23;p13) position (Panagopoulos et al., 2006). Differential analysis of MLL and GAS7 (computational and experimental) may provide a better understanding of MLL-GAS7 fusion protein functions and underlying oncogenic transformation mechanisms.

1.1 Objectives

The aims of this study are:

1. To determine differential cytokines and chemokines (C&Ckines) expression patterns in AML patients and assess diagnostic and/or prognostic possibilities.
2. To carry out computational analysis of TCGA AML data and correlate with C&Ckines expression.
3. To determine the pathogenic role of the MLL-GAS7 fusion protein in AML and differential expression of GAS7 in N2A cell line.

1.2 Research hypothesis

1. There exist differences in plasma protein expression level in AML patients and healthy controls.
2. There exist differential gene expression and DNA methylation pattern in the global AML dataset.
3. There exist differential expressions of MLL and GAS7 protein, responsible for AML pathogenesis.

CHAPTER 2: LITERATURE REVIEW

2.1 Acute myeloid leukemia

2.1.1 Development of AML

Hematopoiesis is a precisely regulated series of complex, dynamic, and integrated processes including self-renewal, development, and differentiation of hematopoietic stem cells (Luczak et al., 2012). After progressively losing self-renewal capacity, those stem cells become committed to a certain lineage (Dexter, 1990; Metcalf, 1999). Development of myeloid lineage depends on the differentiation of pluripotent stem cells into myeloid progenitors; subsequently, develop into myeloid precursors followed by a specific differentiation lineage that leads to the formation of mature neutrophils, eosinophils, basophils and monocytes, erythrocytes, and megakaryocytes. Commitment to a particular lineage is regulated by signal transmission pathways, lineage-specific transcription factors, and miRNAs (Shivdasani, 2006). Small disturbances within any process of myeloid lineage may cause various diseases like myelodysplastic syndromes and myeloid leukemia which are usually characterized by maturation blockage. Myeloid cells can acquire somatic mutations in one or both alleles that result in uncontrolled cell division and inhibition of apoptosis.

Significant progress has been made to understand the regulatory mechanism of myeloid lineage (myelopoiesis) and its deregulations that lead to oncogenic transformation. Genomic and transcriptomic studies have revealed that transcription factors commonly malfunction in AML due to genetic and/or cytogenetic aberrations. Mutations in transcription factors (C/EBPα and PU.1) may lead to the development of AML (Tenen, 2003). Oncofusion proteins often generated by chromosomal translocation are also responsible for the development of AML (Alcalay et al., 2003; C. W. So, Karsunky, et al., 2003). Epigenetic regulation is an important factor of physiological development and first epigenetic aberration in AML was reported by Baylin et al. in 1987

where they noticed hypermethylation of the 5′ regulatory region of the calcitonin gene causes AML (Baylin et al., 1987). Abnormal DNA methylation in p15, p73, E-cadherin, ID 4, and RARbeta2 as well as histone acetylation may lead to the development and progression of AML (Oki & Issa, 2009). Non-coding RNAs, like miRNAs, are important for the regulation oncogene transcripts, whereas deregulation of miRNAs can promote the development of AML (Marcucci, Mrózek, Radmacher, Garzon, & Bloomfield, 2011; Yuan, Kasar, Underbayev, Prakash, & Raveche, 2012). So, the development of AML could progress through multiple ways hence reflecting the heterogeneity of this disease.

2.1.2 Classification

Due to genetic complexity and heterogeneity of multiple acquired genomic lesions, AML classification is always challenging. In the early days, AML diagnosis and classification were performed mainly based on morphological and cytochemical properties of leukemic cells. According to these properties, the French-American-British (FAB) cooperative group proposed eight basic subtypes of AML (Bennett et al., 1976b) from M0 to M7 where M5 is further divided into M5a and M5b (Table 2.1). Based on cytogenetic abnormalities, the prognosis of AML patients can be categorized into three groups: favorable, intermediate, and unfavorable. Almost 25% of AML patients are under the favorable risk group, and mostly consist of t (15;17), t (8;21), and inv (16). Another 25-30% AML patient are under unfavorable risk group with MLL fusion, t (6;9), t (9;22), del (5q), del (7q), inv (3) (q21q26), and other existing complex chromosomal aberrations. The last group (intermediate) consists of 50-60% AML patients with mostly normal karyotype and t(9;11) (Mrózek & Bloomfield, 2006; Schoch et al., 2004; Slovak et al., 2000). According to Slovak et al. five-years overall survival rate for those risk groups are 55, 38, and 11% for favorable, intermediate and unfavorable risk groups, respectively (Slovak et al., 2000).

Table 2.1: French-American-British AML classification (Luczak et al., 2012)

AML subtype	Description
M0	AML with minimal differentiation, 3-5% of cases
M1	AML without maturation, 15-20% of cases
M2	AML with maturation, 25-30% of cases
M3	Acute promyelocytic leukemia (APL), 10-15% of cases
M4	Acute myelomonocytic leukemia, 20-30% of cases
M5a	Acute monoblastic leukemia; 2-7% of cases
M5b	Acute monocytic leukemia; 2-5% of cases
M6	Acute erythroid leukemia; 3-5% of cases
M7	Acute megakaryoblastic leukemia; 3-5% of cases

The most recent intensive AML classification was proposed by the World Health Organization (WHO) in 2008 (Table 2.2) and is widely being used by clinicians (Swerdllow, Campo, & Harris, 2008). This classification system not only considers morphological features of leukemic cells and their immunophenotype but also looks at the reported genetic and cytogenetic properties. Although WHO classification covers the majority of AML variability, the discovery of unclassified AML subtypes is increasing due to rapid improvement of diagnostic techniques. Almost 50% of all AML cases belonging to the intermediate risk group have been reported to lack any cytogenetic abnormalities (Vardiman et al., 2009). Patients without reported genetic/cytogenetic abnormal factors may have distinct epigenetic patterns which have not been considered

yet by any of those two well-known classification methods. Moreover, epigenetic signatures have already been reported in a wide variety of cancers. It can be said that epigenetic patterns/signatures could be an emerging strategy to classify AML patients precisely and accurately. Epigenetic pattern recognition approaches for AML patients have been tried in this study in Chapter 4.

Table 2.2: WHO AML Classification (Döhner et al., 2010)

	AML subtype	Description
1 1	AML with recurrent genetic abnormalities	AML with t(8;21)(q22;q22); RUNX1-RUNX1T1
		AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); CBFB-MYH11
		APL with t(15;17)(q22;q12); PML-RARA
		AML with t(9;11)(p22;q23); MLLT3-MLL
		AML with t(6;9)(p23;q34); DEK-NUP214
		AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1
		AML (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1
		Provisional entity: AML with mutated NPM1
		Provisional entity: AML with mutated CEBPA
2 2	AML with myelodysplasia-related changes	Complex karyotype (3 or more chromosomal abnormalities)
		Unbalanced; balanced changes
3 3	Therapy-related myeloid neoplasms	Alkylating agents; Ionizing radiation therapy; Topoisomerase II inhibitors
4 4	AML, not otherwise specified (NOS)	AML minimally differentiated (FAB – M0)
		AML without maturation (FAB – M1)
		AML with maturation (FAB – M2 and M3)
		Acute myelomonocytic leukemia (FAB – M4)
		Acute monoblastic leukemia and monocytic
		Leukemia (FAB – M5a i M5b)
		Acute erythroleukemia (FAB – M6)
		Acute megakaryioblastic leukemia (FAB – M7)
		Acute basophilic leukemia
		Acute panmyelosis with myelofibrosis
5 5	Myeloid sarcoma	Extramedullary myeloid tumor; granulocytic sarcoma; chloroma
6 6	Myeloid proliferations related to Down syndrome	Transient abnormal myelopoiesis
		Myeloid leukemia associated with Down syndrome
7 7	Blastic plasmacytoid dendritic cell neoplasm	Previously known as natural killer (NK) cell leukemia/lymphoma
8 8	Acute leukemias of ambiguous lineage	Acute undifferentiated leukemia
		Mixed phenotype acute leukemia with t(9;22)(q34;q11.2); BCR-ABL1
		Mixed phenotype acute leukemia with t(v;11q23); MLL rearranged
		Mixed phenotype acute leukemia, B/myeloid, NOS

		Mixed phenotype acute leukemia, T/myeloid, NOS

2.1.3 Population distribution

Cancer is one of the leading causes of death all over the world and the second leading cause of death in the United States (Siegel et al., 2016). It is a major public health problem regardless of countries or economic status. Cancer risk is expected to increase worldwide, particularly in less developed countries, where 82% of world's population resides, due to the gradual increase of risk factors like smoking, poor diet, physical inactivity, and reproductive changes (Torre et al., 2015). According to the US Cancer Statistics Report-2016, the projected number of new cancer cases would be 1,685,210 and cancer-related deaths would be 595,690 in 2016 in the US. The cancer-related death rate has dropped by 23% in the US since 1991 and death rate for leukemia is also decreasing because of the dramatic advances in therapeutics (Siegel et al., 2016). In 2014, AML was the most common type of leukemic diagnosed in both male and female, and the highest estimated leukemic death in the United States (Siegel, Ma, Zou, & Jemal, 2014). Expected numbers of estimated AML death rates are slowly decreasing: 1.78% of total cancer death in 2014 (Siegel et al., 2014), 1.77% of total cancer death in 2015 (Siegel, Miller, & Jemal, 2015), and 1.75% of total cancer death in 2016 in the United States (Siegel et al., 2016). In 2012, an estimated 14.1 million new cancer cases and 8.2 million cancer deaths had occurred worldwide whereas lung cancer in men and breast cancer in women were the most frequently diagnosed cancer in less developed countries (Torre et al., 2015). Overall cancer prevalence and types significantly vary among less developed countries and more developed countries. In general, cancer rates are higher (twice) in more developed regions (Table 2.3) as compared to less developed regions (Torre et al., 2015), but overall incidence will continue to shift to less developed countries due to growth and aging of the population and increasing prevalence of known risk factors (Bray & Møller, 2006).

According to GLOBOCAN 2012, around 786,400 new cancer cases were diagnosed in South-East Asia in 2012 and overall incidence and mortality rates per 100,000 were 138.2 and 94.8, respectively. Fewer mortality rates were noticed in developed countries as compared to developing countries, and this probably reflects differences in cancer profiles and/or the variability of the treatments (Torre et al., 2015).

Table 2.3: Global Incidence and Mortality Rates of Leukemia

		More Developed Countries		Less Developed Countries	
		Incidence (per 100,000)	Mortality (per 100,000)	Incidence (per 100,000)	Mortality (per 100,000)
All Cancers	Males	308.7	138	163	120.1
	Females	240.6	86.2	135.8	79.8
Leukemia	Males	8.8	4.6	4.4	3.7
	Females	5.8	2.8	3.2	2.6

2.1.4 Malaysian context

In Malaysia, leukemia (regardless of subtypes) is the seventh most common cancer, with an incidence of 2.9 per 100,000 populations for all cancers. In males, it is the 7th most common cancer (5.2%), whereas it is the 8th (3.2%) in females. Males are (56.5%) more prone to leukemia than females (43.5%), according to the National Cancer Registry Report, 2007. Leukemia is the most common cancer among children (0-14) in Malaysia. It is also the most predominant in older groups. The number of patients diagnosed in different age groups are as follows: ≤14 years, 12.7% patients; 15–30 years, 25.2% patients; 31–40 years, 15.6% patients; 41–50 years,17.1% patients; 51–60 years, 14.6% patients; and 61 years and above, 14.8% patients (Meng, Noor, Ismail, Ahid, & Zakaria, 2013). Institute for Medical Research in Malaysia reported that chromosome

abnormalities have been detected in 30.4% of AML patients while 69.6% patients have normal karyotypes. According to Cancer Index, 37400 people are newly diagnosed and 21,700 people die from all types of cancer every year in Malaysia. In general, 10-year overall survival (OS) and disease-free survival (DFS) for Malaysian AML patients after allogeneic hematopoietic stem cell transplantation (HSCT) were 63% and 67 %, respectively (Mangantig, Naing, Norsa'adah, & Azlan, 2013).

2.2 Pathogenic role

2.2.1 Deregulated cytokines and chemokines expression in AML

Leukemic cells must acquire multiple abnormalities during their generation (Hanahan & Weinberg, 2000) in order to block the differentiation intensely and to enhance self-renewal capacity of immature blasts. Blockage of myeloid differentiation and blasts formations largely depend on cell surface receptor proteins, transcription factors, and signaling molecules. Cytokines, chemokines, and growth factors (C&Ckines) are a group of signaling molecules that are essential for normal hematopoiesis. They mostly work through autocrine or paracrine modes by binding to cell-surface receptors and by triggering intercellular signaling cascades. C&Ckines mediated aberrant signal transduction pathway (STP) activation is a frequent event in leukemic blasts (Kupsa, Milos Horacek, & Jebavy, 2012), caused by mutant transcription factors expression and are required to block the differentiation process (Gilliland & Tallman, 2002). As aberrant C&Ckines levels in AML patients are reported (Kornblau et al., 2010; Van Etten, 2007), it can be postulated that specific patterns of C&Ckines expression may exist also in Malaysian AML patients. In Chapter 3, C&Ckines expression patterns have been determined for adult Malaysian AML patients.

Regulations of AML proliferation and progression by C&Ckines signaling are important for understanding the underlying mechanism of AML pathogenesis and variability. Sometimes it is difficult to predict the association between leukemic role and individual C&Ckines level because of sharing of receptors, simultaneous production, and autocrine mode of action. C&Ckines binding receptors have multiple transmembrane domains with tyrosine kinase and phosphorylase activities. Some non-receptor tyrosine kinases also have been reported to be involved in the pathogenesis of leukemia and are activated by some growth factors (Pendergast, 2002). After C&Ckines binding to the extracellular domain of the receptors, intracellular domains oligomerization occur which lead to disruption of auto-inhibition and subsequently activation of multiple signaling cascades (Lemmon & Schlessinger, 2010; Schlessinger, 2000).

Interleukin 2 (IL-2), macrophage inhibitory proteins 1-α, IL-1, and IL-1R have shown to have profound impact on leukemic proliferation and progression (Estrov et al., 1992; Estrov et al., 1987; Ferrajoli et al., 1994). Elevated levels of IL-3, IL-6, IL-8, TNF-α, IFN-γ, SCF, and CSF2 have been reported in leukemia patients as compared to controls (Elbaz & Shaltout, 2000; Tao et al., 2000; Tsimberidou et al., 2008). Although deregulated C&Ckines expressions are well documented (Van Etten, 2007), overall C&Ckines expressions reduce after chemotherapy administration and at remission stages (Hsu et al., 2002). Prognostic roles of hepatocyte growth factor (HGF) levels have been reported where AML patients with the lower serum concentration of HGF exhibits better leukemia free survival (Gwang Kim et al., 2005). Abnormal stimulation of leukemic blasts are due to deregulation of C&Ckines and of growth factors through genetically intact receptors and signaling pathways in order to ensure the proliferation of blasts without differentiation and to inhibit apoptosis for survival advantage (Lotem & Sachs, 1996).

2.2.2 Epigenetic landscape of AML

Until recently, whole genome-wide study was limited due to the cost and inefficiency of data analysis and interpretation. As a result of significant advancement of whole genome or exome sequencing platform and advanced computing power, recurrent disease-causing mutation study is presently relatively easy to carry out. Clonally proliferated immature myeloid cells exhibit recurrent mutation in the genome (Mardis et al., 2009; Patel et al., 2012). It has been found that many of those mutated genes are directly or indirectly associated with epigenetic patterns alteration in AML, such as aberrant DNA methylation, histone modification, and chromatin remodeling (Shih, Abdel-Wahab, Patel, & Levine, 2012). Mutation of a DNA methyltransferase (DNMT3A) showed the most prominent association with the DNA methylation pattern by hypomethylating promoter CpG islands (CGIs) of AML genome (Qu et al., 2014). Lysine acetyltransferases (KATs) are an important group of proteins responsible for histone modifications. Nuclear subtypes of KATs are CBP and p300 that have been found recurrently mutated in hematological malignancy (Pasqualucci et al., 2011). Deregulation of chromatin remodeling gene SMARCA5 has been noticed in AML (Stopka et al., 2000). Many of those epigenetic regulator associated mutant genes do not belong to the classes of the "two-hit-hypothesis" genes in leukemia (Figure 2.1) in which class I genes lead to uncontrolled cell proliferation and class II genes lead to inhibition of differentiation (Kelly & Gilliland, 2002). This observations suggest the inadequacy of the two-hit-hypothesis concept in AML (Shih et al., 2012).

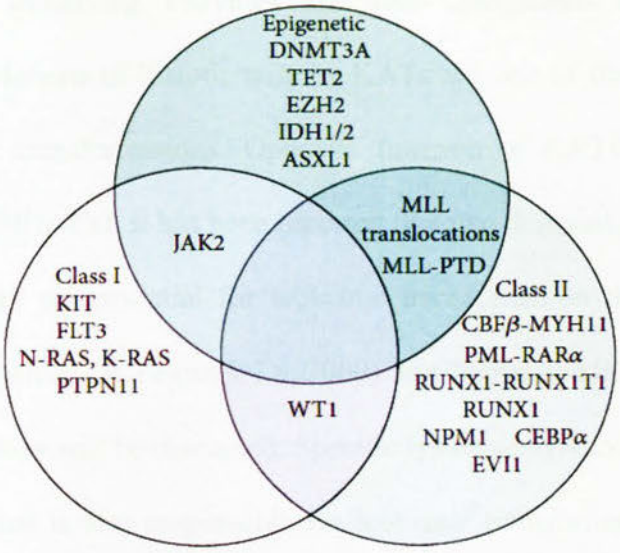


Figure 2.1: Key Mutations in AML (Conway O'Brien et al., 2014)

DNA methyltransferases function in adding a methyl group (-CH3) to the five-carbon position of cytosine bases in CpG dinucleotides and form 5-methylcytosine (5mC). In general, this transformation of cytosine leads to transcriptional silencing to maintain gene expression patterns in normal cells (Bird, 2002). Deregulation of DNMT3A can cause aberrant DNA (hyper or hypo) methylation of promoter CGIs followed by downregulation of tumor suppressor genes and upregulation of oncogenes resulting tumorigenesis. Distinct methylation patterns have been observed for different subtypes of AML (Saied et al., 2012). Such patterns could be useful for AML classification, especially for the intermediate risk group. In Chapter 4, novel DNA methylation patterns have been identified in AML patients. Interestingly, DNA methylation is not limited to CGI only, differentially methylated regions have been found in open seas and shores of the promoters that control cellular differentiation and pluripotency (Doi et al., 2009; Meissner et al., 2008). Deregulated methyltransferase genes mediated hyper-variabilities provide survival advantage of malignant cells through aberrant (oncogenic / tumor suppressive) genes expression (Timp & Feinberg, 2013).

Many chromatin modifying enzymes are also deregulated during oncogenic transformation. Acetylations of histone tails by KATs are one of the common histone modifying epigenetic transformations. Opposite function of KATs is performed by histone deacetylases (HDACs). It has been reported that two domains (KAT domain and bromodomain) of CBP are essential for leukemic transformation through MLL-CBP fusion (Lavau, Du, Thirman, & Zeleznik-Le, 2000). In Chapter 5, MLL fusion-mediated leukemic transformations will be discussed. Specific lysine acetylation activates AML1-ETO fusion protein that is also responsible for leukemic transformation (Wang et al., 2011). Although HDACs mutations are rarely noticed, their aberrant recruitment by oncofusion proteins (PML-RARA, PZLF-RARA) have been reported that leads to aberrant gene silencing (especially tumor suppressor genes) (Grignani et al., 1998; HE LZ & Tribioli, 1998; Lin et al., 1998). Histone modification may also be caused by phosphorylation of serine, threonine and tyrosine residues. Kinases and phosphatases are the two opposite ways for the addition or removal of phosphate residues that lead to changes in the net charge of amino acids of histone proteins. These balances of charges regulate gene transcription, DNA repair, and chromatin condensation. Aberrant kinase activities have been associated with malignant transformation (Hanahan & Weinberg, 2011). C&Ckines-mediated activation of JAK2 phosphorylates histone H3Y41 thereby leading to activation of oncogene LMO2 (Dawson et al., 2009). Therefore, epigenetic regulation is an emerging area to understand differential oncogenic mechanism and inhibition of specific pathway that could be utilized for possible therapeutic intervention.

2.2.3 Polymorphism in AML associated genes

Endogenous and exogenous risk factors are responsible for causing damage to DNA, whereas DNA repair proteins, oxidative stress response proteins, and other protective

mechanism are active in the cells to compensate for possible damages. Any imbalance between risk factors and protective mechanisms may cause genetic lesions like gene mutation and nucleotide polymorphism. Genetic polymorphisms of some key pathway (DNA repair genes and damage-detoxification genes) have frequently been reported in AML cases that lead to genomic instability, chromosomal aberration, and abnormal phenotypic changes. Polymorphism may alter disease susceptibility by affecting protein function, promoter activity, mRNA stability, and miRNA binding capacity (Seedhouse, Faulkner, Ashraf, Das-Gupta, & Russell, 2004). Polymorphisms in miRNA binding sites and encoding sites are also found in hematological malignancy and those polymorphisms are responsible for the deregulation of miRNA network (Dzikiewicz-Krawczyk, 2014; Ramsingh et al., 2010). A single point mutation in miRNA may cause butterfly effect on alteration of global gene expression by cascade mechanism (Lee, Yang, Chen, & Au, 2011). In Chapter 5, novel polymorphism of miRNA encoding genes has been explored in detail. In human, more than 60% of protein-coding genes are directly or indirectly regulated by miRNAs (Friedman, Farh, Burge, & Bartel, 2009). Regulatory roles of miRNAs have been reported for cell proliferation, apoptosis, differentiation, and metabolism (Alvarez-Garcia & Miska, 2005; Bartel, 2004). Mutation of miRNA encoding genes may contribute to oncogenic transformation. Cancer-associated miRNA encoding genes are mostly located to fragile sites of the chromosomes (Calin et al., 2004). miRNA expression patterns vary among leukemic subtypes (AML and ALL) (Mi et al., 2007) and even cytogenetic and molecular subtypes of AML (Dixon-McIver et al., 2008; Jongen-Lavrencic, Sun, Dijkstra, Valk, & Löwenberg, 2008; Li et al., 2008). Those patterns of miRNAs expression may have prognostic impact in cancers, especially in AML. Individual miRNAs (miR-181a) may also supply prognostic information for cancer (Marcucci, Mrózek, et al., 2011). Some miRNAs can work as oncogenic and other as a tumor suppressor. Regulation of miRNA expression, like suppression of oncogenic

miRNAs, may have a therapeutic implication. So, miRNA-based therapeutic approaches in AML would be a potential step to control oncogenic transformation.

Translational repression causing somatic mutation in TNFAIP2 gene has been reported in AML patients (Ramsingh et al., 2010). Polymorphism of IL-17F (rs763780) has been found to be associated with predisposition of AML (Wróbel et al., 2014). Polymorphisms of the homologous recombination repair genes (RAD51and XRCC3) can increase the risk of AML development (Seedhouse et al., 2004). Polymorphism in KRT81 (rs3660, G>C) and XPO5 (rs11077, A>C) may affect multiple miRNA binding that respond to alter protein expression in leukemia (de Larrea et al., 2012). Recently, many acquired genes mutation has been reported (NPM1, CEBPA, RUNX1, WT1, DNMT3A, ASXL1, IDH2 and FLT3) that exhibit prognostic significance (Gaidzik et al., 2011; Ley et al., 2010; Schlenk et al., 2008; Schnittger et al., 2002). Double mutation of CEBPA can predict the best prognosis like PML-RARA and mutation of RUNX1 is associated with unfavorable prognosis outcome (Grossmann et al., 2012). A Large number of intermediate risk AML patients can be reclassified by incorporating genetic mutation data with cytogenetic data, which is a novel initiative for correct risk stratification of AML patients (Hou et al., 2014).

2.2.4 Cytogenetic abnormalities

Chromosomal aberrations are often known as cytogenetic abnormalities that are usually found in leukemic blasts. AML associated cytogenetic and molecular genetic abnormalities are not mutually exclusive and in most of the cases, multiple abnormalities are required to develop cancer. Although cytogenetic abnormalities are considered rather dated in this genomic era, it is still being used routinely in the clinic and are still the best

prognostic factors for induction response and overall survival of AML patients (Grimwade, 2001). Almost 50-60% of newly diagnosed AML cases have shown cytogenetic abnormalities (Martens & Stunnenberg, 2010). One of the popular classifications (FAB) of AML is based on microscopic morphology of the blasts. Recently, the new classification for AML has been proposed that demonstrate better prognostic significance (Grimwade et al., 2010). Different types of cytogenetic abnormalities have been reported in AML such as chromosomal translocation, inversion, deletion, and copy number variation. More than 200 different types of chromosomal translocations and mutations have been reported in AML patients (Löwenberg, Griffin, & Tallman, 2003). Based on TCGA AML dataset, some common cytogenetic abnormalities are inv (16), del (5q), del (7q), trisomy 8, trisomy 21, t(4;11), t(8;21), t(15;17), t(9;11), and t(9;22). In most cases, nonrandom chromosomal translocations result in gene rearrangement and generate fusion or chimeric proteins (like PML-RARA, RUNX1-RUNX1T1, CBFB-MYH11, and MLL-GAS7) that regulate transcription factor to control gene expression and ultimately lead to oncogenic transformation. PML-RARA is the most frequent translocation event in AML. Until 2010, around 85 recurrent translocation events have been reported for mixed-lineage leukemia (MLL) gene alone with 65 partner genes. In chapter 5, the MLL-GAS7 fusion protein study has been described.

The PML-RARA fusion gene created by the translocation event t(15;17)(q22;q12~21) and the fusion protein binds to the HDAC complex, which is a corepressor and is responsible for aberrant chromatin acetylation and alteration of normal gene expression patterns in AML. One of the successful therapeutics against PML-RARA is all-trans-retinoic acid (ATRA) that changes the conformation of the chimera, dissociate HDAC complex and activate downstream target genes expression (Jing, 2004). Since many chemotherapeutic drugs have been administrated by targeting PML-RARA, the overall

survival rate of this translocation positive patient is higher than others. Another complex fusion protein partner is MLL, which is lethal particularly for children. Deregulated MLL can enhance other cooperative mutations that lead to oncogenic transformation. He et al. noticed that a small molecule (MI-2-2) can inhibit menin-MLL integration, an essential step for leukemic transformation, for all partner protein of MLL (He et al., 2015). Genetic and cytogenetic aberrations targeted therapies are required to improve overall patient survival. Also, it is important in categorizing AML patients into different risk groups, not only based on cytogenetic factors but also according to recently identified genetic factors. Improved genetic analysis tools like microarray-based gene expression study and cost effective sequencing techniques are the promising diagnosis and risk stratification prospects for AML patients.

CHAPTER 3: ABNORMAL EXPRESSION OF CYTOKINES AND CHEMOKINES IN ADULT MALAYSIAN AML PATIENTS

3.1 Introduction

Crosstalk between tumor cells and their microenvironment is considerably essential for tumor development, progression, and patient survival (Hanahan & Weinberg, 2000). Mediators of crosstalk are some small molecules such as cytokines, chemokines, and growth factors (C&Ckines) (Ilkow et al., 2015). Tumor microenvironments are rich in inflammatory cells that secrete C&Ckines to promote cellular growth (Coussens & Werb, 2002). Association between inflammatory C&Ckines-mediated innate immunity and cancer are well known (Balkwill & Mantovani, 2001). Both leukemic blasts and immune cells can produce C&Ckines and growth factors for different functions in AML patients. C&Ckines and growth factors bind to cell surface receptors and supply specific information to the targeted cells that cause a specific response. During abnormal conditions (like cancer), C&Ckines mediated information trigger cells and their responses are altered from normal condition to tumor progression. Aberrant signal transduction pathways (STP) mediated oncogenesis are well documented for leukemia (Van Etten, 2007) where a mutation in the receptors or STP components enhance leukemic blasts formation (Hayakawa et al., 2000). But mutations in the receptors and STP components are not prominent in AML (Loriaux et al., 2008; Tomasson et al., 2008), rather abnormal stimulations through intact receptors and STP component are frequent in blasts that ensure proliferation and survival by blocking differentiation and apoptosis (Lotem & Sachs, 1996). Therefore, C&Ckines and growth factors-mediated oncogenic transformation could happen probably without genetic abnormalities in the pathway components that correspond to large intermediate risk AML group.

Since deregulated C&Ckines and growth factors expression levels have been reported in AML plasma, there might be distinct expression patterns which may correspond to

differential diagnostic as well as prognostic outcomes. There might also be expression differences based on gender, age, and ethnicity. Identification of such patterns could help to better understand of AML pathogenesis. Until recently, the most popular way to measure C&Ckines and growth factors levels in plasma or serum was enzyme-linked immunosorbent assay (ELISA) and most studies measured only a few C&Ckines due to technical limitations. As C&Ckines mediated signaling networks are complex and multiple C&Ckines can work for the same pathway, precise and simultaneous detection of large number of C&Ckines are required for association studies. Luminex system based multiplex technology is able to detect up to 100 C&Ckines and growth factor at a time from a single sample (De Jager, te Velthuis, Prakken, Kuis, & Rijkers, 2003). In this case-control study, 32 C&Ckines and growth factors have been quantitatively measured from 38 plasma samples. So far, this is the first integrated C&Ckines and growth factors profiling in adult Malaysian AML patients and healthy controls.

3.2 Literature review

Deregulated C&Ckines and growth factors are often the cause for diseases like leukemia. The presence of IL-6 and IL-1β in the tumor microenvironment is the signature of cancer associated inflammation (Sanchez-Correa et al., 2013) where IL-1β promotes angiogenesis and invasive capacity (Apte et al., 2006) while IL-6 stimulates AML blast growth through an IL-6 receptor (Säily, Koistinen, & Savolainen, 1999). Serum IL-6 level is a prognostic marker for diffuse large cell lymphoma (Preti et al., 1997). Plasma hepatocyte growth factor (HGF) has been reported as a prognostic factor in AML except myelodysplastic syndrome (Verstovsek et al., 2001). Elevated serum prolactin levels have been found in patients with advanced multiple myelomas (Gado et al., 2002). Leukemia patients have been reported to have higher than normal levels of IL-3, granulocyte-

macrophage colony stimulating factor (GM-CSF), IL-6, IL-8, and IFN-γ (Elbaz & Shaltout, 2000; Hsu et al., 2002). AML blast proliferation can be inhibited by IL-10 (Ø Bruserud et al., 1995) whereas IL-2 and sIL-2RA positively stimulate AML blast proliferation (Kornblau et al., 2010). TNF-α also inhibits AML blast proliferation through regulation of CSF3 receptor (Elbaz et al., 1991). Kornblau et al. reported the deregulation of 18 C&Ckines out of 27 in AML serum. Using Bioplex methods, they also showed that the serum expression profile variation between AML and MDS (myelodysplastic syndrome) were statistically insignificant for 24 out of 27 C&Ckines except for IL-8, IL-13, and VEGFA (Kornblau et al., 2010). Many C&Ckines are already established biomarkers for different cancers (like PSA for prostate cancer and MIF for breast cancer). Here, 10 such C&Ckines (which are common for other cancers) were detected for the first time in plasma of AML patients and exploratory analysis has been performed. Publically available peer-reviewed literature have been searched for 32 C&Ckines and growth factors for both AML cases and healthy controls and comparative expression levels are shown in Appendix A (Table A1 and Table A2).

3.3 Materials and methods

3.3.1 Study population

In this study, blood specimens were collected from two major ethnic groups (Malay and Chinese) of Malaysia. Suspected blood cancer patients' specimens were sent to the Hematology Unit at the Institute for Medical Research (IMR), Malaysia. Those were newly diagnosed with AML and no treatments received yet were considered for the analysis (n=22). Patients were excluded in this study if he/she was less than 20 years and older patient was 65 years old. Due to quality control, 3 AML samples were also excluded. In addition, age (20-65 years) and ethnicity matched 19 (M=11, F=8) healthy

controls were also used for this study. Final numbers of the study population were 19 (M=15, F=4). Control subjects were ensured to be free from any types of fever for at least 1 week before testing, had not taken any medications, were not pregnant, and were free from known chronic or acute diseases. Conventional cytogenetic profiling was carried out for all the patients (Table 3.1). Patient's written informed consent was taken for research purpose only. All experimental protocols and medical ethics were approved by the Medical Ethics Committee (NMRR-16-1384-31900). Sample distributions are shown in Figure 3.1.

Table 3.1: Patients cytogenetic profiling.

Cytogenetic	Patients Distribution
t(8;21)	15.72%
inv(16)	10.51%
Tri-11	5.26%
Tri-21	5.26%
t(15;17)	5.26%
Normal	57.89%

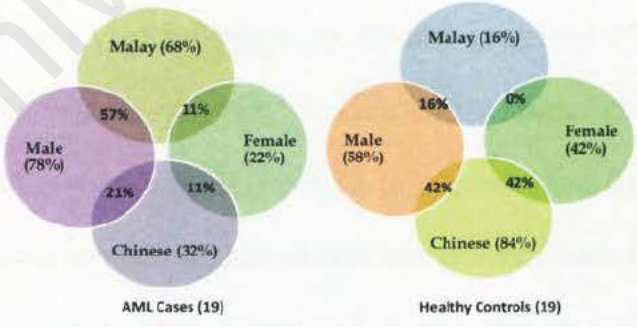


Figure 3.1: Distribution of samples for plasma C&Ckines detection.

3.3.2 Plasma collection

A total of 2-ml peripheral blood was collected from both AML cases and healthy controls in plasma separating (purple top) tubes and stored in the ice box for quick transportation to the lab. Immediately after arrival to the lab, the tubes were centrifuged at 2000 rpm for 10 min to collect fresh plasma. Plasma was aliquoted and frozen at -80^0 C for further analysis.

3.3.3 Cytokines and chemokines profiling

To measure the cytokines and chemokines concentration in plasma, the multiplex analysis was performed using the Luminex 200™ on a Bioplex (Bio-Rad) cytometer which enabled simultaneous detection and quantification of multiple proteins per sample. One antibody specific for a single plasma protein is bound to magnetic beads with distinctive fluorescence. All the antibodies were mixed together and used for each sample. The magnetic beads used for this study were supplied by Millipore. Technical details and procedures are given at www.emdmillipore.com. MILIPLEX MAP Human Circulating Cancer Biomarker Magnetic Bead Panel One and Three (HCCBP1MAG-58K and HCCBP3MAG-58K) were used for this study. Panel-1 contained 25-Plex and panel-3 contained 9-Plex. Due to cross-reactivity concern by the manufacturer however, one antibody was excluded from each panel (free PSA from panel-1 and Catalase from panel-3). A total of 32 proteins (cytokines, chemokines, and growth factors) were measured in each sample. The cytokines, chemokines, and growth factors (C&Ckines) were as follows: AFP, Total PSA, CA 15-3, CA 19-9, MIF, TRAIL, Leptin, IL-6, sFasL, CEA, CA 125, IL-8, HGF, sFas, TNGa, Prolactin, SCF, CYFRA 21-1, OPN, FGF2, bHCG, HE4, TGFa, VEGF; Galectin-3, Cathepsin D, FAP (Fibroblast Activation Protein Alpha), MIA (Melanoma Inhibitory Activity Protein), MIA (Myeloperoxidase), SHBG, IGFBP3,

and Ferritin. Patient plasma was diluted (1:6 for panel-1 and 1:50 for panel-3) and mixed with magnetic beads. Plates were incubated overnight at 4° C with shaking (500 rpm), washed 3X with wash buffer and incubated for 1 hour with detection antibodies. Streptavidin-phycoerythrin was then added and incubated for 30 min, washed 3X and sheath fluid was added for 5 min then read using calibrated Luminex 200™ system on a Bioplex (Bio-Rad) cytometer. Two panels were assayed in two different 96-well plates for the same samples with quality controls (n=2) and standards (n=7). All the samples, quality controls, and standards were run in duplicate and the mean values (after adding dilution factor) of the two readings were used for calculations and subsequent analysis.

3.3.4 Statistical analysis

After measurement of fluorescence intensity by Bioplex cytometer, a digital processor captured the raw data and Bioplex Manager (Version 6.1) was used to analyze the data. Each 96-well plate contained negative controls (duplicate) that were used as background intensity. In order to draw a standard curve, 7-wells (duplicate) of consecutive 3-fold dilution of known standards were used to fit a 6 parameter logistic curve for each C&Ckines. The intensity of the magnetic beads was compared to the corresponding fitted curve to measure the concentration of the proteins in each duplicate samples and mean concentration was calculated from the duplicate. If the intensities of the samples were beyond the standard curve range, the values were considered as out-of-range above (OOR>) or out-of-range below (OOR<) and the values were excluded. Some proteins (IL-6, IL-8, bHCG, TGFa, and VEGF) were below the range (OOR<) mostly in some healthy controls. Outliers (density-based approaches) were removed, and the expression values were normalized to maintain standard range according to literature. Intra and inter panel's units of expression were different. In panel-1, three proteins (CA 15-3, CA 19-9, and CA

125) were in U/ml scale, one protein (bHCG) was in mU/ml scale, and all other proteins were in pg/ml scale. In panel-3, one protein (SHBG) was in nM scale and the rest of them were in ng/ml scale.

R (version 3.1.3) platform (www.r-project.org) and RStudio were used to install and initiate other R/Bioconductor packages for various analyses and calculations. Student t-tests were used to estimate the mean plasma protein concentration differences between AML cases and healthy controls. Pearson correlation, ANOVA, MANOVA, and Chi-squared tests were also performed to analyze and interpret the data from different angles.

To validate the deregulation of plasma protein concentration, differential and exploratory analysis was performed using The Cancer Genome Atlas (TCGA) data (AML) (Network, 2013). Promoter methylation of the corresponding genes were analyzed by RnBeads (http://rnbeads.mpi-inf.mpg.de/) package (Assenov et al., 2014) that runs many statistical tests automatically (Fisher's exact test, Wilcoxon rank-sum test, Kruskal-Wallis one-way analysis or Pearson correlation test) based on data type in order to identify significant associations among the sample annotations and validate the analysis. Correlation between DNA methylation and corresponding genes expression in TCGA AML patients were analyzed with Spearman's rank correlation. In the case of survival analysis based on cytogenetic risk factors, the p-value was adjusted (age and gender) and the significance level of this study was set at $p<0.001$. Enrichment analysis was performed and significant gene ontology (GO) terms were determined by hypergeometric test.

