Roaming Around AML

Reza Asakereh[a]

[a] *CE Department, Sharif University of Technology*

Abstract

AML is a lethal type of blood cancer originating from bone marrow. This article is devoted to analyze microarray data and derive a number of differentially expressed genes, consequently to come up with a handful of information. An immune system interaction with AML in the form of both cause and effect is once again verified.

1. Introduction

This article conducts an unplanned investigation of bunch of data and research results about AML decease. As indicated by Parkin et al., the fifth most affected site in body by cancer is the circulatory system, with about a hundred thousands estimated cases in total of a million cancer in the US, 2002. AML is one of the most suffered and fatal one with almost 35% of the cases resulting into death. Hence, piling up some results in this area will accelerate more experienced scientists toward a working solution. However, this field is not abandoned, and has been worked on extensively. Therefore, they are considered as rich sources of result validation. The process of the investigation begins with a beginner level differential expression analysis resulting in some gene lists with significant gene expression alterations. These lists are used as the input of the next pipeline component, gene enrichment. Extracting some pieces of information in this step will be the basis of some fact finding endeavors.

2. Data Preparation and Validation

The data can be accessed via GSE48558 accession number. Primarily, to verify if data is normalized and the comparisons are mathematically meaningful, a box plot was employed, resulting in Figure 1.

As the data needed no preprocessing, the process of validation was embarked on, which indicated a great clustering and reliability of B Cell data in comparison with other data subgroups. Figure 2 to Figure 10 would depict the claim.

In addition to the clustering property, separability was assessed using two dimension reduction mechanisms, PCA and t-SNE.

In spite of data not being separable and cluster-able satisfyingly while looking as a whole, no portion of it was eliminated and the data analysis was done even once with the whole data, supporting the results acquired in subgroups analyses.

Email address: rasakereh@ce.sharif.edu (Reza Asakereh)

July 13, 2019

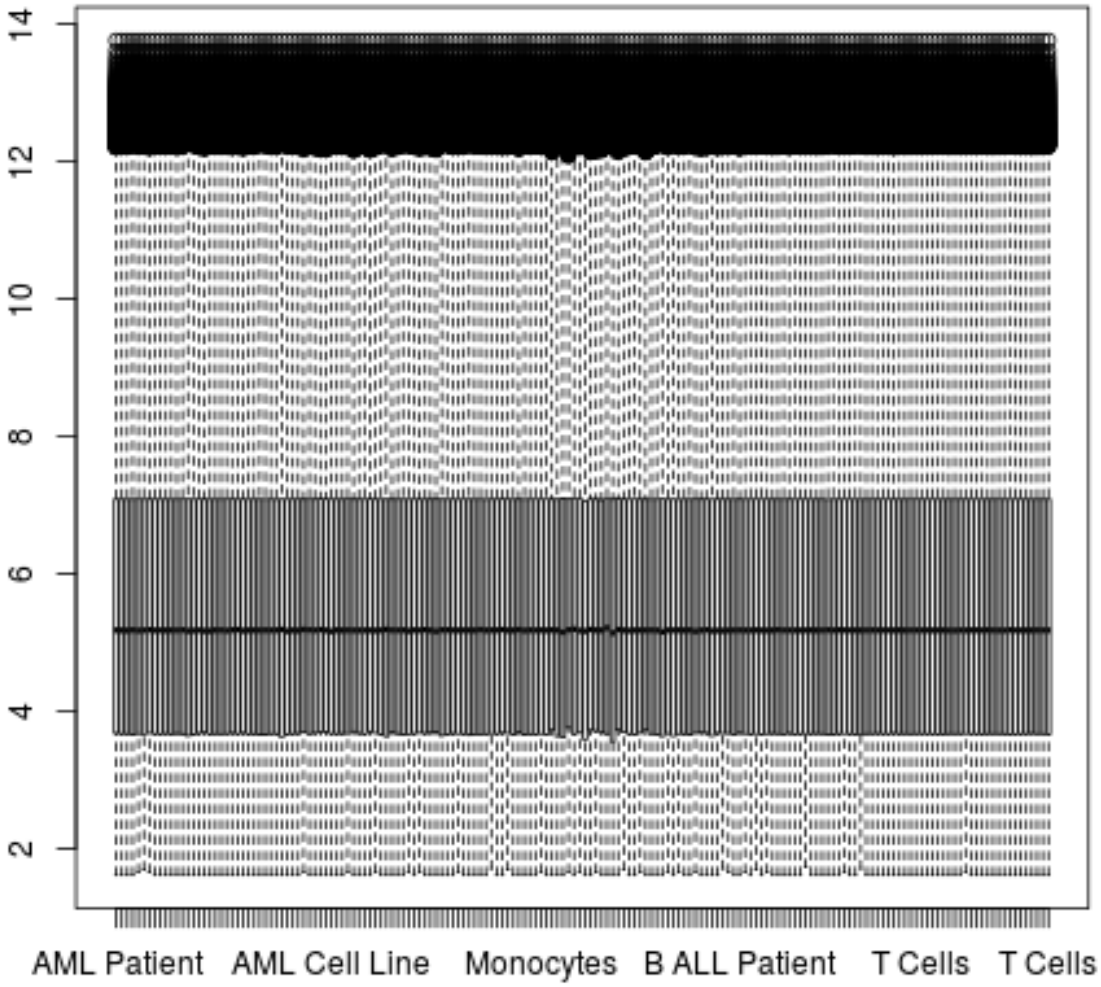

Figure 1: The data needs no additional normalization.

