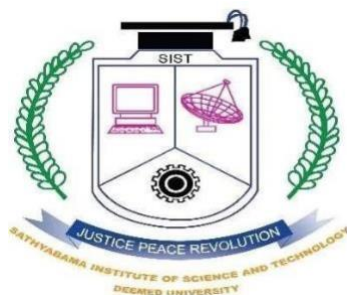


METABOLOMICS OF ACUTE MYELOID LEUKEMIA

Submitted in Partial Fulfilment of the Requirements for the award of Bachelor
of Technology Degree in Bioinformatics

By

G. DEVISHRI (Reg. No. 37220003)



DEPARTMENT OF BIOINFORMATICS

SCHOOL OF BIO AND CHEMICAL ENGINEERING

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with Grade “A” by NAAC

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119

APRIL– 2021



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with "A" grade by NACC

Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600119 www.sathyabama.ac.in



DEPARTMENT OF BIOINFORMATICS

BONAFIDE CERTIFICATE

This is to certify that this project report is the bonafide work **G. DEVISHRI (37220003)** who carried out the project entitled **"METABOLOMICS OF ACUTE MYELOID LEUKEMIA USING METABOLOMIC APPROACHES"** under my supervision from September 2020 to April 2021.

Internal Guide
(Dr. DANIEL ALEX ANAND)

Head of the Department (Dr. H. JEMMY CHRISTY)

Submitted for viva voce Examination held on _____

Internal Examiner

External Examiner

DECLARATION

I, **G.DEVISHRI (37220003)** hereby declare that the project report entitled **“METABOLOMICS OF ACUTE MYELOID LEUKEMIA USING METABOLOMIC APPROACHES”** ” done by me under the guidance of **Dr. DANIEL ALEX ANAND Ph.D.**, is submitted in partial fulfilment of the requirements for the award of Bachelor of Technology in BIOINFORMATICS.

PLACE: CHENNAI.

**DATE:
SIGNATURE OF THE CANDIDATE**

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to Board of Management of SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. H Jemmy Christy, Head of the Department, Dept. of Bioinformatics** for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. Daniel Alex Anand, Department of Bioinformatics, Sathyabama Institute of Science and Technology**, for his valuable guidance, and encouragement which paved way for the successful completion of my project work.

I wish to express my thanks to all Faculty members and Non-teaching staff members of the Department of Biotechnology who were helpful in many ways for the completion of the project.

Heartfelt thanks to my lovable Parents for their endless moral support, care, affection and blessings.

G.DEVISHRI

ABSTRACT

Acute Myeloid Leukaemia (AML) is a complex, heterogeneous haematological malignancy caused by mutations in myeloid differentiation and proliferation. Its response to therapy is based on chromosomal and molecular aberrations. Metabolomics, the global profiling of small molecules in a biological sample, and it emerged as an important tool for studying the metabolic dependency of cancer cells. AML is the most common acute leukaemia affecting adults and increases with age. Traditional diagnosis of AML depends on invasive bone marrow biopsies. Metabolomics study showed that systemic metabolic markers were associated with prognosis in AML. Metabolomic profiling is a tool for understanding the mechanisms of metabolites. Our study is focused on gene expression patterns which correlate with metabolic pathway contribute to suppress the AML cell proliferation, in this metabolome analysis the pathway of interest is associated with the disease is JAK STAT pathway and the beta alanine metabolism is showing the highest impact. Patients with relapsed or refractory acute myeloid leukemia (AML) with mutations in the FMS-like tyrosine kinase 3 gene (FLT3) infrequently have a response to salvage chemotherapy. Gilteritinib is an oral, potent, selective FLT3 inhibitor with single-agent activity in relapsed or refractory FLT3-mutated AML patients.

KEYWORDS: AML, serum metabolites, metabolic markers, metabolomics, cancer gilteritinib.

TABLE OF CONTENTS

CHAPTERNO	TITLE	PAGE NO
	ABSTRACT	I
	LIST OF TALES	II
	LIST OF FIGURES	III
	LIST OF ABBREVIATIONS	IV
1	INTRODUCTION	
	1.1 Metabolomics	1
	1.2 Acute Myeloid Leukaemia	3
	1.3 AML for FLT3, DNMT3, NPM1	6
	1.4 Effect of AML On Metabolites	10
2	LITERATURE SURVEY	21
3	AIM AND SCOPE	24
4	MATERIALS AND METHODS USED	
	4.1 Cbio Portal	25
	4.2 LGA (Leukaemia Gene Atlas)	26
	4.3 HMDB	27
	4.4 Metaboanalyst	28
	4.5 Cytoscape	29
	4.6 Metscape	30
	4.7 KEGG	32
	4.8 PDB	33
	4.9 MMDB	34
	4.10 DrugBank	35

5	RESULT AND DISCUSSION	
	5.1 Metaboanalyst	38
	5.2 Cytoscape	40
	5.3 Drug Bank	45
6	SUMMARY AND CONCLUSION	
	6.1 Summary	50
	6.2 Conclusion	52
	REFERENCES	54

LIST OF TABLES

TABLE#	TITLE	PAGENO
1.1	TYPES OF AML	18
1.2	Showing the copy number alteration in AML	36
1.3	Metabolites involved in AML	37
1.4	Pathways involved in the selected metabolites in metaboanalyst	39
1.5	List of metabolites are recognized and their KEGGID'S	41
1.6	List of compounds, enzymes interaction network by Cytoscape	43
1.7	List of drug molecule which will inhibit the AML.	46

LIST OF FIGURES

TITLE	PAGENO
Phenotype of AML	14
Classification of blood cells to AML	16
Genes involved in AML	17
Overview of pathway showing the maximum impact	38
Network with Labelling in Metscape	40
Showing pathway labelling with Metscape	42
Network variation in AML.	45
PDB visualisation of inhibiting Drug molecules	47
MMDB 3D structure visualisation.	49

LIST OF ABBREVIATIONS

- **AML** Acute myeloid Leukaemia
- **FLT3** FMS- like tyrosine kinase 3
- **DNMT3A** DNA methyltransferase 3A
- **NPM1** Nucleophosmin gene
- **HMDB** Human Metablome Database
- **KEGG** Kyoto Encyclopedia of Genes and Genomes
- **PDB** Protein Data Bank
- **MMDB** Molecular Modelling Database

CHAPTER -1

INTRODUCTION

1.1 METABOLOMICS:

Metabolomics is the large-scale study of small molecules, commonly known as metabolites, within cells, bio fluids, tissues or organisms. Collectively, these small molecules and their interactions within a biological system are known as the metabolome.

Metabolomics is an emerging field of "omics" research specializing in the near global analysis of small molecule metabolites found in living organisms. Metabolomics is a powerful approach because metabolites and their concentrations, unlike other "omics" measures, directly reflect the underlying biochemical activity and state of cells / tissues. Thus, metabolomics best represents the molecular phenotype.

It is the study of substrates and products of metabolism, which are influenced by both genetic and environmental factors. Metabolomics is a powerful approach because metabolites and their concentrations, unlike other "omics" measures, directly reflect the underlying biochemical activity and state of cells / tissues. Thus, metabolomics best represents the molecular phenotype.

Metabolomics usually studies small molecules within a mass range of 50 – 1500 Daltons it identifies and quantifies the identified metabolites using various methods and to extract information for data interpretation. Another important part in metabolomics.

Metabolomics uses various analytical methods, such as spectroscopy, chromatography, and multivariate analysis; for the identification & quantification of various metabolites the analytical methods facilitate the classification of various known & unknown metabolites the comprehensive nature of metabolites enables its application in diverse medical fields thereby facilitating the development of new therapeutic procedures and diagnostics.

There are some tools like NMR, spectroscopy liquid chromatography are used to analyse large metabolite samples.

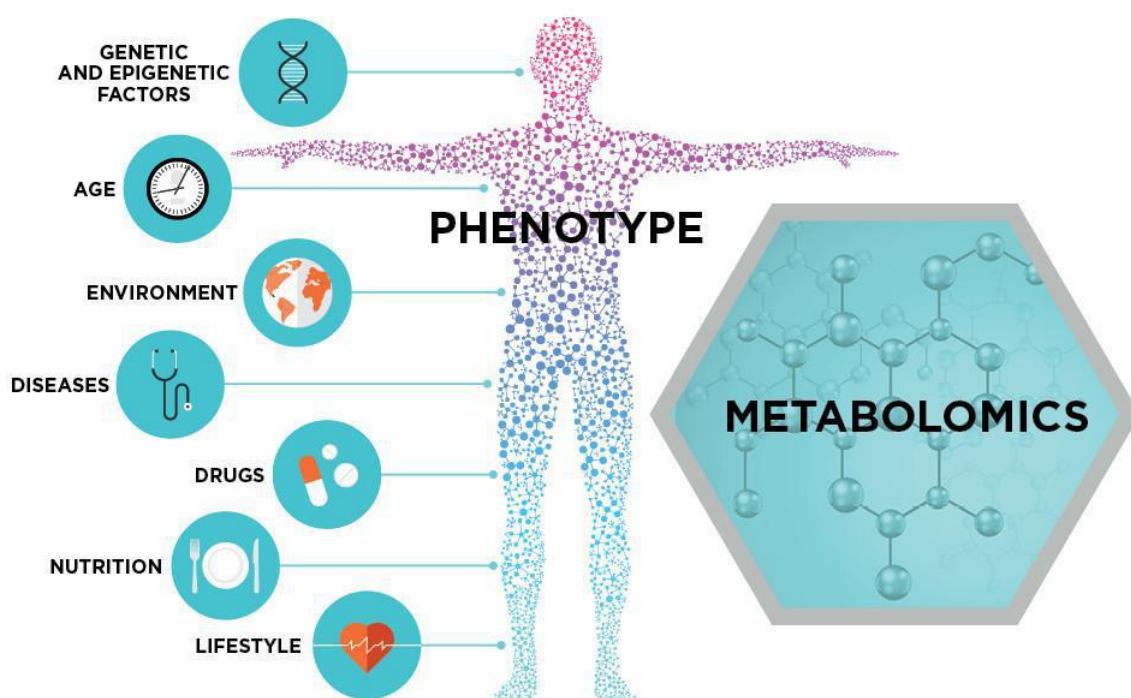


Figure 1.1 phenotype of metabolism (Metabolomics technologies.inc)

1.2 ACUTE MYELOID LEUKEMIA:

Acute myeloid leukaemia (AML) is a complex, heterogeneous haematological malignancy caused by mutations in myeloid differentiation and proliferation. Response to therapy and long-term outcomes vary widely based on chromosomal and molecular aberrations. Many platforms have been used to characterize and stratify AML. Acute myeloid leukaemia (AML) is a form of cancer that is characterized by infiltration of the bone marrow, blood, and other tissues by proliferative, clonal, abnormally differentiated, and occasionally poorly differentiated cells of the hematopoietic system. Acute myeloid leukemia (AML) is a form of cancer that is characterized by infiltration of the bone marrow, blood, and other tissues by proliferative, clonal, abnormally differentiated, and occasionally poorly differentiated cells of the hematopoietic system. Although it was incurable 50 years ago, AML is now cured in 35 to 40% of adult patients who are 60 years of age or younger and in 5 to 15% of patients who are older than 60 years of age.¹ The outcome in older patients who are unable to receive intensive chemotherapy without unacceptable side effects remains dismal, with a median survival of only 5 to 10 months.