One-way ANOVA was used to identify informative C&Ckines and bimodality index (Wang, Wen, Symmans, Pusztai, & Coombes, 2009) to identify whether C&Ckines follow bimodal expression pattern. To cluster the samples and C&Ckines, different strategies (silhouette width, k-means, hierarchical clustering with Euclidean distance and

complete linkage, principal component analysis, multidimensional scaling) were taken. Targeted pathway analysis was performed for AML using different tools (GeneMANIA and KEGG) and attempts were made to correlate findings with existing literature.

3.4 Results

3.4.1 Sample quality and experimental validation

All the samples were collected within 2 months and after plasma extraction, were stored at -80° C until use without freeze/thaw cycles in order to avoid the possibility of the proteins being denatured or be exposed to any other contamination. As plasma was extracted shortly after blood collection, there was no hemolysis and the plasma color was straw-yellow. The same instruments were used for both panels and a single person was handling the samples in order to minimize the batch effect. For the standard, all the expected and observed concentration ratios were within acceptable range (0.8-1.2). Both quality controls (QC) were generally within the expected range of the corresponding proteins. Inter-sample variability was relatively higher, perhaps due to the presence of different AML subtypes.

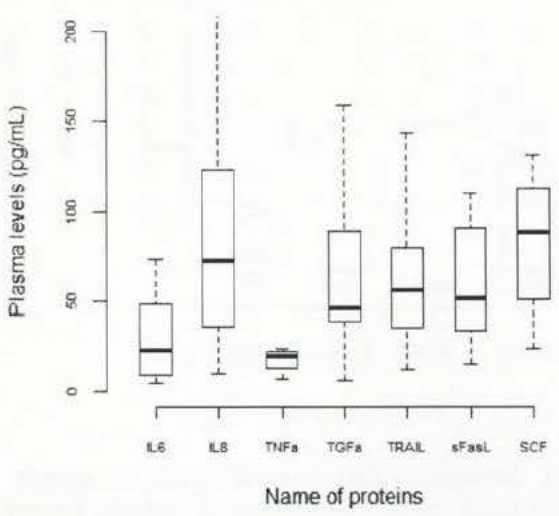
3.4.2 Comparison of expression in healthy controls with literature

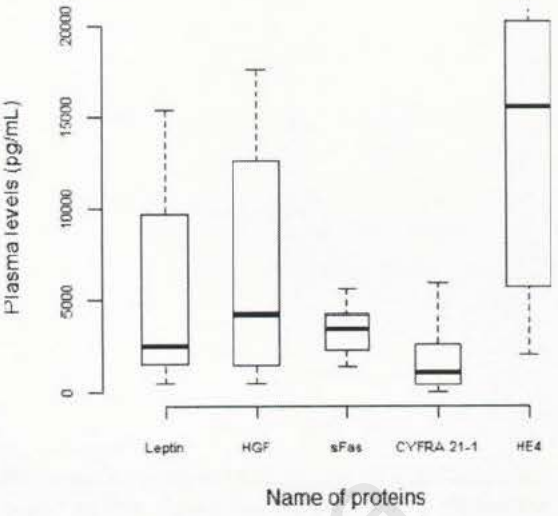
Peer-reviewed literature was searched for the expression level of the selected 32 C&Ckines in plasma (some cases in serum if plasma level were not available) in the healthy control samples (APPENDIX A: Table A1) in order to determine whether this multiplex assay was generating comparable results. Most of the early studies were conducted using ELISA based techniques. Although there were excellent correlations between ELISA and other methods with Multiplex in the case of some specific proteins (Christiansson et al., 2014; De Jager et al., 2003), notable variations were also found (Table A1). The variations of normal plasma proteins level in healthy subjects were found among inter-methods and intra-methods. This heterogeneity of expression level might be

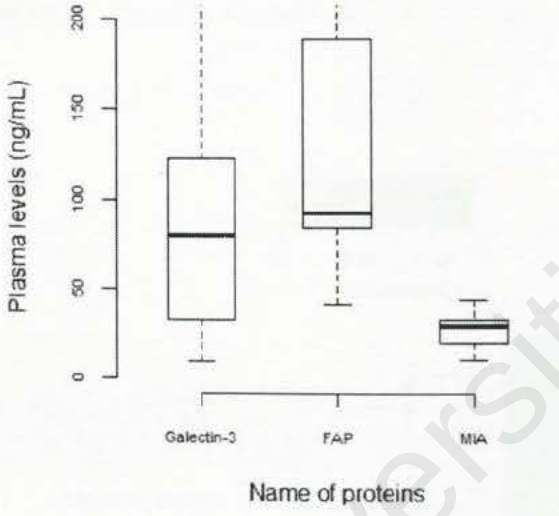
due to several factors including assay sensitivity, age, gender, ethnicity and batch effect. Due to limited data availability in the literature, it was not possible to perform a statistical comparison for all the proteins, rather some selected C&Ckines. Generally, the Bioplex results were more consistent with other studies using Bioplex as compared to ELISA or other methods. Although it was noticed that expression level of some C&Ckines (median) were lower (Total PSA, MIF, sFasL, SCF, and IGFBP3) and higher (CA 19-9, IL-6, Prolactin, FAPa and MIA), in general, the range of expressions overlapped with those in the literature. Standard deviation (SD) was relatively high but it supports other literature (Kleiner, Marcuzzi, Zanin, Monasta, & Zauli, 2013; Kornblau et al., 2010) that had used the same Bioplex methods.

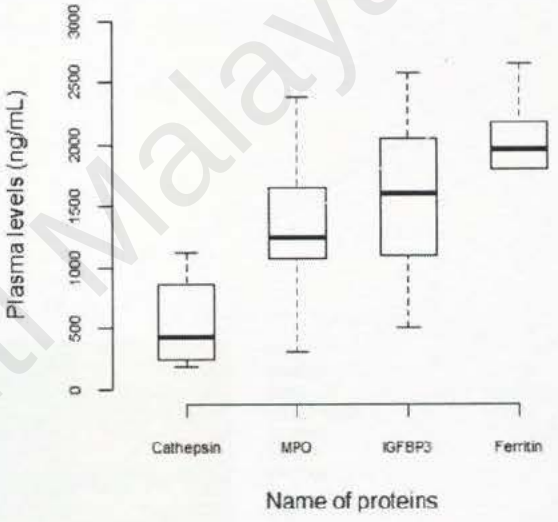
3.4.3 Comparison of expression in AML cases with literature

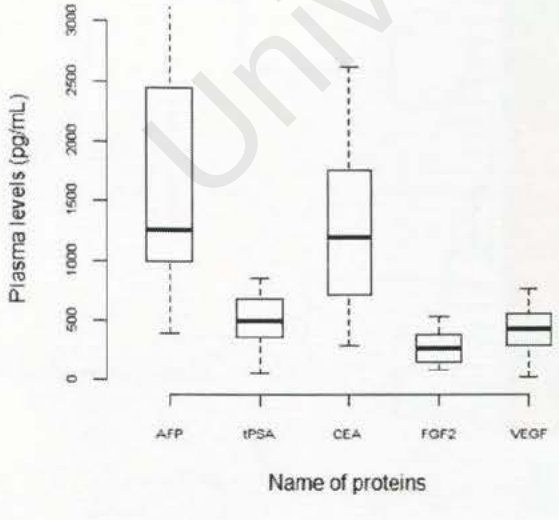
Peer-reviewed literature was also searched for AML cases which measured the expression levels of the selected C&Ckines. However, we noticed many (31.25%) of the listed C&Ckines have never been measured in AML cases. Most of the publications measured a limited number of C&Ckines using ELISA methods and much of the relevant information are still missing. Relevant information was extracted from those papers and these were compared with Bioplex system. The results of this study are shown in APPENDIX A: Table A2. C&Ckines expression ranges have shown as box plot for AML cases (Figure 3.2). Ten C&Ckines (Total PSA, CA 19-9, MIF, TRAIL, CYFRA 21-1, HE4, Cathepsin D, FAPa, MIA, and SHBG) levels that are missing /unknown in literature could be considered as standard AML plasma concentration as detected using Bioplex system (Table A2). For five of those C&Ckines (MIF, TRAIL, Cathepsin D, FAPa, and SHBG), their levels have been found to be significantly deregulated in AML cases as compared to healthy controls.

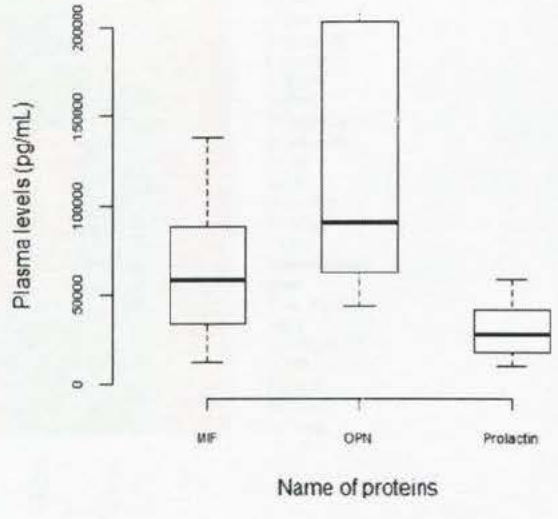












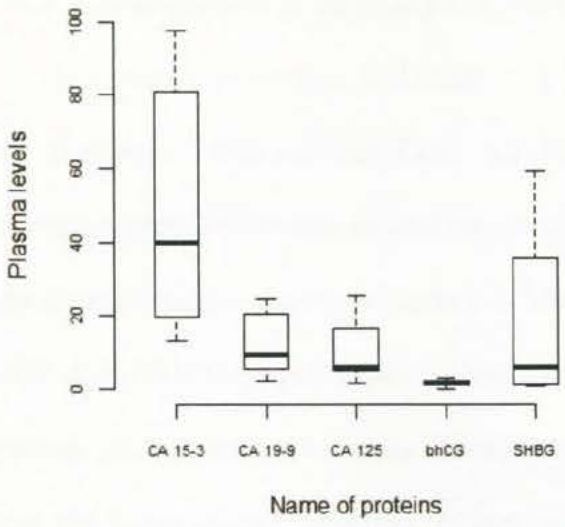


Figure 3.2: C&Ckines expression ranges for AML cases. Closely related ranges of C&Ckines are organized together after removing outliers. Only last panel of this figure has no scale, as it varies for different C&Ckines (CA- U/ml, bHCG- mU/ml, SHBG- nM). Details scales, ranges, and median values are shown in Table A2 in Appendix A.

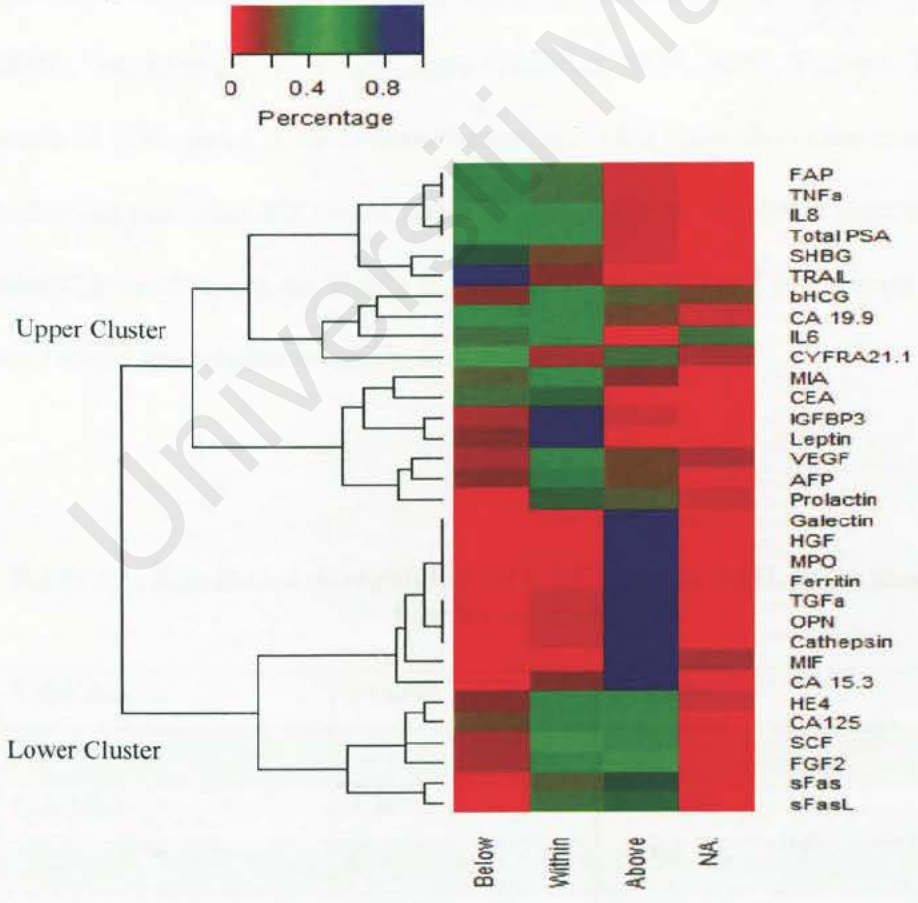


Figure 3.3: C&Ckines expression level compared to normal range. Two big clusters can be noticed where the upper cluster is either within or below the normal range and the lower cluster is either within or above the normal range. Above: above the normal expression range, below: below the normal expression range, within: within normal expression range, and NA: missing values that were out of detection limit.

3.4.4 Comparison of expression in AML cases and healthy controls

The expressions (median and range) of C&Ckines in healthy controls and AML cases are shown in Table A1 and Table A2. Percent of AML cases which show normal expression pattern (healthy control range) for each C&Ckines or deviation (below/above the normal range) is shown in Figure 3.3. Missing values are comparatively higher in case of IL-6 and bHCG (mostly OOR<). Dendrogram shows two distinct clusters of expression pattern of C&Ckines. The upper cluster (n=17) is within normal or below normal range and the lower cluster (n=15) is within normal or above the normal range. Although IGFBP3 and Leptin are shown in the upper cluster, they are distinct from other C&Ckines of this cluster (more that 80% AML cases belong to the normal range of expression for these two proteins). TRAIL, SHBG, and FAPa expression level in AML cases were significantly lower (Table 3.2) as compared with the healthy control group (Student t-test $P<0.0001$). Interestingly, nine C&Ckines (Galectin, HGF, MPO, Ferritin, TGFa, OPN, Cathepsin D, MIF, and CA 15-3) from the lower cluster show above the normal range of expression (in more than 80% AML cases) that could be an important signature for AML prognosis. Seven of them are highly significant ($p<0.0001$) and the remaining two (CA 15-3 and HGF) are relatively significant ($p<0.005$).

Table 3.2: Significant deregulation of C&Ckines in AML cases compared to healthy controls.

C&Ckines	t-value	p-value
Upregulation		
CA 15.3	3.3056	0.00469
MIF	6.5031	1.39E-05
sFasL	12.9469	2.99E-06
IL8	5.247	5.38E-05
HGF	3.4722	0.003412

sFas	6.2874	3.22E-06
TNFa	8.09	2.28E-07
OPN	4.738	0.0001499
TGFa	5.3454	5.76E-05
Galectin-3	5.7781	3.62E-05
Cathepsin D	4.936	0.0001424
MPO	10.503	7.37E-09
Ferritin	6.7283	0.0005088
Downregulation		
TRAIL	-9.7065	2.52E-11
Leptin	-3.361	0.002666
FAPa	-4.5352	6.48E-05
SHBG	-5.0951	3.36E-05

There are no reports before of MIF and Cathepsin D levels been measured in AML cases (Table A2). The role of MIF has been reported for ovarian cancer (Agarwal et al., 2007; Krockenberger et al., 2012; Tas, Karabulut, Serilmez, Ciftci, & Duranyildiz, 2014) and breast cancer (Ciftci et al., 2014; Richard, Kindt, & Saussez, 2015; Verjans et al., 2009; Xu et al., 2008). Deregulation of Cathepsin D level also has been reported in ovarian cancer (Baekelandt, Holm, Trope, Nesland, & Kristensen, 1999; Scambia et al., 1991; Scambia et al., 1994) and breast cancer (Foekens et al., 1999; Garcia et al., 1996; Johnson, Torri, Lippman, & Dickson, 1993). It was checked to find out whether expression of those two proteins were gender specific, but no significant differences were noticed between male and female ($p=0.341$ and $p=0.716$ for MIF and Cathepsin D respectively) in this study. Significant promoter hypomethylation ($\beta<0.25$) of their corresponding genes (MIF and CTSD) were also noticed in 194 AML cases in TCGA (Network, 2013). There was no significant methylation difference found between male and female (Figure 3.11). Gene expression microarray data from TCGA also demonstrated higher expression level for MIF and Cathepsin D (CTSD) (Figure 3.13).

This finding raises a question as to whether any overlap exists between AML pathway with ovarian and breast cancer pathway. In order to answer this, pathway analysis was performed to determined possible pathways that might be similar between the three types of cancers. Interestingly, MIF is involve in breast cancer through the PI3K/Akt pathway (Richard et al., 2015). Receptors for MIF are CD74 and CXCR-4 and other intermediary proteins like Hsp90 stabilize MIF. HIF-1a upregulates MIF level, and overexpression of HER-2 receptor is also reported. Other well-established proteins (EGFR and VEGF) were also considered for this exploratory analysis, and it is believed that these are also possibly involved through PI3K/Akt pathway. Their corresponding genes expression pattern in AML, breast, and ovarian cancer were then checked. Notably, all three cancer types showed almost similar expression pattern for PI3K/Akt pathway associated genes (Figure 3.4) with no significant expression differences noticed (chi-sq=1.393, df=9, and p-value=0.9979). Cathepsin D is involved in breast cancer through M6PR receptor and downregulates Caspase 3 (Rochefort, Capony, & Garcia, 1990) and similar fashion was noticed in AML gene expression data. MIF works in ovarian cancer by downregulating NKG2D (KLRK1 gene) level (Krockenberger et al., 2008) and microarray data supports this notion even in AML cases (Figure 3.4). This data support the hypothesis of possible overlapping pathways.

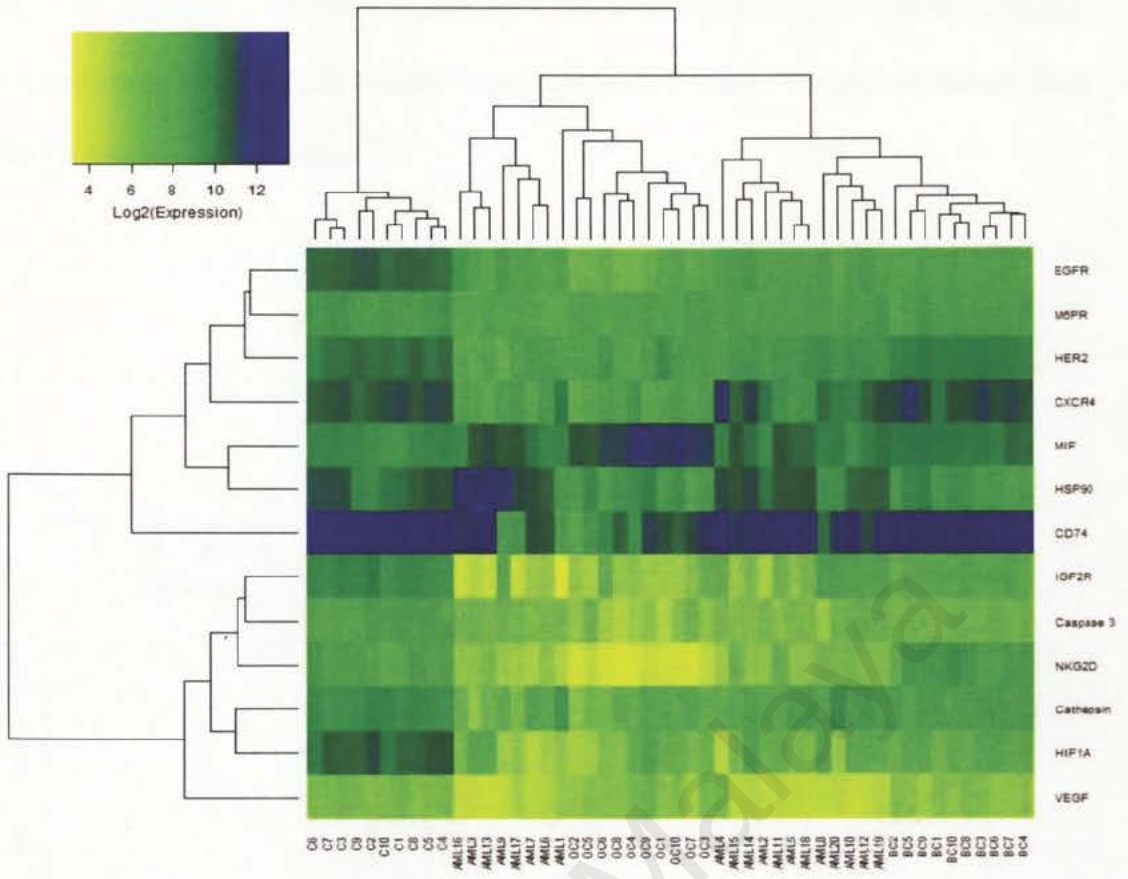


Figure 3.4: MIF and Cathepsin D pathway associated genes expressions. Caspase 3 and M6PR are associated to Cathepsin D and rests of the genes are associated to MIF. Controls (C), cases (AML), ovarian (OC) and breast (BC) cancer data are shown.

3.4.5 Correlated expression analysis

Correlated expressions have been noticed for some C&Ckines in AML cases and healthy controls. Pearson's correlation test was performed and some C&Ckines show statistically significant ($r > \pm 0.8$) expression levels (Table A3). Five pairs of C&Ckines (SCF~HE4, SCF~bHCG, bHCG~HE4, HE4~VEGF, and OPN~Ferritin) in AML cases, one pair of C&Ckines (FGF2~HE4) in controls, and three pairs of C&Ckines (SCF~HE4, bHCG~HE4, and HE4~VEGF) in both cases & controls exhibited strong linear correlation of plasma protein expression (Figure 3.5). Only one pair of C&Ckines (OPN~Ferritin) was showing strong negative correlation ($r = -0.958$) in AML cases. All the pairs were shown to have relatively strong coefficient of determination ($R^2 > 0.7$) and their regression models are also shown in Table A3. Mutually correlated expressions were

noticed for multiple C&Ckines (OPN~SCF~HE4 and bHCG~SCF~HE4). Although correlated expressions are statistically significant in this study, the reasons behind these linear correlations are not clear.

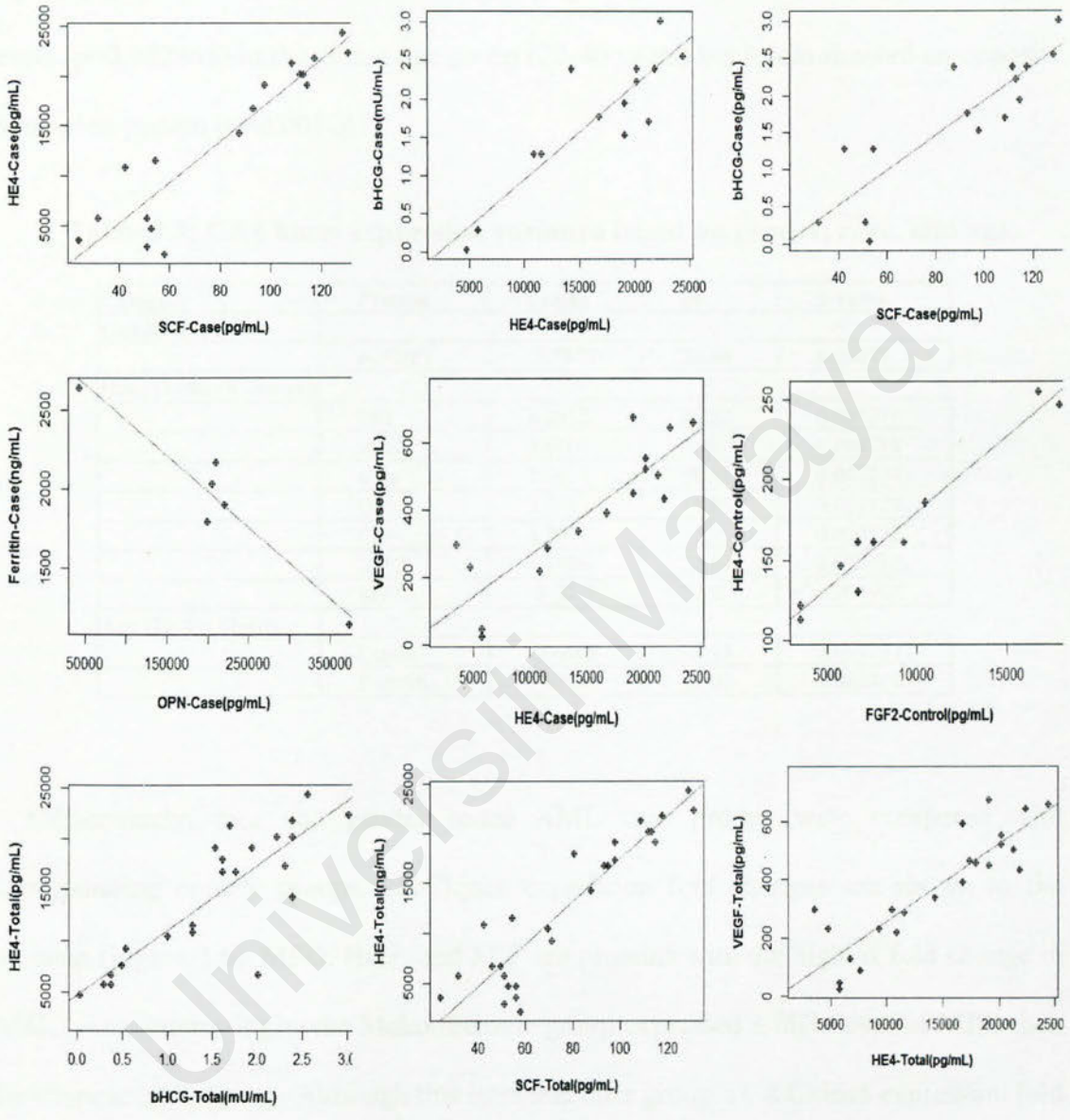


Figure 3.5: Linear correlation of C&Ckines expression with other C&Ckines.

3.4.6 C&Ckines expression variance based on groups

AML cases were divided based on gender (M/F), race (Malay/Chinese), and age (20-40/41-65 years) and the C&Ckines expression was compared between those two groups to see whether there were any significant difference in expression were found (Table 3.3). IGFBP3 expression in male group was significantly lower than female group

(p=0.00078). Six C&Ckines (MIF, sFasL, sFas, OPN, Cathepsin D, MPO, and SHBG) were highly expressed in the Malay ethnic group (p = 0.001016, 0.008768, 0.007034, 0.003129, 0.000572, and 0.000382, respectively) and one C&Ckines (SHBG) were expressed significantly in the Chinese group (p=0.001966). Ferritin expressed higher levels (p=0.002466) in the young age group (20-40 years) but leptin showed an opposite expression pattern (p=0.001267).

Table 3.3: C&Ckines expression variance based on gender, race, and age.

Category	Protein	t value	df	p-value
Gender (M/F)				
	IGFBP3	-3.7877	26.84	0.00078
Race (Malay/Chinese)				
	MIF	4.0012	16.12	0.001016
	sFasL	3.6104	6.93	0.008768
	sFas	2.9037	28.61	0.007034
	OPN	3.508	15.15	0.003129
	Cathepsin D	4.2971	15.75	0.000572
	MPO	4.1696	22.54	0.000382
	SHBG	-3.5017	22.5	0.001966
Age (20-40/ 41-65)				
	Leptin	-3.6685	23.13	0.001267
	Ferritin	3.8973	11.05	0.002466

Subsequently, race and gender based AML case groups were compared with corresponding control groups. C&Ckines expression fold changes are shown in the heatmap (Figure 3.6). MPO, HGF, and MIF are proteins with the highest fold change in AML cases. Interestingly, the Malay-Female group expressed a high level of MPO than the Chinese-Male group. Although this intra and inter group's C&Ckines expression fold change apparently seems interesting in this small cohort, it may not be significant in larger sample size.

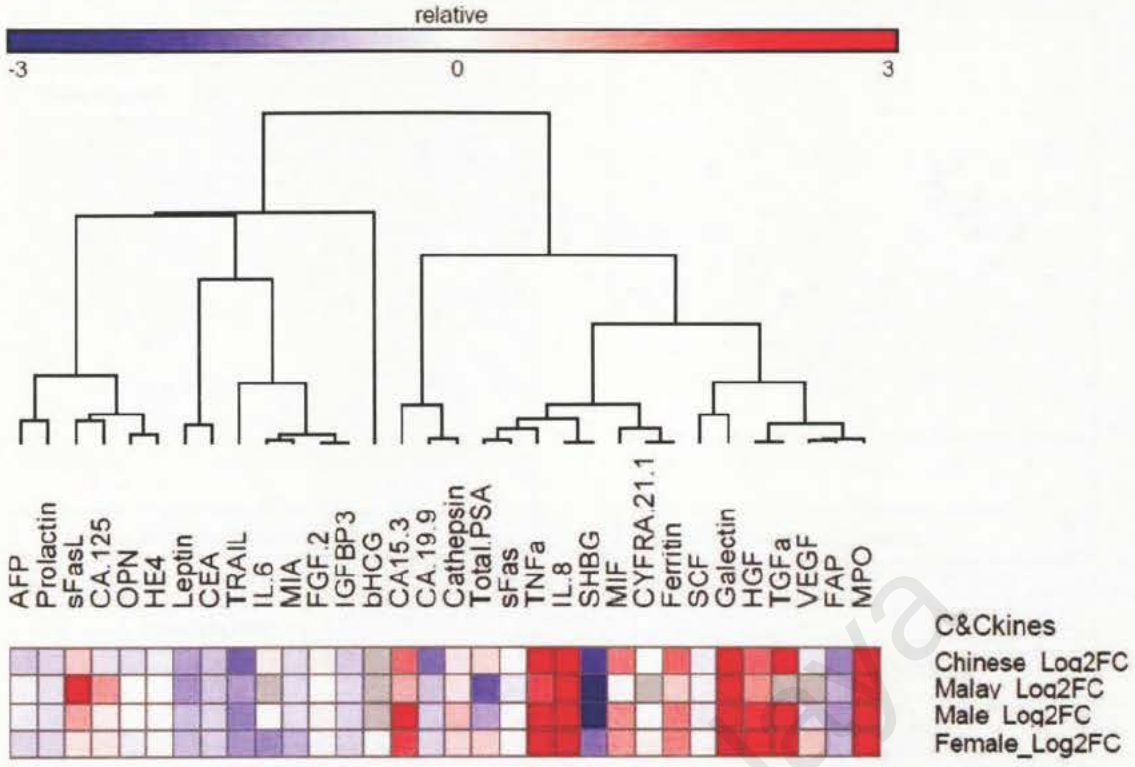


Figure 3.6: Gender and Race based expression fold changes. Gender and Race based expression fold changes. Average expression of Male-Female and Malay-Chinese groups of AML cases are compared with corresponding healthy control groups. MPO, HGF, Galectin-3, and IL8 are showing the highest fold changes for all the groups. sFasL is high is Malay population whereas SHBG is relatively low in Malay and Male groups.

3.4.7 C&Ckines follow distinct expression patterns

After individual C&Ckines expression analysis, attempts were made to determine whether the expression of these C&Ckines follow any distinct pattern. Principal component analysis (PCA) and multidimensional scaling (MDS) were performed for 38 samples based on 32 C&Ckines expression values. As shown in Figure 3.7A and 3.7B, two distinctive clusters were shown that represented AML cases and healthy control groups. The control group was more homogenous and plotted next to each other in both figures.

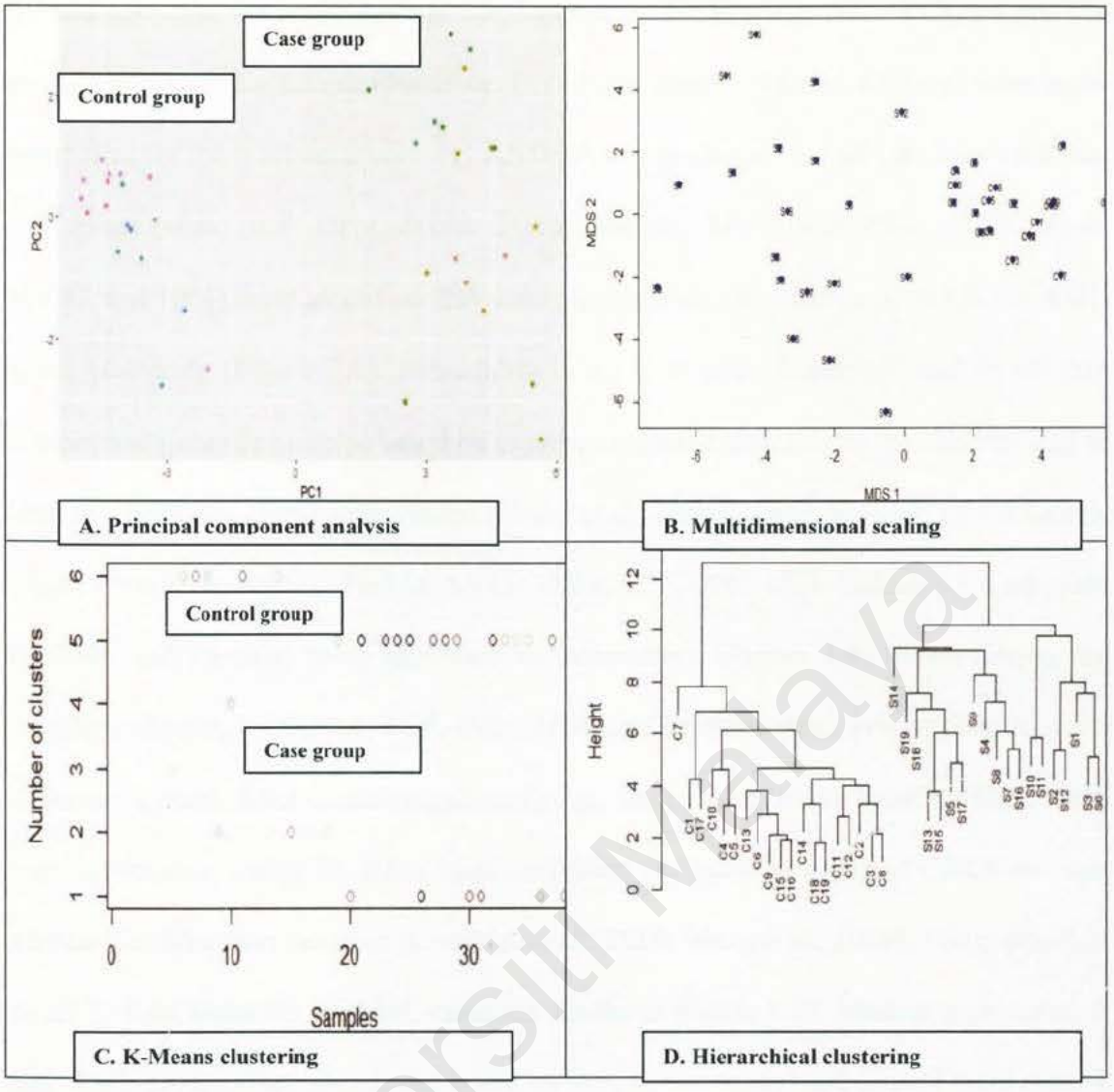


Figure 3.7: Samples clustering based on expression values. Control samples are clustered together in PCA and MDS whereas AML cases are distinct. Hierarchical clustering clearly differentiates cases (S) and controls (C).

Although it was difficult to get sub-cluster of AML cases in MDS plot (Figure 3.7B), at least 2 groups were visible based on PC2 score (positive/negative) in PCA plot (Figure 3.7A). First, three PC can explain more than 70% variation and a heatmap of PC against samples is shown in Figure A1b. For further visualization of clusters, K-means clustering (Figure 3.7C), hierarchical clustering dendrogram (Figure 3.7D), and a silhouette clustering (Figure A1a) were carried out. Based on average silhouette width and other clustering approaches, it can be said that at least five clusters of samples are visible

(please note that the entire normal subjects are in one cluster). All C&Ckines are not equally important for sample clustering. In order to identify which C&Ckines were more instructive for those clusters, one-way ANOVA was performed for all C&Ckines. Eleven C&Ckines (sFas, SCF, OPN, FGF2, TGFa, Ferritin, AFP, Total PSA, CYFRA21-1, bHCG, and HE4) were identified that were significantly informative ($p<0.05$) for AML cases clustering (Figure 3.8). Bimodality index is another important tool to identify informative genes or proteins based on expression profile that divides the samples into at least two normally distributed clusters (Wang et al., 2009). Using this tool, 13 C&Ckines (AFP, Total PSA, Leptin, sFasL, CA 125, TNFa, SCF, OPN, HE4, Galectin-3, Cathepsin D, FAP, and Ferritin) were identified as informative (Figure 3.9). Interestingly, six C&Ckines (AFP, Total PSA, SCF, OPN, HE4, and Ferritin) were overlapping between these two groups. After combining these two groups (union), it was found 18 C&Ckines were informative out of 32. It has been confirmed that bimodal group of C&Ckines was a better classifier than the other (Kornblau et al., 2010; Wang et al., 2009). The expression of all 32 C&Ckines for 19 AML cases are shown in Figure 3.10. Median expression of four AML groups (that were found based on sample clustering) and normal were plotted in Figure A2.

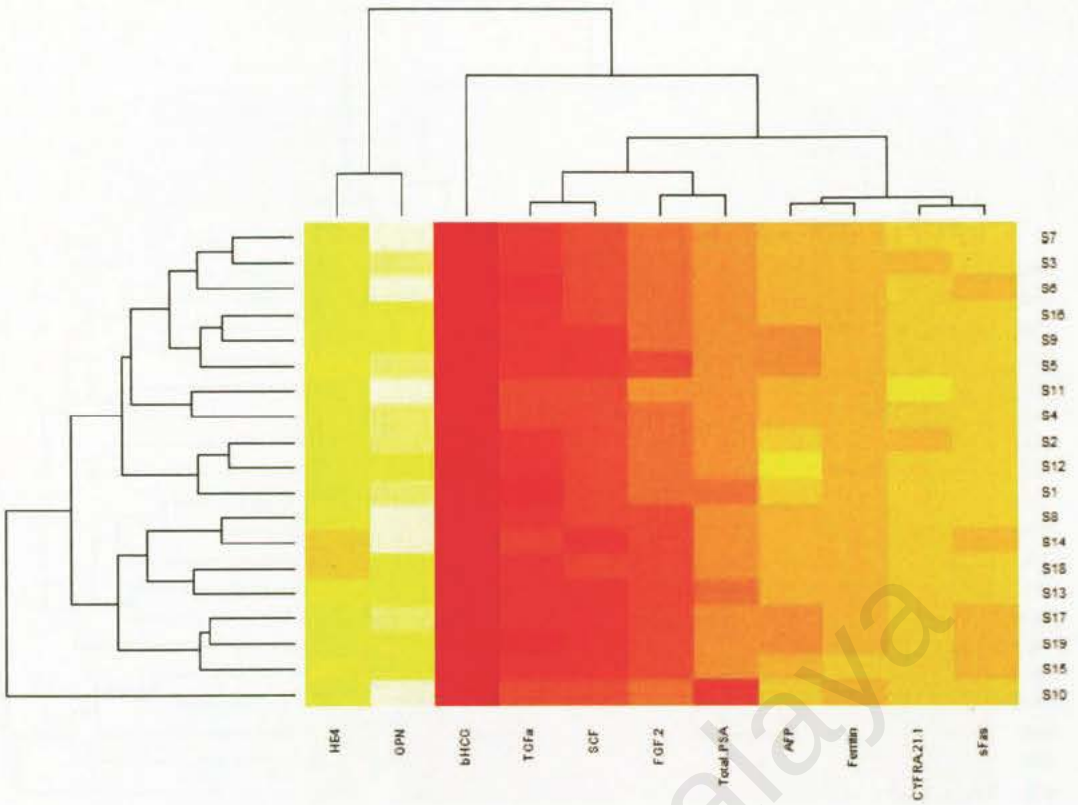


Figure 3.8: One-way ANOVA based AML patients clustering.

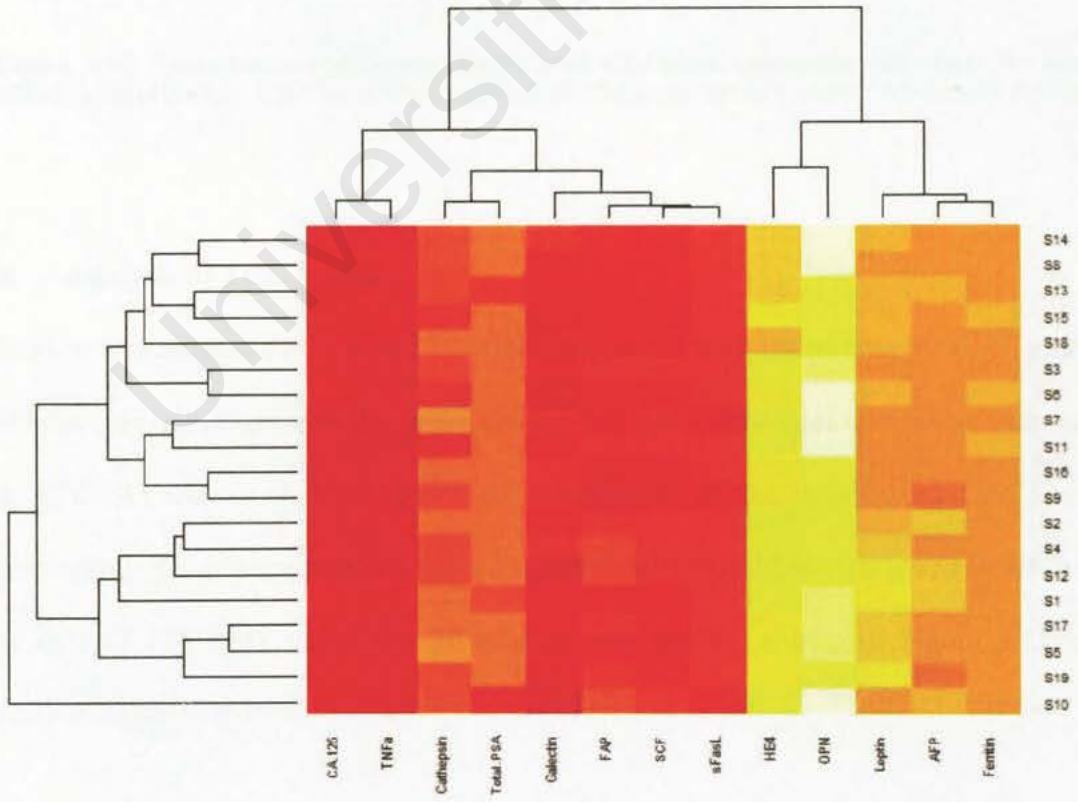


Figure 3.9: Bimodality Index based AML patients clustering.

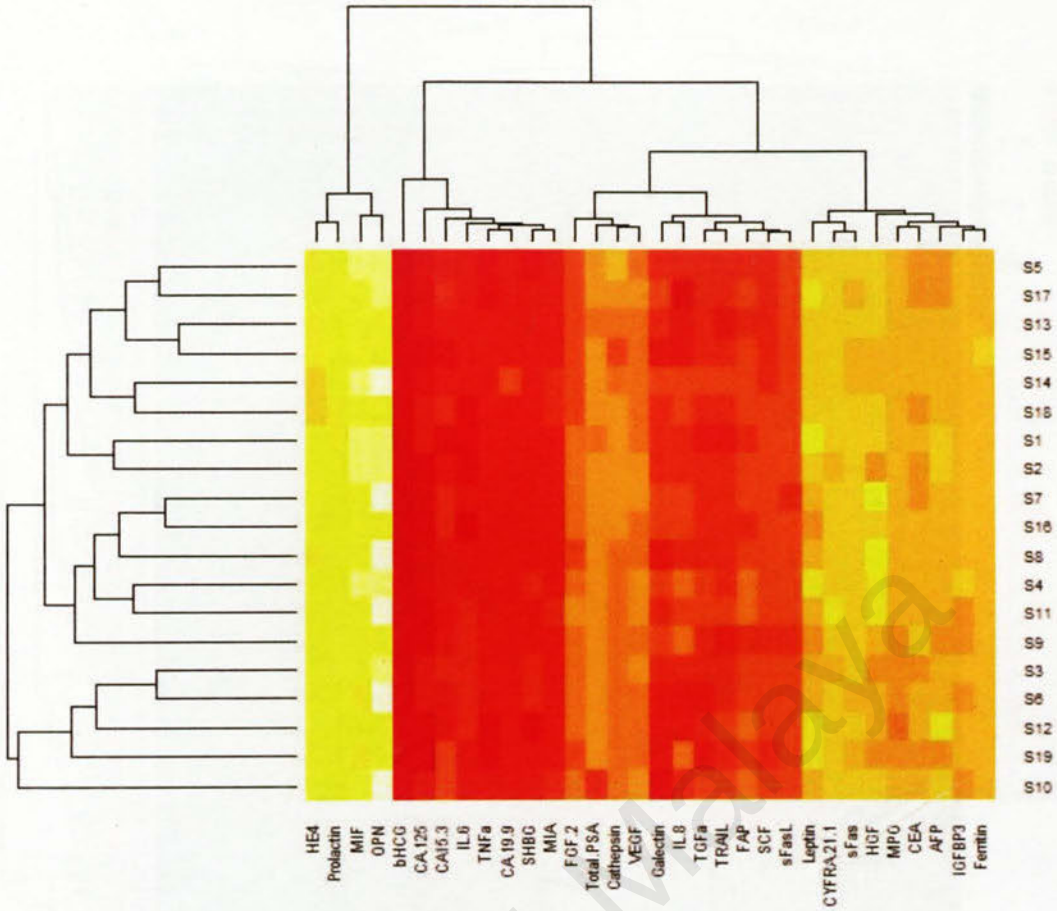


Figure 3.10: Two-dimensional cluster diagram of C&Ckines expression (columns) for each individual patient (rows). There are at least 5 clusters of C&Ckines and 4 clusters of AML cases noticed.

3.4.8 Analysis of TCGA data

Next, it was decided that it would be interesting to see how global data for AML cases were behaving with regards to this study and for this AML data from the Cancer Genome Atlas (TCGA) was used. First, promoter methylation of the targeted 32 C&Ckines corresponding genes were checked using the *Infinium* HumanMethylation450 BeadChip array data of 194 AML cases and 30 healthy controls and shown in Figure 3.11 by Manhattan distance matrix.

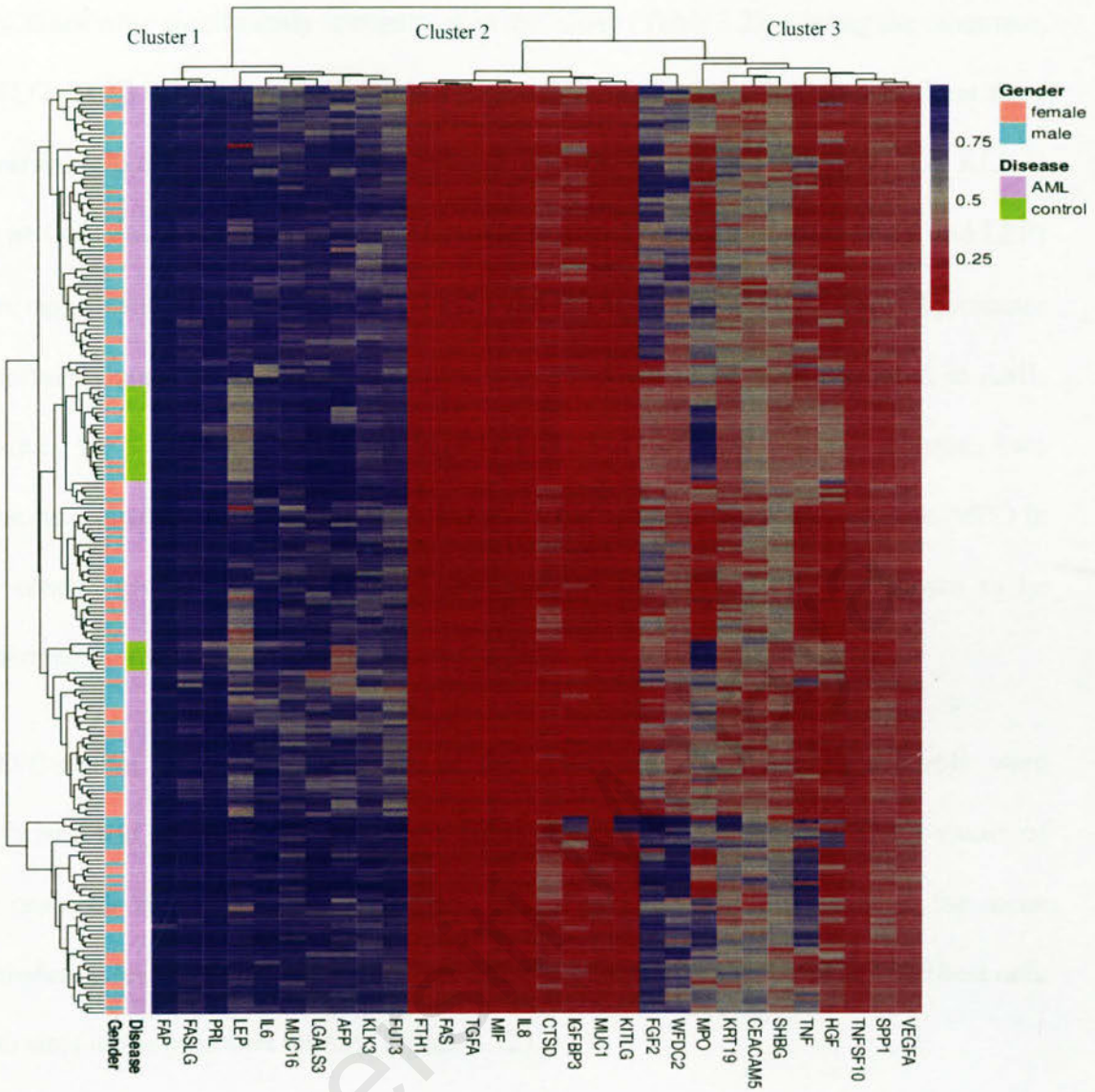


Figure 3.11: Studied C&Ckines corresponding genes promoter methylation patterns.

There were three clusters of C&Ckines: hypermethylated (β>0.75) genes cluster (Cluster 1) which contain 10 C&Ckines corresponding genes (FAP, FASLG, PRL, LEP, IL6, MUC16, LGALS3, AFP, KLK3, and FUT3), hypomethylated (β<0.25) genes cluster (Cluster 3) which contain 9 C&Ckines corresponding genes (FTH1, FAS, TGFA, MIF, IL8, CTSD, IGFBP3, MUC1, and KITLG), and relatively less hypomethylated (β<0.50) genes cluster (Cluster 2) that contain the rest of the C&Ckines corresponding genes (n=13). Interestingly, 7 out of 9 hypomethylated (β<0.25) genes cluster corresponding

C&Ckines were significantly upregulated in this study (Table 3.2). Among the other two, KITLG (SCF) was upregulated and IGFBP3 was downregulated but none of them were significant. In the case of a hypermethylated (β>0.75) cluster, 4 (FAP, LEP, IL6, KLK3) out of 10 genes which were downregulated in this study, only 2 of these (FAP and LEP) were significant. Although FASLG (sFasL) and LGALS3 (Galectin-3) showed promoter hypermethylation in the TCGA data (and was supposed to be downregulated in AML plasma), we found their level significantly higher in our AML patients plasma. Two distinctive clusters were noticed for MPO and LEP between cases and controls. MPO in the control group was shown to be hypermethylated-while LEP was shown to be hypermethylated in the control group.

Differential methylation analysis of the AML cases and healthy controls were performed and plotted as the mean beta value of cases against controls. The p-values of the promoter sites were calculated using limma package. The difference in the mean methylation levels of case and control groups was compared and plotted as 1000 best rank CpG sites in the promoter region (Figure 3.12).

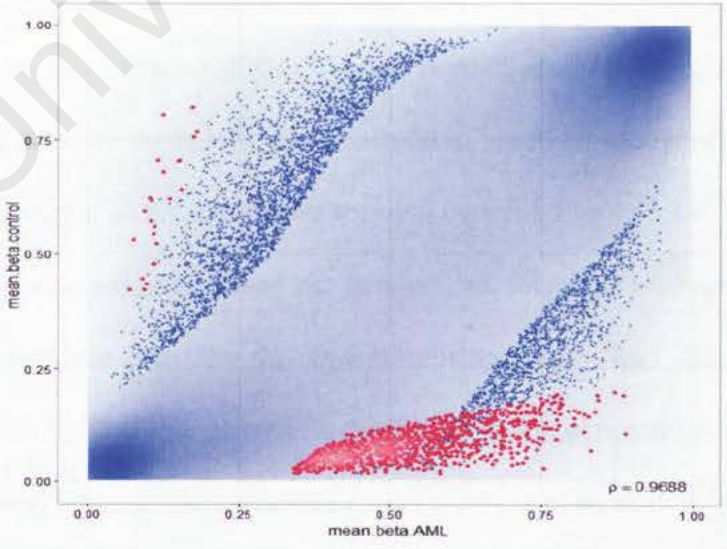


Figure 3.12: Differential methylation value (mean) of cases (x-axis) and controls (y-axis). The density plot is overlaid with a scatter plot in which promoters with FDR-corrected *p* value < 0.05 are depicted as red points.

TCGA AML gene expression microarray (HG-U133 Plus 2.0) data was used to check the pattern of the corresponding genes expression. Gene expression data also followed quite similar fashion like differential promoter methylation (Figure 3.13).

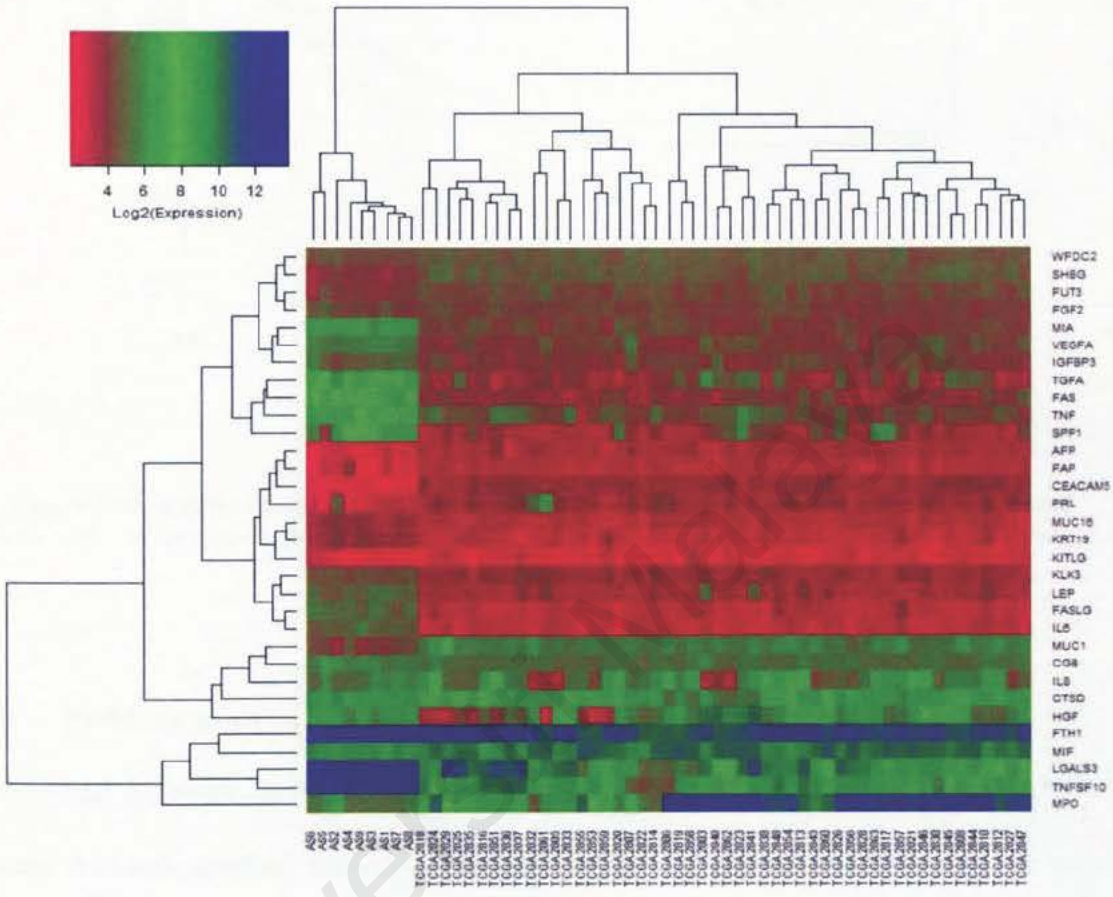


Figure 3.13: Differential (C&Ckines corresponding) genes expression.

Overall 5 years survival probability was calculated based on cytogenetic risk factors of TCGA data and shown as Kaplan-Meier survival curve in Figure 3.14. The p-value of the chi-squired test was adjusted based on gender and found to be highly significant (p=0.0003). The survival rate for the favorable risk group was distinct from the intermediate and the unfavorable groups. Unfavorable group survival probability (50%) was less than 2 years, whereas the favorable group survives more than 5 years. This data support the previously known phenomenon of favorable and unfavorable group survival.

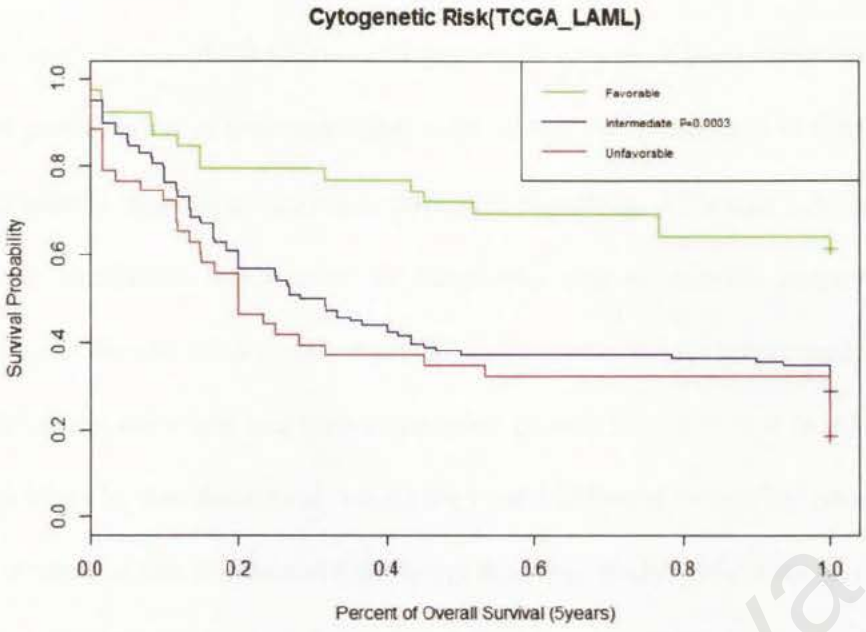


Figure 3.14: Kaplan-Meier survival curve for cytogenetic risk groups. Expected median overall survival (OS) for the unfavorable group was less than 2 years, whereas median OS for exceeded 5 years.

3.4.9 Pathway analysis

Targeted pathway analysis for 32 C&Ckines was performed using GeneMANIA (Figure A3) and checked for AML pathway in KEGG database (Figure A4). It was noticed that 32 C&Ckines interact through 20 possible intermediary proteins and their size represents important functionality (Figure A3). PI3K/Akt signaling pathway is mediated mostly through anti-apoptotic features and the putative role of MIF and Cathepsin D in this pathway has already been discussed in section 3.4.4.

3.5 Discussion

Cytokines network-mediated crosstalk between AML blasts and fibroblasts promote cell proliferation and reduce apoptosis in AML cases (Ryningen, Wergeland, Glenjen, Gjertsen, & Bruserud, 2005). Disruptions of normal C&Ckines and growth factors networks have been demonstrated for leukemia (especially for AML) (Kornblau et al., 2010; Van Etten, 2007) whereby variation in the plasma concentrations were shown to be

of prognostic significance (Tsimberidou et al., 2008). The pathogenic as well as therapeutic implications of C&Ckines and growth factors have been identified (Dranoff, 2004). Self-proliferation of leukemic blast cells occurs by stimulation of C&Ckines that is mediated mostly through autocrine or paracrine signaling. Although C&Ckines levels demonstrated significant implication for diagnostic and prognostic purposes and for understanding of disease etiology, no study has been conducted for larger scale C&Ckines and growth factors detection and their expression pattern identification in AML plasma, except for a study by Kornblau et al. where they used different biomarker panel for AML and MDS serum and less number of C&Ckines than this study. This is the very first time that detection of 32 C&Ckines in plasma of adult Malaysian AML patients was achieved. Ten of them have not been detected before in any AML cases.

Average C&Ckines expression levels for AML cases and healthy controls were comparable to other multiplex based studies. As multiplex is more sensitive than ELISA, some C&Ckines average levels (detected by ELISA) were comparatively higher than the multiplex. All the samples were freshly collected from newly diagnosed patients, stored at -80^0 C until use, and handled by a single person, so there was no chance of plasma protein degradation or batch effect. As the method is well-established and reliable, the chance of reproducibility of this result is high.

Only a small number of C&Ckines expression was significantly ($p<0.005$) dependent on age (Leptin and Ferritin), gender (IGFBP3), and ethnicity (MIF, sFas, sFasL, OPN, Cathepsin D, MPO, and SHBG). Two main ethnic groups (Malay and Chinese) have been selected for this study, as they are less distinct from each other compared to Indian group. Because previous ethnic group based genetic studies revealed that another major group (Indian) is genetically far distinct than those two, so excluded them from this study. The age limit was more than 20 years and the highest age patients were 65 years.

Significant expression differences were noticed for 17 C&Ckines between AML cases and healthy controls, in which 13 C&Ckines were upregulated and 4 C&Ckines were downregulated in AML patients. Healthy controls have similar expression patterns and were clustered together closely in all types of clustering algorithms.

Detected C&Ckines corresponding genes expression and promoters methylation data from TCGA AML patients were mostly correlated with multiplex based C&Ckines expression levels in this study. Small variation among genes expression, promoter methylation, and C&Ckines levels correspond that there is other controlling mechanism (like miRNAs and histone modifications) working in between gene expression and protein translation.

In this study, the highest fold change has been detected for MPO, which is an established biomarker for AML. Other significant fold changes have been noticed for MIF, HGF, Galectin-3, and IL-8 that could possibly be considered to be biomarkers for AML diagnosis. Several studies reported upregulation of TNF-α and its receptor in AML patient, similar to this study, which is seemed to derive from the AML blasts and is associated with the unfavorable risk group (Kornblau et al., 2010; Vinante et al., 1998). It has been shown that TNF-α inhibits proliferation of AML blasts by inhibiting CSF3 receptor expression in the blasts (Elbaz et al., 1991). The adverse prognostic impact of high TNF-α has already been reported (Tsimberidou et al., 2008) as it could decrease chemotherapeutic sensitivity. Etanercept and Infliximab are two approved drugs that target TNF-α. Based on the pathway analysis, it is probable that the involvement of MIF in the pathogenesis of AML is through the PIP3/Akt pathway. However, further studies are required to prove this hypothesis.

One of the most interesting observations in this study was the distinct patterns of C&Ckines expression. Only based on expression level, 5 major groups of C&Ckines and

4 major groups of AML patients were noticed. For clustering, 13 C&Ckines were significantly informative based on 'Bimodality Indexing' and 11 C&Ckines were significantly informative based on one-way analysis of variance (ANOVA), whereas 6 of them were mutually inclusive. Further studies of these clusters could potentially improve AML classification.

A major limitation of this study was the small sample size. There may have other important C&Ckines more relevant to AML pathogenesis or prognosis that have not been detected in this study. Only 2 specific panels have been selected that contain a limited number (n=34) of C&Ckines, 2 of them were not detected due to cross-reactivity concern. There was no customized detection in the multiplex, rather specific panel based study. Although plasma has been used for this study, Stroncek et al. reported that almost 50% C&Ckines and growth factors expression level vary between plasma and serum (Stroncek et al., 2005). The best leukemic microenvironment study sample would be marrow juice, but due to unavailability of marrow, plasma samples have been used for this study.

As leukemia is a heterogeneous and complex disease, understanding of oncogenic transformation is not easy, which require an integrated and a more intensive study. Here deregulation of C&Ckines level can only be detected, but to understand the effect of those C&Ckines on leukemic blasts would require more study. There is even no clear information as to where those C&Ckines came from; it could be from leukemic blasts or fibroblast, the stroma, inflammatory cells, or from somewhere else. In a future study, different sources of samples like serum and marrow juice would be used to elucidate further specific findings.

3.6 Conclusion

According to the literature search, this is the first study of detection of more than 30 different C&Ckines in Malaysian AML patients and probably the first time that the aforementioned 10 C&Ckines were shown to be present in AML plasma. This study gives an overview of C&Ckines expression in AML plasma. In general, the C&Ckines expression in AML patients differs from healthy controls. More than half of the studied C&Ckines have been found to be deregulated in AML cases compared to healthy controls. C&Ckines expression based AML patients clustering could help to improve the conventional classification system. A small number of C&Ckines may have potential diagnostic and prognostic utility. MIF was the most notable cytokine in this study, consistently showing statistical significance. We speculated that MIF could be a possible therapeutic target. To understand the role of C&Ckines in leukemic blasts and underlying molecular mechanism, further studies are required.

CHAPTER 4: DIFFERENTIAL ANALYSIS OF GENETIC, EPIGENETIC, AND CYTOGENETIC ABNORMALITIES IN AML

4.1 Introduction

AML is a hematological disorder characterized by excessive proliferation of immature myeloid cells and disruption of differentiation (Tenen, 2003) in the bone marrow that infiltrate liver, spleen, lymph node, and circulating blood (Mehdipour et al., 2015). This cancer type progresses rapidly and is relatively fatal due to acquired genetic and/or cytogenetic aberrations. In 2014, AML was the most common type of leukemia diagnosed and it accounted for 1.78% of predicted cancer deaths in the United States (Siegel et al., 2014). Overall 5-year survival rate greatly varies based on the cytogenetic risk classification, with 55%, 24%, and 5% overall survival for favorable, intermediate, and adverse risk group patients, respectively (Byrd et al., 2002). Relapse is the major reason for poor survival rate as it could occur in up to 80% of AML patients (Christopeit & Bartholdy, 2014). Almost 50% of all AML cases belonging to the intermediate risk group have been reported to lack any cytogenetic abnormalities (Vardiman et al., 2009). Furthermore, a significant proportion of the patients carry no reported genetic mutations in any known AML associated driver genes (Patel et al., 2012; Shen et al., 2011). These findings clearly indicate that there are other elements predisposing to and driving the disease in the case of cytogenetically normal AML (CN-AML). Epigenetic modifications and RNA-based gene regulations are an interesting area of research which can provide a better understanding of AML.

Historically, chromosomal changes were the first abnormalities identified in AML and are still among the most effective prognostic factors for induction response and overall survival (Grimwade, 2001). Comprehensive genomic profiling of AML has revealed

recurrent mutations in multiple genes (Patel et al., 2012), which are essential for hematopoietic development. It has already been reported that epigenetic modifications are involved in the regulation of hematopoietic development (Bock et al., 2012; Ji et al., 2010). DNA methyltransferases and histone methyltransferases are well known epigenetic modifiers that contribute to cellular identity through regulation of genes expression in myeloid progenitor lineages. They are among a group of frequently mutated genes in AML (Lokody, 2014; Network, 2013) which suggests that epigenetic modification could be a causative agent for disease progression and relapse in CN-AML.

An overwhelming evidence suggests a complex interaction of genetic events contributing to AML pathogenesis (Network, 2013). From the biological point of view, the cancer phenotype is likely to be directly associated with gene expression, which can be potentially driven by genetic, epigenetic, and cytogenetic changes. In order to explore the pathophysiology of this disease, integration of genetic, epigenetic, and cytogenetic data and differential analysis are required. The combination of information on every single molecular parameter into phenotypic signatures would be a robust classifier in disease diagnosis and risk stratification in AML (Eppert et al., 2011; Gentles, Plevritis, Majeti, & Alizadeh, 2010; Schoch et al., 2002). Disease prognosis and classification of AML are important for physicians to decide particular treatment protocol in order to avoid minimal residual disease and risk of relapse (Christopeit, Kröger, Haferlach, & Bacher, 2014). None of the two well-known classifications (French-American-British(Bennett et al., 1976a) and World Health Organization (Vardiman, Harris, & Brunning, 2002) integrate epigenetic signatures with genetic and cytogenetic ones. To facilitates the molecular pathophysiological study of AML, the Cancer Genome Atlas (TCGA) research network has generated a huge amount of data that are available to the public for further analysis and interpretation of AML pathogenesis, classification, and risk stratification (Network, 2013). TCGA has provided the most comprehensive molecular and clinical

data of over 33 different tumor types on 10000 cancer patients (Weinstein et al., 2013). Although TCGA has made many important discoveries and published more than 20 marker papers, remarkable opportunities still exist to analyze the data using novel methods and strategies from dynamic viewpoints.

For this study, the data available in TCGA under study code LAML was used and differentially analyzed using a carefully selected set of online and offline platforms. For the comprehensive analysis of differential DNA methylation, RnBeads was preferentially used (Assenov et al., 2014). In order to integrate genome-wide methylation and mutation data with cytogenetic data, custom R scripts were developed for analyzing every single event and combining events. Circos, ingenuity pathway analysis (IPA) database, and other platforms were also used to organize and visualize the data. Hierarchical clustering and heatmap were used to identify and visualize distinct groups of AML patients and genes. Further, a large collection of diagnostic and phenotypic traits were used to calculate individual and composite risk factors and presented them as Kaplan-Meier curves. This integration strategy provides useful insights on disease subtypes identification, risk measurement, and detection of clinical prognostic markers.

4.2 Literature review

AML is a biologically heterogeneous and complex malignant disorder. Cytogenetic analysis was the first tool to study molecular pathogenesis of AML for more than three decades until the genomic era revolutionized molecular biology. Although cytogenetic abnormalities are still being used for diagnostic and prognostic purposes, it is not sufficient to understand the insight of AML pathogenesis and require genetic and epigenetic analysis. Now it is clear that AML is a complex interplay of genetic, epigenetic and cytogenetic abnormalities where none of them are mutually exclusive. No single causative agent has yet been identified which is capable of oncogenic transformation of

myeloid cells alone. Over the last decade, rapid improvements of genome analysis technologies have made it easier to study large-scale genomic data to figure out genomic lesions. New technologies like microarray and deep sequencing have facilitated studies of genomic and epigenomic deregulation in cancer cells.

The first identified somatic mutations in AML were the fusion of key transcription factors (Mrózek, Heerema, & Bloomfield, 2004; Rowley, 2008). Fusion of specific transcription factors were associated with specific patterns of mRNA expression and in some cases specific patterns of even miRNA expression (Network, 2013). A new subtype of AML patients with distinct epigenetic signature can be found after a combination of mutations with intermediate-risk cytogenetic factors (Network, 2013). A recent study found 16 groups of AML patients based on DNA methylation data and 5 of these were new subtypes (Figueroa, Lugthart, et al., 2010). Figueroa et al. also reported that unique methylation patterns could be found for specific cytogenetic abnormalities like inv (16), t (8;21), t(15;17), and 11q23 (Figueroa, Lugthart, et al., 2010). One of the notable features of epigenetic abnormalities is reversibility that is not like genetic abnormalities. Epigenetic abnormalities means a change in genes expression but the DNA coding sequence is not modified, so if somehow it is possible to reverse the process, this could be utilized as therapeutics.

The leading DNA methyltransferase enzyme is DNMT3A. HOX genes are often hypomethylated in CN-AML if DNMT3A is mutated (Qu et al., 2014). Mutation of other epigenetic regulators, such as TET2 (Delhommeau et al., 2009), MLL (Michaux et al., 2000), IDH1/2 (Figueroa, Abdel-Wahab, et al., 2010), FLT3 (Network, 2013), NPM1 (Network, 2013), RUNX1 (Mendler et al., 2012), and ASXL1 (Chou et al., 2010) have been associated with unfavorable clinical outcomes for AML patients (Patel et al., 2012). Therefore, mutations in epigenetic modifiers as well as alterations of genome-wide

methylation patterns in AML imply epigenetic deregulations as one of the fundamental causal agents in AML pathogenesis. There are no absolutely correct classifications for AML, so the focus of this study would be epigenetic information based patients clustering that would give us hints regarding potential use of epigenetic factors for AML classification in future.

4.3 Materials and methods

4.3.1 Data

TCGA (https://tcga-data.nci.nih.gov/tcga/) data repositories were used as the primary data source for this study. To analyze the AML data generated by TCGA, the raw data was directly accessed and downloaded, using the "Data Matrix"- tool provided by TCGA Data Portal. The clinical, gene expression, and DNA methylation data of all available AML patients were also downloaded. In total, 200 patients with clinical data were found, with 194 DNA methylation data (date of download: August 15th, 2015). The available data type for DNA methylation analysis was Illumina Infinium HumanMethylation450 BeadChip (450K microarray). The human genome (UCSC hg19) was used as reference. Gene definitions were based on Ensembl release 79. A gene promoter was defined as the region spanning 1500 bases upstream and 500 bases downstream of the gene's transcription start site. The number of the promoters with available methylation measurements was 28,642. The mean number of interrogated probes (CpG sites) sites per promoter is 10 in the data set. In the case of multiple CpG sites in a particular promoter, the average methylation (beta) value was assigned as the overall methylation level for the promoter. For survival analysis, survival interval from the date of diagnosis until the date of death or last follow-up was used. Basic characteristics of TCGA AML patients are shown in Table 4.1. The patient cohort was predominantly white Americans with a mean

age 55.50 years. Detected genetic abnormalities were higher than cytogenetic abnormalities.

Table 4.1: Characteristics of TCGA AML patients.

Patient Characteristics	Values (units)
Gender (M/F)	109/91
Age at Diagnosis	55.50±16.07 (Years)
Race(Black/White/NA)	15/183/2
Ethnicity (Non-Hispanic/Hispanic/NA)	194/3/3
Vital Status (Dead/Alive)	133/67
Blasts in Peripheral Blood(M/F)1	67.76/64.86 (%)
WBC counts after 24h of Storage 2	37.246 (k/mm^3)
Platelets Counts	65.98 (cells/µL)
Blasts Cell Counts	36.28 (cells/µL)
Neutrophils Counts after 24hr of Storage	12.16 (k/mm^3)
Lymphocytes Counts	27.50 (cells/µL)
Monocytes Counts	12.34 (cells/µL)
No. of Patients Cytogenetically Abnormal	106
No. of Patients Genetically Abnormal	160

Male=M, Female=F. NA = Not Available. [1]Blasts= Immature white blood cells. More than 20% of blasts is generally required for a diagnosis of AML. [2]Number of WBCs reduce after long time storage.

4.3.2 Software

R platform (version 3.1.3) and RStudio were used to install and initiate other R/Bioconductor packages. RnBeads (http://rnbeads.mpi-inf.mpg.de/) was used to analyze and visualize DNA methylation data. RnBeads is a software tool written in the R programming language (http://www.r-project.org/) for large-scale analysis and interpretation of genome-wide data sets, particularly epigenomics data (bisulfite sequencing and Infinium microarrays) with a user-friendly and customizable analysis

pipeline. It provides self-configuring workflow at the data input, quality control, preprocessing, and tracking and table stages. The output was inspected, in particular at the exploratory analysis and differential methylation analysis stages, available as interactive hypertext report with high quality figures and tables (Assenov et al., 2014).

4.3.3 Statistics

The data files were preprocessed using the R platform for statistical computing and Bioconductor. DNA methylation data was analyzed by the RnBeads pipeline which performs the following statistical tests: Fisher's exact test, Wilcoxon rank-sum test, Kruskal-Wallis one-way analysis or Pearson correlation test - based on the data type in order to identify significant associations among the sample annotations and validate the analysis. Differentially methylated sites and regions were identified using limma tests and the Fisher's method for the combination of p-values. In the case of survival analysis, the effect of age and gender were corrected, the resulting p-values were adjusted using the Benjamini-Hochberg method, and a significance threshold of 0.05 was then applied.

4.3.4 Analysis and visualization

The analysis presented here can be broadly separated into the molecular and clinical analysis. Later on, molecular and clinical data were combined to study the survival curves. In the case of clinical data, the reprocessed txt files were uploaded to RStudio and every column of available patient information were analyzed and tests performed, the result was summarized, and presented in a barplot in Figure 4.1. Patient characteristics are also presented in Table 4.1. The gene ontology, functional genes, and miRNAs networks were analyzed using the IPA (Ingenuity Systems, Mountain View, CA). For multivariate analysis, we applied principal component analysis (PCA) and multidimensional scaling (MDS) through RnBeads. It is important to note that in many

cases, the patient's clinical data were incomplete. Although all available data were analyzed, many patients' information was missing in the TCGA data file that is presented separately and some information was not presented in this study as it was irrelevant to this study. Some genetic and cytogenetic abnormalities are presented and visualized for the clinical data analysis (Figure 4.2). We also applied various techniques for batch effect detection and dimension reduction, as implemented in RnBeads.