Although it was incurable 50 years ago, AML is now cured in 35 to 40% of adult patients who are 60 years of age or younger and in 5 to 15% of patients who are older than 60 years of age.¹ The outcome in older patients who are unable to receive intensive chemotherapy without unacceptable side effects remains dismal, with a median survival of only 5 to 10 months.

AML is classified according to the World Health Organization (WHO) Classification of Tumours of Haematopoietic and Lymphoid Tissues, 2 which was last updated in 2008. The major categories of the current classification include AML with recurrent genetic abnormalities, AML with myelodysplasia related changes, therapy related AML. AML is the most common acute leukaemia in adults, and AML increases with age. Nearly 1.3 million people are affected by AML. The patients were diagnosed with the clinical prognosis and the identified genes include FLT3, DNMT3A, CEBPA, RAS, BRCL, IDH. The mutation in FLT3, DNMT3, has more effect on the growth of white blood cells and induces too many tumour cells. The mutational pattern may be indicative of AML ontogeny (that is, whether the AML was newly diagnosed as primary disease or as a secondary disorder after an antecedent myeloid disorder such as a myelodysplastic syndrome). In a recent study, the presence of mutations in SRSF2, SF3B1, U2AF1, ZRSR2, ASXL1, EZH2, BCOR, or STAG2 defined a distinct genetic subtype of AML that shares clinicopathologic features with clinically confirmed secondary AML.¹⁹

The mutational pattern may be indicative of AML ontogeny (that is, whether the AML was newly diagnosed as primary disease or as a secondary disorder after an antecedent myeloid disorder such as a myelodysplastic syndrome). In a recent study, the presence of mutations in SRSF2, SF3B1, U2AF1, ZRSR2, ASXL1, EZH2, BCOR, or STAG2 defined a distinct genetic subtype of AML that shares clinicopathologic features with clinically confirmed secondary AML.

The frequent occurrence of mutations in receptor tyrosine kinase genes (FLT3 and KIT) has generated interest in the development of tyrosine kinase inhibitors. (Döhner H, Estey EH, Amadori S, et al.)

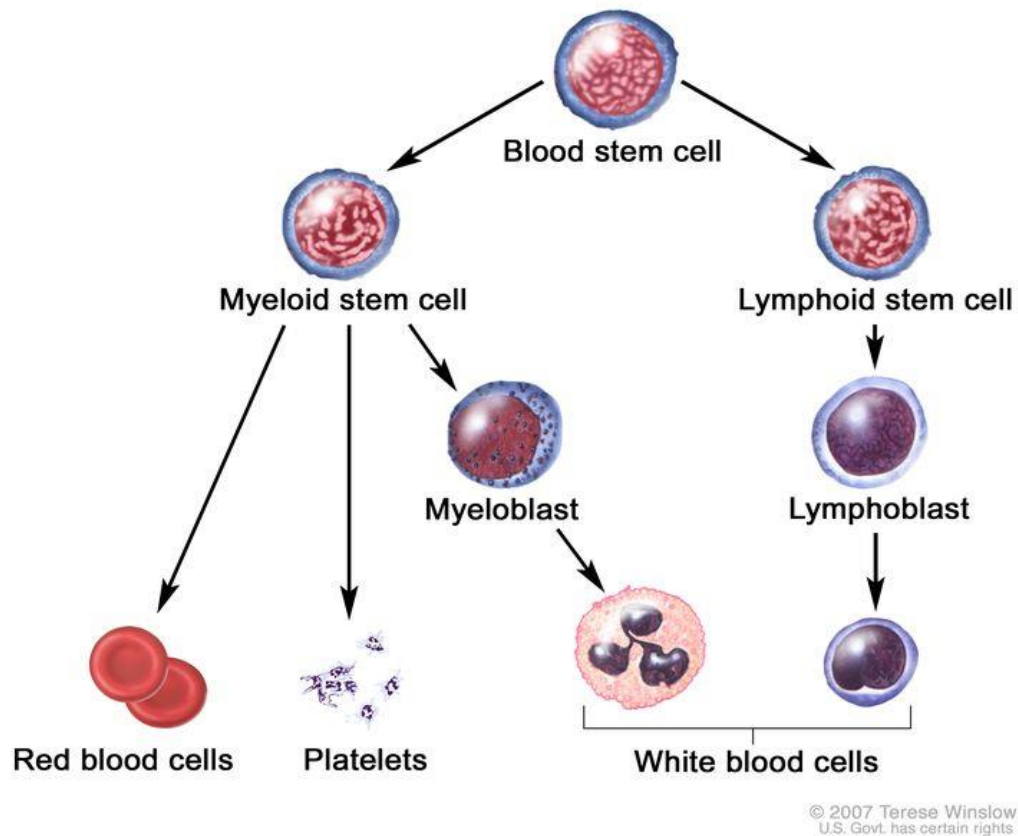


Figure 1.2 shows the proliferation of AML
(Brunning RD, Matutes E, Harris NL, et al.: Acute myeloid leukaemia)

AML is a heterogeneous group of blood cancers that arise as a result of clonal expansion of myeloid hematopoietic precursors in the bone marrow. Not only are circulating leukemia cells (also called blasts) seen in the peripheral blood, but granulocytopenia, anemia, and thrombocytopenia are also common as proliferating leukemia cells interfere with normal hematopoiesis.

Mutations in signaling genes such as the class III tyrosine kinase receptor gene *FLT3* confer a proliferative advantage through the RAS–RAF, JAK–STAT, and PI3K–AKT signaling pathways (upper left box). Mutations in myeloid transcription factors such as *RUNX1* and transcription factor fusions by chromosomal rearrangements, such as $t(8;21)(q22;q22);RUNX1-RUNX1T1$, lead to transcriptional deregulation and impaired hematopoietic differentiation (center left box).

In the nucleophosmin (NPM1) gene, encoding a multifunctional nucleocytoplasmic shuttling protein, mutations result in the aberrant cytoplasmic localization of NPM1 and NPM1-interacting proteins (lower left box). Mutations of spliceosome-complex genes such as SRSF2, SF3B1, U2AF1, and ZRSR2 are involved in deregulated RNA processing (lower right box). Cohesin-complex gene mutations, such as STAG2 and RAD21, might impair accurate chromosome segregation and transcriptional regulation (center middle box). Mutations of genes involved in the epigenetic homeostasis of cells, such as mutations of ASXL1 and EZH2, lead to deregulation of chromatin modification (e.g., methylation of histones H3 and H2A on lysine residues K79, K27, and K119, respectively), as well as KMT2A–MLLT3 gene fusion, which can impair other methyltransferases such as DOT1L (DOT1-like histone H3K79 methyltransferase) (center right box). DNMT3A and TET2 mutations, as well as IDH1 and IDH2 mutations, acting through the 2-hydroxyglutarate (2HG) oncometabolite production, can lead to the deregulation of DNA methylation (hmC denotes 5-hydroxymethylcytosine, and mC 5-methylcytosine) (upper right box). In tumor-suppressor genes such as TP53, mutations can lead to transcriptional deregulation and impaired degradation through the mouse double minute 2 homologue (MDM2) and the phosphatase and tensin homologue (PTEN) (upper middle box)

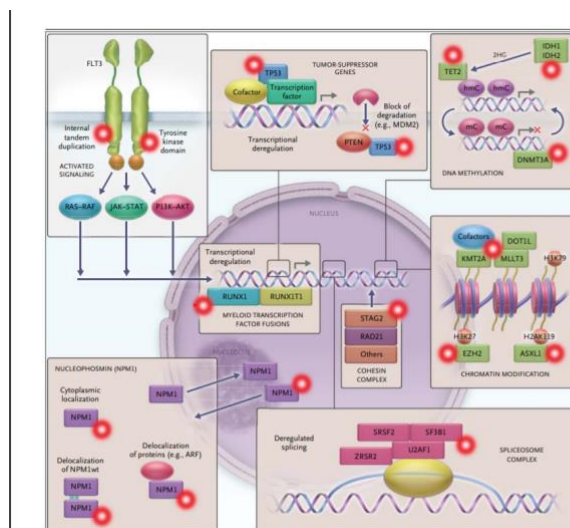


Figure 1.3 shows the genes involved in AML
(acute myeloid leukaemia. N Engl J Med 2013)

TABLE 1.1 CLASSIFICATIONS BASED ON RISK FACTORS

RISK PROFILE	CHROMOSOME	GENES INVOLVED
Favorable	t(8;21)(q22;q22), (p13.1q22) or t(16;16)(p13.1;q22)	RUNX1-RUNX1T1, ; RUNX1-RUNX1T1
Intermediate-I†	t(8;21)(q22;q22	Mutated NPM1 and FLT3-ITD (normal karyotype) Wild-type NPM1 and FLT3-ITD (normal karyotype) Wild- type NPM1 without FLT3-ITD
Intermediate-II	t(9;11)(p22;q23);	MLLT3-KMT2A
Adverse	(q21q26.2) or t(3;3)(q21;q26.2); 6;9)(p23;q34	GATA2-MECOM, DEK- NUP214, KMT2A

First, cases of AML with mutated CEBPA are now restricted to cases with biallelic CEBPA mutations.⁴ Second, the molecular designation of inv (3)(q21q26.2) or t(3;3)(q21;q26.2) has been changed to GATA2-MECOM (EVI1).³ Finally, for MLL, the official gene symbol KMT2A (lysine [K]-specific methyltransferase 2A) has been adopted.

This category includes all cases of AML with a normal karyotype except for those included in the favorable subgroup.

Adequate numbers of most abnormalities have not been studied to draw firm conclusions regarding their prognostic significance. (reported by Döhner et al.¹).

1.3 TABLE SHOWING THE SUBTYPES OF AML

M0 – Undifferentiated acute myeloblastic leukemia
M1 – Acute myeloblastic leukemia with minimal maturation
M2 - Acute myeloblastic leukemia with maturation
M3 –Acute promyelocytic leukemia with maturation
M4 –Acute myelomonocytic leukemia with eosinophilia
M5 –Acute erythroid leukemia
M6 –Acute megakaryoblastic leukemia

Subtypes M0 through M5 all start in immature forms of white blood cells. M6 AML starts in very immature forms of red blood cells, while M7 AML starts in immature forms of cells that make platelets.

1.4 TABLE SHOWING FREQUENCY OF MUTATION IN AML

MUTATED GENES	FREQUENCY
NPM1	25–35
RUNX1	5–15
DNMT3A	18–22
FLT3-ITD	20
ASXL1	5 --17
TET2	7--25

CHAPTER -2

LITERATURE SURVEY

Acute myeloid leukemia (AML) is a cancer of the blood and bone marrow. AML is the most common type of acute leukemia in American adults and the average age of an AML patient is 67. In 2010, 12,330 people were estimated to have been diagnosed with AML in the United States, with an estimated 8,950 deaths. Acute myeloid leukemia (AML) and play a key role in the pathogenesis and evolution of the disease, particularly in the absence of AML-associated recurrent cytogenetic abnormalities. DNMT3A, FLT3, and NPM1 mutations are among the most common genomic alterations. DNMT3A encodes a DNA methyltransferase that catalyzes the addition of a methyl group to cytosine residues in CpG islands resulting in reduced expression of downstream genes. DNMT3A mutations, most commonly DNMT3A are significantly enriched in patients with intermediate risk cytogenetics. AML patients in which DNMT3A, FLT3, and NPM1 mutations coexist at a higher frequency than would be expected for a chance occurrence. The clinical implications of FLT3 and NPM1 mutations in de novo AML are largely well established, the impact of DNMT3A mutations on NPM1 and/or FLT3 mutations remains poorly understood especially in patients with AML. (J Hematol Oncol 7, 74 (2014).