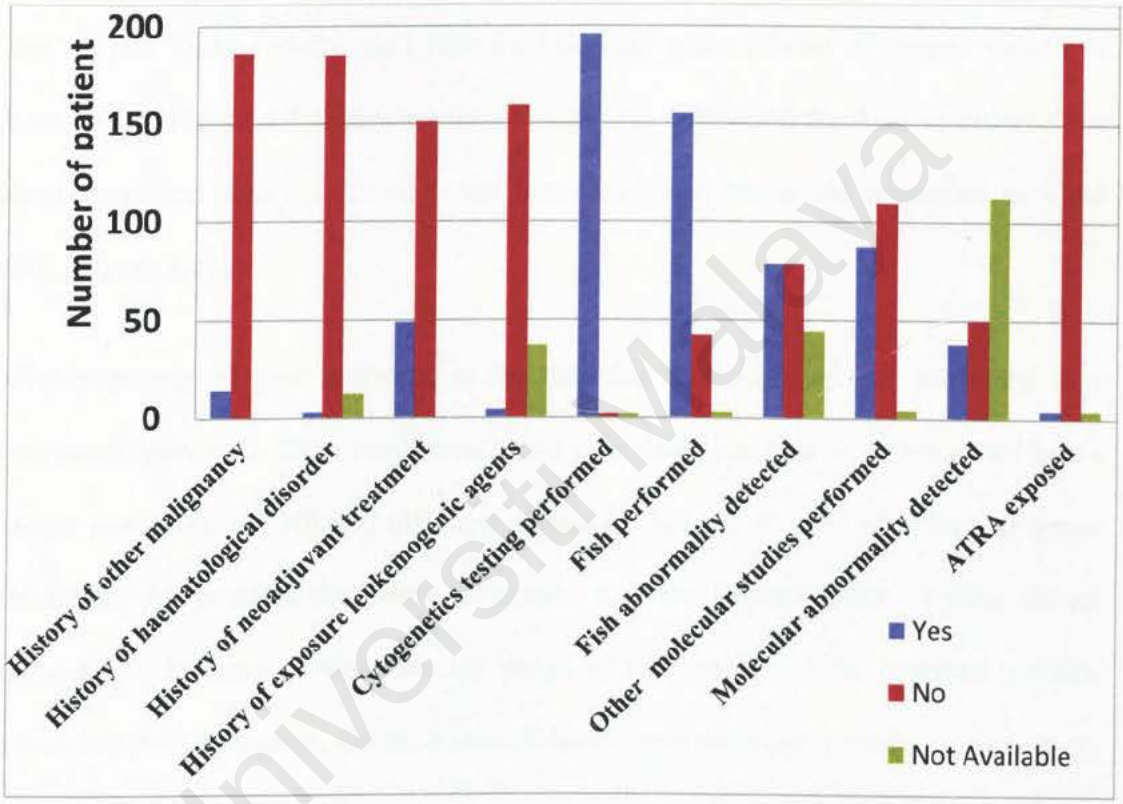


Figure 4.1: Summary of patient's clinical tests performed. X-axis represents the number of patients and Y-axis represents the clinical information. Different clinical observations are shown in the plot. No epigenetic testing performed in the clinic. Cytogenetic test and Fluorescence in situ hybridization (FISH) test are performed for most of the patient cohort. 'History of other malignancy' means malignancy other than haematological malignancy. One fourth of the total patients received neoadjuvant treatment and only handful number of patients show haematological disorder other that AML.

Most of the molecular analysis includes genetic and epigenetic data. RnBeads performed exploratory and differential analysis. Density plots (Figure 4.5) showcased a slight deviation from the global distribution (bimodal shape) of DNA methylation. Overall methylation at different CpG island (CGI) position and promoter region of the epigenome usually vary. Scatter plots also presented the variability and displayed

clustered samples based on mean methylation value. Heatmaps were drawn based on hierarchical clustering that provides a global view of AML subtypes. Sample annotation and DNA methylation data analysis were performed in a systematic way in this study. Combined differential DNA methylation of individual CpGs as well as extended genomic regions increased statistical power, interpretability, and reproducibility (Bock, 2012; Bock, Walter, Paulsen, & Lengauer, 2008). In the case of differential DNA methylation analysis, I not only considered statistical significance, but also the biological significance of the output. Differentially methylated AML data were plotted as scatter plots and volcano plots (Figure 4.7). Enrichment analysis was conducted for over-represented or under-represented gene sets using Gene Ontology (GO) terms and presented as word clouds (Figure 4.7).

The frequency of gene mutation in the data file was analyzed and presented as a histogram (Figure 4.2). DNA methylation and gene mutation data were integrated into a heatmap plot (Figure 4.10) and different groups of patients as well as groups of genes were found. All possible data were combined to draw Kaplan-Meier survival curves (Figure 4.12). In order to visualize the result of the analyzed data, heatmap, cluster, ggplot2, bar plot, histogram, circus, Kaplan-Meier curve and other visualizing tools were used mostly through RnBeads.

4.4 Results

4.4.1 Characteristics of the patient cohort and clinical testing

In the primary study cohort, a total of 200 AML patients were listed in the TCGA database. The mean age at diagnosis was 55.50 years; 109 (54.5%) patients were male and 91 (45.5%) female. The majority of the patients were white American (91.50%) and Non-Hispanic (97%). Only 106 patients were cytogenetically abnormal. Overall genetic abnormalities were around 80% of the cohort and patients with both genetic and

cytogenetic abnormalities were more prone to death compared to single type abnormality. Additional information, such as average blood cell counts, is shown in Table 4.1.

Cytogenetic tests were performed for almost every patient (196) and FISH test was performed for 155 patients. Only three patients had other previous hematological disorders and 14 patients had other malignancies. Neoadjuvant treatment was prescribed to 24.50% of the patients. The vast majority of the patients (79.5%) had never been exposed to leukemogenic agents before; all-trans retinoic acid (ATRA) induced apoptosis was reported in only 4 patients. Exposure data were missing for 37 patients. The number of patients who were positive for FISH test was equal to the number of patients (78) who were negative and no data were available for 44 patients. Molecular abnormalities were detected in only 38 patients and data was not available for 112 patients. Figure 4.1 summarizes the clinical traits in the form of bar plot.

4.4.2 Genetic and cytogenetic abnormalities

Figure 4.2 presents 18 types of genetic and cytogenetic abnormalities. Almost 30% (58) of the patients showed FLT3 mutation, whereas only one patient showed BCR-ABL fusion and T (8; 21). The second highest mutation rate (23%) was found for NPMc. A total of 21 patients were positive for both trisomy 8 and del (7q). IDH1 R132 and T (15; 17) positive patients were only 18 (9%); IDH1 R140 and del (5q) positive patients were 15 (7.50%) and activating RAS was reported for only 5.5% of the patients. The rest of the genetic and cytogenetic abnormalities were reported among <5% of the patients. Unknown cytogenetic abnormalities were significantly higher than unknown genetic abnormalities. The highest negative test was reported for IDH1 R172 (96.5%). One notable observation in the clinical data file was that patients exhibiting multiple abnormalities tended to have shorter survival time. A statistical test was not significant for many reported abnormalities due to small positive sample size.

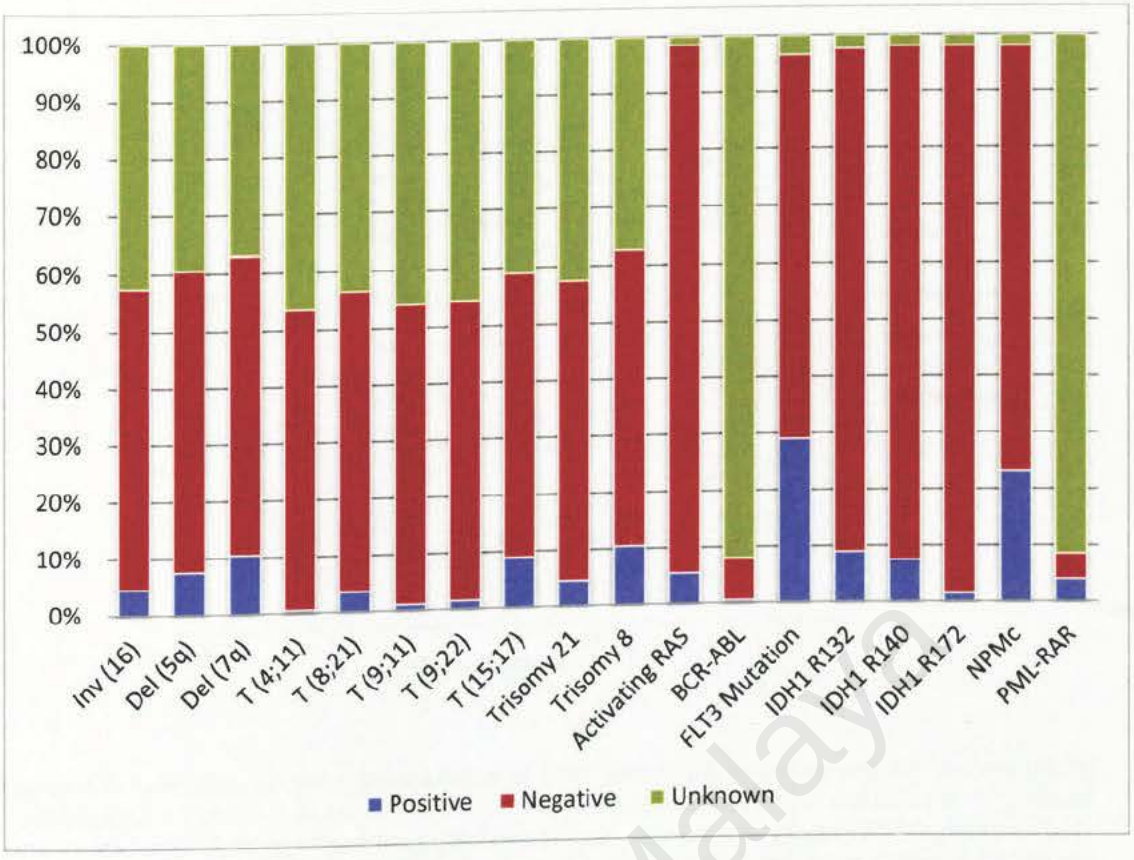


Figure 4.2: Genetic and cytogenetic abnormalities in AML. Eighteen genetic and cytogenetic traits with positive, negative, and unknown data among patients are shown. Many patients shown more than one abnormality.

4.4.3 Low-dimensional representation

In order to visually inspect the dataset, RnBeads implies two methods called multidimensional scaling (MDS) and principal component analysis (PCA). After ignoring incomplete features due to missing methylation values, the number of CpGs used for MDS was 439751, and for PCA- 436441. Similarly, promoters used for MDS were 28642, and PCA-28547. The scatter plot of Figure 4.3 shows the sample coordinates of the second and third principal components. Point type is based on IDH1 R172 status, and color denoted NPMc status. IDH1 R172 negative and NPMc positive samples occupy the upper side of the plot, and double negative samples are located on the lower side of the plot. Only one sample (circle) does not conform to this separation. Principal components 2 and 3 together can explain around 95 % of the total variance, and can be estimated from the shown cumulative distribution function(s).

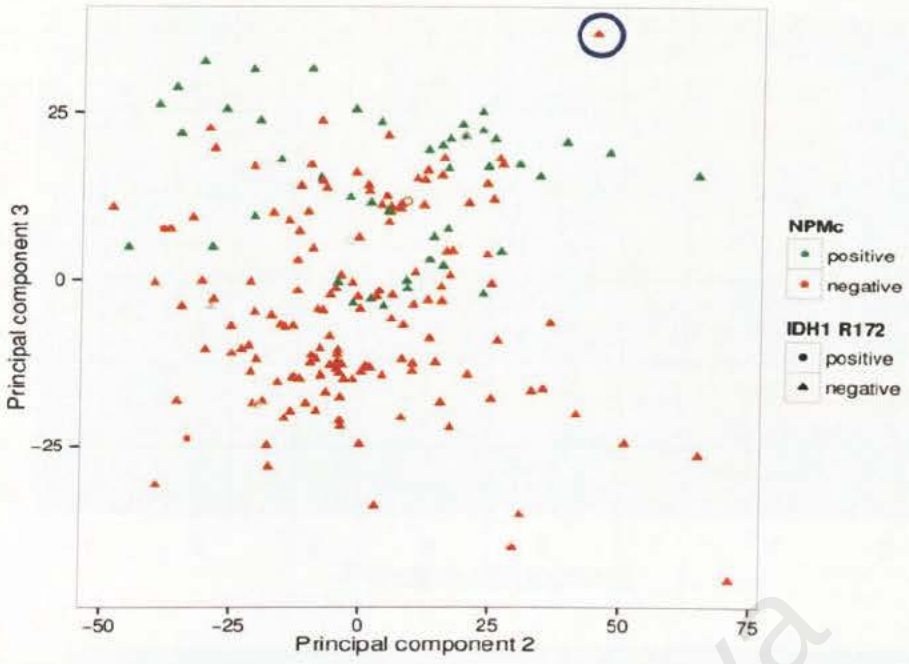


Figure 4.3: Low-dimensional representation of AML dataset. Principal component analysis (PCA) test performed for dimension reduction. Scatter plot shows coordinates of NPMc and IDH1 R172 traits on principal components. Although there are some homogenous samples, two groups (NPMc positive-IDH1 R172 negative and NPMc negative-IDH1 R172 negative) are clearly clustered in the picture at the top and bottom side. Gray color represents missing value.

4.4.4 Batch effects

Different properties of the dataset were tested for significant associations. All genetic and cytogenetic traits were tested against principal components (Figure 4.4 a&b) as well as between traits. Two statistical tests were performed: Wilcoxon rank-sum test and Fisher's exact test were carried out and showed that the significant p-values were less than 0.01. Second and third principal components show the strongest association with the traits. In the case of association between traits, del (5q) and del (7q) were more associated than other genetic traits (data not shown) and cytogenetic traits were more associated than genetic traits. Many association tests could not produce reliable results due to small sample size.

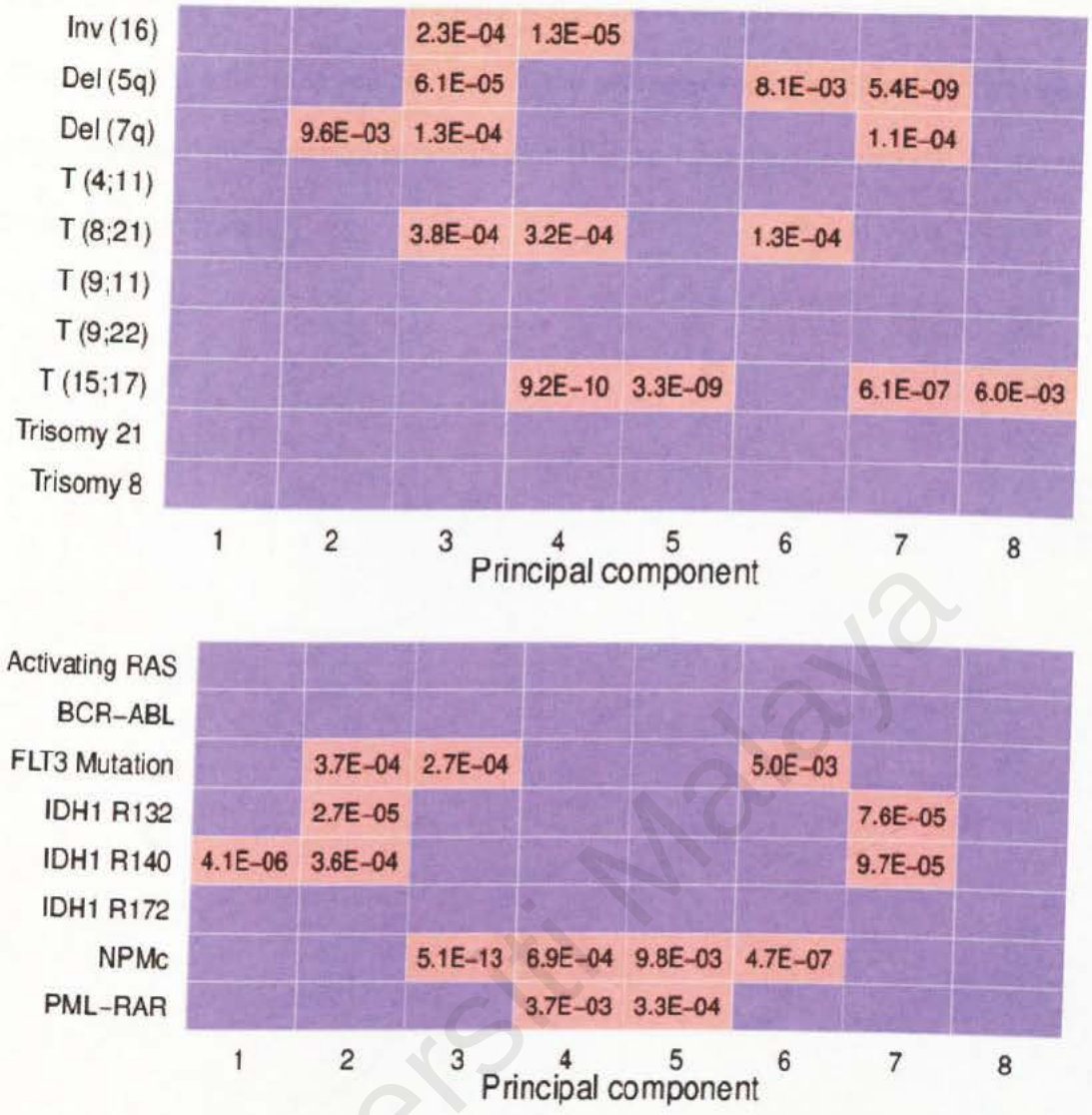


Figure 4.4: Association studies between traits. Wilcoxon rank sum test performed for first eight principal components (PC) across all the traits that together can explain more than 99% of total variance across the patient cohort. Significant p-values (less than 0.01) are printed in pink boxes. Non-significant values are represented by blue boxes. PC1 is not significant for most of the traits except IDH1 R140. We used PC2 and PC3 for scatter plot in Figure 2 and S3 that together can explain the highest variances across the patient cohort. PC3 is highly significant for NPMc (p=5.1E-13) and FLT3 (p=2.7E-4).

4.4.5 Regional sites and methylation value distribution in the whole genome

After annotation and necessary adjustment, it was found that there were 28,642 promoter sites in the TCGA AML dataset. Total samples were divided based on abnormality types (genetic and cytogenetic) and among each group, there were positive and negative patients. An average number of CpG sites per promoter were around 10.

Relatively more CpG sites were found at the end of the promoter region whereas some sites were at the flanking region. The overall promoter methylation pattern was like the genome-wide CpG methylation distribution (Figure 4.5a), i.e. it is bimodal, however, the peaks differ in height.

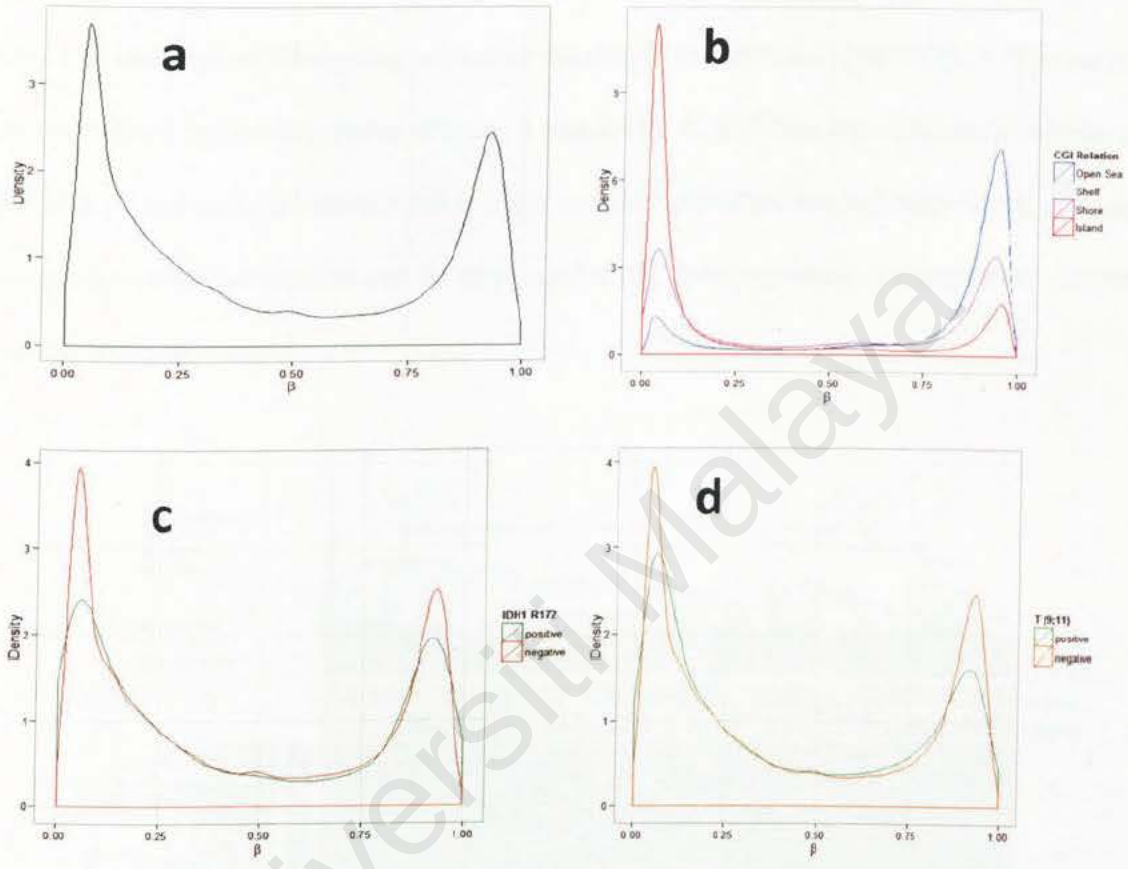


Figure 4.5: **Methylation distribution in genomes**. (a) Methylation distribution (beta value) across all promoters that follow bimodal distribution pattern. (b) Methylation value density estimation in all samples at different CpG region. (c) IDH1 R172-positive and IDH1 R172-negative patient's promoter methylation distribution. Negative patient distribution is similar with (a), but positive patients show lower density. (d) T (9; 11)-positive and T (9; 11) negative patient's promoter methylation distribution, and here, also positive patients show lower density.

T (9;11) seems to be a unique aberration in the sense that promoter methylation of the positive patients is lower compared to the methylation of negative patients, including FLT3, IDH1 R140, trisomy 8, trisomy 21, and inv (16) (data not shown). The maximum number of methylated genomic regions was found at shelf (2-4 kb from CGI) and then open sea (>4 kb from CGI) but less at the shore (up to 2 kb from CGI) and island (Figure 4.5b). IDH1 R172 and T (9; 11) positive and negative patient's promoter methylation vary as shown in Figure 4.5c and Figure 4.5d, respectively.

4.4.6 Clustering of samples

Samples were clustered hierarchically based on methylation values using correlation-based distance metric and visualized as heatmaps. Only 2 traits showed up (Figure 4.6). Heatmap displayed only sites with the highest variance across all samples and complete linkage strategy. Most of the traits did not show good clustering except Del (5q) and T (15; 17). Del (5q) positive group which showed hypermethylation but T (15; 17) positive group showed hypomethylation (Figure 4.6a&4.6b). It is difficult to determine whether the lack of association between other traits and the identified methylation-based patient subgroups (data not shown) can be attributed to the heterogeneity of the disease or the incomplete annotation.

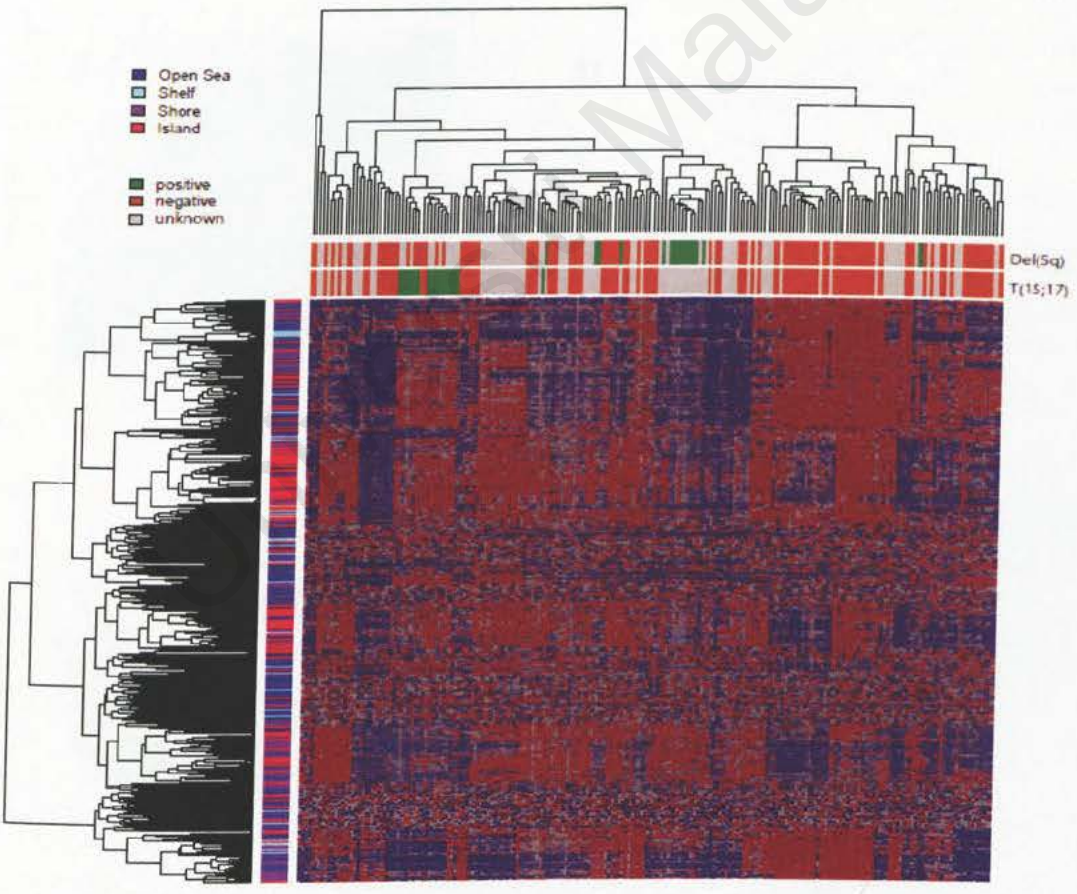
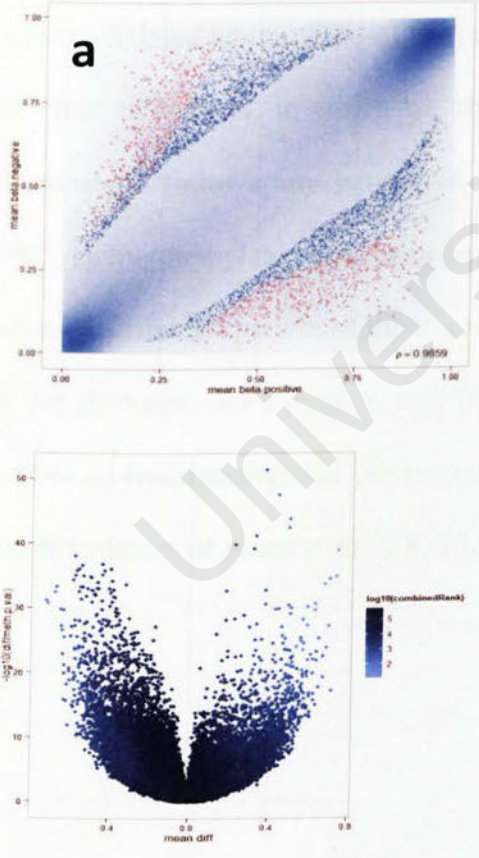


Figure 4.6: **Hierarchical clustering of AML samples.** Heatmap displayed only sites with the highest variance across all samples and complete linkage strategy. x-axis represents the number of patients, and y-axis represents different CpG regions. Gray color in across x-axis side bar represents unknown for particular trait. Color key represents beta value (red means hypomethylated and blue means hypermethylated). Del (5q)-positive patients show mostly hypermethylation across the genomic regions. T (15; 17)-positive patients show hypomethylation across the genomic regions.

4.4.7 Differential methylation

Differential methylation of CGI sites was computed based upon different metrics and limma statistical test was performed for validation. False discovery rate adjusted p-value was <0.05 and 1000 promoter methylation sites were shown in a scatter plot (Figure 4.7a) for T (15; 17). This plot represents the significant difference of the promoter methylation (mean beta value) between positive and negative patients. Figure 4.7b represented volcano plot for T (15; 17) trait that also represents the difference between positive and negative patients. Word clouds for Gene Ontology (GO) enrichment analysis of T (15; 17) showed the carboxylic acid binding problem due to promoter hypomethylation (Figure 4.7c).



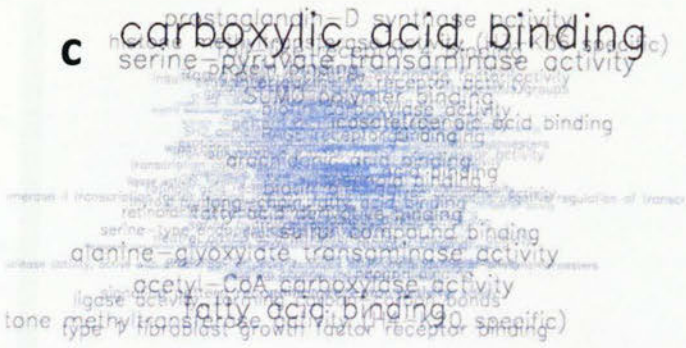


Figure 4.7: Differential methylation of samples based on traits and GO enrichment. (a) Combined differential methylation measure among 1000 best-ranking promoter sites for T (15; 17) positive and negative samples. Red color means positive and blue color means negative samples. (b) Volcano plot for T (15; 17) indicates that positive and negative patients are segregated. (c) Word cloud for T (15; 17) promoter hypomethylation that affects (mainly) carboxylic acid binding.

4.4.8 Highly methylated genes in AML

More than 1300 highly methylated gene promoters were found in the AML dataset. Around 1100 genes were hypermethylated for only 5% of the patients, while 200 were hypermethylated among 90% of the patients (Figure 4.8). Those 200 genes were one of our areas of interest, in which we aim to determine whether their distribution in the chromosome follows any pattern. A Circus plot was generated for those frequently (>90%) hypermethylated genes with associated chromosomal position (Figure 4.9). There was no distribution pattern of those genes; rather they were randomly distributed among all the chromosomes except Y and 18. Chromosome 1 showed a comparatively large number of hypermethylated genes and there were some chromosomes with an only a handful number of genes (Chr-4, 8, 20, 21, 22).

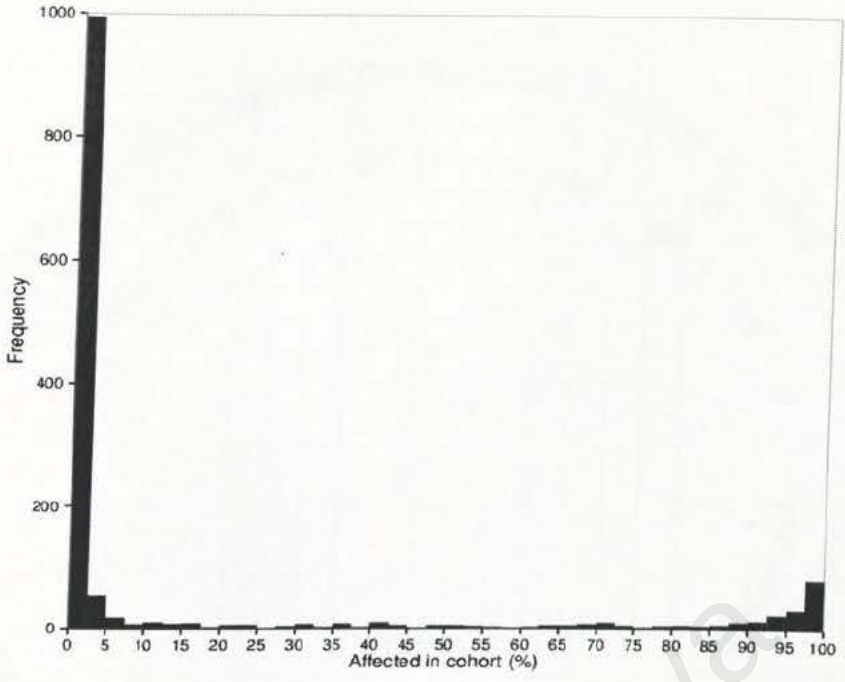


Figure 4.8: Highly methylated gene promoters in the dataset. X-axis is the number of genes and y-axis is the number of patients. Almost 200 gene promoters (right side) are highly methylated in 90% AML patients. We checked the distribution of those genes across the chromosomes in Figure 4.9.

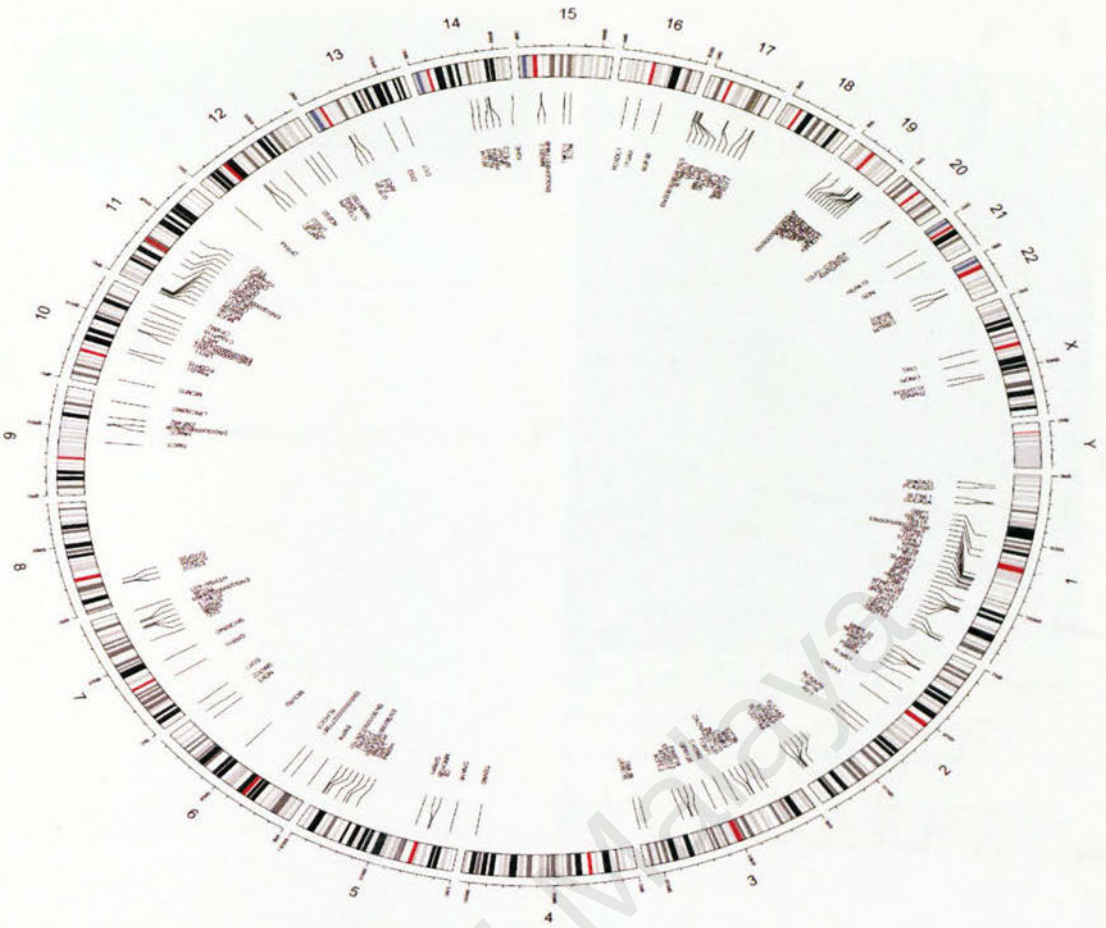


Figure 4.9: Distribution of frequently hypermethylated genes in chromosomes. Commonly hypermethylated Genes in AML are randomly distributed among chromosomes.

4.4.9 Gene mutation and hypermethylation pattern

All possible mutated genes and hypermethylated genes of 200 AML cases were plotted in a single heatmap (Figure 4.10). X-axis indicated patients and Y-axis indicated genes. Genes that were both mutated and hypermethylated were indicated in different colors. Mutation or hypermethylation are also shown in a different color. Figure 4.10a indicated a densely packed group of genes while the rest of the plot was empty. When the part of the densely packed gene was further zoomed (Figure 4.10b), some interesting pattern were obtained. There were at least three groups of patients and two groups of genes that were clearly visible (arrow). Only mutated genes were at the bottom. This is quite an interesting pattern that could be useful in future for better AML classification. The specialties among those groups are still unknown and would be a future area of research interest.

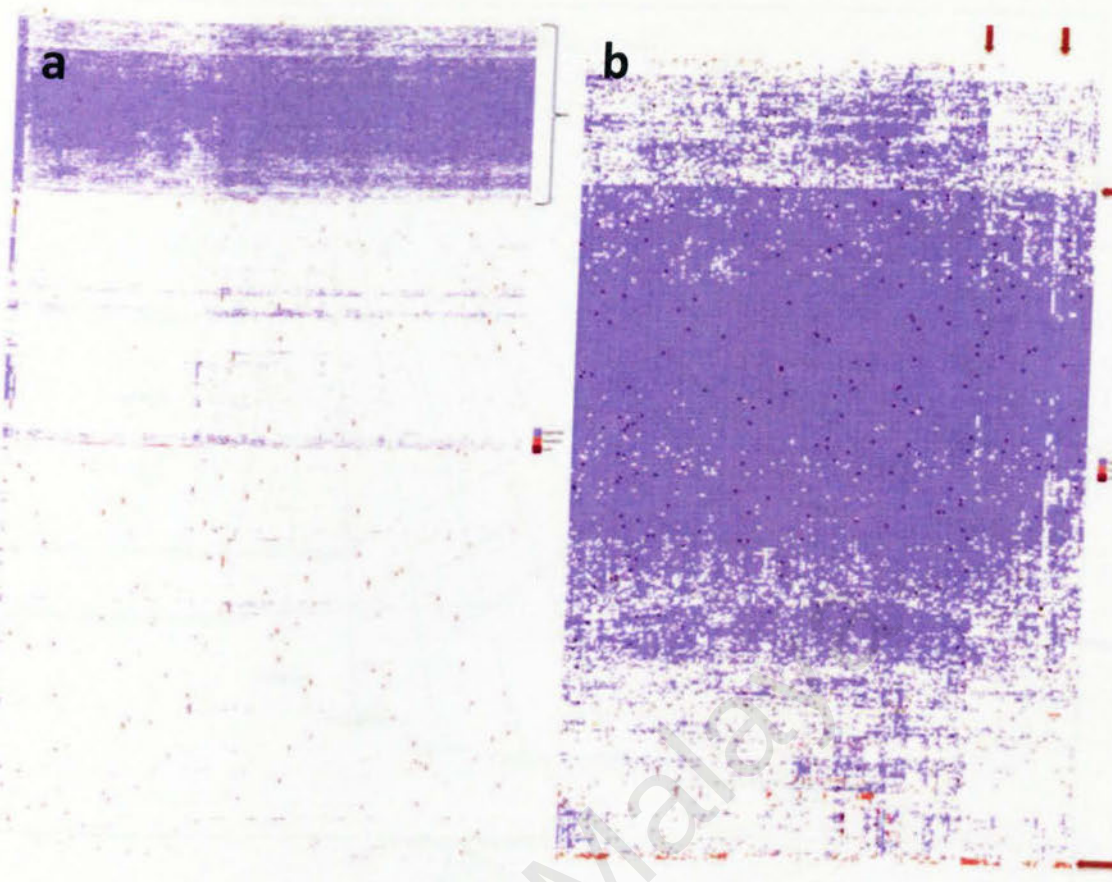


Figure 4.10: Pattern recognition after a combination of methylation and mutation. (a) X-axis represents a number of patients and Y-axis represents the number of genes. (b) The magnifying map shows three distinct groups of patients (arrows) and two distinct groups of genes. One group of the gene is only mutated (bottom) and genes that are both mutated and hypermethylated are scattered randomly.

4.4.10 Pathway analysis

Integrated Pathway analysis for AML was done using IPA online database and the interactive networks are presented in Figure 4.11. The total number of factors/ parameters (genes, transcription factors, and RNAs) affecting AML was 1083 (checked on Sep 23, 2105). The total numbers of upregulated genes were 64 and downregulated genes were 29. Proteasome subunit beta type-9 (PSMB9) and other PSMBs were downregulated and hampered assemble of the proteasome complex in AML. Only 15 miRNAs were found to be directly affecting different AML subtypes and while miR-10 affected multiple AML subtypes.

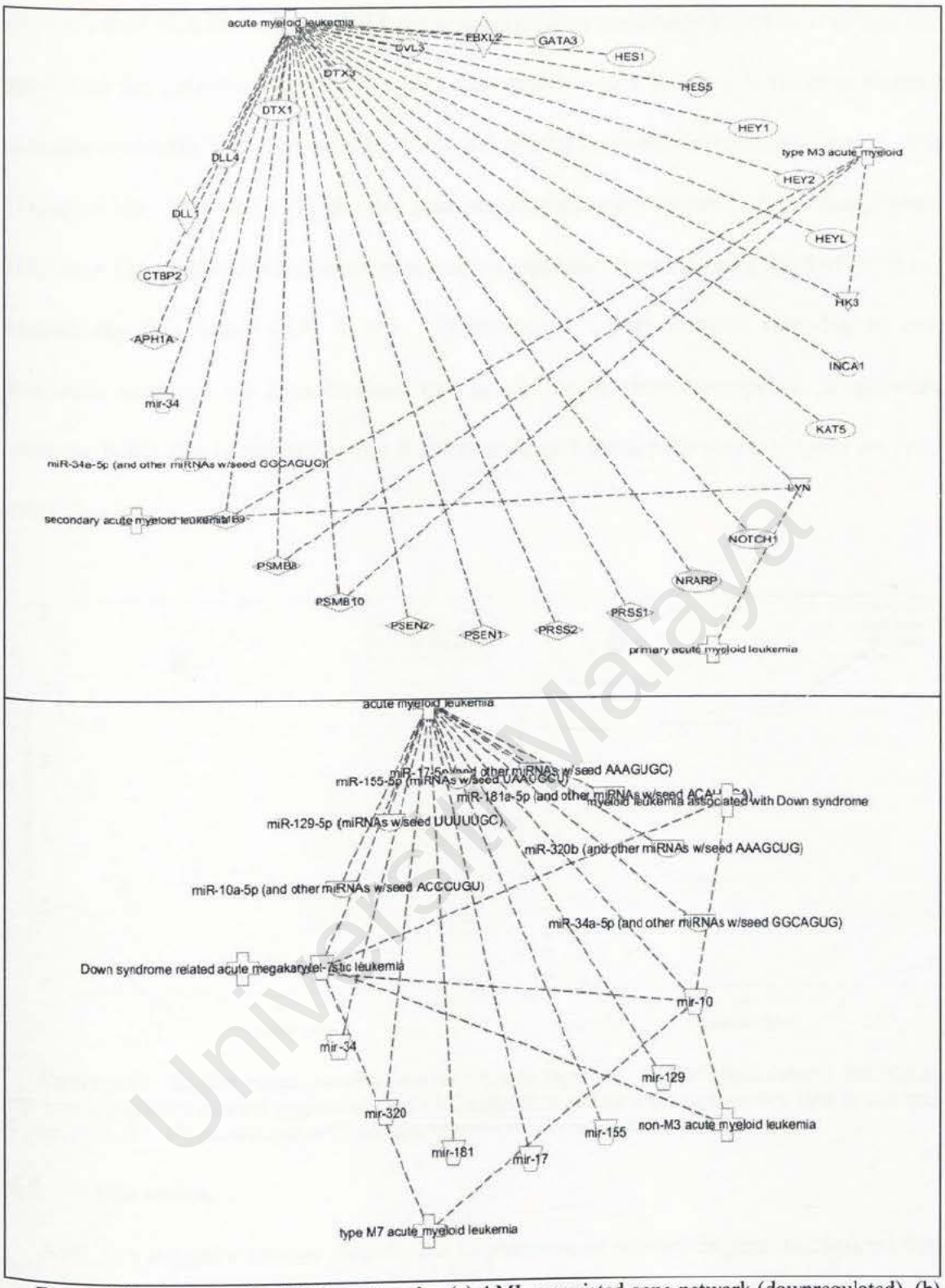


Figure 4.11: Genes and miRNAs networks. (a) AML associated gene network (downregulated). (b) AML associated miRNAs network.

4.4.11 Survival curves

All possible risk factors and traits were combined in order to produce the Kaplan-Meier survival curve (Figure 4.12). P-values were calculated for all possible traits (Table 4.2) and only del (5q), T (9; 11), T (9; 22), and T (15; 17) showed statistical significance

(0.0149, 0.0233, 0.0443 and 0.0457 respectively) after necessary adjustment of age and gender of the patients. Patients with del (5q) positive and T (15; 17) positive showed distinctive results. Del (5q) positive patients' average survival rate of less than 1 year (Figure 4.12a). But T (15; 17) positive patients' average survival rate of more than 3 years (Figure 4.12b). Other traits did not much affect patients' survival, with most of them not statistically significant. Only T (15; 17) showed a higher survival rate due to new treatment strategy. We hypothesized that maybe some chemotherapeutic drugs were working better due to this particular translocation and those patients had higher survival rates than others.

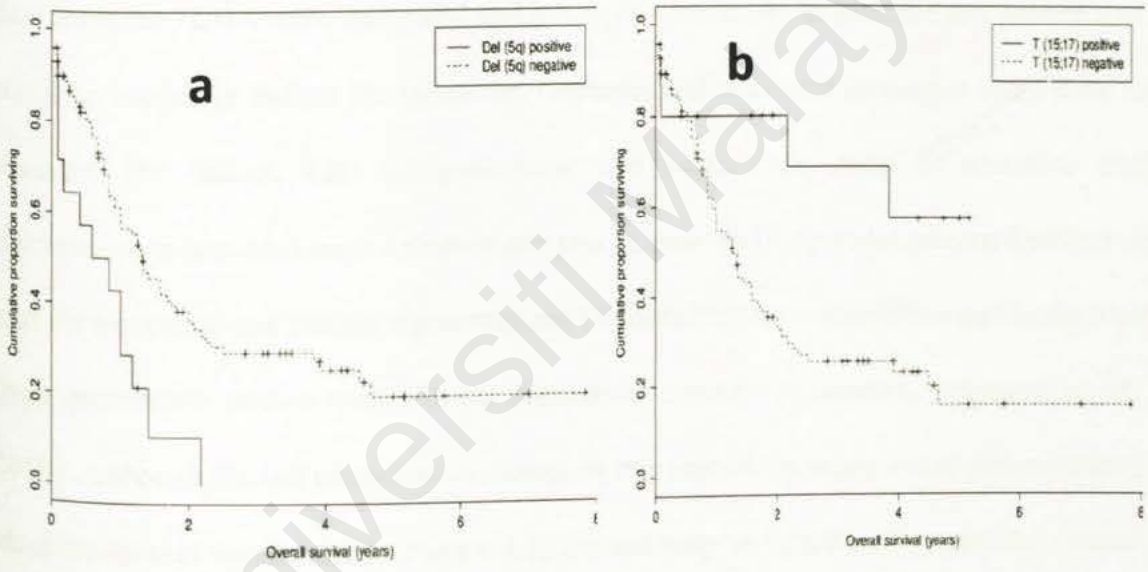


Figure 4.12: Kaplan-Meier survival curves. X-axis represent survival time (years) and Y-axis represent cumulative survival proportion. (a) Del (5q) positive patient's average survival time is less than 1 year. (b) T (15; 17) positive patient's average survival time is more than 3 years.

4.5 Discussion

AML is a complex genetic disease due to presence of numerous genomic lesions that regulate expression of oncogenes as well as tumor suppressor genes. Chromosomal aberrations that cause gene fusions were mostly considered as markers for diagnosis, prognosis, and classification in the early days (Grimwade et al., 2010; Hehlmann et al., 2010; Marcucci, Haferlach, & Döhner, 2011). There were some problems in case of classification based on cytogenetic factors because AML patients with normal karyotypes

(CN-AML) were classified as intermediate risk group, and constitutes more than half of AML cases (Lahortiga & Cools, 2012). Although this heterogeneous group was further classified based on genetic factors, many cases had been reported without even genetic abnormalities. Incorporation of epigenetic factors with genetic and cytogenetic factors might be helpful for better classification as well as diagnostic and prognostic risk stratification.

Classification of AML which takes into consideration all possible genetic, epigenetic, and cytogenetic aberrations is not only a complex and challenging problem, but is also a matter of clinical validation that has never been done before. In this study, all available data from the TCGA were integrated and attempts were made to look for any pattern that could be useful for patient classification. Different software and strategies were used to visualize the dataset. This comprehensive and large-scale study of mutation and methylation in human disease demonstrates that genetic and epigenetic pattern distributed into the biological and clinical signatures and DNA methylation classifiers can be derived from population studies with clinical predictive capacity (Figueroa, Lugthart, et al., 2010). Although the full classification based on this part of the study-is not yet complete, identification of some pattern (Figure 4.10) could help to better understand the overall data structure through which clinically implacable classification might be possible in the future. The analyses made in this study demonstrate that systematic integration of gene expression and DNA methylation profile could improve the classification methods, especially CN-AML could be understood better.

Although there is a huge genomic region outside of the promoter, most of the DNA methylation studies in AML have focused mainly on CGI promoter region (Deneberg et al., 2010; Yalcin et al., 2013). There are almost 28 million CpG dinucleotides in the human genome (M. M. Suzuki & Bird, 2008) and only 2% of them are densely located to

the promoter region (Qu et al., 2014). Significant statistical and biological evidence of non-CGI methylation are now increasing and are reported to play an important role in the regulation of gene expression (De et al., 2013; M. Suzuki et al., 2011). In this study, total numbers of methylation distributions are high at open sea and shore (Figure 4.5b) as compared to the promoter region; which might be due to the relatively large genomic region outside of promoters. In CLL, methylation alterations have been reported in the non-CGI region (Cahill et al., 2013). DNA methylations in shore regions are strongly correlated to gene expression changes in colorectal cancer (Akalin et al., 2012; Irizarry et al., 2009). Non-CGI DNA methylation can be used as a biomarker for differentiation of pluripotent stem cells (Butcher et al., 2016). As bona fide leukemogenic mutation (FLT3, NPM1, CEBPA, and DNMT3A) are not sufficient to explain diverse clinical AML subtypes, epigenetic alterations have been found to be abundant and are common and might explain the biology behind various AML groups.

It has been reported that some genes (FLT3, NPM1, and DNMT3A) are recurrently mutated (more than 20%) in AML (Network, 2013). It is however, not clear yet which genes are particularly involved for what specific methylation pattern. Mutations of NPM1 were associated with four slightly distinct epigenetic signatures that could not be explained by concurrent FLT3-ITD mutation (Figueroa, Lugthart, et al., 2010). This suggests that there might be some unrecognized mechanisms to determine different epigenetic groups. This study also suggests that almost 200 genes are recurrently hypermethylated (Figure 4.8) in 90% AML cases and are randomly distributed among chromosomes (Figure 4.9).

As DNA methyltransferases are responsible for *de novo* methylation, DNMT3A is a crucial gene for epigenetic alteration of AML. DNMT3A mutated patients show global hypomethylation (Qu et al., 2014) because impaired catalytic activity has been recently

reported to be due to this mutation (Holz-Schietinger, Matje, & Reich, 2012). There is a different pattern of DNMT3A methylation in AML cases. In bone marrow, mononucleotide cells of AML showed significant hypomethylation of DNMT3A (Zhang et al., 2015), but peripheral blood cells DNMT3A were found to be hypermethylated (Jost et al., 2014). It can be said that aberrant DNMT3A methylation would be an independent negative prognostic factor in AML. In this analysis, many traits have been associated with abnormal methylation patterns. T (15; 17) positive patients showed global hypomethylation but del (5q) represent hypermethylation (Figure 4.6). Other traits were not showing any significant cluster or pattern in heatmap but this may be due to insufficient sample numbers or molecular heterogeneity between inter and intra traits.

One of the notable findings of this study was the identification of patterns in heatmap (Figure 4.10) with a combination of all probable genetic and epigenetic signatures. Three groups of patients were clearly visible at the zoom in a part of the figure, whereas at least two groups of genes were able to be identified. Only mutated genes were clustered at the bottom of the figure and genes that were both methylated and mutated had no clear pattern. The common or differentiating features of those groups are not clear yet. This could be a further area of study before considering epigenetic patterns as a factor for AML patient classification.

Another important observation from the AML dataset is the role of traits in patients' survival. Del (5q) positive patients survive less than one year. Only T (15; 17) shows higher survival rate (>3 years) and this is quite exceptional when compared to other traits (Figure 4.12b). Perhaps some chemotherapeutic drugs were working better in those patients with this translocation, as many drugs have recently been introduced in which PML-RARA translocation is targeted. There is no prescribed drug information in the TCGA (LAML) clinical data files. A large cohort of AML patient with detailed clinical

information is required to validate the hypothesis. Survival curves (Kaplan-Meier) for many traits (Table 4.2) were not statistically significant (0.05), most probably due to small positive sample size and some traits (like FLT3 and NPMc) which showed the same survival pattern for both positive and negative patients.

Table 4.2: Statistical test results for survival curves.

Aberration	Positive	Negative	P-value (adjusted)
Inv (16)	9	103	0.1671
Del (5q)	15	103	**0.0149**
Del (7q)	21	102	0.0656
T (4;11)	1	103	0.8124
T (8;21)	7	103	0.3686
T (9;11)	2	103	**0.0233**
T (9;22)	3	103	**0.0443**
T (15;17)	16	98	**0.0457**
Trisomy 21	8	103	0.1900
Trisomy 8	20	101	0.7451
Activating RAS	11	180	0.7481
BCR-ABL	1	15	n/a
FLT3 Mutation	57	130	0.0744
IDH1 R132	18	171	0.9118
IDH1 R140	15	175	0.2802
IDH1 R172	2	188	0.3275
NPMc	45	145	0.2866
PML-RAR	8	9	0.0675
Cytogenetically normal	97	97	0.9888

There are some limitations of this study that might have to be considered before clinically implementable decisions can be put into practice. Numbers of samples were collected during a long period of time and all clinical tests were not performed equally. Many routine tests have now already been changed in the clinic and new genetic testing assigned. Some data were missing from the clinical data file and there were some batch effect in microarray data. There was also a challenge to use RnBeads to analyze, interpret, and visualize TCGA data. Feasibility of the findings has not been tested in the clinical context. Although the statistically significant data were presented here, statistical significance might not always be translated into clinical practice.

Finally, this secondary data analysis revealed global pictures of AML genomes that could be useful for further broad spectrum analysis and interpretation of cancer epigenomics data. These integrated analysis approaches show some interesting findings like epigenetic patterns and survival curves. How individual trait is associated with global methylation changes were shown in this study.

4.6 Conclusion

In this study, TCGA AML data have been comprehensively analyzed to understand the underlying mechanism of genetic, epigenetic, and cytogenetic abnormalities in AML patients. In previous chapter (Chapter 3), we showed how TCGA AML data correspond to C&Ckines expression data in Malaysian population (Figure 3.11 to 3.14). Many AML cases (40-50%) have been reported to have normal karyotypes, but the hypothesis is that there might be distinctive features of genetic and epigenetic abnormalities. To identify any features or patterns, 200 AML cases have been selected from a publicly available cohort, and then differentially analyzed for genetic, epigenetic and cytogenetic abnormalities. Three genes (FLT3, DNMT3A, and NPMc) are found to be predominantly mutated. Pathway analysis identified 93 AML associated genes, of which 64 were

upregulated and 29 downregulated. Almost 200 genes are highly methylated in 90% of AML cases and equally distributed among all chromosomes. Del (5q) and T (15; 17) positive patients show a distinct pattern in the heatmap plot. Principal component analysis shows a clear distinction between NPMc positive and IDH1 R172 negative patients. IDH1 R172 hypomethylation is associated with abnormal cellular adhesion. Three distinct groups of patients and two distinct groups of genes were found to be hypermethylated. Four aberrations-Del (5q), T (15; 17), T (9; 22), and T (9; 11)-were significantly ($p<0.05$) associated with patient survival. In particular, Del (5q) positive patients have an average survival of less than 1 year, whereas T (15; 17) positive patients have significantly higher mean survival time (>3 years). From this observation, it can be hypothesized that currently prescribed chemotherapeutic drugs are more efficient in the presence of this translocation (T (15; 17)). A combination of genetic, epigenetic, and cytogenetic data could further expand our understanding of the biology of AML. This analysis would motivate other groups to increase the sample size to confirm our findings as well as to include epigenetic signatures for classification of AML. Larger sample sizes would be needed to confirm our findings. We hope our analysis might motivate clinicians to eventually include epigenetic signatures in the classification of AML and inspire them to investigate the association with patient survival, T (15; 17), and in particular, with drug responses.

CHAPTER 5: DIFFERENTIAL ANALYSIS OF MLL-GAS7 FUSION PROTEIN FOR UNDERSTANDING OF AML PATHOGENESIS

5.1 Introduction

Chromosomal translocations are common cytogenetic abnormalities in acute myeloid leukemia (AML) that are often linked to the generation of oncogenic fusion proteins. Different types of translocations have been reported in AML and they are distinct based on pathogenic severity and describing frequency among patients. The study of specific chromosomal translocation resulting fusion protein may provide relevant diagnostic and prognostic information that can increase our knowledge pertaining to the underlying pathogenesis of the disease (Alcalay et al., 2001). One of the frequently fused genes in leukemia is mixed-lineage leukemia (MLL) which causes a chromosomal rearrangement in 10% of adult and 80% of infant acute leukemia (So, Karsunky, Wong, Weissman, & Cleary, 2004). In general, AML patients with MLL translocation exhibit poor prognosis (Eguchi, Eguchi-Ishimae, & Greaves, 2005). MLL fuses to more than 50 different partner proteins in various leukemia lineages (DiMartino & Cleary, 1999; So, Karsunky, et al., 2003), which is in contrast to most other chimeric oncoproteins in leukemia, which are not so variable and are relatively specific for a single lineage. The leukemia blasts of the mixed-lineage frequently express surface markers for both myeloid and lymphoid lineages. It has been reported that, patients who were initially diagnosed with acute lymphoblastic leukemia (ALL) can switch the lineage and relapsed as AML (Stass et al., 1984), hence the term mixed-lineage leukemia (Mirro et al., 1986). The main causal agent of this lineage switching leukemia is the MLL gene that fuses with different partner genes.

There are no reported common functions among MLL fusion proteins, whereas the partner proteins work as an effective unit for transactivation (Yokoyama, Lin, Naresh, Kitabayashi, & Cleary, 2010). Cooperatively they reprogram transcriptional signatures (mostly epigenetic) and alter the normal function of the transcription factors. Fusion

proteins frequently interfere with the expression of target genes necessary for myeloid development and differentiation that leads to leukemic transformation (Alcalay et al., 2001). Potential targeting of these interfering pathways could be a major focus for the development of novel therapeutics. One successful example of such therapeutic is ATRA in acute promyelocytic leukemia (APL) fusion protein (Melnick & Licht, 1999). Five-year survival rates of non-mixed lineage leukemia in children approached 90% (Pui et al., 2004), but mixed lineage leukemia survival only shows 40% in children (Pui et al., 2004; Tomizawa et al., 2007).

Normal MLL gene encodes a protein that functions as DNA methyltransferase, which is necessary for efficient transcription. The MLL gene encodes a homolog of the Drosophila fly gene called trithorax (Trx) and the finding of its involvement in leukemia had initially influenced the name to be HRT (human trithorax), ALL-1 (acute lymphocytic leukemia-1) or MLL (mixed lineage leukemia) but later it was agreed that the term MLL be used (Slany, 2009). MLL protein is post-translationally modified by proteolytic cleavage to form N-terminus (320kDa) and C-terminus (180kDa) fragments (Hsieh, Ernst, Erdjument-Bromage, Tempst, & Korsmeyer, 2003). The C-terminus of MLL contains a highly conserved SET domain that methylates at H3k4 (Milne et al., 2002). Another protein (MOF) associated with C-terminus, loosens up chromatin folding by charge neutralization through acetylating at H4k16 (Dou et al., 2005). A nucleosome remodeling complex has also been identified in this terminus (Ernst, Wang, Huang, Goodman, & Korsmeyer, 2001). The N-terminus portion of the MLL directs the protein complex to specific loci through DNA binding (Ayton, Chen, & Cleary, 2004; Slany, Lavau, & Cleary, 1998). It also contains a binding site for a tumor suppressor protein called menin (multiple endocrine neoplasias) and together with MLL they form an interactive surface for LEDGF (lens epithelium-derived growth factor) that binds to chromatin via PWWP domain (Yokoyama & Cleary, 2008). Evidence have shown that

transcription factors (p53 and β-catenin) have been associated with MLL and these probably recruit the MLL complex to initiate RNA synthesis (Dou et al., 2005; Sierra, Yoshida, Joazeiro, & Jones, 2006).

Translocation of MLL (11q23) is one of the causal agents for many AML cases and it is most likely that the aberrant non-homologous end joining of the double-strand breaks is responsible for the translocation. This DNA repair mechanism is also known from the generation of antibodies and T-cell receptors (Slany, 2009). Due to translocation at 11q23, MLL gene loses normal histone methyltransferase functions and is replaced by heterogeneous function contributed by corresponding fusion proteins. One of the striking features of MLL is the diversity of fusion proteins, generated through 85 recurrent translocation and 66 partner genes (*http://atlasgeneticsoncology.org/Genes/MLL.html; last update 02/2010*). All partner proteins follow two rules, (1) they fuse in frame to N-terminus fragment, upstream of the PHD fingers and (2) the most frequently reported proteins (AF4, AF9, ENL, AF10, ELL, AF6) exhibit nuclear localization (except AF6, located in the cytoplasm) (Meyer et al., 2009). Several studies have demonstrated that the LEDGF-menin binding motif and the CxxC domain were absolutely required for MLL fusion protein to function (Ayton et al., 2004; Slany et al., 1998). There are four different ways in which those fusion proteins work for leukemic blast formation: H3 lysine methylation, H4 arginine methylation, histone acetylation, and dimerization (Slany, 2009).

One of the partners for MLL is growth arrest-specific 7 (GAS7) in the myeloid lineage resulting from t(11;17) (q23;p13) and generate chimeric protein (MLL-GAS7). GAS7 is a relatively less frequent partner protein and shows cytoplasmic localization. It is evolutionally conserved in mammals and encodes for more than one protein isoform (Chang, Kuo, Lin-Chao, & Chao, 2005; Lazakovitch et al., 1999). From evolutionary,

genetic, and disease point of view, GAS7 is important because of protein sequence among various species and conservation of orthologous DNA (Su et al., 2000). It also becomes crucial to understand the activation mechanism of MLL-GAS7, one of the responsible factors for AML pathogenesis. Gas7 is located on human chromosome 17 which also hosts some important genes, like BRCA1 (17q21) for breast cancer and TP53 (17p13.1) for DNA damage repair (Zody et al., 2006). GAS7 is 48kDa protein, usually expressed from terminally differentiated brain cells at G_0 phase and regulate morphological differentiation and neuritogenesis (Ju et al., 1998). Human GAS7 isoform-c possesses N-terminus SH3 domain, WW domain, Fes/CIP4 homology (FCH) domain, and C-terminus coiled-coil (CC) region (Aspenstrøm, 2009). The SH3 domain binds to proline-rich ligands and is found in the cytoskeleton regulatory signaling pathway (Mayer, 2001). The WW domain provides a versatile platform for protein-protein interaction (Ingham et al., 2005) and in GAS7 may function to regulate actin dynamics. The FCH and CC domains are structurally similar to Bin/amphiphysin/RVS (BAR) domains and function in endocytosis, actin cytoskeleton reorganization that leads to membrane protrusions, vesicular trafficking, cell migration, and signaling (Heath & Insall, 2008). The coiled-coil domain of GAS7 is necessary and sufficient for activation of MLL oncogenic potential (So, Lin, Ayton, Chen, & Cleary, 2003).

Although CC domain of GAS7 is necessary for homodimerization of MLL-GAS7 and sufficient for leukemogenic transformation (So, Lin, et al., 2003), it is unknown how dimerized fusion proteins activate target genes that lead to abnormal differentiation of hematopoietic progenitors. Hematopoietic progenitors transduced by MLL-GAS7 exhibited very different Hox genes expression profile (So, Lin, et al., 2003), which is a downstream target of wild-type MLL (Benjamin, Hess, Horning, Brown, & Korsmeyer, 1995). Differential expression of Hoxa7 and Hoxa9 are associated with normal

hematopoietic progenitor homeostasis but they are not necessary for MLL-GAS7 mediated leukemogenesis (So et al., 2004).

To get a better sense of how MLL-GAS7 fusion protein functions and of the underlying oncogenic transformation mechanisms, MLL and GAS7 downstream target genes expression have been differentially analyzed from the Cancer Genome Atlas (TCGA) database. As abnormal MLL usually co-function with other chimeric proteins, the most frequently fused genes expression pattern in AML patients have been checked. GAS7 associated proteins network were drawn and their expression pattern in AML patients were checked. To investigate the possible role of GAS7 as MLL fusion protein partner, GAS7 expression in N2A cell line were differentially analyzed. Computational biology approaches were also used to find out MLL-GAS7 mediated leukemogenesis mechanism.

5.2 Literature review

MLL is one of the major fusion proteins in AML and can fuse with more than 66 partner genes, hence demonstrating the complexity of this disease. Although some studies reported approximately 10% of adult therapy-related leukemia represents MLL translocation-related abnormalities, clinical data from TCGA AML paper noted that around 30% AML patients are harboring MLL abnormalities. It is more common in infant leukemia (>70%) (Biondi, Cimino, Pieters, & Pui, 2000). DNA binding protein encoded by normal MLL gene could methylate histone at H3K4 and positively regulates gene expression, whereas fusion transcripts of MLL lose their transferase activities (Krivtsov & Armstrong, 2007). A hallmark of MLL mediated leukemia is fusion with GAS7 and formation of MLL-GAS7 chimeric protein (So, Karsunky, et al., 2003). MLL-GAS7 works for the deregulation of many downstream genes especially HOX genes and ultimately leads to leukemic transformation. But the mechanism of MLL-GAS7 is yet to

be clearly understood and this study is designed to reveal some insight into the features of MLL-GAS7 and AML pathogenesis.

5.3 Materials and methods

5.3.1 Data analysis and visualization

TCGA (https://tcga-data.nci.nih.gov/tcga/) data repository was used as primary data source for differential gene expression analysis. To analyze the AML data generated by TCGA, the raw data were directly accessed and downloaded, using the "data matrix" tool provided by TCGA data portal (date of download: May 25th, 2016). Clinical data were available for 200 patients and downloaded as a .txt file and reported gene fusions were analyzed. Affymetrix U133 Plus 2 platform was used by TCGA for 197 AML samples for genes expression profiling (Network, 2013). Fifty .CEL files were randomly selected from TCGA for AML cases and nine .CEL files were downloaded from GEO (GSE38511) for control groups. R platform (version 3.1.3) was used to analyze the data files and other R/Bioconductor packages were used to visualize the data.

5.3.2 Cell culture and chemicals

For differential GAS7 expression analysis, neuroblastoma cell line, Neuro 2A (ATCC® CCL-131™; American Type Culture Collection, Manassas, VA), was cultured at 37^{0}C in 5% CO_2 incubator in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 1% sodium Pyruvate, and 1% penicillin (Gibco, Rockville, MD).

5.3.3 Overexpression of GAS7

Gas7 shRNA (shGas7) was expressed in N2A cells using cloning vector (pLAS3w.PeGFP-I2-Puro). Plasmid construction, DNA preparation, and subsequent

transfection were performed according to established protocol (Sambrook, Fritsch, & Maniatis, 1989). A green fluorescence protein (GFP) was tagged at the 3'-end of the Gas7 fused plasmid. TRC Clone ID was Gas7-C-SF-TAP-3 and plasmid size was 12kb. Lipofectamine 2000 reagents were used for transfection. N2A cells were seeded at 8×10^5 cells/well in 500 µl media. DNA-lipofectamine 2000 complex was prepared without serum (Opti-MEM medium) and added to each well that contained N2A cells and incubated for 48h. Cells that over-expressed GAS7 containing plasmid were examined using a fluorescence microscope. Cells were grown in coverslip and the image was taken after fixing the cells in glass slide to observe GFP expression that indicated GAS7 co-expression in N2A cells.

5.3.4 Detection of GAS7

After incubation (48h), whole proteins were extracted and separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) from both wild-type cells (N2A-P7) and transfected type cells (N2A-EGFP-GAS7), prior to transfer onto polyvinylidene difluoride (PVDF) membranes (Immobilon; Millipore, Bedford, MA). The membranes were blocked with 5% non-fat dried milk for 1 h. Membranes were blotted for 15 min at room temperature with a rabbit anti-Gas7 primary antibody diluted 1:3,000 and horseradish peroxidase conjugated secondary antibody diluted 1:6000 (Santa Cruz Biotechnology, Santa Cruz, CA). After that, the signal was detected by western blot on x-ray film by the ECL system (Amersham Bioscience, Little Chalfont, UK).

5.3.5 Cells growth rate estimation

Two types of cells (N2A-P7 and N2A-EGFP-GAS7) were seeded in 25-cm Petrie dish plate with DMEM medium. As N2A cells attested to the plate, Trypsin-EDTA treatment was required for cell counting. Cells were counted (triplet) after every 12 hours (0h, 12h, 24h, 36h, 48h, and 72h) using both manual and automated cell counter (Invitrogen).

5.3.6 Stress and serum dose responses

To determine the stress response of the two types of cells, the same plates were treated with Trypsin-EDTA every 12h, the cell numbers counted, replated again up to 72h, and the growth rate observed during continuous stress. Cells were grown without serum (0.5 % BSA) and neuritis formations were noticed under a microscope (0.5 h, 2 h, 6 h, 24 h). Neuritis formations were also noticed due to different serum concentrations (10%, 1%, 0.1%, and 0.01%).

5.3.7 Network analysis

GAS7 associated genes and miRNAs were selected (20-10-2015) from Ingenuity Pathway Analysis (IPA, www.ingenuity.com). Using GeneMania (Warde-Farley et al., 2010) and MirOB (http://mirob.interactome.ru/) software, gene network and miRNA network were drawn respectively.

5.4 Results

5.4.1 MLL-GAS7 fusion event and tissue distributions

Although DNA topoisomerase II inhibitors are popular chemotherapeutic drugs for primary cancer treatment, complications have been reported for leukemia with MLL gene translocation (Megonigal et al., 2000). Megonigal *et al.* investigated the fusion event of MLL gene with a partner gene (Gas7) at an early stage of chemotherapeutic treatment

using panhandle PCR-based methods. Two different MLL-GAS7 fusion transcripts were found where breakpoint in MLL was either intron-8 or intron-7 but breakpoint in GAS7 was always exon-1. Fifty-three nucleotide base pairs were lost due to intron-8 and exon-1 fusion event (Figure 5.1)(Megonigal et al., 2000).

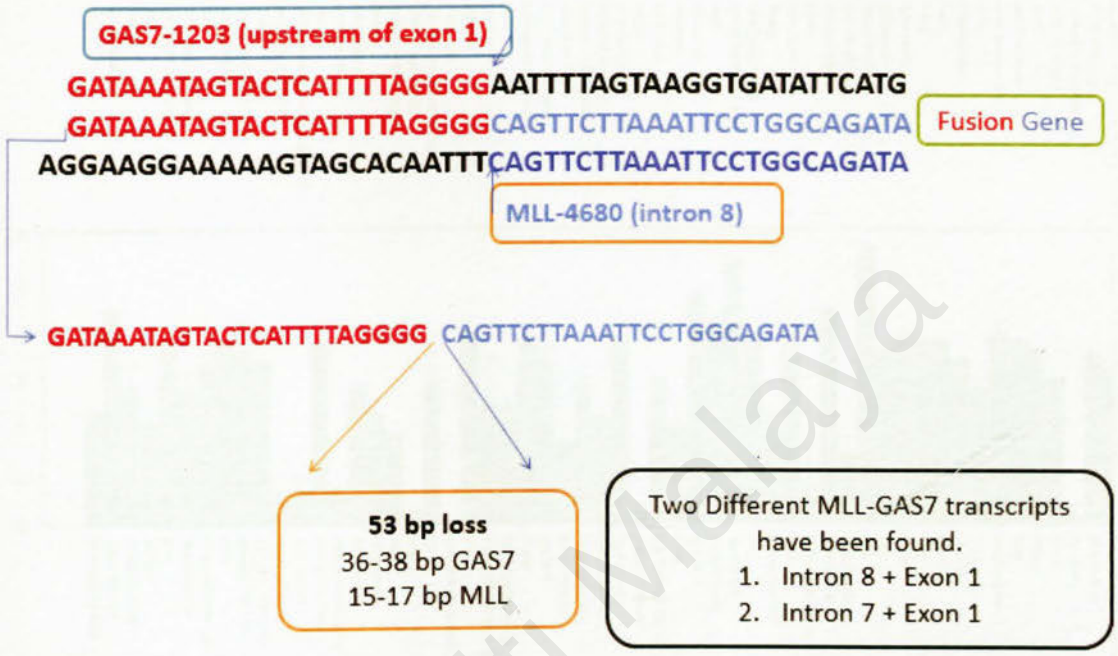


Figure 5.1: **Fusion event between MLL and GAS7.** Intron 8 of MLL and exon 1 of GAS7 fused together and lost 53 base pairs (bp). Break points have been shown by the arrow. Red color means GAS7 and green color means MLL.

The expression level of MLL and GAS7 vary based on tissue types. Expression values for different genes in different tissues were summarized by GeneProf (Halbritter, Kousa, & Tomlinson, 2014) from 250 public datasets. The average expression of MLL was 129.87 RPM (reads per million) ranging from 2.76 RPM to 635.22 RPM. The overall mean GAS7 expression was (63.02 RPM) almost half of MLL ranging from 0.0 RPM to 772.50 RPM. Interestingly, brain tissues were noticed to be the most highly expressed for both MLL and GAS7 (Figure 5.2).

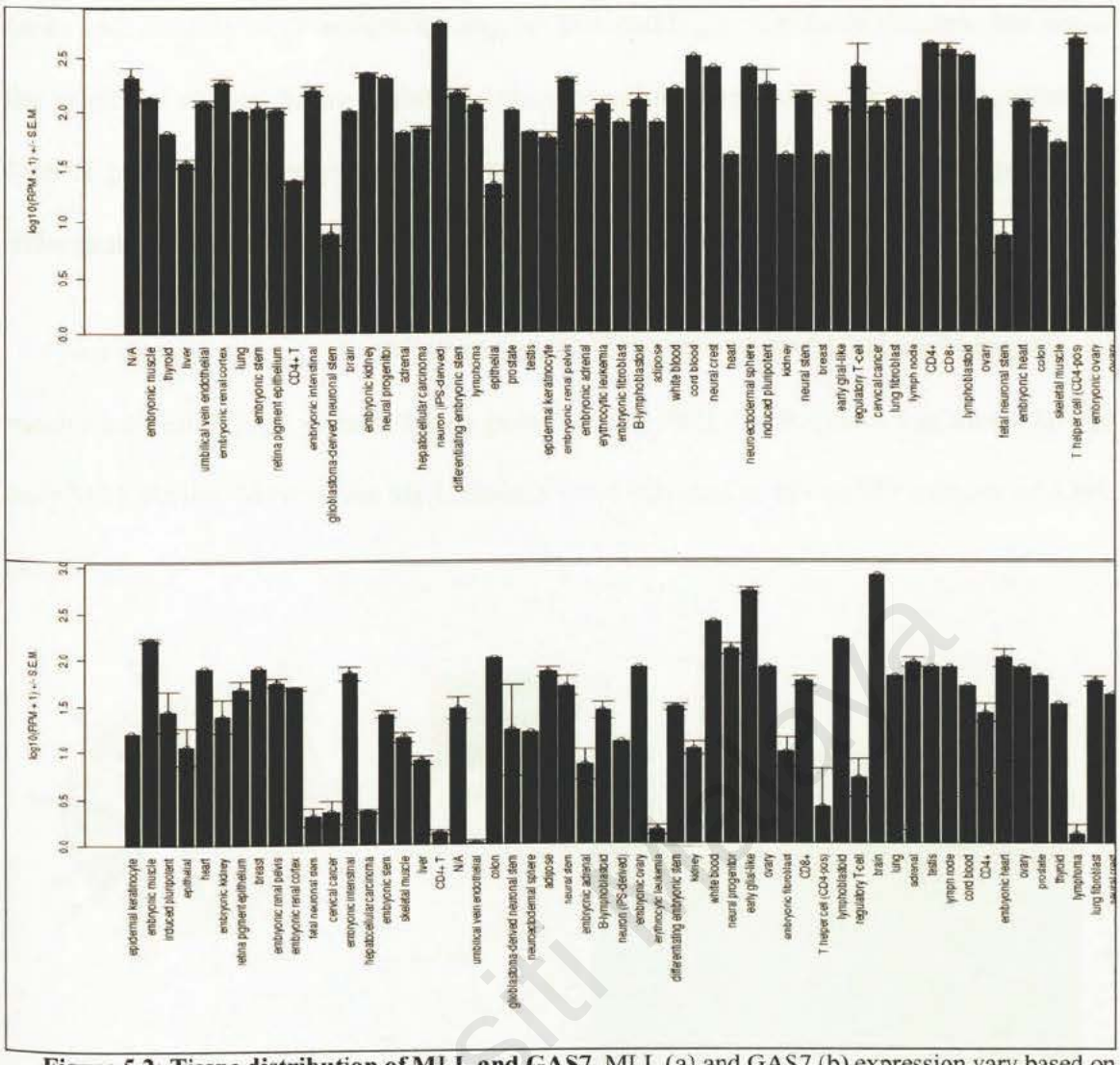


Figure 5.2: Tissue distribution of MLL and GAS7. MLL (a) and GAS7 (b) expression vary based on tissues. Both expressed higher level in the brain. Expression values have shown log10 of RNA transcript reads per million.