FLT3 inhibitors that could be alternatively used to circumvent drug resistance. The crosstalk between major downstream signaling pathways, such as PI3K/PTEN/Akt/mTOR, RAS/Raf/MEK/ERK, and Jak/STAT, can be exploited for therapeutic purposes by targeting key signaling molecules with selective inhibitors. Cytokine receptor–derived signals are predominantly mediated through the **JAK/STAT pathway**, which plays a major role in maintaining hematopoietic balance. STAT transcription factors play a critical role in mediating the effects of cytokines on myeloid cells. As STAT target genes control key processes such as survival, proliferation and self-renewal, it is not surprising that constitutive activation of STATs, particularly STAT3 and STAT5, are common events in many myeloid tumours. STATs are activated both by mutant tyrosine kinases as well as other pathogenic events, and continued activation of STATs is common in the setting of resistance to kinase inhibitors.

STAT-mediated gene expression is involved in many normal physiological processes, such as proliferation, survival and differentiation. In hematopoietic cells, cytokines whose effects are transduced by STATs play a central role in regulating the production of red blood cells, platelets and the full spectrum of white blood cells. Several small molecule FLT3-tyrosine kinase inhibitors (TKI) have been developed and examined in AML patients as single agents or in combination with chemotherapy. The induction of cytotoxicity by FLT3 inhibitors is closely correlated with deactivation of STAT5, while resistance to FLT3 inhibition is associated with persistent activation of STAT5. STAT pathway as a means of cancer therapy will be enhanced, thus contributing to a more personalized approach of treating patients. (Darnell JE. Jr STATs and gene regulation. *Science*).

Acute myeloid leukemia (AML) is a life-threatening hematological malignancy. Traditional diagnosis of AML depends on invasive bone marrow biopsies. To recognize the metabolic characteristics related with AML and search for early non-invasive biomarkers of AML, in this work we applied ultra-high-performance liquid chromatography quadrupole time-of-flight tandem mass spectrometry (UHPLC-Q-TOFMS)-based metabolome method were used. We observed AML-related metabolic differences mainly involved in alanine, aspartate and glutamate metabolism; d-Glutamine and d-glutamate metabolism; tryptophan metabolism; taurine and hypotaurine metabolism; and phenylalanine metabolism as well as fatty acid metabolism. (Dong Wang- pub 2019 Jan 1). Our study focused on the metabolomics pathway analysis were beta alanine metabolic pathways has showing the highest impact of P value < 0.05 by the utilisation of metaboanalyst software were used to separate the metabolites based on the p value significance and the suitable drug target to inhibit the FLT3, DNMT3A, NPM1 mutations. Gilteritinib is an oral, potent, selective FLT3 inhibitor with single-agent activity in relapsed or refractory FLT3-mutated AML. Gilteritinib is a new, highly selective, oral FLT3 inhibitor with activity against both FLT3 mutation subtypes (ITD and TKD) and weak activity against c-Kit. Gilteritinib also inhibits the tyrosine kinase AXL, which is implicated in FLT3 inhibitor resistance.

In a phase 1–2 study, single-agent gilteritinib therapy resulted in sustained inhibition of FLT3 autophosphorylation and, at doses of at least 80 mg per day, led to 41% of the patients with relapsed or refractory FLT3-mutated AML having a composite complete remission. Gilteritinib resulted in significantly longer survival and higher percentages of patients with remission than salvage chemotherapy among patients with relapsed or refractory *FLT3*-mutated AML patients. (N Engl J Med 2019).

CHAPTER -3

AIM AND SCOPE

The main objective of the study is to find the suitable metabolite to inhibit the AML based on pathway analysis.

1. To Identify a suitable Metabolite.
2. To find a pathway analysis for the targeted metabolite compound.
3. To obtain a network of interaction of AML.
4. To identify a suitable drug for the selected metabolite compound.

CHAPTER - 4

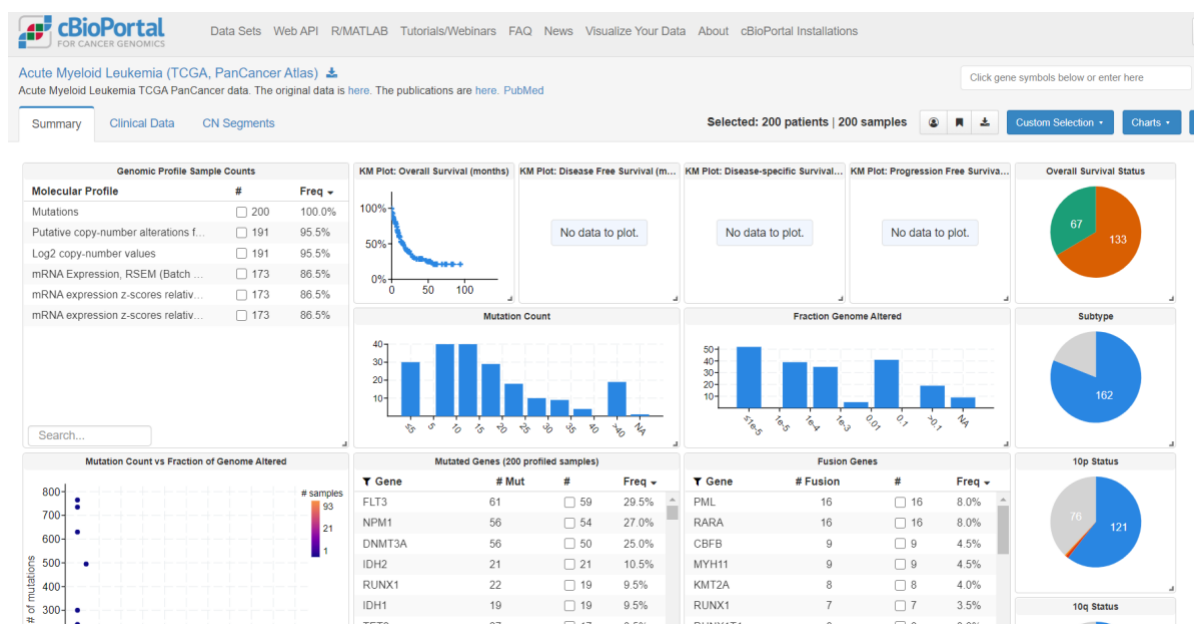
MATERIALS AND METHODS USED

4.1 Use of bioinformatics data set collection:

For gene expression studies analysis collection of datasets is very important to understand the biological system. Though dataset collection is a huge task and it contain a very large number of datasets and it need the computational support.

4.2 CBIOPORTAL

The cBioPortal for Cancer Genomics was originally developed at Memorial Sloan Kettering Cancer Center (MSK). The **CBIO Portal** is an exploratory analysis tool for exploring large-scale cancer genomic data sets that hosts data from large consortium efforts, like TCGA and TARGET. Currently providing access to data from more than 5,000 tumor samples from 20 cancer studies. The cBio Cancer Genomics Portal significantly lowers the barriers between complex genomic data and cancer researchers who want rapid, intuitive, and high-quality access to molecular profiles and clinical attributes from large-scale cancer genomics projects and empowers researchers to translate these rich data sets into biologic insights and clinical applications.



In our study this database is used to collect the dataset of AML were 191 genes were analysed.

4.3 LEUKAEMIA GENE ATLAS (LGA)

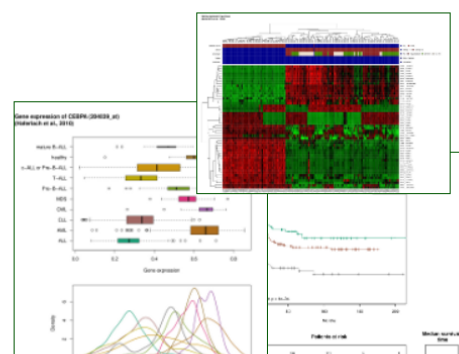
Leukaemia is exceptionally well studied at the molecular level and a wealth of high-throughput data has been published. But further utilization of these data by researchers is severely hampered by the lack of accessible integrative tools for viewing and analysis.

The Leukaemia Gene Atlas (LGA) is a public platform designed to support research and analysis of diverse genomic **data** published in the field of **leukaemia**. With respect to **leukaemia** research, the LGA is an unique resource with comprehensive search and browse functions. GA provides a novel comprehensive easy and fast access to published genome wide data sets in haematopoiesis and haematological malignancies. It provides extensive analysis and visualization tools for various types of molecular data. Currently, its database contains data from more than 5,800 leukemia and hematopoiesis samples generated by microarray gene expression, DNA methylation, SNP and next generation sequencing analyses. The LGA allows easy retrieval of large published data sets and thus helps to avoid redundant investigations. The Leukemia Gene Atlas (LGA), a novel online bioinformatics tool that provides comprehensive, easy and fast access to

published genome wide data sets in hematopoiesis and hematological malignancies. In the following section we describe the architecture of the LGA paying attention to the database and the data stored therein. The primary purpose of the LGA is to support translational research and biomarker discovery in haematology.



The Leukemia Gene Atlas (LGA) is a public platform that supports the research and analysis of molecular data of leukemias. In addition to the molecular data the LGA stores the corresponding clinical and laboratory data. The LGA supports comprehensive queries of data, the visualization and analysis as well as download and upload. Currently, there are *gene expression, methylation, sequencing data and genotypes of over 6000 samples* in the LGA. To analyze data the user may choose between a set of tools: *cluster analyses, statistical tests, data distribution plots and the Integrative Genomics Viewer (IGV)*.



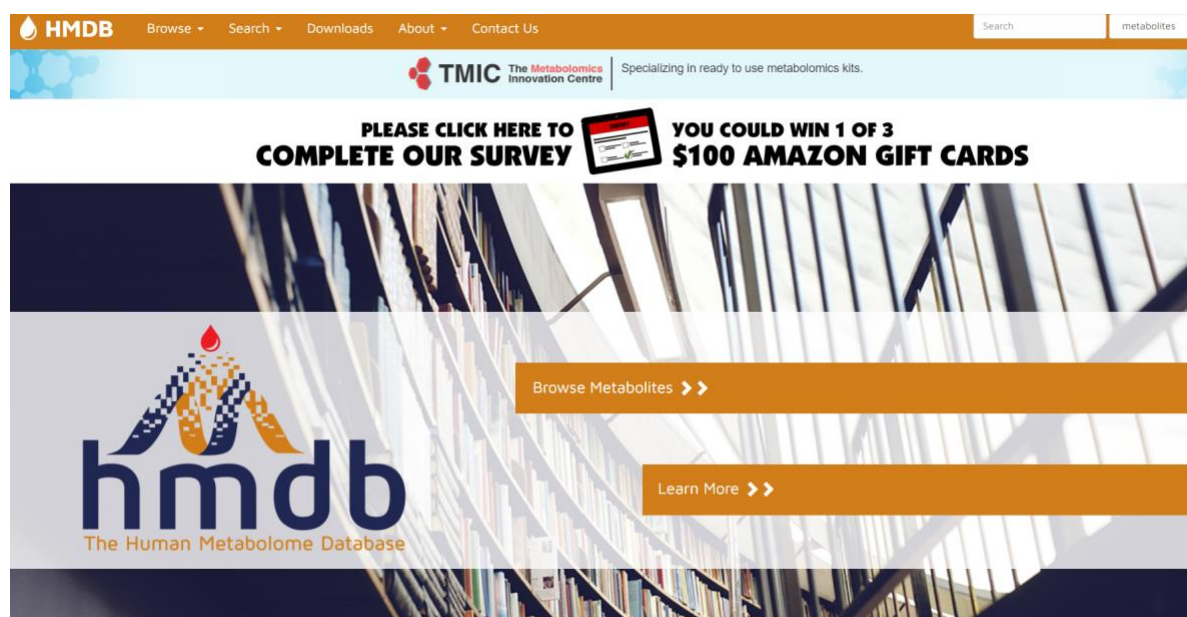
Leukemia gene atlas is one of the oncology databases & is used in our study to collect the gene list in this platform.