5.4.2 Frequently fused genes in AML and expression pattern

In the TCGA AML cohort, frequently detected in-frame (green) and out-frame (orange) fusion genes were shown in a circos plot (Figure 5.3a) where ribbon widths are proportional to the frequency of a fusion event. Each of those particular fusion genes expression patterns was detected from Affymetrix U133 Plus 2 expression array of AML patients (Network, 2013) and plotted as heatmap (Figure 5.3b). Three clusters of fusion genes were observed based on log2 expression values: low expression (GAS6, FAM70B, MYH11, MLLT4, MLLT3, RUNX1, RUNX1T1, and GPR128), moderate expression (RARA, NSD1, PML, ELL, ABL, NUP98, and MLLT10), and high expression (TGF, CBFB, PMAP1, MLL, GAS7, BCR, and PICALM) clusters. Notable variation among

cases and controls were noticed among the low and high expression clusters, but not in the moderate cluster. Some of those genes also exhibited a similar expression pattern in control groups (especially moderate expression cluster) that indicates fusion genes and individual partner genes which follow distinct expression patterns.

From clinical data of 200 AML patients, 62 (31%, M=36, F=26) showed MLL fusion, mostly with other predominant fusion proteins (like PML-RARA) and 5 of them showed only MLL fusion. Most of the MLL fusions were reported at M4 or M5 subtype of AML (43.54%).

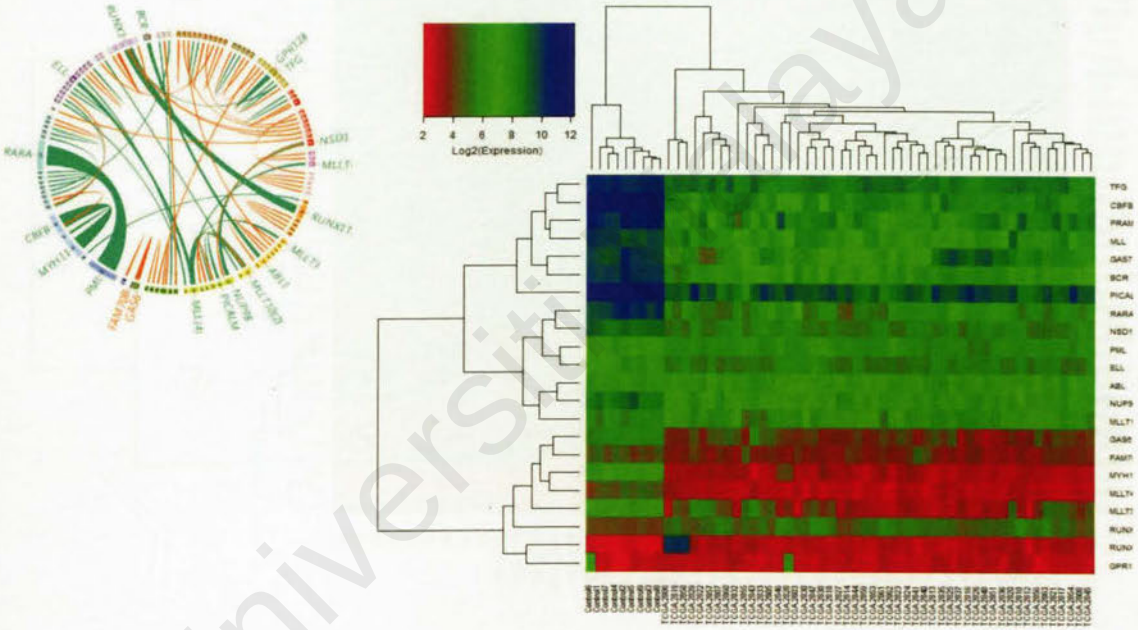


Figure 5.3: Frequently fused genes in AML and their expression pattern. Frequently fused genes shown as circus plot (3a) where ribbon widths are representing the frequency of fusion in TCGA AML dataset (modified from TCGA paper). The green color is in-frame fusion and orange color is out-of-frame fusion.

5.4.3 MLL and GAS7 downstream target genes expression

As the N-terminus fragment of the wild-type MLL were thought to be involved in target selection that fused with GAS7, it seems likely that the chimeric protein will target and regulate many of the MLL and GAS7 downstream genes expression. In order to confirm this assumption, 38 genes were identified from literature as downstream target

genes and Affymetrix U133 Plus 2 array of AML patients (Network, 2013) was considered for expression levels. Seven AML patients with MLL fusion, 25 AML patients without MLL fusion and 9 controls were used for comparative analysis and visualized as heatmap (Figure 5.4).

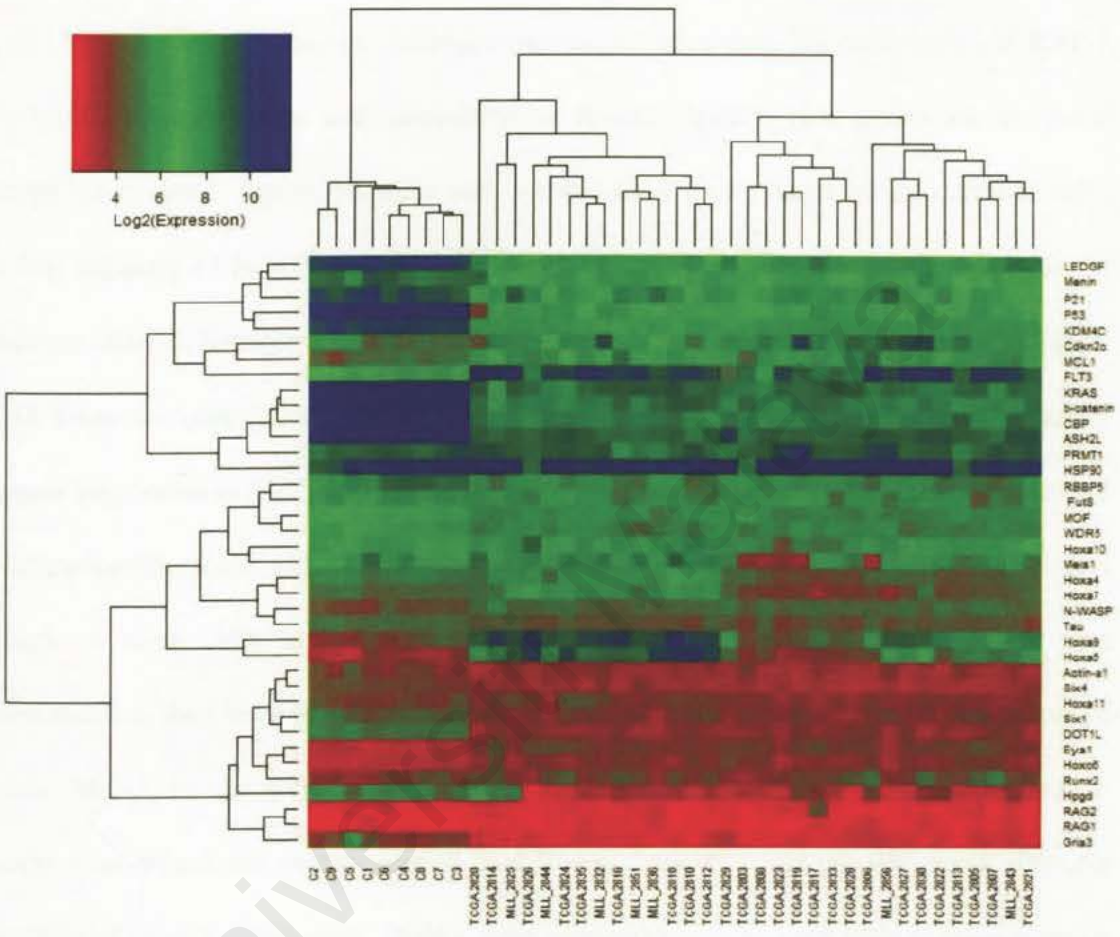


Figure 5.4: MLL and GAS7 downstream target genes expression. Based on TCGA expression data, commonly known downstream target genes expressions are shown. Only 9 controls, 7 AML patients with MLL fusion and 25 patients without MLL fusion are included. Notable differences noticed between cases and controls.

Differential expressions of HOX genes were noticed. Hoxa9 and Hoxa5 expressions, in particular, were significantly higher in AML patients (notable variation noticed between patients with and without MLL fusion). FLT3, HSP90, Eya1, and Hoxc6 expressions were higher, but P53, P21, KDM4C, and LEDGF expression were lower in AML patients as compared to healthy controls. MLL chimeric proteins activation influence secondary translocation and/or mutagenesis (Eguchi et al., 2006) due to

inhibition of the tumor suppressor protein p53 (Megonigal, Rappaport, Nowell, Lange, & Felix, 1998; Wiederschain, Kawai, Shilatifard, & Yuan, 2005). Expression differences were also noticed for kRAS, b-Catenin, CBP, and ASH2L. Low expression of Six1, DOT1L, Gria3, and Hpgd was observed in AML cases. Wang et al., however, reported that they found both Eya1 and Six1 to be overexpressed in acute leukemia (Q.-f. Wang et al., 2011), although the subtype of leukemia was not specified. No differences in RAG1 and RAG2 between cases and controls were found. Clearly, Hox genes are the most susceptible of MLL fusion proteins and ectopic Hox genes expressions enhance self-renewal capacity of hematopoietic cells and block their differentiation (Ernst, Mabon, Davidson, Zon, & Korsmeyer, 2004; Li et al., 2009; Milne, Martin, Brock, Slany, & Hess, 2005). From the experiments for MLL fusion controlled transcript, only a small number of genes responded to MLL fusion and many genes were resistant to MLL fusion induced modification (Horton et al., 2005; Zeisig et al., 2004). Although the Hox genes expression is high in stem cells and early precursors where they control development and differentiation, they have to be downregulated for maturation. Hoxa9 and its dimerization partner Meis1 levels were higher in AML cases and they are among the major oncoproteins which are overexpressed in different leukemia and usually, work through activation of c-Myb (Hess et al., 2006). As compared to MLL, very little is known about the regulatory role of GAS7. Arp2/3 complex-mediates actin polymerization is activated by N-WASP, that is found to be upregulated in AML patients and GAS7 then interact with N-WASP through N-terminus WW domain.

5.4.4 GAS7 associated proteins network and expression pattern

Reported proteins that are associated with GAS7 were identified from UniProt and presented as interconnected proteins network (Figure 5.5). Ten proteins were found that were directly or indirectly related to GAS7. Major functions of all the interconnected proteins are listed in Table 5.1. Interestingly, most of the proteins are associated with the

cell cytoskeleton related functions that indicate that the possible function of GAS7 could be cells maturation and differentiation through cytoskeleton development. So, it can be hypothesized that MLL-GAS7 may play a role in cytoskeleton formation in AML blast.

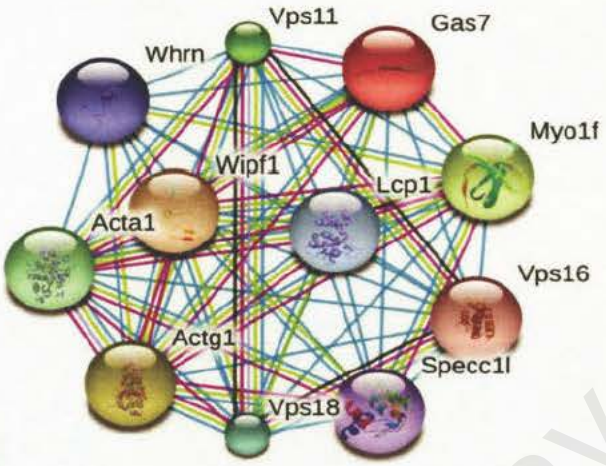


Figure 5.5: GAS7 associated proteins network. Ten proteins are found closely associated to GAS7 in UniProt database and their functions are related to cell cytoskeletons development.

Table 5.1: GAS7 associated proteins functions.

Protein Symbol	Ensembl ID	Major Functions
Vps11	ENSMUSP00000034644	Vesicle-mediated protein trafficking
Vps16	ENSMUSP00000028900	Vesicle-mediated protein trafficking
Vps18	ENSMUSP00000036915	Vesicle-mediated protein trafficking
Whrn	ENSMUSP00000081557	Interacts with calmodulin-dependent serine kinase
Acta1	ENSMUSP00000034453	Cell motility
Actg1	ENSMUSP00000071486	Cell motility
Specc1l	ENSMUSP00000045099	Required for proper cell adhesion and migration
Myo1f	ENSMUSP00000084887	Intracellular movements
Lcp1	ENSMUSP00000116271	Modulates the cell surface expression of IL2RA/CD25 and CD69
Wipf1	ENSMUSP00000092268	Reorganization of the actin cytoskeleton, Plays a role in the formation of cell ruffles
Gas7	ENSMUSP00000104322	Promoting maturation and morphological differentiation

Those proteins corresponding genes expression profiles were examined in AML patients (Figure 5.6). Four genes (Vps16, Wipf1, Myo1f, and Lcp1) were significantly downregulated in AML cases compared to healthy controls. The notable differences were also noticed for other genes expression patterns (except Actg1 and Acta1).

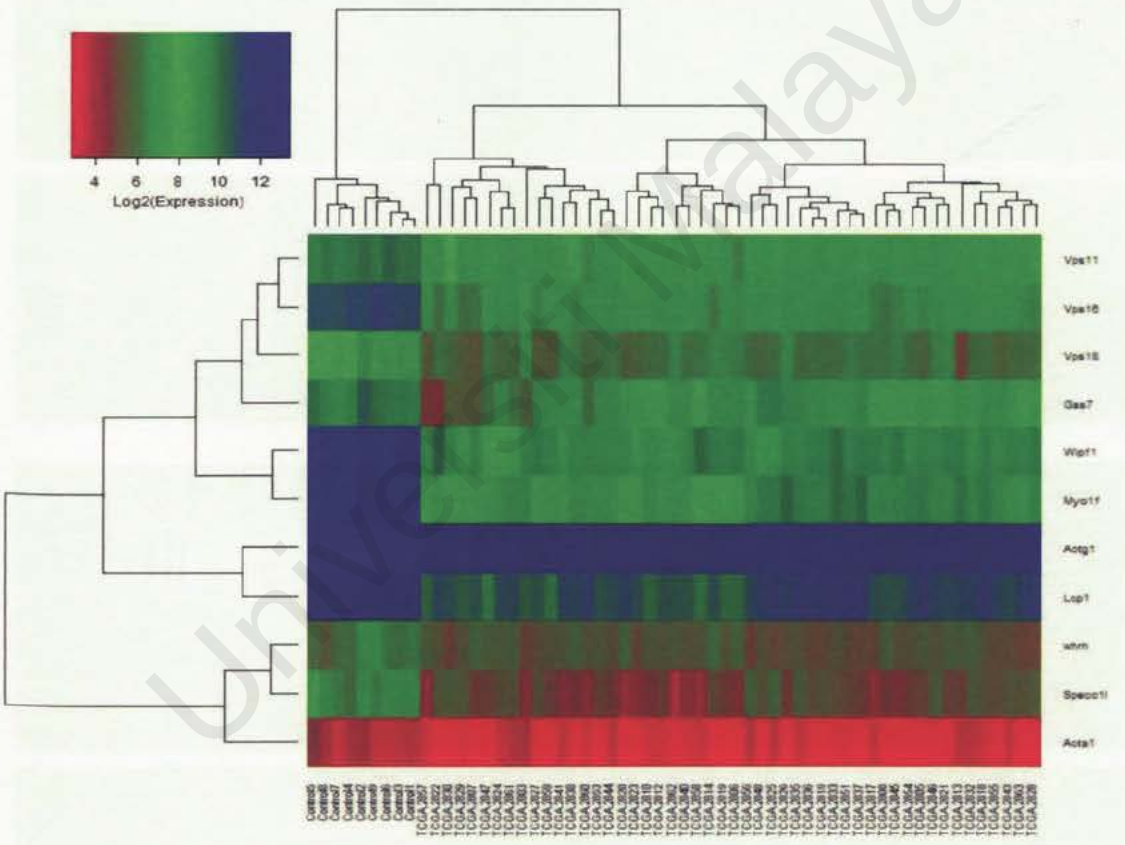
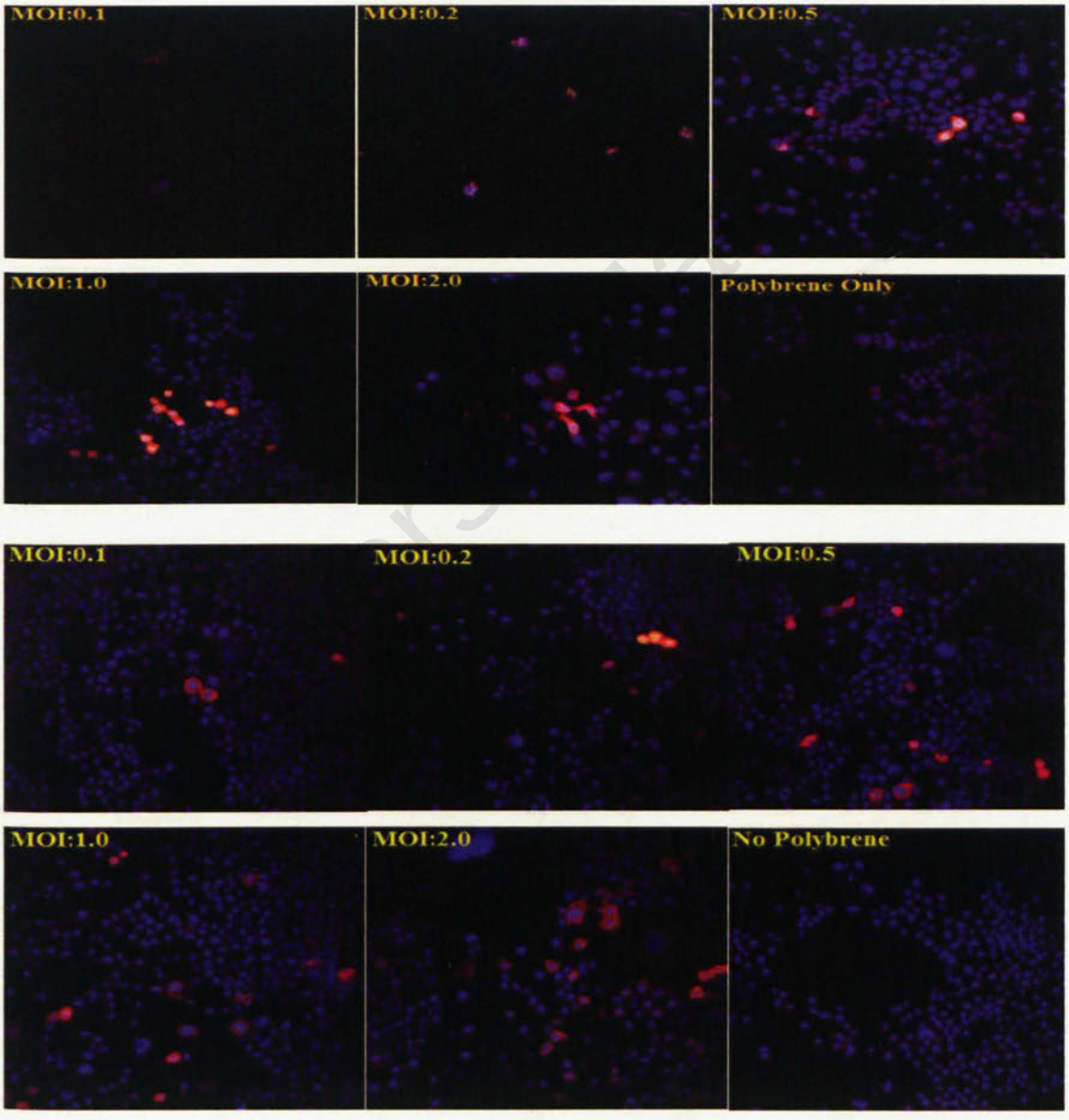


Figure 5.6: GAS7 associated protein's corresponding genes expression. Reported protein's corresponding genes expression in AML patients are shown in heatmap where most of them are showing higher expression level.

5.4.5 Generation and characterization of GAS7-overexpressed N2A cell line

To study the differential role of the GAS7 protein *in vivo*, shGas7 containing external plasmids were transfected into N2A cell line using lipofectamine 2000 reagents. The

plasmid vector (12kb) also contained green-fluorescence protein (GFP) with Gas7 as marker and polybrene were used to facilitate the transfection events. Transfected cell images were taken at different multiplicity of infection (MOI). In order to characterize successful transfection and overgrowth of GAS7, western blot was performed using GAS7-specific antibody. As GAS7 was translated with GFP, total band size was 75kDa (GAS7=48kDa and GFP=27kDa) and no band was observed in the control group (Figure 5.7).



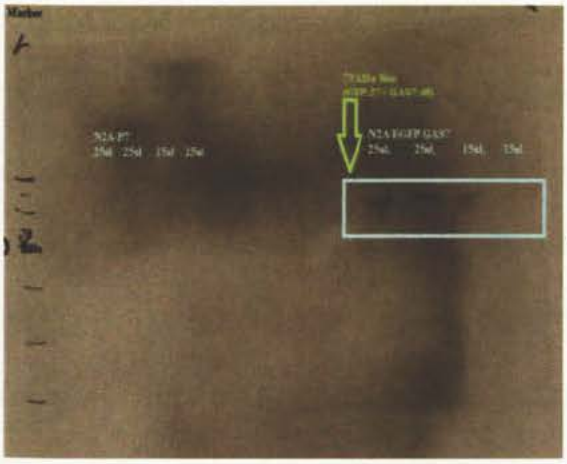


Figure 5.7: Transfection of shGAS7 in N2A cells and western blot. GAS7 gene was transfected using lipofectamine 2000 (a-polybrene and b-for case-no polybrene) and confirmed by western blot (c).

5.4.6 Effect of GAS7 on growth curves and stress response

To identify the effect of GAS7 on cellular growth, normal growth curves were drawn for wild-type (N2A-P7) and transfected N2A cells (N2A-EGFP-GAS7). Notable changes were observed regarding cell growth patterns and doubling time (Figure 5.8). As normal GAS7 arrests cell growth, the transfected cell growth rate was limited compared to that of the wild type.

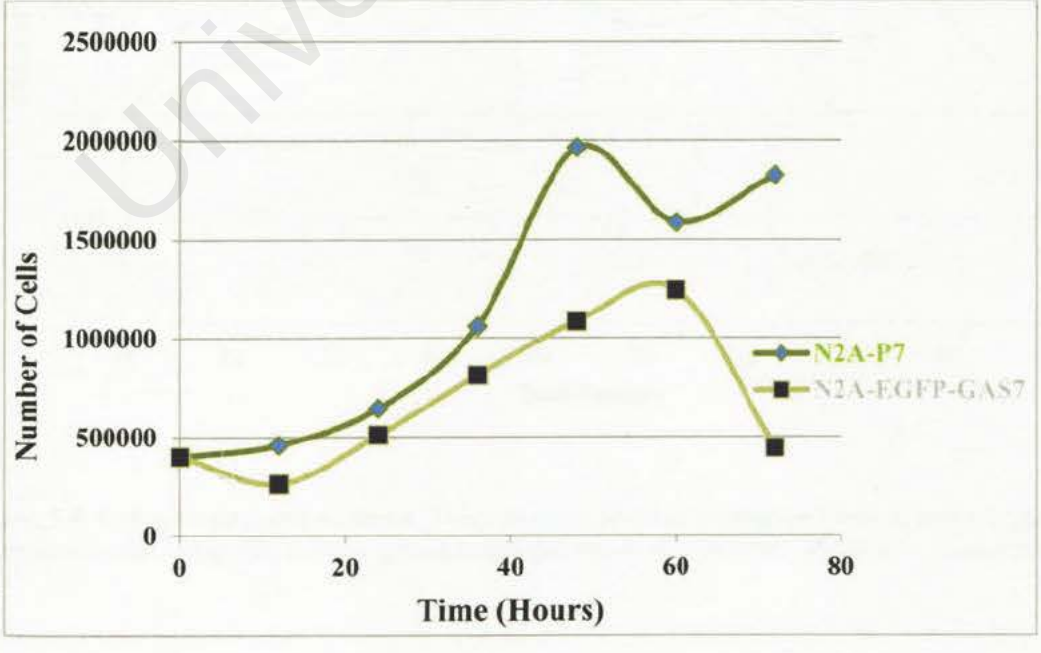


Figure 5.8: Cell growth curves. Wild-type (N2A-P7) and transfected (N2A-EGFP-GAS7) cells growth curves are drawn and significant difference noticed due to GAS7 expression.

How those cells respond to stress, was examined by treating the same plate with Trypsin-EDTA after every 12 hours and the cell number was counted. As N2A cells adhered to the plates and Trypsin-EDTA chopped down the adhesion proteins in cell-cell and cell-matrix interaction before every cell division cycle, this constitutes a stressful condition. The role of GAS7 to overcome the stress was observed by counting the cell numbers at different time intervals. An unusual (curly shape) growth pattern was noticed (Figure 5.9). Although it is difficult to explain the observation, it can be hypothesized that GAS7 assists the N2A-EGFP-GAS7 cells to overcome the stress up to the first 40h through regeneration of cell cytoskeletons. GAS7 in chimeric protein (MLL-GAS7) might also be important for AML blast to overcome different stresses during uncontrolled growth.

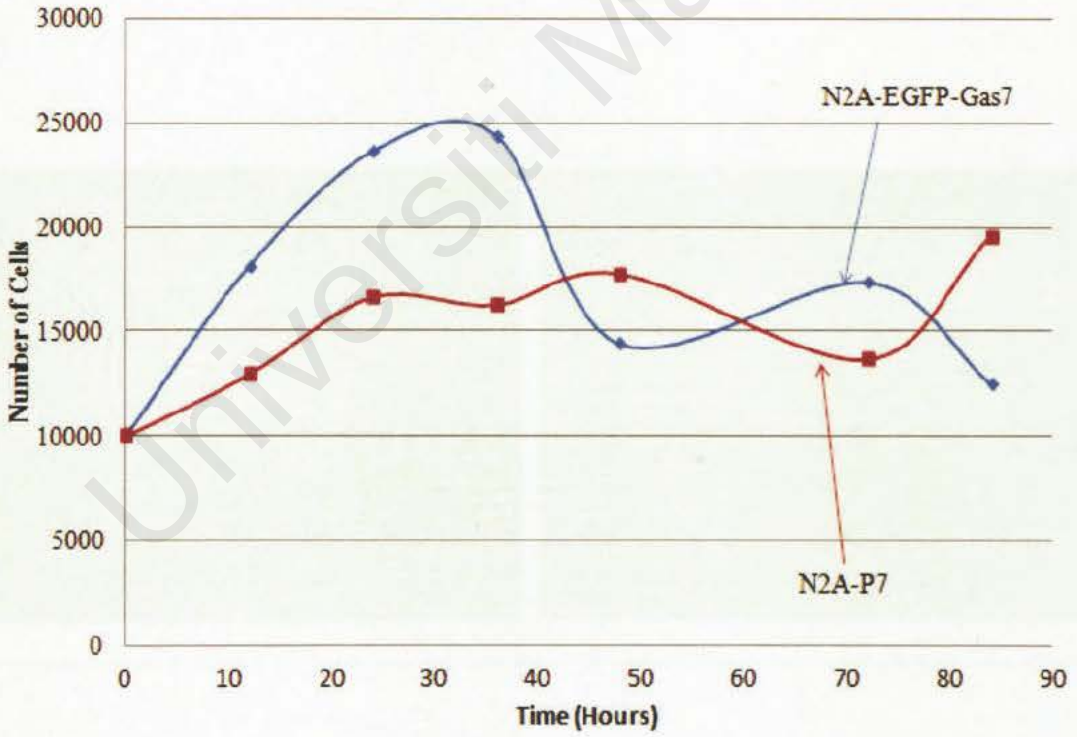
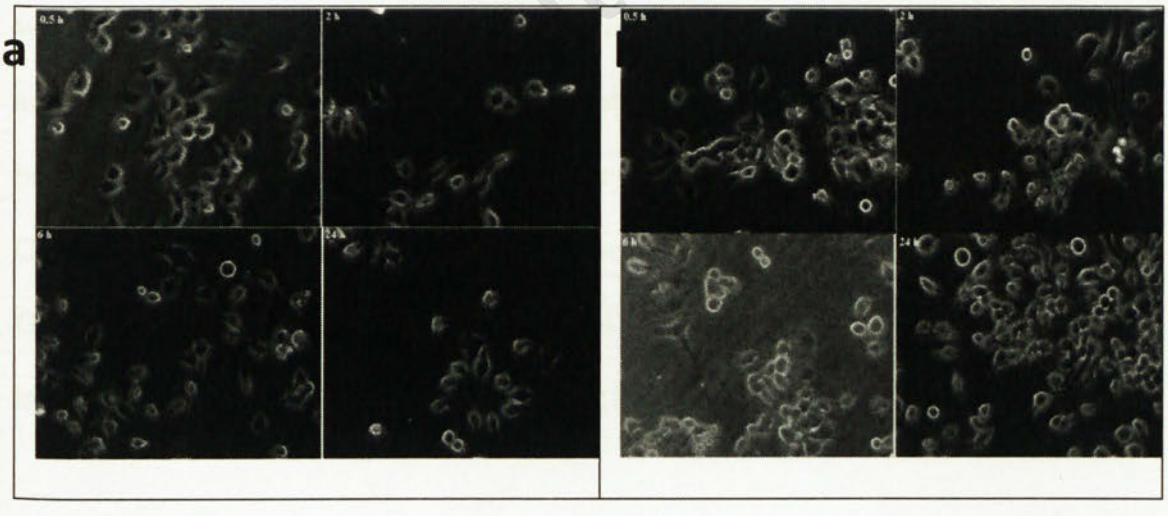


Figure 5.9: Cell growth curve at stress. Two cell types stress responses are checked using Trypsin-EDTA treatment after every 12h. GAS7 expressed cells can overcome cell matrix chop-down related stress.

5.4.7 Serum response to neurite formation

To investigate the effect of fetal bovine serum on N2A cell growth, neurites formations were observed at a different time interval (0.5h, 2h, 6h, and 24h) using serum-free growth medium. GAS7 expressed cells were found to be the highest in neurite formation at 24h compared to wild-type cells (Figure 5.10a), and that therefore indicate that the differential neurite formation was due to the increased GAS7 expression. Following that, the cells were treated with different serum concentration medium (10%, 1%, 0.1%, and 0.01%) and neurites formations were observed. It was noticed that the lesser the serum concentration in the medium, the greater were the neuritis formation. Neurites are packed with microtubules, and GAS7 probably induces neurites outgrowth through helping microtubule development. As microtubules are necessary for cell migration, MLL-GAS7 protein might also play a role in migration of AML cells.



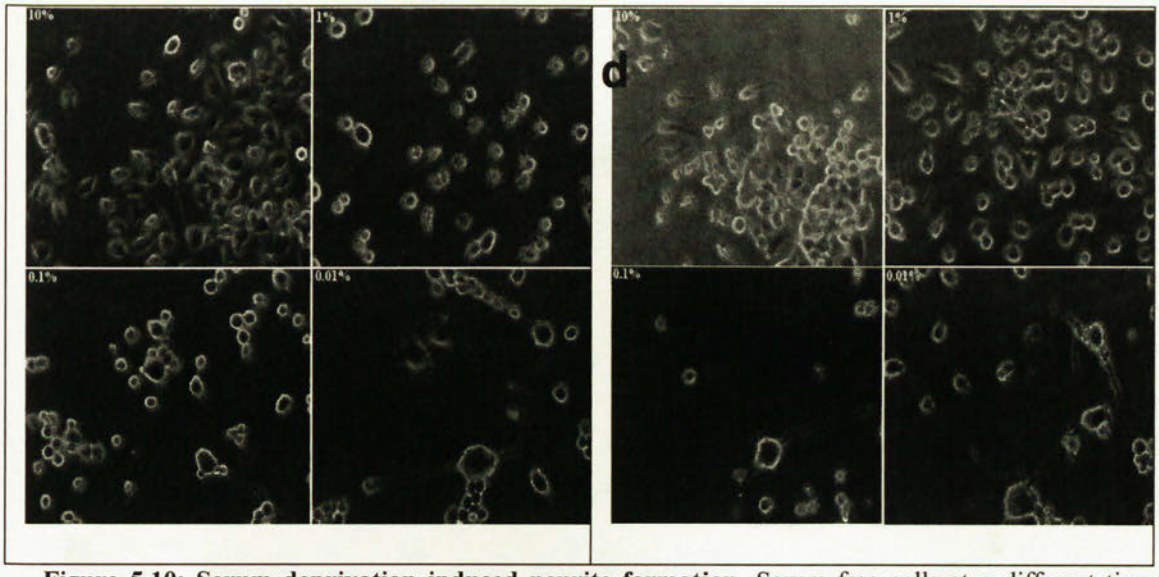


Figure 5.10: Serum deprivation induced neurite formation. Serum free cells at a different time interval (0.5h, 2h, 6h, and 24h) are shown at 5.10a and 5.10b where 5.10a is for N2A-P7 and 5.10b is for N2A-EGFP-GAS7. Serum dose responses at different concentration (10%, 1%, 0.1%, and 0.01%) are shown at 5.10c and 5.10d where 5.10c is for N2A-P7 and 5.10d is for N2A-EGFP-GAS7. Different serum percentages stimulate neurite formation differentially.

5.4.8 Network

Using online tools, GAS7 associated genes and miRNAs networks were developed. Directly related (black) and indirectly related (gray) genes are shown in Figure 5.11a whereas Figure 5.11b indicated GAS7 centric network (query genes were found from IPA and literature). Figure legends were marked separately, indicating the diverse roles between interlinked proteins. The GAS7 centric network that was developed using GeneMANIA was closely related to the network developed using UniProt (Figure 5.5). GAS7 associated transcription factors and miRNAs were also found from IPA and the interactive network was drawn using MirOB (Figure 5.12).

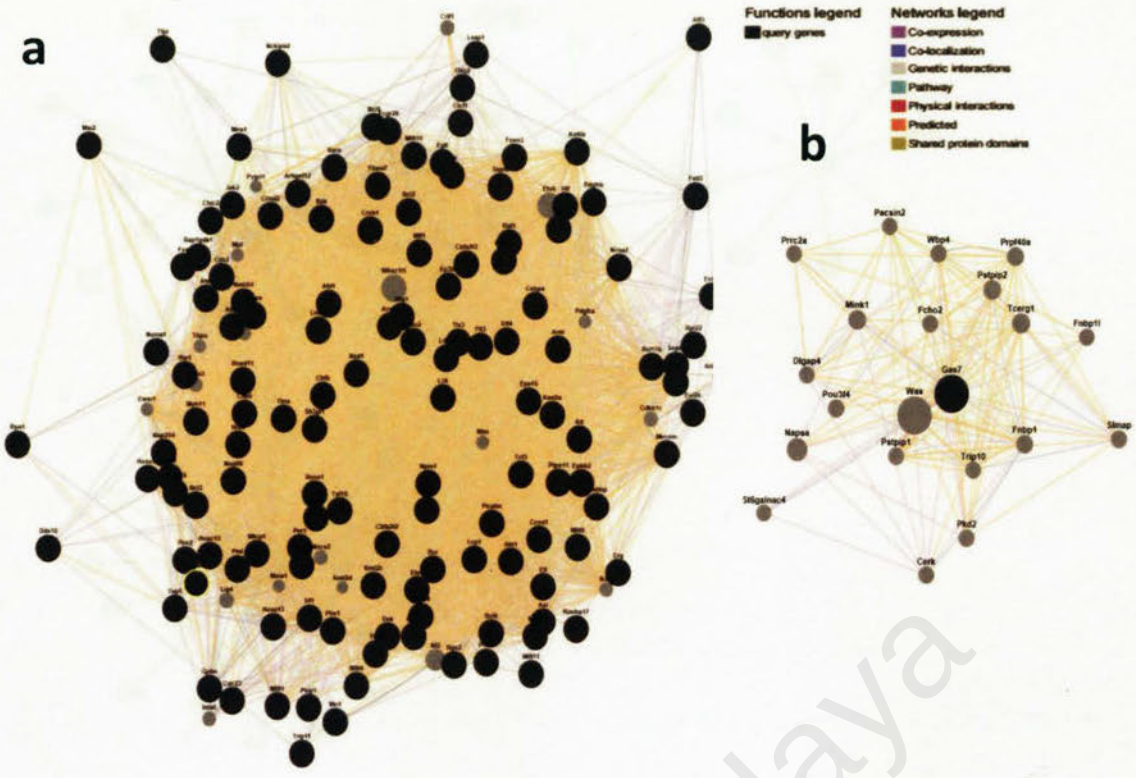


Figure 5.11: Network of GAS7 associated genes. Black color indicates closely related genes and gray color indicate distinctly related genes (5.11a). The right side is a Gas7 centric network (5.11b). (Generated by GeneMania).

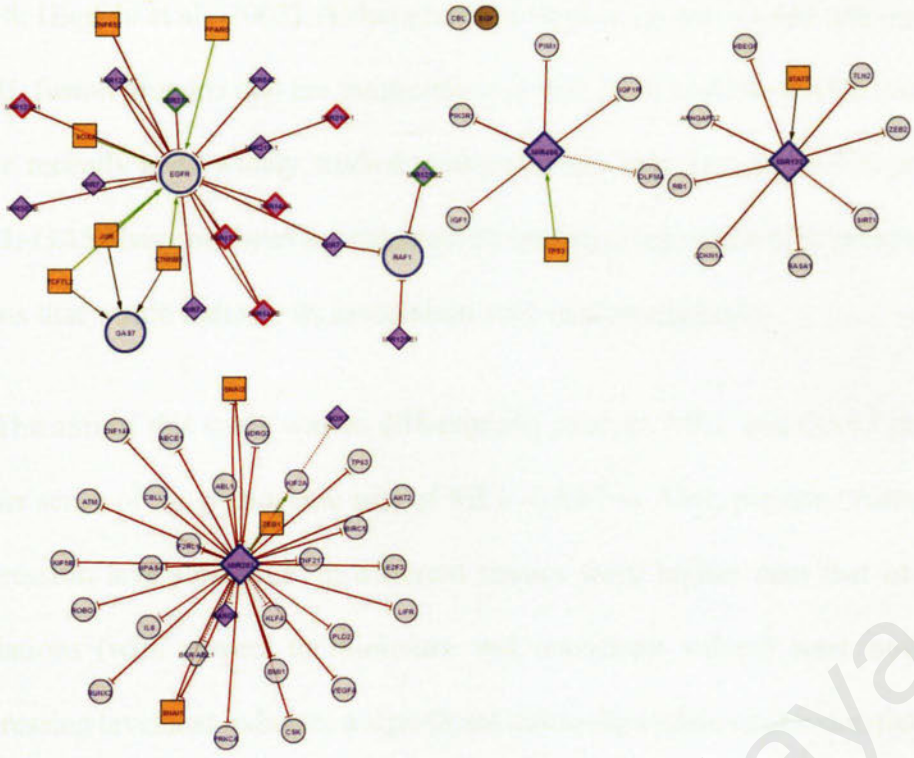


Figure 5.12: GAS7 associated miRNAs centric network. Different miRNAs, genes and transcription factors interaction are shown. (Generated by MirOB).

5.5 Discussion

Chromosomal translocation-mediated fusion proteins in AML interfere with the function of their native partners and ultimately proceed with myeloid differentiation arrest. Functions of partner proteins are heterogeneous and their underlying oncogenic mechanisms are often not clear. Many of these proteins, however, might enhance expression of anti-apoptotic genes and stimulate the machinery for the self-renewal capacity of undifferentiated hematopoietic stem cells through regulation of transcription factors or cofactors. Some AML fusion proteins may repress the expression of several DNA repair genes (homologous recombination and double-strand break-repair pathway genes) that lead to genomic instability followed by predisposition to cancer (Hoeijmakers, 2001). Transgenic mice model studies suggest that AML fusion proteins might also induce preleukemic state to gain necessary potential for oncogenesis (HE LZ & Tribioli,

1998; Higuchi et al., 2002). Although the biological properties and pathogenesis of some AML fusion proteins that are frequently reported, such as PML-RARA and AML1-ETO, have recently been widely studied, comparatively less frequent fusion proteins such as MLL-GAS7 have not been investigated adequately to associate their presence in leukemic blasts that would indicate its association with leukemogenesis.

The aim of this study was to differentially analyze MLL and GAS7 proteins to get a better sense of the pathogenic role of MLL-GAS7 in AML patients. Although the mean expression levels of MLL in different tissues were higher than that of GAS7, larger variations (with respect to minimum and maximum values) were noticed in GAS7 expression level that indicates a significant tissue-dependent expression pattern. MLL and GAS7 were both found to be among the high expression fusion gene cluster. Evidence suggests those fusion genes transcripts and their partner genes transcripts levels were not the same. MLL fusion events were largely found at the M4/M5 stages. Variations were noticed in wild-type MLL and GAS7 downstream regulatory genes expression level in AML patients, especially within the Hox genes cluster that dictates cell fate. Knockout mice model experiment demonstrated that many of these Hox genes (like Hoxa7 and Hoxa9) are not necessary for MLL-GAS7 mediated leukemogenesis, although their deregulation may affect disease latency (So et al., 2004). Notable differences were not observed between AML patients with MLL fusion and without MLL fusions. Interestingly, 10 proteins that were closely associated with GAS7, directly or indirectly, were involved in regulation of cell cytoskeletons and among the highly expressed ones; four of them were related to cellular movement. This finding gave the notion that GAS7 fused chimeric proteins might also help in migration of leukemic blasts. GAS7 overexpressed N2A cells also exhibited significant differences with regard to cell growth rate, doubling time, plating and attachment efficiency, stress response, and neurite formation as compared to wild type. Differential role of GAS7 supported the idea of

multiple regulatory genes alteration in MLL-GAS7 positive patients. But GAS7 itself might not have intrinsic transcriptional potential (So, Lin, et al., 2003) and that the fusion event was necessary to gain that capability. Network analysis showed that many miRNAs were also associated with this MLL-GAS7 mediated leukemogenesis and distinct miRNA expression profiles were identified from different experiments (Li et al., 2008; Lu et al., 2005). Deregulation of miR-191 (Nakamura, Canaani, & Croce, 2007) and overexpression of miR-17-92 cluster (Li et al., 2009) and miR-196b (Popovic et al., 2009) were reported in AMLs with MLL fusion. So, this fusion protein is important for AML pathogenesis. However, more studies are required to understand the detailed mechanism.

5.6 Conclusion

MLL can be fused with growth-arrest specific 7 (GAS7) to generate an oncofusion protein (MLL-GAS7) which is a potent driver of leukemic transformation. However, it is still not clear what roles individual proteins are playing and how this oncogenic transformation occurs. The aim of this study was to understand the pathogenic roles of MLL-GAS7 in acute myeloid leukemia (AML). MLL and GAS7 downstream targeted genes expression levels were determined in a case-control study using the Cancer Genome Atlas (TCGA) AML data repository. Differential expression pattern were noticed for frequently fused genes and GAS7 associated genes in AML patients. Variation of genes expression demonstrated that multiple regulatory genes alteration occurred in MLL-GAS7 positive patients. Notable differences were found in GAS7 overexpressed N2A cells growth patterns, stress responses, and neurites formations compared to wild-type N2A cells. Functional study of GAS7 and network analysis reveals the notion that MLL-GAS7 might play a role in leukemic blast migration and stress responses. Although much important advancement has already been made to understand MLL-GAS7 mediated pathogenesis, quite a few are important for therapeutic implication. The 'omics' based integrated approach can lead to better understanding of MLL-GAS7 pathogenesis that can

help to develop targeted therapeutics. Due to unfavorable prognosis of MLL fused leukemia, it is a long-standing quest to develop therapeutics targeting MLL fusion protein. Recently, it was found that a small molecule (MI-2-2) works as an inhibitor of menin-MLL, and could therefore prevent oncogenic transformation of MLL regardless of any partner proteins (He et al., 2015). There are still, many missing links with regard to the normal biological and pathological phenomenon that is unique to MLL and MLL fusion protein in order to develop successful therapeutics.

CHAPTER 6: GENERAL CONCLUSION

Acute myeloid leukemia (AML) is a hematological malignancy characterized by heterogeneous abnormalities in hematopoietic progenitors leading to the formation of excessive undifferentiated white blood cells (WBC). Abnormal genetic and/or cytogenetic changes in progenitors would lead to deregulation of protein expression. Differential protein expression patterns have not been studied intensively before in Malaysian AML patients. Comprehensive profiling of expression of plasma proteins (cytokines, chemokines, and growth factors) in AML patients and analysis of molecular abnormalities would provide greater insight into the disease etiology as well as identification of prospective diagnostic and prognostic biomarkers. It would also be interesting to compare global protein expression and molecular abnormalities with Malaysian AML patients to see whether there are any significant deviation. One of the recent advancement of biological research is multiplex array technology that allows measurement of many proteins precisely and simultaneously in biological samples like plasma for descriptive and predictive research. In a case-control study, 38 (case=19, control=19) plasma samples were analyzed and 32 cytokines and chemokines (C&Ckines) were quantitatively measured in each sample. Genetic and cytogenetic aberrations were also detected by conventional methods and comparative analysis was performed with the Cancer Genome Atlas (TCGA) AML data. Whether C&Ckines expression profiles correlate with the corresponding DNA methylation and gene expression profiles were confirmed by TCGA data. Chromosomal translocation is one of the common molecular aberrations in AML that generate fusion proteins. Mix lineage leukemia (MLL) fused with growth arrest specific gene-7 (GAS7) at t(11;17)(q23;p13) and produce a fusion protein that disrupts hematopoietic differentiation in AML.

Differential expression pattern of GAS7 was identified in neuro-2a (N2A) cell line. C&Ckines expression level in AML plasma differed significantly ($p<0.005$) from healthy controls, in which 13 were upregulated and 4 were downregulated. The expression of some of the C&Ckines correlated with age (leptin and ferritin), gender (IGFBP3), and ethnicity (MIF, sFas, sFasL, OPN, Cathepsin D, MPO, and SHBG). Five pairs of C&Ckines in AML cases, one pair of healthy controls and three pairs of C&Ckines regardless of disease (combined with cases and controls) were shown to be significantly correlated with expression (Pearson's correlation $r>0.8$ and adjusted $R^2>0.7$). Different clustering algorithms showed distinct patterns in the C&Ckines and AML patients based on plasma C&Ckines expression levels. Four groups of AML cases and five groups of C&Ckines patterns were observed. In general, DNA methylation and gene expression patterns were matched with plasma C&Ckines levels. Ten C&Ckines expression levels were detected for the first time in AML plasma and out of them 2 (MIF and Cathepsin D) were identified to have prospective diagnostic impact. Based on the literature, plasma protein level, DNA methylation, gene expression, and pathway analysis, it can be hypothesized that MIF promotes cancer cell survival in AML patients through CD74 and CXCR4 receptors in the PI3K/Akt pathway. HGF and Galectin-3 also showed diagnostic possibility. In spite of some variation of earlier methods (like ELISA) base plasma C&Ckines level, overall expression data follow global expression pattern. Based on TCGA data analysis, three genes (FLT3, DNMT3A, and NPMc) were found to be predominantly mutated in AML cases. Overall, DNA methylation followed a bimodal distribution. Methylation distribution in promoter region was quite low compared to the open sea and shore region. Almost 200 genes were hypermethylated in 90% of AML cases and were equally distributed among all chromosomes. Three aberrations (Del (5q), T (15; 17), and T (9; 11)) were significantly ($p<0.05$) associated with patient survival. Patients with Del (5q) positive, average survival was <1 year, whereas T (15; 17) positive

patients showed significantly higher survival rate (>3 years) in Kaplan-Meier curve that lead to a new hypothesis, which is that there is a possibility that some chemotherapeutic drugs were more effective due to this translocation (T (15; 17)). Three distinct groups of patients and two distinct groups of genes have been found to be hypermethylated in TCGA, and this may be considered to be informative for epigenetic-based AML classification approaches. Notable differences had been noticed in N2A cell line for attachment and plating efficiency as well as the growth curve due to the expression of GAS7 which is a partner protein of MLL, one of the reported causative factors in AML. Serum deprivation enhanced neuritis formation in N2A cell that expresses very small GAS7 after 6 to 24 h and dose responses of fetal calf serum (FBS) were also determined.

One of the major limitations of this study is small sample size. Due to budget limitation for this project, we only detected 38 samples' plasma C&Ckines. Ethnic distribution was biased in this study, as there was no Indian Ethnic population in this study. Comprehensive analysis has not been done for the same patients, as there is no C&Ckines information in the TCGA AML dataset. Some fusion proteins are more important than MLL-GAS7, but due to laboratory accessibility, we performed this protein's differential analysis.

In future, we would increase sample size with bigger cytokine panel. Also multivariate analysis would be performed for cytogenetic abnormalities with C&Ckines expression level. More extensive analysis would be performed using other available genomic data (AML) from GEO and TCGA.

In conclusion, C&Ckines expression profile in AML patients differs from that of the control unaffected individuals and form a distinct pattern of expression. Genetic and cytogenetic abnormalities follow a global pattern. Further integrated analysis of copy

number variation, SNPs, gene expression, DNA methylation, and fusion protein expression study could reveal insight into the pathogenesis of AML.

REFERENCES

Agarwal, R., Whang, D. H., Alvero, A. B., Visintin, I., Lai, Y., Segal, E. A., . . . Mor, G. (2007). Macrophage migration inhibitory factor expression in ovarian cancer. *American journal of obstetrics and gynecology, 196*(4), 348. e341-348. e345.

Aguayo, A., Kantarjian, H., Manshouri, T., Gidel, C., Estey, E., Thomas, D., . . . Keating, M. (2000). Angiogenesis in acute and chronic leukemias and myelodysplastic syndromes. *Blood, 96*(6), 2240-2245.

Akalin, A., Garrett-Bakelman, F. E., Kormaksson, M., Busuttil, J., Zhang, L., Khrebtukova, I., . . . Hess, J. L. (2012). Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet, 8*(6), e1002781.

Al Maskari, M. Y., & Alnaqdy, A. A. (2006). Correlation between serum leptin levels, body mass index and obesity in omanis. *Sultan Qaboos University Medical Journal, 6*(2), 27.

Alcalay, M., Meani, N., Gelmetti, V., Fantozzi, A., Fagioli, M., Orleth, A., . . . Casciari, C. (2003). Acute myeloid leukemia fusion proteins deregulate genes involved in stem cell maintenance and DNA repair. *The Journal of clinical investigation, 112*(11), 1751-1761.

Alcalay, M., Orleth, A., Sebastiani, C., Meani, N., Chiaradonna, F., Casciari, C., . . . Minucci, S. (2001). Common themes in the pathogenesis of acute myeloid leukemia. *Oncogene, 20*(40), 5680-5694.

Alvarez-Garcia, I., & Miska, E. A. (2005). MicroRNA functions in animal development and human disease. *Development, 132*(21), 4653-4662.

Apte, R. N., Dotan, S., Elkabets, M., White, M. R., Reich, E., Carmi, Y., . . . Voronov, E. (2006). The involvement of IL-1 in tumorigenesis, tumor invasiveness, metastasis and tumor-host interactions. *Cancer and Metastasis Reviews, 25*(3), 387-408.

Arican, O., Aral, M., Sasmaz, S., & Ciragil, P. (2005). Serum levels of TNF-α, IFN-γ, IL-6, IL-8, IL-12, IL-17, and IL-18 in patients with active psoriasis and correlation with disease severity. *Mediators of inflammation, 2005*(5), 273-279.

Aspenstrøm, P. (2009). *The pombe Cdc15 homology proteins*: Landes Bioscience.

Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T., & Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nature methods, 11*(11), 1138-1140.

Aw, S. S., Tar-Choon. (2000). Age-related reference intervals for free and total prostate-specific antigen in a Singaporean population. *Pathology, 32*(4), 245-249.

Ayaz, T., Sahin, S. B., Sahin, O. Z., Cure, M. C., Sumer, F., & Ilkkilic, K. (2014). Serum macrophage migration inhibitory factor levels in Hashimoto's thyroiditis; a case control study.

Ayton, P. M., Chen, E. H., & Cleary, M. L. (2004). Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. *Molecular and cellular biology, 24*(23), 10470-10478.

Azab, N. A., Rady, H. M., & Marzouk, S. A. (2012). Elevated serum TRAIL levels in scleroderma patients and its possible association with pulmonary involvement. *Clinical rheumatology, 31*(9), 1359-1364.

Baekelandt, M., Holm, R., Trope, C., Nesland, J., & Kristensen, G. (1999). The significance of metastasis-related factors cathepsin-D and nm23 in advanced ovarian cancer. *Annals of oncology, 10*(11), 1335-1341.

Balkwill, F., & Mantovani, A. (2001). Inflammation and cancer: back to Virchow? *The Lancet, 357*(9255), 539-545.

Ball, D., Rose, E., & Alpert, E. (1992). Alpha-fetoprotein levels in normal adults. *The American journal of the medical sciences, 303*(3), 157-159.

Barderas, M. G., Dardé, V. M., de la Cuesta, F., Martin-Ventura, J. L., Blanco-Colio, L. M., Jiménez-Narcher, J., . . . Egido, J. (2008). Proteomic analysis of circulating monocytes identifies cathepsin D as a potential novel plasma marker of acute coronary syndromes. *Clinical Medicine Insights. Cardiology, 1*, 191.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *cell, 116*(2), 281-297.

Baylin, S. B., Fearon, E. R., Vogelstein, B., de Bustros, A., Sharkis, S. J., Burke, P. J., . . . Nelkin, B. D. (1987). Hypermethylation of the 5'region of the calcitonin gene is a property of human lymphoid and acute myeloid malignancies. *Blood, 70*(2), 412-417.

Beiras-Fernandez, A., Weis, F., Rothkopf, J., Kaczmarek, I., Ledderose, C., Dick, A., . . . Kreth, S. (2012). Local expression of myocardial galectin-3 does not correlate with its serum levels in patients undergoing heart transplantation. *Annals of transplantation: quarterly of the Polish Transplantation Society, 18*, 643-650.

Benjamin, D. Y., Hess, J. L., Horning, S. E., Brown, G. A., & Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature, 378*(6556), 505-508.

Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D., Gralnick, H. R., & Sultan, C. (1976a). Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *British journal of haematology, 33*(4), 451-458.

Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., & Sultan, C. (1976b). Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *British journal of haematology, 33*(4), 451-458.

Biondi, A., Cimino, G., Pieters, R., & Pui, C.-H. (2000). Biological and therapeutic aspects of infant leukemia. *Blood, 96*(1), 24-33.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & development, 16*(1), 6-21.

Bird, C. L., Witte, J. S., Swendseid, M. E., Shikany, J. M., Hunt, I. F., Frankl, H. D., . . . Haile, R. W. (1996). Plasma Ferrtin, Iron Intake, and the Risk of Colorectal Polyps. *American journal of epidemiology, 144*(1), 34-41.

Blum, W. F., Englaro, P., Hanitsch, S., Juul, A., Hertel, N. T., Müller, J., . . . Attanasio, A. M. (1997). Plasma leptin levels in healthy children and adolescents: Dependence on body mass index, body fat mass, gender, pubertal stage, and testosterone 1. *The Journal of Clinical Endocrinology & Metabolism, 82*(9), 2904-2910.

Bock, C. (2012). Analysing and interpreting DNA methylation data. *Nature Reviews Genetics, 13*(10), 705-719.

Bock, C., Beerman, I., Lien, W.-H., Smith, Z. D., Gu, H., Boyle, P., . . . Meissner, A. (2012). DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Molecular cell, 47*(4), 633-647.

Bock, C., Walter, J., Paulsen, M., & Lengauer, T. (2008). Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic acids research, 36*(10), e55-e55.

Bray, F., & Møller, B. (2006). Predicting the future burden of cancer. *Nature reviews cancer, 6*(1), 63-74.

Bruserud, Ø., Tore, G. B., Brustugun, O. T., Bassøe, C., Nesthus, I., Espen, A. P., . . . Pawelec, G. (1995). Effects of interleukin 10 on blast cells derived from patients with acute myelogenous leukemia. *Leukemia, 9*(11), 1910-1920.

Bruserud, Ø., & Ulvestad, E. (2000). Soluble Fas/Apo-1 (CD95) levels during T cell activation in the presence of acute myelogenous leukemia accessory cells; contributions from local release and variations in systemic levels. *Cancer Immunology, Immunotherapy, 49*(7), 377-387.

Butcher, L. M., Ito, M., Brimpari, M., Morris, T. J., Soares, F. A., Ährlund-Richter, L., . . . Beck, S. (2016). Non-CG DNA methylation is a biomarker for assessing endodermal differentiation capacity in pluripotent stem cells. *Nature communications, 7*.

Byrd, J. C., Mrózek, K., Dodge, R. K., Carroll, A. J., Edwards, C. G., Arthur, D. C., . . . Watson, M. S. (2002). Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood, 100*(13), 4325-4336.

Cahill, N., Bergh, A.-C., Kanduri, M., Göransson-Kultima, H., Mansouri, L., Isaksson, A., . . . Sundström, C. (2013). 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. *Leukemia, 27*(1), 150-158.

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., . . . Negrini, M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A, 101*(9), 2999-3004.

Capitanio, U., Perrotte, P., Zini, L., Suardi, N., Antebi, E., Cloutier, V., . . . Arjane, P. (2009). Population-based analysis of normal total PSA and percentage of free/total PSA values: results from screening cohort. *Urology, 73*(6), 1323-1327.

Chang, P. Y., Kuo, J. T., Lin-Chao, S., & Chao, C. C. K. (2005). Identification of rat Gas7 isoforms differentially expressed in brain and regulated following kainate-induced neuronal injury. *Journal of neuroscience research, 79*(6), 788-797.

Chou, W.-C., Huang, H.-H., Hou, H.-A., Chen, C.-Y., Tang, J.-L., Yao, M., . . . Huang, S.-Y. (2010). Distinct clinical and biological features of de novo acute myeloid leukemia with additional sex comb-like 1 (ASXL1) mutations. *Blood, 116*(20), 4086-4094.

Christiansen, M., Høgdall, C., Andersen, J., & Nørgaard-Pedersen, B. (2001). Alpha-fetoprotein in plasma and serum of healthy adults: preanalytical, analytical and biological sources of variation and construction of age-dependent reference intervals. *Scandinavian journal of clinical and laboratory investigation, 61*(3), 205-215.

Christiansson, L., Mustjoki, S., Simonsson, B., Olsson-Strömberg, U., Loskog, A. S., & Mangsbo, S. M. (2014). The use of multiplex platforms for absolute and relative protein quantification of clinical material. *EuPA Open Proteomics, 3*, 37-47.

Christopeit, M., & Bartholdy, B. (2014). Epigenetic signatures as prognostic tools in acute myeloid leukemia and myelodysplastic syndromes. *Epigenomics, 6*(4), 371-374.

Christopeit, M., Kröger, N., Haferlach, T., & Bacher, U. (2014). Relapse assessment following allogeneic SCT in patients with MDS and AML. *Annals of hematology, 93*(7), 1097-1110.

Ciftci, R., Tas, F., Aksit, E., Vatansever, S., Karabulut, S., Sen, F., . . . Kilic, L. (2014). *Clinical significance of serum macrophage migration inhibitory factor (MIF) level in breast cancer.* Paper presented at the ASCO Annual Meeting Proceedings.

Clendenen, T. V., Arslan, A. A., Lokshin, A. E., Liu, M., Lundin, E., Koenig, K. L., . . . Krogh, V. (2013). Circulating prolactin levels and risk of epithelial ovarian cancer. *Cancer Causes & Control, 24*(4), 741-748.

Colomer, R., Ruibal, A., Genollk, J., & Salvador, L. Circulating CA 15-3 antigen levels in non-mammary malignancies.

Conway O'Brien, E., Prideaux, S., & Chevassut, T. (2014). The epigenetic landscape of acute myeloid leukemia. *Advances in hematology, 2014*.

Coussens, L. M., & Werb, Z. (2002). Inflammation and cancer. *Nature, 420*(6917), 860-867.

Dawson, M. A., Bannister, A. J., Göttgens, B., Foster, S. D., Bartke, T., Green, A. R., & Kouzarides, T. (2009). JAK2 phosphorylates histone H3Y41 and excludes HP1&agr; from chromatin. *Nature, 461*(7265), 819-822.

De Jager, W., te Velthuis, H., Prakken, B. J., Kuis, W., & Rijkers, G. T. (2003). Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells. *Clinical and diagnostic laboratory immunology, 10*(1), 133-139.

de Larrea, C. F., Navarro, A., Tejero, R., Tovar, N., Díaz, T., Cibeira, M. T., . . . Rozman, M. (2012). Impact of MiRSNPs on survival and progression in patients with multiple myeloma undergoing autologous stem cell transplantation. *Clinical Cancer Research, 18*(13), 3697-3704.

De, S., Shaknovich, R., Riester, M., Elemento, O., Geng, H., Kormaksson, M., . . . Polo, J. M. (2013). Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. *PLoS Genet, 9*(1), e1003137.

de Willige, S. U., Malfliet, J. J., Deckers, J. W., Dippel, D. W., Leebeek, F. W., & Rijken, D. C. (2015). Plasma levels of soluble fibroblast activation protein in arterial thrombosis; determinants and cleavage of its substrate alpha-2-antiplasmin. *International journal of cardiology, 178*, 105-110.

Delhommeau, F., Dupont, S., Valle, V. D., James, C., Trannoy, S., Masse, A., . . . Alberdi, A. (2009). Mutation in TET2 in myeloid cancers. *New England Journal of Medicine, 360*(22), 2289-2301.

Deneberg, S., Grövdal, M., Karimi, M., Jansson, M., Nahi, H., Corbacioglu, A., . . . Ekström, T. (2010). Gene-specific and global methylation patterns predict outcome in patients with acute myeloid leukemia. *Leukemia, 24*(5), 932-941.

Dexter, T. M. (1990). Introduction to the haemopoietic system. *Cancer surveys*.

Dilek, I., Ayakta, H., Demir, C., Meral, C., & Ozturk, M. (2005). CA 125 levels in patients with non-Hodgkin lymphoma and other hematologic malignancies. *Clinical & Laboratory Haematology, 27*(1), 51-55.

DiMartino, J. F., & Cleary, M. L. (1999). Mll rearrangements in haematological malignancies: lessons from clinical and biological studies. *British journal of haematology, 106*(3), 614-626.

Distler, O., Del Rosso, A., Giacomelli, R., Cipriani, P., Conforti, M. L., Guiducci, S., . . . Michel, B. A. (2002). Angiogenic and angiostatic factors in systemic sclerosis: increased levels of vascular endothelial growth factor are a feature of the earliest disease stages and are associated with the absence of fingertip ulcers. *Arthritis Res, 4*(6), R11.

Dixon-McIver, A., East, P., Mein, C. A., Cazier, J.-B., Molloy, G., Chaplin, T., . . . Debernardi, S. (2008). Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. *PloS one, 3*(5), e2141.

Döhner, H., Estey, E. H., Amadori, S., Appelbaum, F. R., Büchner, T., Burnett, A. K., . . . Larson, R. A. (2010). Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood, 115*(3), 453-474.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M. J., Irizarry, R., . . . Loewer, S. (2009). Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature genetics, 41*(12), 1350-1353.

Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., . . . Roeder, R. G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *cell, 121*(6), 873-885.

Dranoff, G. (2004). Cytokines in cancer pathogenesis and cancer therapy. *Nature reviews cancer, 4*(1), 11-22.

Dzikiewicz-Krawczyk, A. (2014). MicroRNA-binding site polymorphisms in hematological malignancies. *J Hematol Oncol, 7*(1), 83.

Eguchi, M., Eguchi-Ishimae, M., & Greaves, M. (2005). Molecular pathogenesis of MLL-associated leukemias. *International journal of hematology, 82*(1), 9-20.

Eguchi, M., Eguchi-Ishimae, M., Knight, D., Kearney, L., Slany, R., & Greaves, M. (2006). MLL chimeric protein activation renders cells vulnerable to chromosomal damage: an explanation for the very short latency of infant leukemia. *Genes, Chromosomes and Cancer, 45*(8), 754-760.

El Maksoud, N. A., Ragab, H. M., El Latif, M. M. A., & Abdalla, S. (2010). Prognostic Impact of Elevated Serum Hyulronic Acid, Ferritin and Interleukin-6 in Patients with Acute Myeloid Leukemia. *Journal of American Science, 6*(10), 532-541.

Elbaz, O., Budel, L., Hoogerbrugge, H., Touw, I., Delwel, R., Mahmoud, L., & Löwenberg, B. (1991). Tumor necrosis factor downregulates granulocyte-colony-stimulating factor receptor expression on human acute myeloid leukemia cells and granulocytes. *Journal of Clinical Investigation, 87*(3), 838.

Elbaz, O., & Shaltout, A. (2000). Implication of granulocyte-macrophage colony stimulating factor (GM-CSF) and interleukin-3 (IL-3) in children with acute myeloid leukaemia (AML); malignancy. *Hematology (Amsterdam, Netherlands), 5*(5), 383-388.

Emmertsen, F., Glenthøj, A., Sønderskov, J., Kampmann, P., Sengeløv, H., & Borregaard, N. (2014). ProHNPs are specific markers of normal myelopoiesis.

Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., . . . Beyene, J. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nature medicine, 17*(9), 1086-1093.

Erden, G., Barazi, A., Tezcan, G., & Yildirimkaya, M. (2008). Biological variation and reference change values of CA 19-9, CEA, AFP in serum of healthy individuals. *Scandinavian journal of clinical and laboratory investigation, 68*(3), 212-218.

Ernst, P., Mabon, M., Davidson, A. J., Zon, L. I., & Korsmeyer, S. J. (2004). An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Current biology, 14*(22), 2063-2069.

Ernst, P., Wang, J., Huang, M., Goodman, R. H., & Korsmeyer, S. J. (2001). MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Molecular and cellular biology, 21*(7), 2249-2258.

Estrov, Z., Kurzrock, R., Estey, E., Wetzler, M., Ferrajoli, A., Harris, D., . . . Talpaz, M. (1992). Inhibition of acute myelogenous leukemia blast proliferation by interleukin-1 (IL-1) receptor antagonist and soluble IL-1 receptors. *Blood, 79*(8), 1938-1945.

Estrov, Z., Roifman, C., Mills, G., Grunberger, T., Gelfand, E. W., & Freedman, M. H. (1987). The regulatory role of interleukin 2-responsive T lymphocytes on human marrow granulopoiesis. *Blood, 69*(4), 1161-1166.

Fei, F., Abdel-Azim, H., Lim, M., Arutyunyan, A., von Itzstein, M., Groffen, J., & Heisterkamp, N. (2013). Galectin-3 in pre-B acute lymphoblastic leukemia. *Leukemia, 27*(12), 2385.

Ferrajoli, A., Talpaz, M., Zipf, T., Hirsch-Ginsberg, C., Estey, E., Wolpe, S., & Estrov, Z. (1994). Inhibition of acute myelogenous leukemia progenitor proliferation by macrophage inflammatory protein 1-alpha. *Leukemia, 8*(5), 798-805.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., . . . Fernandez, H. F. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer cell, 18*(6), 553-567.

Figueroa, M. E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P. J., . . . Skrabanek, L. (2010). DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer cell, 17*(1), 13-27.

Finnerty, C. C., Herndon, D. N., Przkora, R., Pereira, C. T., Oliveira, H. M., Queiroz, D. M., . . . Jeschke, M. G. (2006). Cytokine expression profile over time in severely burned pediatric patients. *Shock, 26*(1), 13-19.

Foekens, J., Look, M., Bolt-de Vries, J., Meijer-van Gelder, M., Van Putten, W., & Klijn, J. (1999). Cathepsin-D in primary breast cancer: prognostic evaluation involving 2810 patients. *British journal of cancer, 79*(2), 300.

Foss, B. (1999). Serum levels of thrombopoietin and stem cell factor in acute leukemia patients with chemotherapy-induced cytopenia and complicating infections. *Platelets, 10*(1), 17-23.

Friedman, R. C., Farh, K. K.-H., Burge, C. B., & Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome research, 19*(1), 92-105.

Gado, K., Rimanoczi, E., Hasitz, A., Gigler, G., bor, a., Toth, B., . . . Domján, G. (2002). Elevated levels of serum prolactin in patients with advanced multiple myeloma. *Neuroimmunomodulation, 9*(4), 231-236.

Gaidzik, V. I., Bullinger, L., Schlenk, R. F., Zimmermann, A. S., Röck, J., Paschka, P., . . . Ganser, A. (2011). RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. *Journal of Clinical Oncology, 29*(10), 1364-1372.

Garcia, M., Platet, N., Liaudet, E., Laurent, V., Derocq, D., Brouillet, J. P., & Rochefort, H. (1996). Biological and clinical significance of cathepsin D in breast cancer metastasis. *Stem Cells, 14*(6), 642-650.

Gentles, A. J., Plevritis, S. K., Majeti, R., & Alizadeh, A. A. (2010). Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. *Jama, 304*(24), 2706-2715.

Gilliland, D. G., & Tallman, M. S. (2002). Focus on acute leukemias. *Cancer cell, 1*(5), 417-420.

Glenjen, N., Mosevoll, K. A., & Bruserud, Ø. (2002). Serum levels of angiogenin, basic fibroblast growth factor and endostatin in patients receiving intensive chemotherapy for acute myelogenous leukemia. *International journal of cancer, 101*(1), 86-94.

Góes, L. S. P. d., Lopes, R. I., Campos, O. H. A., Oliveira, L. C. N. d., & Srougi, M. (2014). Recurrence of acute myeloid leukemia in cryptorchid testis: case report. *Einstein (São Paulo), 12*(1), 109-111.

Gonçalves, I., Hultman, K., Dunér, P., Edsfeldt, A., Hedblad, B., Fredrikson, G. N., . . . Bengtsson, E. (2016). High levels of cathepsin D and cystatin B are associated with increased risk of coronary events. *Open heart, 3*(1), e000353.

Gorelik, E., Landsittel, D. P., Marrangoni, A. M., Modugno, F., Velikokhatnaya, L., Winans, M. T., . . . Lokshin, A. E. (2005). Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. *Cancer Epidemiology Biomarkers & Prevention, 14*(4), 981-987.

Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., . . . Zamir, I. (1998). Fusion proteins of the retinoic acid receptor-α recruit histone deacetylase in promyelocytic leukaemia. *Nature, 391*(6669), 815-818.

Grimwade, D. (2001). The clinical significance of cytogenetic abnormalities in acute myeloid leukaemia. *Best Practice & Research Clinical Haematology, 14*(3), 497-529.

Grimwade, D., Hills, R. K., Moorman, A. V., Walker, H., Chatters, S., Goldstone, A. H., . . . Group, N. C. R. I. A. L. W. (2010). Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood, 116*(3), 354-365.

Grossmann, V., Schnittger, S., Kohlmann, A., Eder, C., Roller, A., Dicker, F., . . . Serve, H. (2012). A novel hierarchical prognostic model of AML solely based on molecular mutations. *Blood, 120*(15), 2963-2972.

Gwang Kim, J., Kyun Sohn, S., Hwan Kim, D., Ho Baek, J., Young Lee, N., Soo Suh, J., . . . Bo Lee, K. (2005). Clinical implications of angiogenic factors in patients with acute or chronic leukemia: hepatocyte growth factor levels have prognostic impact, especially in patients with acute myeloid leukemia. *Leukemia & lymphoma, 46*(6), 885-891.

Halbritter, F., Kousa, A. I., & Tomlinson, S. R. (2014). GeneProf data: a resource of curated, integrated and reusable high-throughput genomics experiments. *Nucleic acids research, 42*(D1), D851-D858.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *cell, 100*(1), 57-70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *cell, 144*(5), 646-674.

Hartozińska, A., Noworolska, A., & Richter, R. (1985). Nonspecific Cross-Reacting Antigen as a Marker of Myelocytic Leukemias in Individual Stages of Myelocytic Cell Differentiation *Modern Trends in Human Leukemia VI New Results in Clinical and Biological Research Including Pediatric Oncology* (pp. 194-196): Springer.

Hathaway, B., Landsittel, D. P., Gooding, W., Whiteside, T. L., Grandis, J. R., Siegfried, J. M., . . . Ferris, R. L. (2005). Multiplexed analysis of serum cytokines as biomarkers in squamous cell carcinoma of the head and neck patients. *The Laryngoscope, 115*(3), 522-527.

Hayakawa, F., Towatari, M., Kiyoi, H., Tanimoto, M., Kitamura, T., Saito, H., & Naoe, T. (2000). Tandem-duplicated Flt3 constitutively activates STAT5 and MAP kinase and introduces autonomous cell growth in IL-3-dependent cell lines. *Oncogene, 19*(5).

HE LZ, G. F., & Tribioli, C. (1998). Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to retinoic acid in acute promyelocytic leukemia. *Nature Genet, 18*, 126-135.

He, S., Malik, B., Borkin, D., Miao, H., Shukla, S., Kempinska, K., . . . Parkin, B. (2015). Menin-MLL inhibitors block oncogenic transformation by MLL-fusion proteins in a fusion partner-independent manner. *Leukemia*.

Heath, R. J., & Insall, R. H. (2008). F-BAR domains: multifunctional regulators of membrane curvature. *Journal of cell science, 121*(12), 1951-1954.

Hehlmann, R., Grimwade, D., Simonsson, B., Apperley, J., Baccarani, M., Barbui, T., . . . Berger, U. (2010). The European LeukemiaNet-Achievements and perspectives. *Haematologica*, haematol. 2010.032979.

Hess, J. L., Bittner, C. B., Zeisig, D. T., Bach, C., Fuchs, U., Borkhardt, A., . . . Slany, R. K. (2006). c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. *Blood, 108*(1), 297-304.

Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E.-J., & Downing, J. R. (2002). Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t (8; 21) acute myeloid leukemia. *Cancer cell, 1*(1), 63-74.

Hoeijmakers, J. H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature, 411*(6835), 366-374.

Hoffmann, T. K., Sonkoly, E., Homey, B., Scheckenbach, K., Gwosdz, C., Bas, M., . . . Whiteside, T. L. (2007). Aberrant cytokine expression in serum of patients with adenoid cystic carcinoma and squamous cell carcinoma of the head and neck. *Head & neck, 29*(5), 472-478.

Holz-Schietinger, C., Matje, D. M., & Reich, N. O. (2012). Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *Journal of Biological Chemistry, 287*(37), 30941-30951.

Horton, S. J., Grier, D. G., McGonigle, G. J., Thompson, A., Morrow, M., De Silva, I., . . . Brady, H. J. (2005). Continuous MLL-ENL expression is necessary to establish a "Hox Code" and maintain immortalization of hematopoietic progenitor cells. *Cancer Research, 65*(20), 9245-9252.

Hou, H., Lin, C., Chou, W., Liu, C., Chen, C., Tang, J., . . . Chiang, Y. (2014). Integration of cytogenetic and molecular alterations in risk stratification of 318 patients with de novo non-M3 acute myeloid leukemia. *Leukemia, 28*(1), 50-58.

Hsieh, J. J.-D., Ernst, P., Erdjument-Bromage, H., Tempst, P., & Korsmeyer, S. J. (2003). Proteolytic cleavage of MLL generates a complex of N-and C-terminal fragments that confers protein stability and subnuclear localization. *Molecular and cellular biology, 23*(1), 186-194.

Hsu, H.-C., Lee, Y.-M., Tsai, W.-H., Jiang, M.-L., Ho, C.-H., Ho, C.-K., & Wang, S.-Y. (2002). Circulating levels of thrombopoietic and inflammatory cytokines in patients with acute myeloblastic leukemia and myelodysplastic syndrome. *Oncology, 63*(1), 64-69.