4.4 **HMDB-HUMAN MOLECULAR INTERACTOME: [HMDB]**

The Human Metabolism **Database (HMDB)** is freely available as electronic **database** containing detailed information about small molecule metabolites found in the human body. It is intended to be used for applications in metabolomics, clinical chemistry and bio-marker discovery. The database is designed to contain or link three kinds of data: 1) chemical data, 2) clinical data, and 3) molecular biology/biochemistry data. The database contains 114,304 metabolite entries including both water-soluble and lipid soluble metabolites as well as metabolites that would be regarded as either abundant ($> 1 \mu\text{M}$) or relatively rare ($< 1 \text{ nM}$). Additionally, 5,702 protein sequences are linked to these metabolite entries. Each MetaboCard entry contains 130 data fields with 2/3 of the information being devoted to chemical/clinical data and the other 1/3 devoted to enzymatic or biochemical data. Many data fields are hyperlinked to other databases (KEGG, PubChem, MetaCyc, ChEBI, PDB, UniProt, and GenBank) and a variety of

structure and pathway viewing applets. The HMDB database supports extensive text, sequence, chemical structure and relational query searches. Four additional databases, DrugBank, T3DB, SMPDB and FooDB are also part of the HMDB suite of databases.

Drug Bank contains equivalent information on ~2280 drug and drug metabolites, T3DB contains information on ~3670 common toxins and environmental pollutants, SMPDB contains pathway diagrams for ~25,000 human metabolic and disease pathways.



Human Metabolome database is used to collect the compound metabolites for our particular study of interest in leukemia cancer. These metabolite compounds have been selected to study the pathway interaction and the eleven metabolite compounds have been collected from this database.

4.5 KEGG DATABASE:

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. The KEGG database has been in development by Kanehisa Laboratories since 1995, and is now a prominent reference knowledge base for integration and interpretation of large-scale molecular data sets generated by genome sequencing and other high-throughput experimental technologies.



KEGG PATHWAY Database

Wiring diagrams of molecular interactions, reactions and relations

[Menu](#) [PATHWAY](#) [BRITE](#) [MODULE](#) [KO](#) [GENES](#) [LIGAND](#) [NETWORK](#) [DISEASE](#) [DRUG](#) [DBGET](#)

Select prefix

map

Organism

Enter keywords

AML

Go

[Help](#)

[\[New pathway maps | Update history \]](#)

Pathway Maps

KEGG PATHWAY is a collection of manually drawn [pathway maps](#) representing our knowledge of the molecular interaction, reaction and relation networks for:

1. Metabolism

Global/overview Carbohydrate Energy Lipid Nucleotide Amino acid Other amino Glycan
Cofactor/vitamin Terpenoid/PK Other secondary metabolite Xenobiotics Chemical structure

2. Genetic Information Processing

3. Environmental Information Processing

4. Cellular Processes

5. Organismal Systems

6. Human Diseases

7. Drug Development

KEGG PATHWAY is the reference database for pathway mapping in [KEGG Mapper](#).

Pathway Identifiers

Each pathway map is identified by the combination of 2-4 letter prefix code and 5 digit number (see [KEGG Identifier](#)). The prefix has the following meaning:

map manually drawn reference pathway
ko reference pathway highlighting KOs
ec reference metabolic pathway highlighting EC numbers
rn reference metabolic pathway highlighting reactions
<org> organism-specific pathway generated by converting KOs to gene identifiers

and the numbers starting with the following:

011 global map (lines linked to KOs)
012 overview map (lines linked to KOs)
010 chemical structure map (no KO expansion)
07 drug structure map (no KO expansion)
other regular map (boxes linked to KOs)

are used for different types of maps.

The KEGG database is used in our study to find the pathway analysis in acute myeloid leukaemia.

The most unique data object in KEGG is the molecular networks -- molecular interaction, reaction and relation networks representing systemic functions of the cell and the organism. Experimental knowledge on such systemic functions is captured from literature and organized in the following forms:

- Pathway map - in KEGG PATHWAY
- Brite hierarchy and table - in KEGG BRITE
- Membership (logical expression) - in KEGG MODULE
- Membership (simple list) - in KEGG DISEASE

4.6 **METABOANALYST**

MetaboAnalyst is the web interface data analysis software and it was created by Wishart Research Group. It was the implementations based on R packages for batch processing. It includes HMDB, Drug bank, Toxin and toxin targeted database, small molecule pathway database. MetaboAnalyst generates the PDF for each analysis and that can be visualize in a graphical format.

It includes the following set of parameters such as

- Raw data processing
- Statistical analysis
- Ms peaks to pathway analysis
- Enrichment analysis
- Integrated pathway analysis
- Network import module
- Time series or two factor design
- Power analysis module.
- The data should be uploaded in the .CDF, xML formats.
- The data shows the pathways affected by the metabolites.

MetaboAnalyst 5.0 - user-friendly, streamlined metabolomics data analysis

News & Updates

- **MetaboAnalyst 5.0** is launched! 🎉🎉
- Please help complete the [User Survey on behalf of TMIC](#). You could win \$100 Amazon Gift Card! 🎉🎉
- Added support for merging technical replicates with different mathematical models and QC measures (02/18/2021); 🎉🎉
- Updated KEGG IDs in internal compound databases (02/09/2021); 🎉🎉
- Enhanced name handling to better accommodate special characters in compound/lipid names (02/09/2021); 🎉🎉
- Added [seven tutorials](#) introducing new features in MetaboAnalyst 5.0 (01/15/2021); 🎉🎉
- Users can perform meta-analysis of global metabolomics data using [Functional Meta-analysis](#) module as illustrated in our [COVID-19 paper](#) (01/09/2021); 🎉🎉
- Updated the KEGG global metabolic map for better metabolome coverage (01/08/2021); 🎉🎉
- Users can now enter **Study IDs** to analyze data deposited in the [Metabolomics Workbench](#) (12/20/2020); 🎉🎉
- Users can [Start New Journey](#) from the [Download page](#) to explore other modules at the end of each session (12/14/2020); 🎉🎉
- Expanded compound database for lipids (>150,000 lipids) together with a smart compound name matching algorithm (12/10/2020); 🎉🎉
- Upgraded all main interactive plots using ChartJS; enhanced Enrichment module and Pathway Analysis module (12/07/2020); 🎉🎉

[Click here to start](#)

The metaboanalyst software is used in our study for the enrichment pathway analysis and found that beta alanine metabolism has the highest impact.

4.7 CYTOSCAPE

Cytoscape is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data. Additional features are available as plugins. Plugins are available for network and molecular profiling analyses, new layouts, additional file format support and connection with databases and searching in large networks. Cytoscape was originally created at the Institute of Systems Biology in Seattle in 2002. Now, it is developed by an international consortium of open source developers. While Cytoscape is most used for biological research applications, it is agnostic in terms of usage. Cytoscape can be used to visualize and analyze network graphs of any kind involving nodes and edges.

4.8 METSCAPE:

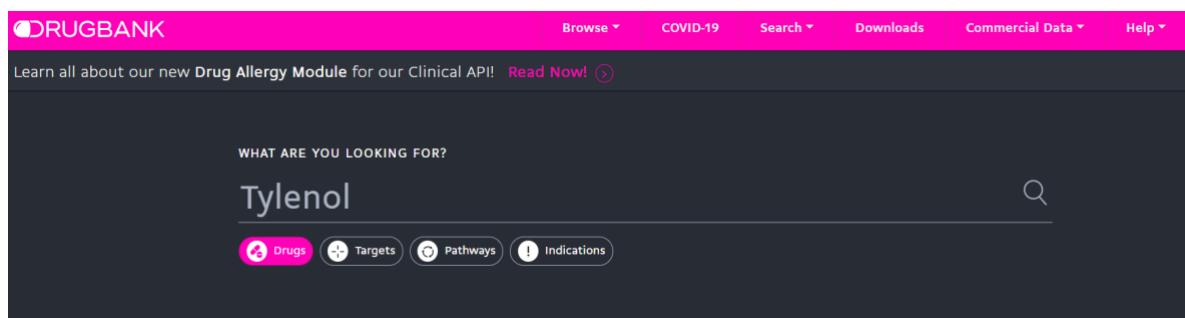
The MetScape 3 App for Cytoscape provides a bioinformatics framework for the visualization and interpretation of metabolomic and expression profiling data in the context of human metabolism. It allows users to build and analyze networks of genes and compounds, identify enriched pathways from expression profiling data, and visualize changes in metabolite data. A Cytoscape plugin used to visualize and analyse metabolomic data. It uses the data from Edinburgh Human Metabolic Network reconstruction. MetScape facilitates visualization of complex networks and displays related information about reactions, enzymes, and pathways.

4.9 DRUG DATA BANK.

The original DrugBank project was funded by the the Canadian Institutes of Health Research (award #111062), Alberta Innovates - Health Solutions, and by The Metabolomics Innovation Centre (TMIC), a nationally funded research and core facility that supports a wide range of cutting-edge metabolomic studies.

DrugBank Online is a comprehensive, free-to-access, online database containing information on drugs and drug targets. DrugBank Online is widely used by the drug industry, medicinal chemists, pharmacists, physicians, and the general public.

The figure below shows the home page of drug bank database which is used to find a suitable drug target molecule.



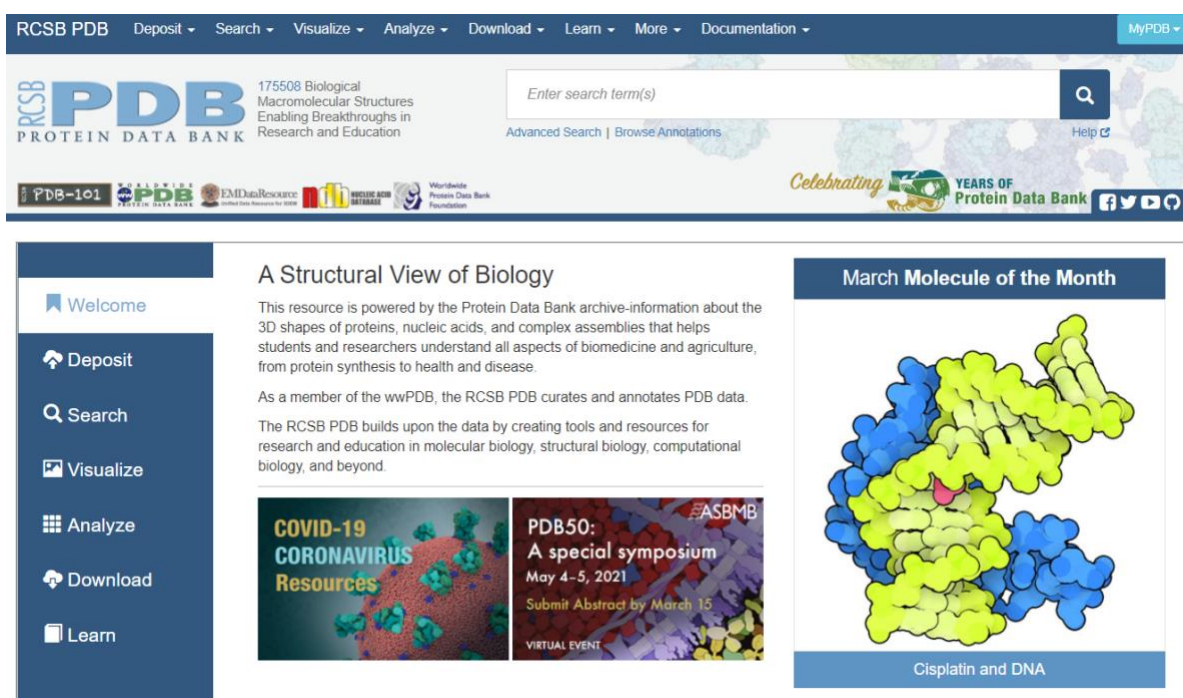
5.0 PDB [Protein Data Bank].

The Protein Data Bank (PDB) is a database for the three-dimensional structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography, NMR spectroscopy, or, increasingly, cryo-electron microscopy, and submitted by biologists and biochemists from around the world, are freely accessible on the Internet via the websites of its member organisations (PDB [PDB RCSB] and BMRB). The PDB is overseen by an organization called the Worldwide Protein Data Bank.

The PDB is a key in areas of structural biology, such as structural genomics. Most major scientific journals and some funding agencies now require scientists to submit their structure data to the PDB. Many other databases use protein structures deposited in the PDB. For example, SCOP and CATH classify protein structures, while PDBsum provides a graphic overview of PDB entries using information from other sources, such as Gene ontology.

The structure files may be viewed using one of several free and open source computer programs, including Jmol, Pymol, VMD, and Rasmol. Other non-free, shareware programs include ICM-Browser, MDL Chime, UCSF Chimera, Swiss-PDB Viewer.

The figure below shows the PDB database which is used to find a 3D structure for the targeted drug molecule.



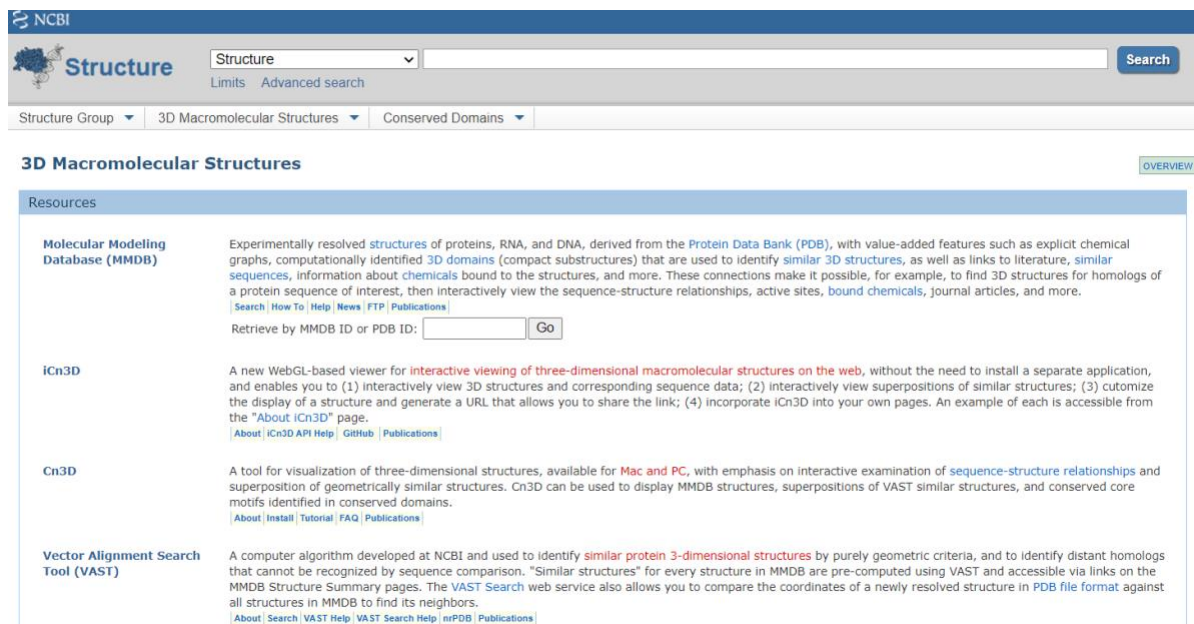
5.1 MMDB- [Molecular Modelling Database]

The molecular modelling database is a 3D structural database it is experimentally resolved structures of proteins, RNA, and DNA, derived from the Protein Data Bank (PDB), with value-added features such as explicit chemical graphs, computationally identified 3D domains (compact substructures) that are used to identify similar 3D structures, as well as links to literature, similar sequences, information about chemicals bound to the structures, and more. It has various resources include

iCn3D is a WebGL-based viewer for interactive viewing of three-dimensional macromolecular structures and chemicals on the web, without the need to install a separate application. It can be accessed from the "molecular graphic" that appears on the structure summary page for any record in the Molecular Modelling Database (MMDB). users can now view the electrostatic potential map for any subset of 3D structures within 30,000 atoms. The potential is calculated using the DelPhi program by solving linear Poisson-Boltzmann equation. Users can show the potential on surface or show equipotential map.

The potential map shows the effect of charges on molecular interactions qualitatively.

VAST and VAST+ are applied on every protein in the Molecular Modelling Database (MMDB) during MMDB data processing in order to identify similar 3D structures.



The screenshot displays the NCBI Structure website interface. At the top, there is a navigation bar with the NCBI logo and the word "Structure". Below this, a search bar is visible with the text "Structure" and a "Search" button. The main content area is titled "3D Macromolecular Structures" and includes a sub-section "Resources". This section lists several tools and databases:

- Molecular Modeling Database (MMDB)**: Described as a database of experimentally resolved structures of proteins, RNA, and DNA, derived from the Protein Data Bank (PDB). It includes value-added features like explicit chemical graphs and computationally identified 3D domains. A search bar is provided with the text "Retrieve by MMDB ID or PDB ID:" and a "Go" button.
- iCn3D**: A new WebGL-based viewer for interactive viewing of three-dimensional macromolecular structures on the web, without the need to install a separate application. It enables users to (1) interactively view 3D structures and corresponding sequence data; (2) interactively view superpositions of similar structures; (3) customize the display of a structure and generate a URL that allows you to share the link; (4) incorporate iCn3D into your own pages. An example of each is accessible from the "About iCn3D" page.
- Cn3D**: A tool for visualization of three-dimensional structures, available for Mac and PC, with emphasis on interactive examination of sequence-structure relationships and superposition of geometrically similar structures. Cn3D can be used to display MMDB structures, superpositions of VAST similar structures, and conserved core motifs identified in conserved domains.
- Vector Alignment Search Tool (VAST)**: A computer algorithm developed at NCBI and used to identify similar protein 3-dimensional structures by purely geometric criteria, and to identify distant homologs that cannot be recognized by sequence comparison. "Similar structures" for every structure in MMDB are pre-computed using VAST and accessible via links on the MMDB Structure Summary pages. The VAST Search web service also allows you to compare the coordinates of a newly resolved structure in PDB file format against all structures in MMDB to find its neighbors.

This database is used for the secondary structure visualisation of the molecule and it will obtain the structure from PDB. It will give detailed information about the domain regions in the secondary structure prediction.

CHAPTER -5

RESULT AND DISCUSSION

5.1 COLLECTION OF CANCER DATA FROM CBIO PORTAL:

From the data collected from the cBio- portal as shown the cancer type of AML and the sample were collected from 200 patients and in that 191 with the mutation and the molecular profiling were analysed in that mutation, copy number alterations, mRNA expression were analysed and its showing with 100% frequency .In the most total number of mutations were identified as FLT3, NPM1, DNMT3A these genes have high number of mutation.

From the molecular profile of these out of 200 samples 95 samples have more number of mutated genes in it.

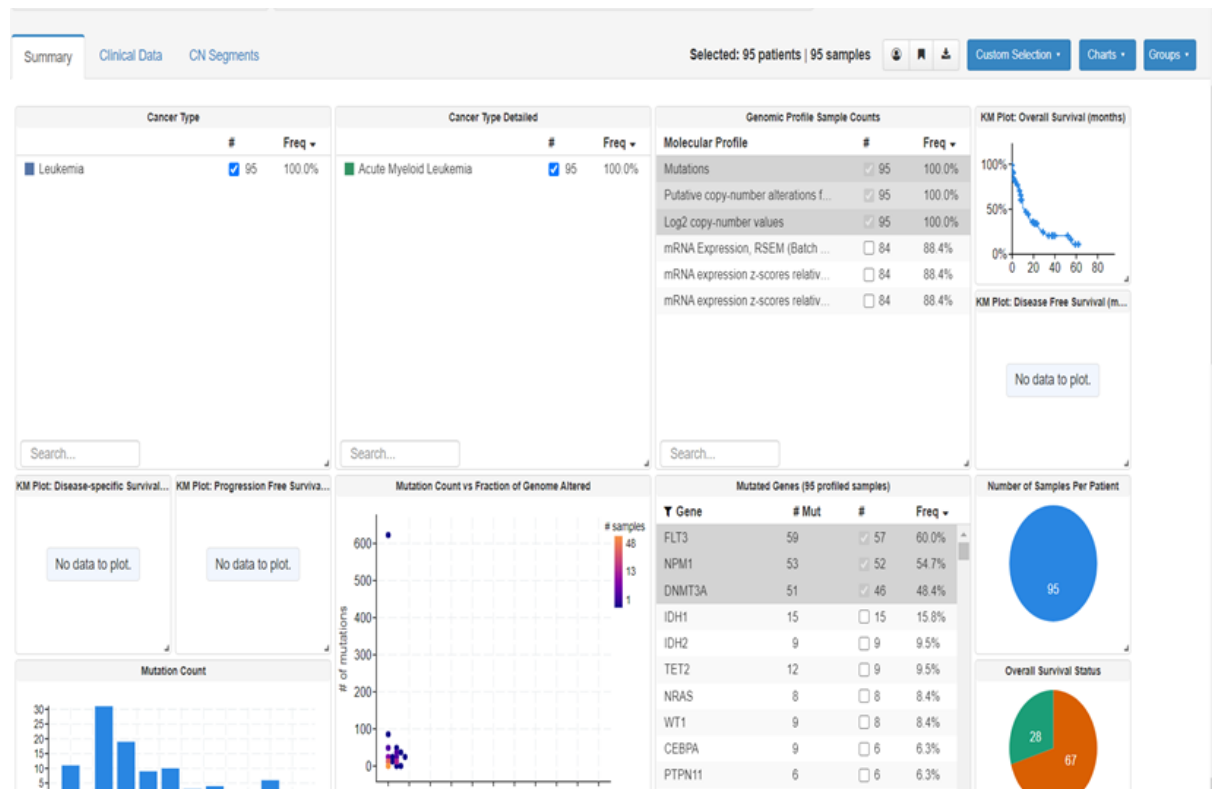


Figure 5.1 showing the sample genes from cbiportal

5.1.2 CNA –COPY NUMBER ALTERATION:

It means changes in the chromosomal arrangements of genes which may lead to mutation. In this the genes include the copy number variations include alterations and deep deletion where the genes where highly altered include ERG, MNX1 , RHEB, XRCC2 which include oncogenes ,tumour suppressor genes and it estimate the 50% survival, Which includes amplification,deep-deletion,translocation etc.