Ilkow, C. S., Marguerie, M., Batenchuk, C., Mayer, J., Neriah, D. B., Cousineau, S., . . . Bellamy, D. (2015). Reciprocal cellular cross-talk within the tumor microenvironment promotes oncolytic virus activity. *Nature medicine, 21*(5), 530-536.

Ingham, R. J., Colwill, K., Howard, C., Dettwiler, S., Lim, C. S., Yu, J., . . . Mbamalu, G. (2005). WW domains provide a platform for the assembly of multiprotein networks. *Molecular and cellular biology, 25*(16), 7092-7106.

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., . . . Webster, M. (2009). The human colon cancer methylome shows similar hypo-and

hypermethylation at conserved tissue-specific CpG island shores. *Nature genetics, 41*(2), 178-186.

Ji, H., Ehrlich, L. I., Seita, J., Murakami, P., Doi, A., Lindau, P., . . . Kim, K. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature, 467*(7313), 338-342.

Jing, Y. (2004). The PML-RARα fusion protein and targeted therapy for acute promyelocytic leukemia. *Leukemia & lymphoma, 45*(4), 639-648.

Johnson, M. D., Torri, J. A., Lippman, M. E., & Dickson, R. B. (1993). The role of cathepsin D in the invasiveness of human breast cancer cells. *Cancer Research, 53*(4), 873-877.

Jongen-Lavrencic, M., Sun, S. M., Dijkstra, M. K., Valk, P. J., & Löwenberg, B. (2008). MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia. *Blood, 111*(10), 5078-5085.

Jost, E., Lin, Q., Weidner, C. I., Wilop, S., Hoffmann, M., Walenda, T., . . . Brümmendorf, T. (2014). Epimutations mimic genomic mutations of DNMT3A in acute myeloid leukemia. *Leukemia, 28*(6), 1227-1234.

Ju, Y.-T., Chang, A. C., She, B.-R., Tsaur, M.-L., Hwang, H.-M., Chao, C. C.-K., . . . Lin-Chao, S. (1998). gas7: A gene expressed preferentially in growth-arrested fibroblasts and terminally differentiated Purkinje neurons affects neurite formation. *Proceedings of the National Academy of Sciences, 95*(19), 11423-11428.

Kadota, J., Mizunoe, S., Mito, K., Mukae, H., Yoshioka, S., Kawakami, K., . . . Kohno, S. (2005). High plasma concentrations of osteopontin in patients with interstitial pneumonia. *Respiratory medicine, 99*(1), 111-117.

Kalu, E., Sumar, N., Giannopoulos, T., Patel, P., Croucher, C., Sherriff, E., & Bansal, A. (2007). Cytokine profiles in serum and peritoneal fluid from infertile women with and without endometriosis. *Journal of Obstetrics and Gynaecology Research, 33*(4), 490-495.

Kaya, M. G., Yalcin, R., Okyay, K., Poyraz, F., Bayraktar, N., Pasaoglu, H., . . . Cengel, A. (2012). Potential role of plasma myeloperoxidase level in predicting long-term outcome of acute myocardial infarction. *Texas Heart Institute Journal, 39*(4), 500.

Kelly, L. M., & Gilliland, D. G. (2002). Genetics of myeloid leukemias. *Annual review of genomics and human genetics, 3*(1), 179-198.

Kitoh, T., Ishikawa, H., Ishii, T., & Nakagawa, S. (1998). Elevated SCF levels in the serum of patients with chronic renal failure. *British journal of haematology, 102*(5), 1151-1156.

Kleiner, G., Marcuzzi, A., Zanin, V., Monasta, L., & Zauli, G. (2013). Cytokine levels in the serum of healthy subjects. *Mediators of inflammation, 2013*.

Kolomecki, K., Maciaszczyk, P., Stepien, H., Stepien, T., Kuzdak, K., & Ulanska, J. (2005). P53 concentration and soluble FasL (sFasL) serum level as indicators of apoptosis in patients with benign and malignant thyroid tumors. *Bratislavské lekárske listy, 106*(10), 297.

Kornblau, S. M., McCue, D., Singh, N., Chen, W., Estrov, Z., & Coombes, K. R. (2010). Recurrent expression signatures of cytokines and chemokines are present and are independently prognostic in acute myelogenous leukemia and myelodysplasia. *Blood, 116*(20), 4251-4261.

Krivtsov, A. V., & Armstrong, S. A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. *Nature reviews cancer, 7*(11), 823-833.

Krockenberger, M., Dombrowski, Y., Weidler, C., Ossadnik, M., Hönig, A., Häusler, S., . . . Steinle, A. (2008). Macrophage migration inhibitory factor contributes to the immune escape of ovarian cancer by down-regulating NKG2D. *The Journal of Immunology, 180*(11), 7338-7348.

Krockenberger, M., Kranke, P., Haeusler, S., Engel, J. B., Horn, E., Nuernberger, K., . . . Hoenig, A. (2012). Macrophage migration-inhibitory factor levels in serum of patients with ovarian cancer correlates with poor prognosis. *Anticancer research, 32*(12), 5233-5238.

Kupsa, T., Milos Horacek, J., & Jebavy, L. (2012). The role of cytokines in acute myeloid leukemia: a systematic review. *Biomedical Papers, 156*(4), 291-301.

Lahortiga, I., & Cools, J. (2012). HEMATOLOGY NEWS.

Lambeck, A. J., Crijns, A. P., Leffers, N., Sluiter, W. J., Klaske, A., Braid, M., . . . Kast, W. M. (2007). Serum cytokine profiling as a diagnostic and prognostic tool in ovarian cancer: a potential role for interleukin 7. *Clinical Cancer Research, 13*(8), 2385-2391.

Lavau, C., Du, C., Thirman, M., & Zeleznik-Le, N. (2000). Chromatin related properties of CBP fused to MLL generate a myelodysplastic like syndrome that evolves into myeloid leukemia. *The EMBO Journal, 19*(17), 4655-4664.

Ławicki, S., Będkowska, G. E., Gacuta-Szumarska, E., & Szmitkowski, M. (2013). The plasma concentration of VEGF, HE4 and CA125 as a new biomarkers panel in different stages and sub-types of epithelial ovarian tumors. *Journal of ovarian research, 6*(1), 1.

Lazakovitch, E. M., She, B.-R., Lien, C.-L., Woo, W.-M., Ju, Y.-T., & Lin-Chao, S. (1999). The Gas7 gene encodes two protein isoforms differentially expressed within the brain. *Genomics, 61*(3), 298-306.

Lee, H.-C., Yang, C.-W., Chen, C.-Y., & Au, L.-C. (2011). Single point mutation of microRNA may cause butterfly effect on alteration of global gene expression. *Biochemical and biophysical research communications, 404*(4), 1065-1069.

Lee, J. H. (2013). Clinical Usefulness of Serum CYFRA 21–1 in Patients with Colorectal Cancer. *Nuclear medicine and molecular imaging, 47*(3), 181-187.

Lee, S. E., Kwak, C., Park, M. S., Lee, C. H., Kang, W., & Oh, S.-J. (2000). Ethnic differences in the age-related distribution of serum prostate-specific antigen values: a study in a healthy Korean male population. *Urology, 56*(6), 1007-1010.

Lemmon, M. A., & Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *cell, 141*(7), 1117-1134.

Lemos-Gonzalez, Y., Rodriguez-Berrocal, F., Cordero, O., Gomez, C., & de la Cadena, M. P. (2007). Alteration of the serum levels of the epidermal growth factor receptor and its ligands in patients with non-small cell lung cancer and head and neck carcinoma. *British journal of cancer, 96*(10), 1569-1578.

Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., . . . Welch, J. (2010). DNMT3A mutations in acute myeloid leukemia. *New England Journal of Medicine, 363*(25), 2424-2433.

Leyhe, T., Hoffmann, N., Stransky, E., & Laske, C. (2009). Increase of SCF plasma concentration during donepezil treatment of patients with early Alzheimer's disease. *International Journal of Neuropsychopharmacology, 12*(10), 1319-1326.

Li, Z., Lu, J., Sun, M., Mi, S., Zhang, H., Luo, R. T., . . . Qian, Z. (2008). Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proceedings of the National Academy of Sciences, 105*(40), 15535-15540.

Li, Z., Luo, R. T., Mi, S., Sun, M., Chen, P., Bao, J., . . . Wang, L. (2009). Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias. *Cancer Research, 69*(3), 1109-1116.

Lin, R. J., Nagy, L., Inoue, S., Shao, W., Miller, W. H., & Evans, R. M. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature, 391*(6669), 811-814.

Lippi, G., De Vita, F., Salvagno, G. L., Montagnana, M., Targher, G., & Guidi, G. C. (2009). Serum prolactin in professional soccer players. *Biochemia Medica, 19*(2), 177-181.

Liu, Z.-Y., Sun, Y.-H., Xu, C.-L., Gao, X., Zhang, L.-M., & Ren, S.-C. (2009). Age-specific PSA reference ranges in Chinese men without prostate cancer. *Asian J Androl, 11*(1), 100-103.

Lokody, I. (2014). Epigenetics: histone methyltransferase mutations promote leukaemia. *Nature reviews cancer, 14*(4), 214-215.

Loriaux, M. M., Levine, R. L., Tyner, J. W., Fröhling, S., Scholl, C., Stoffregen, E. P., . . . Berger, R. (2008). High-throughput sequence analysis of the tyrosine kinome in acute myeloid leukemia. *Blood, 111*(9), 4788-4796.

Lotem, J., & Sachs, L. (1996). Control of apoptosis in hematopoiesis and leukemia by cytokines, tumor suppressor and oncogenes. *Leukemia, 10*(6), 925-931.

Löwenberg, B., Griffin, J. D., & Tallman, M. S. (2003). Acute myeloid leukemia and acute promyelocytic leukemia. *ASH Education Program Book, 2003*(1), 82-101.

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., . . . Ferrando, A. A. (2005). MicroRNA expression profiles classify human cancers. *Nature, 435*(7043), 834-838.

Luczak, M., Kazmierczak, M., Hadschuh, L., Lewandowski, K., Komarnicki, M., & Figlerowicz, M. (2012). Comparative proteomics in acute myeloid leukemia. *Współczesna Onkologia, 16*(2), 95.

Malati, T., & Kumari, G. R. (2004). Racial and ethnic variation of PSA in global population: Age specific reference intervals for serum prostate specific antigen in healthy south Indian males. *Indian Journal of Clinical Biochemistry, 19*(1), 132-137.

Mangantig, E., Naing, N. N., Norsa'adah, B., & Azlan, H. (2013). Survival and prognostic factors in Malaysian acute myeloid leukemia patients after allogeneic haematopoietic stem cell transplantation. *International journal of hematology, 98*(2), 197-205.

Marcucci, G., Haferlach, T., & Döhner, H. (2011). Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *Journal of Clinical Oncology*, JCO. 2010.2030. 2554.

Marcucci, G., Mrózek, K., Radmacher, M. D., Garzon, R., & Bloomfield, C. D. (2011). The prognostic and functional role of microRNAs in acute myeloid leukemia. *Blood, 117*(4), 1121-1129.

Mardis, E. R., Ding, L., Dooling, D. J., Larson, D. E., McLellan, M. D., Chen, K., . . . McGrath, S. D. (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. *New England Journal of Medicine, 361*(11), 1058-1066.

Martens, J. H., & Stunnenberg, H. G. (2010). The molecular signature of oncofusion proteins in acute myeloid leukemia. *FEBS letters, 584*(12), 2662-2669.

Mayer, B. J. (2001). SH3 domains: complexity in moderation. *Journal of cell science, 114*(7), 1253-1263.

Megonigal, M. D., Cheung, N.-K. V., Rappaport, E. F., Nowell, P. C., Wilson, R. B., Jones, D. H., . . . Williams, T. M. (2000). Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. *Proceedings of the National Academy of Sciences, 97*(6), 2814-2819.

Megonigal, M. D., Rappaport, E. F., Nowell, P. C., Lange, B. J., & Felix, C. A. (1998). Potential role for wild-type p53 in leukemias with MLL gene translocations. *Oncogene, 16*(10).

Mehdipour, P., Santoro, F., & Minucci, S. (2015). Epigenetic alterations in acute myeloid leukemias. *FEBS Journal, 282*(9), 1786-1800.

Meijers, W. C., van der Velde, A. R., Ruifrok, W. P., Schroten, N. F., Dokter, M. M., Damman, K., . . . van Gilst, W. H. (2014). Renal handling of galectin-3 in the general population, chronic heart failure, and hemodialysis. *Journal of the American Heart Association, 3*(5), e000962.

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., . . . Jaffe, D. B. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature, 454*(7205), 766-770.

Melnick, A., & Licht, J. D. (1999). Deconstructing a Disease: RAR ・, Its Fusion Partners, and Their Roles in the Pathogenesis of Acute Promyelocytic Leukemia. *Blood, 93*(10), 3167-3215.

Mendler, J. H., Maharry, K., Radmacher, M. D., Mrózek, K., Becker, H., Metzeler, K. H., . . . Kohlschmidt, J. (2012). RUNX1 mutations are associated with poor outcome in younger and older patients with cytogenetically normal acute myeloid leukemia and with distinct gene and MicroRNA expression signatures. *Journal of Clinical Oncology, 30*(25), 3109-3118.

Meng, C. Y., Noor, P. J., Ismail, A., Ahid, M. F. M., & Zakaria, Z. (2013). Cytogenetic profile of de novo acute myeloid leukemia patients in Malaysia. *International journal of biomedical science: IJBS, 9*(1), 26.

Metcalf, D. (1999). Stem Cells, Pre-Progenitor Cells and Lineage-Committed Cells: Are Our Dogmas Correct? *Annals of the New York Academy of Sciences, 872*(1), 289-304.

Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., Trka, J., . . . De Braekeleer, M. (2009). New insights to the MLL recombinome of acute leukemias. *Leukemia, 23*(8), 1490-1499.

Mi, S., Lu, J., Sun, M., Li, Z., Zhang, H., Neilly, M. B., . . . Zhang, Y. (2007). MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. *Proceedings of the National Academy of Sciences, 104*(50), 19971-19976.

Michaux, L., Wlodarska, I., Stul, M., Dierlamm, J., Mugneret, F., Herens, C., . . . Verhoef, G. (2000). MLL amplification in myeloid leukemias: a study of 14 cases with multiple copies of 11q23. *Genes, Chromosomes and Cancer, 29*(1), 40-47.

Milewicz, T., Krzysiek, J., Sztefko, K., Radowicki, S., & Krzyczkowska-Sendrakowska, M. (2005). 17β-estradiol regulation of human growth hormone (hGH), insulin-like growth factor-I (IGF-I) and insulin-like growth factor binding protein-3 (IGFBP-3) axis in hypoestrogenic, hypergonadotropic women. *Endokrynologia Polska, 56*(6), 876-882.

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., & Hess, J. L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular cell, 10*(5), 1107-1117.

Milne, T. A., Martin, M. E., Brock, H. W., Slany, R. K., & Hess, J. L. (2005). Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9

locus, promoting transcription and multiple histone modifications. *Cancer Research, 65*(24), 11367-11374.

Mirro, J., Kitchingman, G., Williams, D., Murphy, S., Zipf, T., & Stass, S. (1986). Mixed lineage leukemia: the implications for hematopoietic differentiation [letter]. *Blood, 68*(2), 597-599.

Mocatta, T. J., Pilbrow, A. P., Cameron, V. A., Senthilmohan, R., Frampton, C. M., Richards, A. M., & Winterbourn, C. C. (2007). Plasma concentrations of myeloperoxidase predict mortality after myocardial infarction. *Journal of the American College of Cardiology, 49*(20), 1993-2000.

Moskal, T. L., Huang, S., Ellis, L. M., Fritsche, H., & Chakrabarty, S. (1995). Serum levels of transforming growth factor alpha in gastrointestinal cancer patients. *Cancer Epidemiology Biomarkers & Prevention, 4*(2), 127-131.

Mrózek, K., & Bloomfield, C. D. (2006). Chromosome aberrations, gene mutations and expression changes, and prognosis in adult acute myeloid leukemia. *ASH Education Program Book, 2006*(1), 169-177.

Mrózek, K., Heerema, N. A., & Bloomfield, C. D. (2004). Cytogenetics in acute leukemia. *Blood reviews, 18*(2), 115-136.

Nagy, B., Bhattoa, H. P., Steiber, Z., Csobán, M., Szilasi, M., Méhes, G., . . . Antal-Szalmás, P. (2014). Serum human epididymis protein 4 (HE4) as a tumor marker in men with lung cancer. *Clinical Chemistry and Laboratory Medicine (CCLM), 52*(11), 1639-1648.

Nakamura, T., Canaani, E., & Croce, C. M. (2007). Oncogenic All1 fusion proteins target Drosha-mediated microRNA processing. *Proceedings of the National Academy of Sciences, 104*(26), 10980-10985.

Naumnik, W., Izycki, T., Ossolinska, M., & Chyczewska, E. (2007). Serum levels of sFas and sFasL during chemotherapy of lung cancer. *Exp Oncol, 29*(2), 132-136.

Network, C. G. A. R. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England journal of medicine, 368*(22), 2059.

Odashiro, M., Hans Filho, G., Pereira, P. R., Castro, A. R. C. M., Stief, A. C., Pontes, E. R. J. C., & Odashiro, A. N. (2015). Melanoma inhibitory activity in Brazilian patients with cutaneous melanoma. *Anais brasileiros de dermatologia, 90*(3), 327-332.

Oesterling, J., Kumamoto, Y., Tsukamoto, T., Girman, C., Guess, H., Masumori, N., . . . Lieber, M. (1995). Serum prostate-specific antigen in a community-based population of healthy Japanese men: lower values than for similarly aged white men. *British journal of urology, 75*(3), 347-353.

Oki, Y., & Issa, J.-P. J. (2009). Epigenetic mechanisms in AML–a target for therapy *Acute Myelogenous Leukemia* (pp. 19-40): Springer.

Panagopoulos, I., Lilljebjorn, H., Strombeck, B., Hjorth, L., Olofsson, T., & Johansson, B. (2006). MLL/GAS7 fusion in a pediatric case of t (11; 17)(q23; p13)-positive precursor B-cell acute lymphoblastic leukemia. *Haematologica, 91*(9), 1287-1288.

Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., . . . Ma, J. (2011). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature, 471*(7337), 189-195.

Patel, J. P., Gönen, M., Figueroa, M. E., Fernandez, H., Sun, Z., Racevskis, J., . . . Aminova, O. (2012). Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *New England Journal of Medicine, 366*(12), 1079-1089.

Pendergast, A. M. (2002). The Abl family kinases: mechanisms of regulation and signaling. *Advances in cancer research, 85*, 51-100.

Petersen, P. M., Skakkebæk, N. E., Vistisen, K., Rørth, M., & Giwercman, A. (1999). Semen quality and reproductive hormones before orchiectomy in men with testicular cancer. *Journal of Clinical Oncology, 17*(3), 941-941.

Popovic, R., Riesbeck, L. E., Velu, C. S., Chaubey, A., Zhang, J., Achille, N. J., . . . Grimes, H. L. (2009). Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. *Blood, 113*(14), 3314-3322.

Preti, H. A., Cabanillas, F., Talpaz, M., Tucker, S. L., Seymour, J. F., & Kurzrock, R. (1997). Prognostic value of serum interleukin-6 in diffuse large-cell lymphoma. *Annals of Internal Medicine, 127*(3), 186-194.

Pui, C., Sandlund, J., Pei, D., Campana, D., Rivera, G., Ribeiro, R., . . . Hudson, M. (2004). Total Therapy Study XIIIB at St Jude Children's Research Hospital. Improved outcome for children with acute lymphoblastic leukemia: results of total therapy Study XIIIB at St Jude Children's Research Hospital. *Blood, 104*(9), 2690-2696.

Qu, Y., Lennartsson, A., Gaidzik, V. I., Deneberg, S., Karimi, M., Bengtzén, S., . . . Lehmann, S. (2014). Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics, 9*(8), 1108-1119.

Ramsingh, G., Koboldt, D. C., Trissal, M., Chiappinelli, K. B., Wylie, T., Koul, S., . . . Goodfellow, P. (2010). Complete characterization of the microRNAome in a patient with acute myeloid leukemia. *Blood, 116*(24), 5316-5326.

Refik, M., Mehmet, N., Durmaz, R., & Ersoy, Y. (2004). Cytokine profile and nitric oxide levels in sera from patients with brucellosis. *Brazilian journal of medical and biological research, 37*(11), 1659-1663.

Richard, V., Kindt, N., & Saussez, S. (2015). Macrophage migration inhibitory factor involvement in breast cancer (Review). *International journal of oncology, 47*(5), 1627-1633.

Rochefort, H., Capony, F., & Garcia, M. (1990). Cathepsin D: a protease involved in breast cancer metastasis. *Cancer and Metastasis Reviews, 9*(4), 321-331.

Rowley, J. D. (2008). Chromosomal translocations: revisited yet again. *Blood, 112*(6), 2183-2189.

Ryningen, A., Wergeland, L., Glenjen, N., Gjertsen, B. T., & Bruserud, Ø. (2005). In vitro crosstalk between fibroblasts and native human acute myelogenous leukemia (AML) blasts via local cytokine networks results in increased proliferation and decreased apoptosis of AML cells as well as increased levels of proangiogenic Interleukin 8. *Leukemia research, 29*(2), 185-196.

Saied, M. H., Marzec, J., Khalid, S., Smith, P., Down, T. A., Rakyan, V. K., . . . Young, B. D. (2012). Genome wide analysis of acute myeloid leukemia reveal leukemia specific methylome and subtype specific hypomethylation of repeats. *PloS one, 7*(3), e33213.

Säily, M., Koistinen, P., & Savolainen, E.-R. (1999). The soluble form of interleukin-6 receptor modulates cell proliferation by acute myeloblastic leukemia blast cells. *Annals of hematology, 78*(4), 173-179.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning* (Vol. 2): Cold spring harbor laboratory press New York.

Sanchez-Correa, B., Bergua, J. M., Campos, C., Gayoso, I., Arcos, M. J., Bañas, H., . . . Tarazona, R. (2013). Cytokine profiles in acute myeloid leukemia patients at diagnosis: survival is inversely correlated with IL-6 and directly correlated with IL-10 levels. *Cytokine, 61*(3), 885-891.

SARS, B. G. o. N. R. P. f. (2003). Dynamic changes in blood cytokine levels as clinical indicators in severe acute respiratory syndrome. *Chinese medical journal, 116*(9), 1283.

Scambia, G., Benedetti, P., Ferrandina, G., Battaglia, F., Baiocchi, G., & Mancuso, S. (1991). Cathepsin D assay in ovarian cancer: correlation with pathological features and receptors for oestrogen, progesterone and epidermal growth factor. *British journal of cancer, 64*(1), 182.

Scambia, G., Panici, P. B., Ferrandina, G., Salerno, G., D'Agostino, G., Distefano, M., . . . Mancuso, S. (1994). Clinical significance of cathepsin D in primary ovarian cancer. *European Journal of Cancer, 30*(7), 935-940.

Schlenk, R. F., Döhner, K., Krauter, J., Fröhling, S., Corbacioglu, A., Bullinger, L., . . . Benner, A. (2008). Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *New England Journal of Medicine, 358*(18), 1909-1918.

Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *cell, 103*(2), 211-225.

Schnittger, S., Schoch, C., Dugas, M., Kern, W., Staib, P., Wuchter, C., . . . Büchner, T. (2002). Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG

study and usefulness as a marker for the detection of minimal residual disease. *Blood, 100*(1), 59-66.

Schoch, C., Kern, W., Schnittger, S., Buchner, T., Hiddemann, W., & Haferlach, T. (2004). The influence of age on prognosis of de novo acute myeloid leukemia differs according to cytogenetic subgroups. *Haematologica, 89*(9), 1082-1090.

Schoch, C., Kohlmann, A., Schnittger, S., Brors, B., Dugas, M., Mergenthaler, S., . . . Haferlach, T. (2002). Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proceedings of the National Academy of Sciences, 99*(15), 10008-10013.

Secchiero, P., Corallini, F., Ceconi, C., Parrinello, G., Volpato, S., Ferrari, R., & Zauli, G. (2009). Potential prognostic significance of decreased serum levels of TRAIL after acute myocardial infarction. *PloS one, 4*(2), e4442.

Seedhouse, C., Faulkner, R., Ashraf, N., Das-Gupta, E., & Russell, N. (2004). Polymorphisms in genes involved in homologous recombination repair interact to increase the risk of developing acute myeloid leukemia. *Clinical Cancer Research, 10*(8), 2675-2680.

Shen, Y., Zhu, Y.-M., Fan, X., Shi, J.-Y., Wang, Q.-R., Yan, X.-J., . . . Jiang, C.-L. (2011). Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood, 118*(20), 5593-5603.

Shih, A. H., Abdel-Wahab, O., Patel, J. P., & Levine, R. L. (2012). The role of mutations in epigenetic regulators in myeloid malignancies. *Nature reviews cancer, 12*(9), 599-612.

Shivdasani, R. A. (2006). MicroRNAs: regulators of gene expression and cell differentiation. *Blood, 108*(12), 3646-3653.

Siegel, R., Ma, J., Zou, Z., & Jemal, A. (2014). Cancer statistics, 2014. *CA: a cancer journal for clinicians, 64*(1), 9-29.

Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA: a cancer journal for clinicians, 65*(1), 5-29.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: a cancer journal for clinicians, 66*(1), 7-30.

Sierra, J., Yoshida, T., Joazeiro, C. A., & Jones, K. A. (2006). The APC tumor suppressor counteracts β-catenin activation and H3K4 methylation at Wnt target genes. *Genes & development, 20*(5), 586-600.

Slany, R. K. (2009). The molecular biology of mixed lineage leukemia. *Haematologica, 94*(7), 984-993.

Slany, R. K., Lavau, C., & Cleary, M. L. (1998). The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Molecular and cellular biology, 18*(1), 122-129.

Slovak, M. L., Kopecky, K. J., Cassileth, P. A., Harrington, D. H., Theil, K. S., Mohamed, A., . . . Rowe, J. M. (2000). Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study. *Blood, 96*(13), 4075-4083.

So, C. W., Karsunky, H., Passegué, E., Cozzio, A., Weissman, I. L., & Cleary, M. L. (2003). MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. *Cancer cell, 3*(2), 161-171.

So, C. W., Karsunky, H., Wong, P., Weissman, I. L., & Cleary, M. L. (2004). Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. *Blood, 103*(8), 3192-3199.

So, C. W., Lin, M., Ayton, P. M., Chen, E. H., & Cleary, M. L. (2003). Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. *Cancer cell, 4*(2), 99-110.

So, H. J., Hong, S. I., Lee, J. K., Chang, Y. H., Kang, S. J., & Hong, Y. J. (2014). Comparison of the serum fibrin-fibrinogen degradation products with cytokeratin 19 fragment as biomarkers in patients with lung cancer. *Biomedical reports, 2*(5), 737-742.

Stabler, T., Piette, J. C., Chevalier, X., Marini-Portugal, A., & Kraus, V. B. (2004). Serum cytokine profiles in relapsing polychondritis suggest monocyte/macrophage activation. *Arthritis & Rheumatism, 50*(11), 3663-3667.

Standal, T., Hjorth-Hansen, H., Rasmussen, T., Dahl, I., Lenhoff, S., Brenne, A.-T., . . . Borset, M. (2004). Osteopontin is an adhesive factor for myeloma cells and is found in increased levels in plasma from patients with multiple myeloma. *Haematologica, 89*(2), 174-182.

Stass, S., Mirro, J., Melvin, S., Pui, C.-H., Murphy, S. B., & Williams, D. (1984). Lineage switch in acute leukemia. *Blood, 64*(3), 701-706.

Stopka, T., Zakova, D., Fuchs, O., Kubrova, O., Blafkova, J., Jelinek, J., . . . Zivny, J. (2000). Chromatin remodeling gene SMARCA5 is dysregulated in primitive hematopoietic cells of acute leukemia. *Leukemia, 14*(7), 1247-1252.

Stroncek, D., Slezak, S., Khuu, H., Basil, C., Tisdale, J., Leitman, S. F., . . . Panelli, M. C. (2005). Proteomic signature of myeloproliferation and neutrophilia: analysis of serum and plasma from healthy subjects given granulocyte colony-stimulating factor. *Experimental hematology, 33*(10), 1109-1117.

Su, W. B., You, J.-J., Huang, B.-T., Sivakumar, V., Yeh, S.-D., & Lin-Chao, S. (2000). Gas7.

Suzuki, M., Oda, M., Ramos, M.-P., Pascual, M., Lau, K., Stasiek, E., . . . Jing, Q. (2011). Late-replicating heterochromatin is characterized by decreased cytosine methylation in the human genome. *Genome research, 21*(11), 1833-1840.

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics, 9*(6), 465-476.

Swerdllow, S., Campo, E., & Harris, N. L. (2008). *WHO classification of tumours of haematopoietic and lymphoid tissues*: France: IARC Press, 2008.

Tao, M., Li, B., Nayini, J., Andrews, C. B., Huang, R.-W., Devemy, E., . . . Preisler, H. D. (2000). SCF, IL-1β, IL-1ra and GM-CSF in the bone marrow and serum of normal individuals and of AML and CML patients. *Cytokine, 12*(6), 699-707.

Tas, F., Bilgin, E., Tastekin, D., Erturk, K., & Duranyildiz, D. (2016). Serum IGF-1 and IGFBP-3 levels as clinical markers for patients with lung cancer. *Biomedical reports, 4*(5), 609-614.

Tas, F., Karabulut, S., Serilmez, M., Ciftci, R., & Duranyildiz, D. (2014). Serum levels of macrophage migration-inhibitory factor (MIF) have diagnostic, predictive and prognostic roles in epithelial ovarian cancer patients. *Tumor Biology, 35*(4), 3327-3331.

Tenen, D. G. (2003). Disruption of differentiation in human cancer: AML shows the way. *Nature reviews cancer, 3*(2), 89-101.

Timp, W., & Feinberg, A. P. (2013). Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nature reviews cancer, 13*(7), 497-510.

Tomasson, M. H., Xiang, Z., Walgren, R., Zhao, Y., Kasai, Y., Miner, T., . . . McLellan, M. D. (2008). Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with de novo acute myeloid leukemia. *Blood, 111*(9), 4797-4808.

Tomizawa, D., Koh, K., Sato, T., Kinukawa, N., Morimoto, A., Isoyama, K., . . . Hayashi, Y. (2007). Outcome of risk-based therapy for infant acute lymphoblastic leukemia with or without an MLL gene rearrangement, with emphasis on late effects: a final report of two consecutive studies, MLL96 and MLL98, of the Japan Infant Leukemia Study Group. *Leukemia, 21*(11), 2258-2263.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA: a cancer journal for clinicians, 65*(2), 87-108.

Tripatara, A., Srichana, N., Lamool, P., Amnuaykan, S., Hongart, P., & Jetsrisuparb, A. (2012). Relationship between Plasma Ferritin Level and Siderocyte Number in Splenectomized β-Thalassemia/HbE Patients. *Anemia, 2012*.

Tsimberidou, A. M., Estey, E., Wen, S., Pierce, S., Kantarjian, H., Albitar, M., & Kurzrock, R. (2008). The prognostic significance of cytokine levels in newly diagnosed acute myeloid leukemia and high-risk myelodysplastic syndromes. *Cancer, 113*(7), 1605-1613.

Ugurel, S., Rappl, G., Tilgen, W., & Reinhold, U. (2001). Increased soluble CD95 (sFas/CD95) serum level correlates with poor prognosis in melanoma patients. *Clinical Cancer Research, 7*(5), 1282-1286.

Uitte de Willige, S., Malfliet, J., Janssen, H., Leebeek, F., & Rijken, D. (2013). Increased N-terminal cleavage of alpha 2 antiplasmin in patients with liver cirrhosis. *Journal of Thrombosis and Haemostasis, 11*(11), 2029-2036.

Van Etten, R. A. (2007). Aberrant cytokine signaling in leukemia. *Oncogene, 26*(47), 6738-6749.

Vardiman, J. W., Harris, N. L., & Brunning, R. D. (2002). The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood, 100*(7), 2292-2302.

Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., . . . Tefferi, A. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood, 114*(5), 937-951.

Verjans, E., Noetzel, E., Bektas, N., Schütz, A. K., Lue, H., Lennartz, B., . . . Bernhagen, J. (2009). Dual role of macrophage migration inhibitory factor (MIF) in human breast cancer. *BMC cancer, 9*(1), 1.

Vermeulen, A., Kaufman, J., & Giagulli, V. (1996). Influence of some biological indexes on sex hormone-binding globulin and androgen levels in aging or obese males. *The Journal of Clinical Endocrinology & Metabolism, 81*(5), 1821-1826.

Verstovsek, S., Kantarjian, H., Estey, E., Aguayo, A., Giles, F., Manshouri, T., . . . Keating, M. (2001). Plasma hepatocyte growth factor is a prognostic factor in patients with acute myeloid leukemia but not in patients with myelodysplastic syndrome. *Leukemia, 15*(8), 1165-1170.

Vinante, F., Rigo, A., Tecchio, C., Morosato, L., Nadali, G., Ricetti, M. M., . . . Gallati, H. (1998). Serum levels of p55 and p75 soluble TNF receptors in adult acute leukaemia at diagnosis: correlation with clinical and biological features and outcome. *British journal of haematology, 102*(4), 1025-1034.

Vucetic, B., Rogan, S. A., Hrabac, P., Hudorovic, N., Cupic, H., Lukinac, L., . . . Zekan, M. (2008). Biological value of melanoma inhibitory activity serum concentration in patients with primary skin melanoma. *Melanoma research, 18*(3), 201-207.

Wang, J., Wen, S., Symmans, W. F., Pusztai, L., & Coombes, K. R. (2009). The bimodality index: a criterion for discovering and ranking bimodal signatures from cancer gene expression profiling data. *Cancer informatics, 7*, 199.

Wang, L., Gural, A., Sun, X.-J., Zhao, X., Perna, F., Huang, G., . . . Xu, H. (2011). The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science, 333*(6043), 765-769.

Wang, Q.-f., Wu, G., Mi, S., He, F., Wu, J., Dong, J., . . . Prabhakar, S. (2011). MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood, 117*(25), 6895-6905.

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., . . . Lopes, C. T. (2010). The GeneMANIA prediction server: biological network

integration for gene prioritization and predicting gene function. *Nucleic acids research, 38*(suppl 2), W214-W220.

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., . . . Network, C. G. A. R. (2013). The cancer genome atlas pan-cancer analysis project. *Nature genetics, 45*(10), 1113-1120.

Westhoff, U., Fox, C., & Otto, F. (1997). Quantification of cathepsin D in plasma of patients with malignant melanoma. *Anticancer research, 18*(5B), 3785-3788.

Wiederschain, D., Kawai, H., Shilatifard, A., & Yuan, Z.-M. (2005). Multiple mixed lineage leukemia (MLL) fusion proteins suppress p53-mediated response to DNA damage. *Journal of Biological Chemistry, 280*(26), 24315-24321.

Wilson, H. P., Lesnikov, V., Plymate, S., Ward, J., & Deeg, H. (2005). High IGFBP-3 levels in marrow plasma in early-stage MDS: effects on apoptosis and hemopoiesis. *Leukemia, 19*(4), 580-585.

Wróbel, T., Gębura, K., Wysoczańska, B., Jaźwiec, B., Dobrzyńska, O., Mazur, G., . . . Bogunia-Kubik, K. (2014). IL-17F gene polymorphism is associated with susceptibility to acute myeloid leukemia. *Journal of cancer research and clinical oncology, 140*(9), 1551-1555.

Wu, X.-Y., Li, S.-N., Wen, S.-N., Nie, J.-G., Deng, W.-N., Bai, R., . . . Du, X. (2015). Plasma galectin-3 predicts clinical outcomes after catheter ablation in persistent atrial fibrillation patients without structural heart disease. *Europace, 17*(10), 1541-1547.

Xu, X., Wang, B., Ye, C., Yao, C., Lin, Y., Huang, X., . . . Wang, S. (2008). Overexpression of macrophage migration inhibitory factor induces angiogenesis in human breast cancer. *Cancer letters, 261*(2), 147-157.

Yalcin, A., Kreutz, C., Pfeifer, D., Abdelkarim, M., Klaus, G., Timmer, J., . . . Hackanson, B. (2013). MeDIP coupled with a promoter tiling array as a platform to investigate global DNA methylation patterns in AML cells. *Leukemia research, 37*(1), 102-111.

Yilmaz, M., Kis, C., Ceylan, N., Okan, V., Pehlivan, M., Kucukosmanoglu, E., . . . Tarakcioglu, M. (2008). Serum leptin level in acute myeloid leukemia patients. *Hematology, 13*(1), 21-23.

Yokoyama, A., & Cleary, M. L. (2008). Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer cell, 14*(1), 36-46.

Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I., & Cleary, M. L. (2010). A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer cell, 17*(2), 198-212.

Yuan, Y., Kasar, S., Underbayev, C., Prakash, S., & Raveche, E. (2012). MicroRNAs in acute myeloid leukemia and other blood disorders. *Leukemia research and treatment, 2012*.

Yurkovetsky, Z. R., Kirkwood, J. M., Edington, H. D., Marrangoni, A. M., Velikokhatnaya, L., Winans, M. T., . . . Lokshin, A. E. (2007). Multiplex analysis of serum cytokines in melanoma patients treated with interferon-α2b. *Clinical Cancer Research, 13*(8), 2422-2428.

Zeh, H., Winikoff, S., Landsittel, D., Gorelik, E., Marrangoni, A., Velikokhatnaya, L., . . . Bartlett, D. (2005). Multianalyte profiling of serum cytokines for detection of pancreatic cancer. *Cancer Biomarkers, 1*(6), 259-269.

Zeisig, B. B., Milne, T., García-Cuéllar, M.-P., Schreiner, S., Martin, M.-E., Fuchs, U., . . . Soden, R. (2004). Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Molecular and cellular biology, 24*(2), 617-628.

Zhang, Y.-y., Yao, D.-m., Zhu, X.-w., Zhou, J.-d., Ma, J.-c., Yang, J., . . . Qian, J. (2015). DNMT3A intragenic hypomethylation is associated with adverse prognosis in acute myeloid leukemia. *Leukemia research, 39*(10), 1041-1047.

Zody, M. C., Garber, M., Adams, D. J., Sharpe, T., Harrow, J., Lupski, J. R., . . . Young, S. K. (2006). DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. *Nature, 440*(7087), 1045-1049.

Zurawski, V. R., Orjaseter, H., Andersen, A., & Jellum, E. (1988). Elevated serum CA 125 levels prior to diagnosis of ovarian neoplasia: relevance for early detection of ovarian cancer. *International journal of cancer, 42*(5), 677-680.