Data table shows how the copy number alteration is sorted:

Gene	MutSig(Q-value)	Mutation	Total number of samples.	Profiled Samples
FLT3	59	57	191	29.8%, Yes
NPM1	53	52	191	27.2, Yes
DNMT3A	54	46	191	24.2, yes
IDH2	20	20	191	10.5, yes
RUNX1	22	19	191	9.9, yes
IDH1	19	19	191	9.4, yes
NRAS	16	16	191	8.4, yes
TP53	20	16	191	8.4, yes
TET2	20	16	191	8.4, yes
TTN	14	13	191	6.8, no
WT1	14	13	191	6.8, yes
BPIFC	12	11	191	5.8, yes
CEBPA	17	12	191	6.3, yes

Table 5.2 shows the excel visualisation of mutated gene list.

MUTATED GENES DATA TABLE:

Here from the mutated genes were categorised and analyse for further understanding and visualisation and it is sorted order is based from highest to lowest category.

5.2 HMDB-[HUMAN METABOLOME DATABASE]

Collecting the list of metabolites which is involved in the FLT3,DNMT3A,NPM1 mutated genes by using metabolome database is the open source database for the metabolomic data .We have collected nearly 11 metabolite compounds which is affecting the leukemic pathways.

The listed metabolites and their reference KEGG id where shown in the below table.

HMDB0000001	1-Methylhistidine	HMDB0000001	92105	C01152
HMDB0000002	1,3-Diaminopropane	HMDB0000002	428	C00986
HMDB0000005	2-Ketobutyric acid	HMDB0000005	58	C00109
HMDB0000010	2-Methoxyestrone	HMDB0000010	440624	C05299
HMDB0000011	(R)-3-Hydroxybutyric acid	HMDB0000011	92135	C01089
HMDB0000012	Deoxyuridine	HMDB0000012	13712	C00526
HMDB0000014	Deoxycytidine	HMDB0000014	13711	C00881
HMDB0000015	Cortexolone	HMDB0000015	53477676	C05488
HMDB0000016	Deoxycorticosterone	HMDB0000016	6166	C03205
HMDB0000017	4-Pyridoxic acid	HMDB0000017	6723	C00847

HMDB00000019	Alpha-ketoisovaleric acid	HMDB00000019	49	C00141
--------------	---------------------------	--------------	----	--------

Table 5.2 shows the list of metabolites obtain from HMDB.

5.3 METABOANALYST:

MetaboAnalyst is a set of online tools for metabolomic data analysis and interpretation, created by members of the Wishart Research Group at the University of Alberta. It was first released in May 2009 and version 2.0 was released in January 2012. MetaboAnalyst provides a variety of analysis methods that have been tailored for metabolomic data. These methods include metabolomic data processing, normalization, multivariate statistical analysis, and data annotation. MetaboAnalyst supports a wide variety of data input types commonly generated by metabolomic studies including GC/LC-MS raw spectra, MS/NMR peak lists, NMR/MS peak intensity table, NMR/MS spectral bins, and metabolite concentrations.

MetaboAnalyst has four modules:

1. Data processing
2. Statistical analysis (one-factor, two-factor, and time-series data)
3. Functional enrichment analysis
4. Metabolic pathway analysis.

Valine, leucine and isoleucine biosynthesis	0.00125	2.902	0.10528	0.10528	0	KEGG SMP
Steroid hormone biosynthesis	0.01907	1.7197	1	0.7356	0.04143	KEGG SMP
Pyrimidine metabolism	0.02941	1.5316	1	0.7356	0.06234	KEGG SMP
Synthesis and degradation of ketone bodies	0.03503	1.4556	1	0.7356	0	KEGG SMP
Vitamin B6 metabolism	0.06224	1.2059	1	1	0	KEGG SMP
Butanoate metabolism	0.10176	0.99242	1	1	0	KEGG SMP
Histidine metabolism	0.1082	0.96578	1	1	0	KEGG SMP
Pantothenate and CoA biosynthesis	0.12726	0.89532	1	1	0	KEGG SMP
beta-Alanine metabolism	0.13976	0.85463	1	1	0	KEGG SMP
Propanoate metabolism	0.15209	0.81789	1	1	0.04061	KEGG SMP
Glycine, serine and threonine metabolism	0.2114	0.67489	1	1	0	KEGG SMP

Cysteine and methionine metabolism	0.2114	0.67489	1	1	0.05983	KEGG SMP SMP
Valine, leucine and isoleucine degradation	0.25064	0.60094	1	1	0.01084	KEGG SMP

Table 5.3 shows the metabolites showing the highest impact.

The T is the total number of compounds in the pathway, hits are the actually matched data, raw p value is the original p value, Holm p value is the p value adjusted by Holm-Bonfemoni method, and the FDR p is the p value adjusted using false discovery rate, impact is the pathway impact value calculated from the pathway topology analysis.

GRAPHICAL VIEW OF METABLOME PATHWAY

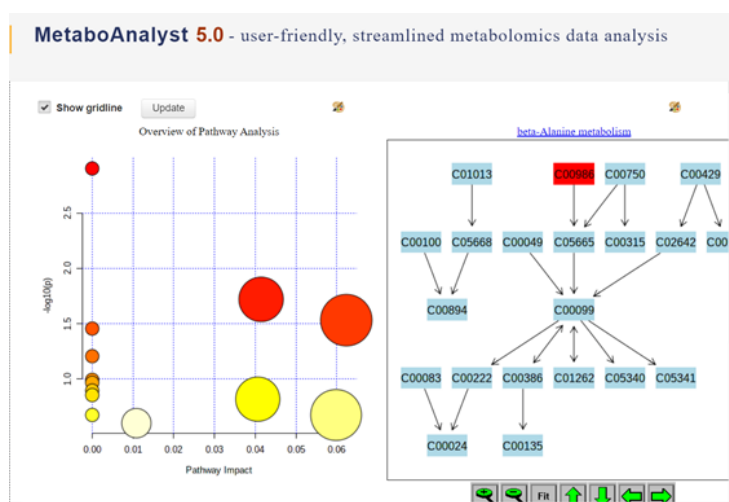


Figure 5.4 shows the graphical visualisation of pathway analysis.

The pathway analysis shown here where each dot represent different pathways and the red colour shows the maximum impact in this beta alanine has the highest impact. The results from the pathway analysis are represented graphically as well as in detailed table. An interactive visualization system was implemented to

facilitate data. The graphical output contains three levels metabolome view, pathway view, and the compound view. Here only the metabolome view is shown.

5.4 **CYTOSCAPE**

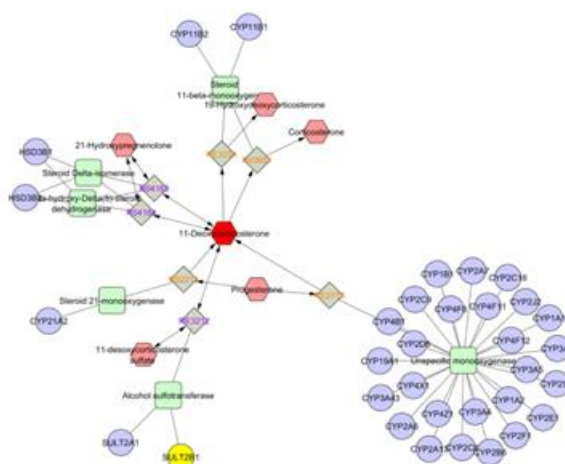
Cytoscape is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data. Additional features are available as plugins. Plugins are available for network and molecular profiling analyses, new layouts, additional file format support and connection with databases and searching in large networks. Cytoscape was originally created at the Institute of Systems Biology in Seattle in 2002. Now, it is developed by an international consortium of open source developers. While Cytoscape is most commonly used for biological research applications, it is agnostic in terms of usage. Cytoscape can be used to visualize and analyze network graphs of any kind involving nodes and edges.

All the metabolites identified are need to be analysed in cytoscape along with the reference KEGG ids.

1,3-Diaminopropane	C00986
2-Ketobutyric acid	C00109
2-Methoxyestrone	C05299
(R)-3-Hydroxybutyric acid	C01089
Deoxyuridine	C00526
Deoxycytidine	C00881
Cortexolone	C05488
Deoxycorticosterone	C03205
4-Pyridoxic acid	C00847
Alpha-ketoisovaleric acid	C00141

The metabolites are cytoscape where only 7 were recognised by the metscape and these metabolites are analysed and built network with reaction, enzymes, genes, and compound with these metabolites.

Building a network in the metscape with these metabolites and running with a compound reaction-enzyme gene analysis. In this analysis every compounds, reaction were related to the query metabolites and the connectivity between them are shown.



Here the list shows below were the enzymes which are interacting in the metabolomic network and there were 39 enzymes were included in the network. The table shows the node data of all reaction, compounds, genes, and enzymes involved in the network. In this there are about 38 enzymes, 67 compounds, 39 reactions, and 9 pathways are involved but it mainly affecting the tyrosine kinase JAK STAT pathway.

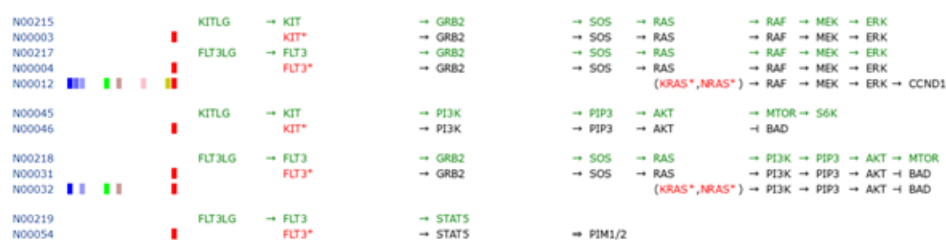
category	canonical name	Enzyme.ecnum	enzyme name
Enzyme	Steroid Delta-isomerase	5.3.3.1	Steroid Delta-isomerase
Enzyme	Cystathionine gamma-lyase	4.4.1.1	Cystathionine gamma-lyase
Enzyme	dUTP diphosphatase	3.6.1.23	dUTP diphosphatase
Enzyme	Nucleoside-triphosphate diphosphatase	3.6.1.19	Nucleoside-triphosphate diphosphatase
Enzyme	Cytidine deaminase	3.5.4.5	Cytidine deaminase
Enzyme	dCMP deaminase	3.5.4.12	dCMP deaminase
Enzyme	5'-nucleotidase	3.1.3.5	5'-nucleotidase
Enzyme	Alcohol sulfotransferase	2.8.2.2	Alcohol sulfotransferase
Enzyme	dTMP kinase	2.7.4.9	dTMP kinase
Enzyme	Deoxycytidine kinase	2.7.1.74	Deoxycytidine kinase
Enzyme	Thymidine kinase	2.7.1.21	Thymidine kinase
Enzyme	Alanine transaminase	2.6.1.2	Alanine transaminase
Enzyme	Deoxyhypusine synthase	2.5.1.46	Deoxyhypusine synthase
Enzyme	Glucuronosyltransferase	2.4.1.17	Glucuronosyltransferase
Enzyme	Acetolactate synthase	2.2.1.6	Acetolactate synthase
Enzyme	Catechol O-methyltransferase	2.1.1.6	Catechol O-methyltransferase
Enzyme	Thymidylate synthase	2.1.1.45	Thymidylate synthase
Enzyme	Amine oxidase (copper-containing)	1.4.3.6	Amine oxidase (copper-containing)
Enzyme	Aldehyde oxidase	1.2.3.1	Aldehyde oxidase
Enzyme	Steroid 21-monooxygenase	1.14.99.10	Steroid 21-monooxygenase

Table 5.5 shows the list of enzymes involved in the metabolomic network.

5.5.1 **PATHWAY ANALYSIS USING KEGG:**

5.5.2 NETWORK VARIATION IN AML:

The network analysis of FLT3 it's a tyrosine kinase 3 ligand in acute myeloid leukemia and its pathway related to JAK STAK pathway and the elements in the pathway include duplication or the mutation activated by FLT3 to JAK STAT pathway.



Here it shows the FLT3 mutated gene variation of acute myeloid leukemia here the red colour indicate it is highly mutated one.

JAK STAT PATHWAY REPRESENTATION:

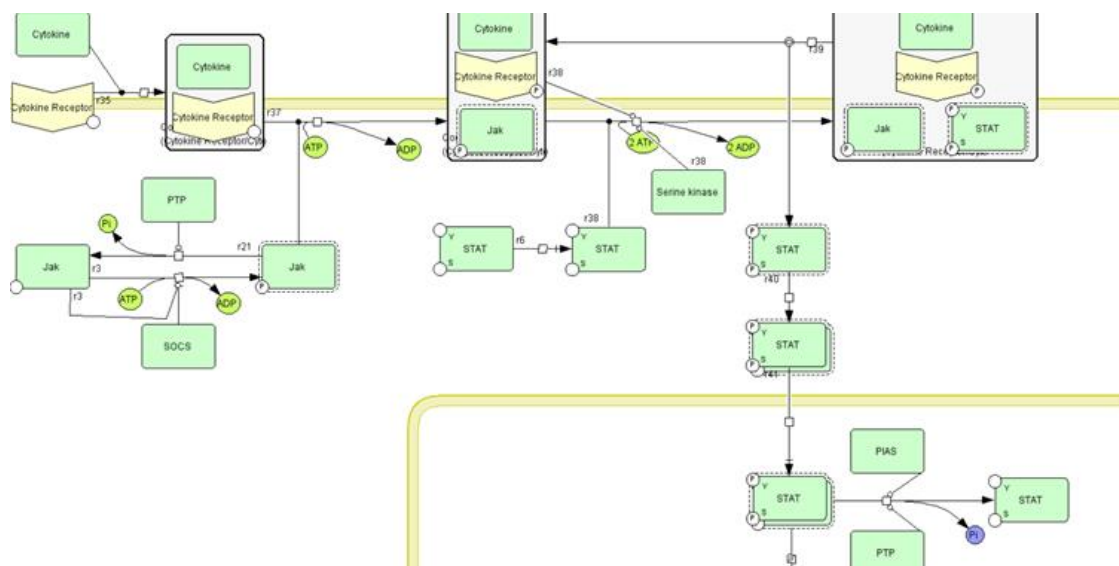


Figure 5.5.3 shows the JAK STAT pathway in AML. (KEGG)

5.6 DRUG BANK

DrugBank Online is a comprehensive, free-to-access, online database containing information on drugs and drug targets. DrugBank Online is widely used by the drug industry, medicinal chemists, pharmacists, physicians, and the general public.

The suitable drug compounds were collected from the drug bank database and the particular Gilteritinib were selected because it is particularly inhibiting the FLT3 mutation.

Gilteritinib, also known as ASP2215, is a small molecule part of the FLT3 tyrosine kinase inhibitors that presented a greater selectivity and potency when compared with other agents from this group.¹ It is a pyrazinecarboxamide derivative that showed high selectivity to FLT3 preventing the c-Kit -driven myelosuppression observed in other therapies.

DRUG MOLECULES	ACCESION NUMBER
Dasatinib	DB01254
Cedazuridine	DB15694
Nilotinib	DB04868
Methylprednisone	DB12952
Ivosidenib	DB14568
Busulfan	DB01008
Gilteritinib	DB12141

MECHANISM OF ACTION:

Gilteritinib is a potent selective inhibitor of both of the mutations, internal tandem duplication (ITD) and tyrosine kinase domain (TKD), of the FLT3 receptor.² In the same note, gilteritinib also inhibits AXL and ALK tyrosine kinases.³ FLT3 and AXL are molecules involved in the growth of cancer cells.

The interest in FLT3 trans membrane tyrosine kinases was raised when studies reported that approximately 30% of the patients with acute myeloid leukaemia.

Presented a mutational activated isoform.¹ As well, the mutation ITD is associated with poor patient outcomes while the mutation TKD produces a resistance mechanism to FLT3 tyrosine kinase inhibitors and the AXL tyrosine kinase tends to produce a resistance mechanism to chemotherapies.

Tyrosine kinase inhibitors FLT3 inhibitors

- ❖ First-generation Midostaurin NCT00651261;
- ❖ NCT01477606 Sunitinib NCT00783653
- ❖ Sorafenib NCT00373373, NCT00893373
- ❖ Second-generation
- ❖ Quizartinib NCT02039726
- ❖ Crenolanib NCT01657682;
- ❖ NCT02298166 ASP2215 NCT02014558
- ❖ KIT inhibitors Dasatinib NCT02013648
- ❖ NCT01238211 Midostaurin NCT01830361.

TARGET IDENTIFICATION

The particular drug has four targets and is mainly based on the tyrosine kinase protein receptor FLT3 it act as the inhibitor. It is highly bonded with the plasma protein region.

METABOLIC ROLE:

Gilteritinib is primarily metabolized in the liver by the activity of CYP3A4. Its metabolism is driven by reactions of N-DE alkylation and oxidation which forms the metabolite M17, M16 and M10 from the plasma concentration.

PDB – STRUCTURE FOR THE FLT3 RECEPTOR:

The structure of the protein has been shown below where the structure is obtained from the X- RAY diffraction at a resolution of 2.10 it is a tyrosine kinase protein and the gene FLT3 has affected in the chromosome of location 13 on the locus of 13q12.

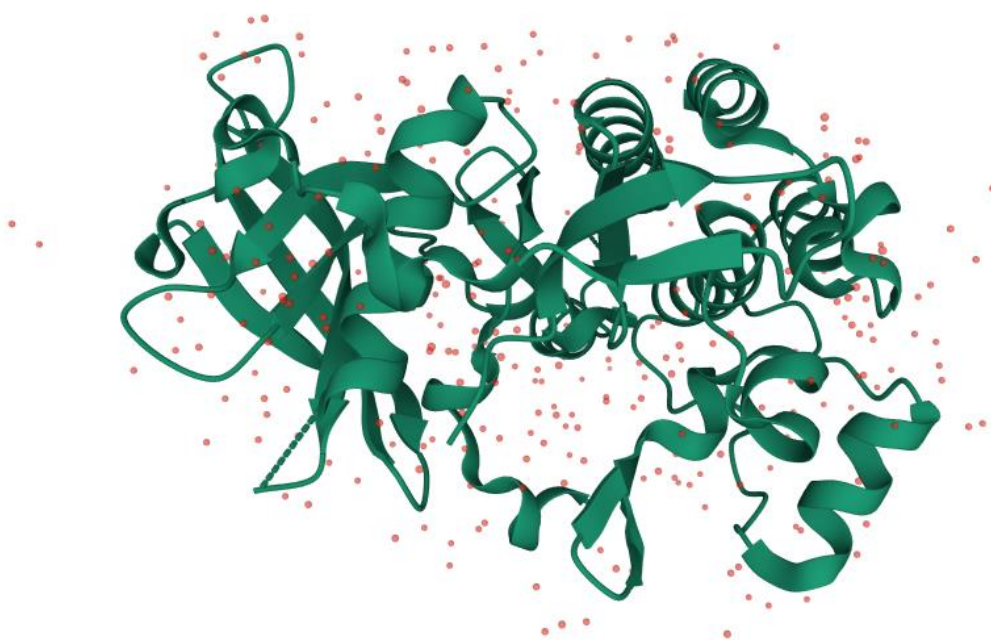


Figure 5.6 shows the 3D structure of the drug molecule.(PDB)

5.6.1 MMDB [Macro molecular structural database]

It is a 3D molecular structural macromolecular interaction. It gives the 3D interaction with the molecule with the molecule Experimentally resolved structures of proteins, RNA, and DNA, derived from the Protein Data Bank (PDB), with value-added features such as explicit chemical graphs, computationally identified 3D domains (compact substructures) that are used to identify similar 3D structures, as well as links to literature, similar sequences, information about chemicals bound to the structures are represented.

STRUCTURAL REPRESENTATION OF THE MOLECULE:

The macromolecular structure was obtained from PDB is 6QJR it's a tyrosine kinase receptor for FLT3 gene and is taken from the source organism of homo sapiens. Its the structure of FLT3 in complex with gilteritinib.

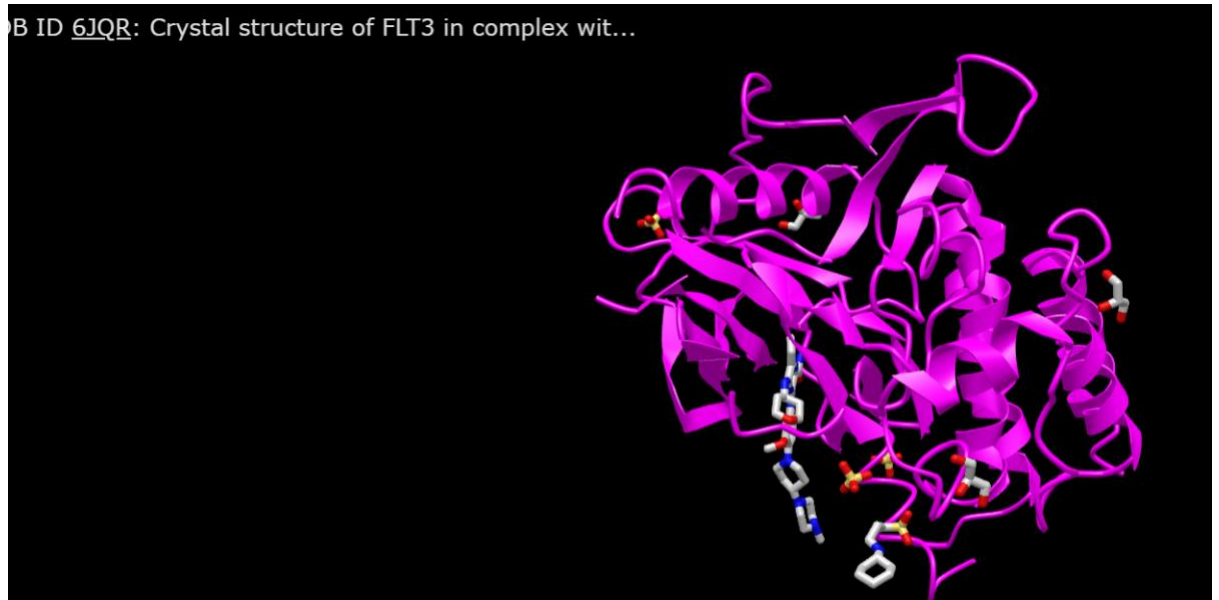


Figure 5.5.3 shows the macromolecular structure if FLT3 complex with gilteritinib.

Here the structure has the 2-domain region and the white colour indicate the side chains and the red one is the mutation occurred and the coiled region has the secondary structure in it.

Therapeutic effects of FLT3 inhibitors have been reported in acute myeloid leukemia (AML) with constitutively activating FLT3 mutations, including internal tandem duplication (ITD) and point mutation, which are found in approximately one-third of AML patients. One of the critical issues of treatment with FLT3 inhibitors in FLT3-mutated AML is drug resistance. FLT3 ligand (FL) represents a mechanism of resistance to FLT3 inhibitors, including quizartinib, midostaurin, and sorafenib, in AML cells harboring both wild-type and mutant FLT3.

The effect of FL on the efficacy of gilteritinib, a FLT3 inhibitor, in AML-derived cells in vitro and in mice. In contrast to other FLT3 inhibitors, FL stimulation had little

effect on growth inhibition or apoptosis induction by gilteritinib. In the FLT3 signaling analyses, gilteritinib inhibited FLT3wt and FLT3-ITD to a similar degree in HEK293 and Ba/F3 cells, and similarly suppressed FLT3 downstream signaling molecules (including ERK1/2 and STAT5) in both the presence and absence of FL in MOLM-13 cells.

Co-crystal structure analysis showed that gilteritinib bound to the ATP-binding pocket of FLT3. These results suggest that gilteritinib has therapeutic potential in FLT3-mutated AML patients with FL overexpression.

CHAPTER -6

SUMMARY AND CONCLUSION.

Based on the results obtained upon analysing the genes which are involved in acute myeloid leukemia where some of the genes which is showing high mutation frequency rate such as FLT3, DNMT3A, NPM1 where the copy number alterations are involved in this gene set. To inhibit the gene function specific metabolite compounds are collected from HMDB and pathway analysis of the particular mutated genes and some seven pathways were shown based on the p value

significance the metabolic pathways are selected where it is mostly affected with the beta alanine metabolic pathway and to know the interaction between the nodes and edges it was shown that it has nearly 345 nodes and 249 edges were connected with one another and the compound enzyme reaction were analysed for the gene list in that nearly 39 enzymes are mostly included in the metabolic pathway profiling.

Metscape is used for the visualisation of the metabolic networks where FLT3 genes have the high mutation frequency and the mutation is located in the chromosome 13th location and the respective pathways are visualised.

To inhibit the mutation frequency Gilteritinib it act as the suitable antimetabolite which will inhibit the FLT3 mutation by interacting with the tyrosine kinase receptors and it will mainly act on the plasma membrane and clofarabine is the antimetabolite is also used in the treatment of leukemia and it prevents cells from making DNA and RNA by interfering with the synthesis of nucleic acids, thus stopping the growth of cancer cells. and it has the FDA approved drug molecule (PDB) and it has to be taken via injection.

Gilteritinib resulted in significantly longer survival and higher percentages of patients with remission than salvage chemotherapy among patients with relapsed or refractory FLT3-mutated AML.

patients with acute myeloid leukemia (AML) whose disease is refractory to, or relapses after, induction chemotherapy have a dismal prognosis with standard chemotherapy .FMS-like tyrosine kinase 3 (FLT3), a cytokine receptor tyrosine kinase that is expressed in early hematopoietic stem and progenitor cells, regulates their proliferation and differentiation. FLT3-activating mutations occur in approximately 30% of patients with AML. Gilteritinib is a new, highly selective, oral

FLT3 inhibitor with activity against both FLT3 mutation subtypes (ITD and TKD) and weak activity against c-Kit. Gilteritinib also inhibits the tyrosine kinase AXL, which is implicated in FLT3 inhibitor resistance. In a phase 1–2 study, single-agent gilteritinib therapy resulted in sustained inhibition of FLT3 autophosphorylation and, at doses of at least 80 mg per day, led to 41% of the patients with relapsed or refractory FLT3-mutated AML having a composite complete remission (complete remission with or without normal hematologic recovery); a starting dose of 120 mg per day was recommended for further study. To investigate the clinical benefit of gilteritinib in the treatment of relapsed or refractory FLT3-mutated AML, we conducted a multicenter, randomized trial comparing gilteritinib with conventional salvage chemotherapy regimens.

Finally, the metabolic pathways involved in cancer are only a small information which is being analysed and further designing of drug molecules and their interactions have to be studied.

CONCLUSION:

The cancer genomic data was collected from the cbio portal platform. The reference study was about the acute myeloid leukaemia cancer which is taken from the tcga atlas 2003) were selected and the genes sets were analysed and identified the highest frequency of mutation are identified in FLT3,DNMT3,NPM1.The gene list were downloaded and visualise in excel sheet, and identified the CNA copy number alterations. Then for this mutated gene list suitable compound metabolites were identified from the HMDB database followed by pathway analysis by using metaboanalyst and their reference ids are obtained from the kegg database. It was found that beta alanine metabolic pathway has showing higher impact, then

cytoscape is used to visualise the network view by using metscape which is a plugin of cytoscape to analysis the enzyme, gene, interactions and the pathway can be visualised. Finally, drug bank data was used to find a suitable drug mechanism to inhibit the AML cancer with already existing data.

6.2 SUMMARY:

The project ACUTE MYELOID LEUKEMIA by using METABOLOMICS APPROACH is to find a suitable drug target by involving various pathway analysis by using Metaboanalyst software and the survival status of the patients were analysed by using cbio portal. From the analysis by using metaboanalyst JAK STAT pathway is affected in FLT3 mutated in patient samples were observed and followed by the network analysis of the pathway shows beta analine metabolism has the highest impact on FLT3 mutation. To inhibit the effect of FLT3 gene proliferation a gilteritnib a antimetabolite has the used treat AML patients. Gilteritinib resulted in superior

overall survival and percentages of remission as compared with salvage chemotherapy in AML.

REFERENCES:

1. Roboz GJ, Rosenblat T, Arellano M, et al. International randomized phase III study of elacytarabine versus investigator choice in patients with relapsed/refractory acute myeloid leukemia.
2. Small D, Levenstein M, Kim E, et al. STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells.
3. Shen Y, Zhu YM, Fan X, et al. Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood*. 2011;118(20):5593–5603. [PubMed]
4. Schlenk RF, Döhner K, Krauter J, et al. German-Austrian Acute Myeloid Leukemia Study Group. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *N Engl J Med*. 2008;358(18):1909–1918. [PubMed]
5. Mrózek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. *Blood Rev*. 2004;18(2):115–136. [PubMed]
6. O'Donnell MR, Abboud CN, Altman J, et al. Acute myeloid leukemia. *J Natl Compr Canc Netw*. 2012; 10(8):984–1021. [PubMed]

7. Wald D, Vermaat JM, Peleg G, Tse W. Genetic abnormalities in acute myelogenous leukemia with normal cytogenetics. *Curr Hematol Malig Rep.* 2008;3(2):83-88
8. . Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med.* 2012;366(12):1079–1089. [PMC [PubMed]]
9. Marcucci G, Haferlach T, Döhner H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *J Clin Oncol.* 2011
10. Burnett AK, Russell NH, Hills RK, et al. Optimization of chemotherapy for younger patients with acute myeloid leukemia: results of the medical research council AML15 trial. *J Clin Oncol* 2013.
11. Litzow MR, Tarima S, Pérez WS, et al. Allogeneic transplantation for therapy-related myelodysplastic syndrome and acute myeloid leukemia. *Blood* 2010.
12. Hills RK, Castaigne S, Appelbaum FR, et al. Acute myeloid leukaemia: a metaanalysis of individual patient data from randomised controlled trials. *Lancet Oncol* 2014.
13. Rohatiner AZ, Gregory WM, Bassan R, et al.: Short-term therapy for acute myelogenous leukemia. *J Clin Oncol* 6 (2): 218-26.
14. Goldstone AH, Burnett AK, Wheatley K, et al.: Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical Research Council AML11 trial. *Blood* 98 (5): 1302-11, 2001.
15. Schlenk RF, Weber D, Fiedler W, et al.: Midostaurin added to chemotherapy and continued single-agent maintenance therapy in acute myeloid leukemia with FLT3-ITD. *Blood* 133 (8): 840-851, 2019.
16. Mayer RJ, Davis RB, Schiffer CA, et al.: Intensive postremission chemotherapy in adults with acute myeloid leukemia. Cancer and Leukemia Group B. *N Engl J Med* 331 (14): 896-903, 1994.
17. Ravandi F, Kantarjian H, Faderl S, et al. Outcome of patients with FLT3-mutated acute myeloid leukemia in first relapse. *Leuk Res* 2010.
18. Champlin R, Gajewski J, Nimer S, et al.: Postremission chemotherapy for adults with acute myelogenous leukemia: improved survival with high-dose cytarabine and daunorubicin consolidation treatment. *J Clin Oncol* 8 (7): 1199-206, 1990.
19. Small D, Levenstein M, Kim E, et al. STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells. *Proc Natl Acad Sci U S A* 1994
20. Nakao M, Yokota S, Iwai T, et al. Internal tandem duplication of the flt3 gene found in acute myeloid leukemia. *Leukemia* 1996.
21. Kiyoi H, Towatari M, Yokota S, et al. Internal tandem duplication of the FLT3 gene is a novel modality of elongation mutation which causes constitutive activation of the product. *Leukaemia*.
22. Chevallier P, Labopin M, Turlure P, et al. A new Leukemia Prognostic Scoring System for refractory/relapsed adult acute myelogenous leukaemia patients: a GOELAMS study. *Leukemia* 2011.

23. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med* 2016.
24. Smith BD, Levis M, Beran M, et al. Single-agent CEP-701, a novel FLT3 inhibitor, shows biologic and clinical activity in patients with relapsed or refractory acute myeloid leukemia.
25. Cortes J, Perl AE, Döhner H, et al. Quizartinib, an FLT3 inhibitor, as monotherapy in patients with relapsed or refractory acute myeloid leukaemia.
26. Galanis A, Ma H, Rajkhowa T, et al. Crenolanib is a potent inhibitor of FLT3 with activity against resistance-conferring point mutants. *Blood* 2014.
27. Galanis A, Levis M. Inhibition of c-Kit by tyrosine kinase inhibitors. *Haematological* 2015.
28. Mori M, Kaneko N, Ueno Y, et al. Gilteritinib, a FLT3/AXL inhibitor, shows antileukemic activity in mouse models of FLT3 mutated acute myeloid leukemia.
29. Weisberg E, Boulton C, Kelly LM, Manley P, Fabbro D, Meyer T, Gilliland DG, Griffin JD (2002) Inhibition of mutant FLT3 receptors in leukemia cells by the small molecule tyrosine kinase inhibitor PKC412. *Cancer Cell*.
30. Grunwald MR, Levis MJ (2015) FLT3 tyrosine kinase inhibition as a paradigm for targeted drug development in acute myeloid leukemia.