3. Differential Expression Analysis

This article has employed Cramer-Morales et al. (2013) data set and Limma tools (Ritchie et al. (2015).) The data set primarily consists of 170 samples having their expressions measured by 32321 probes and classified in two categories, having self-explanatory labels, "Normal" and "Leukemia." These samples were first examined in three prospectives, respectively independent of the cell source, cells of type T, cells of type B(Table 1.)

After having the data in this format, a basic visualization was employed to form an

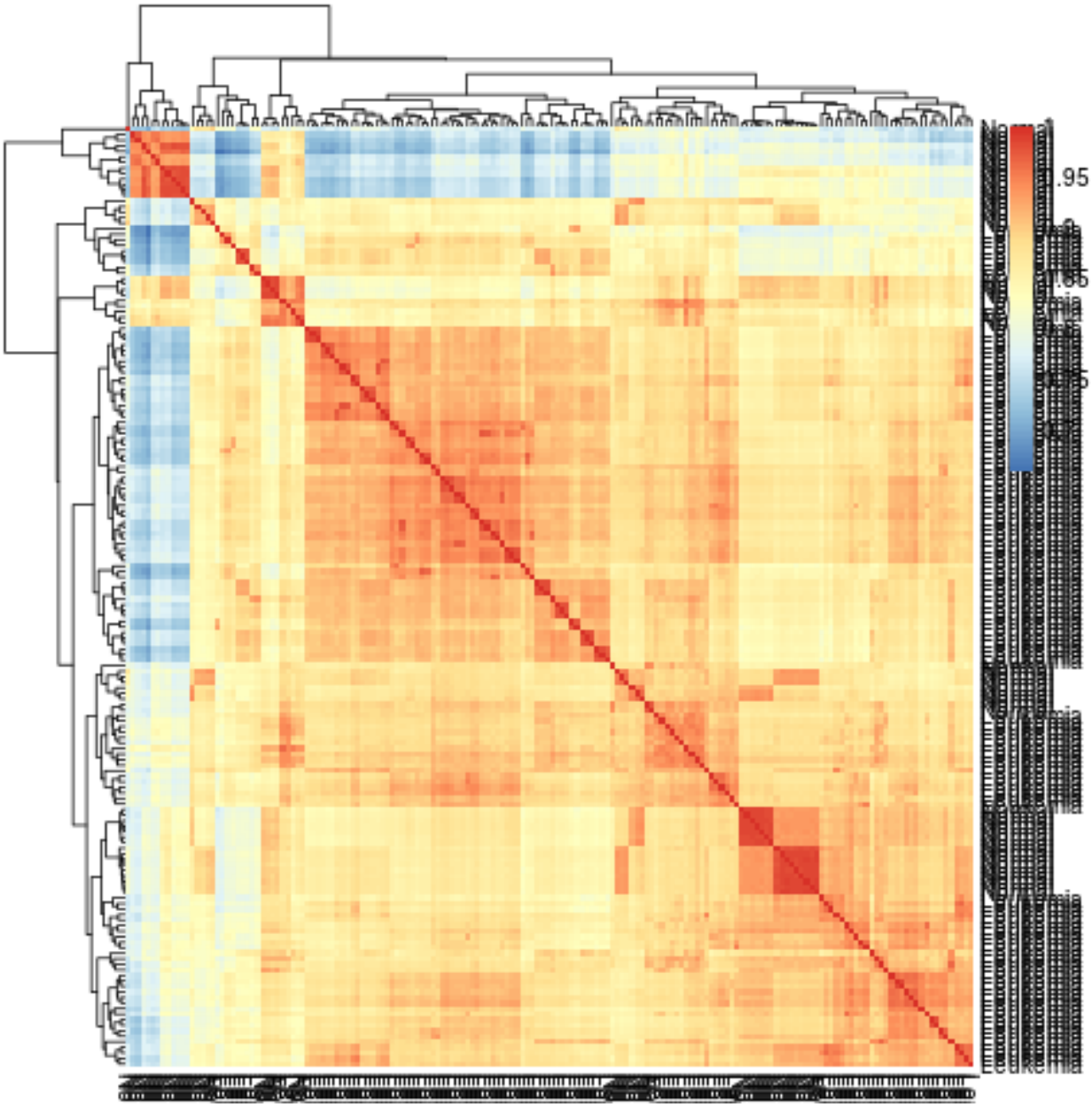

Figure 2: The heat map for the mixture of cells, showing poor clustering. Clustering score was not calculated due to the god damn deadline.

July 13, 2019

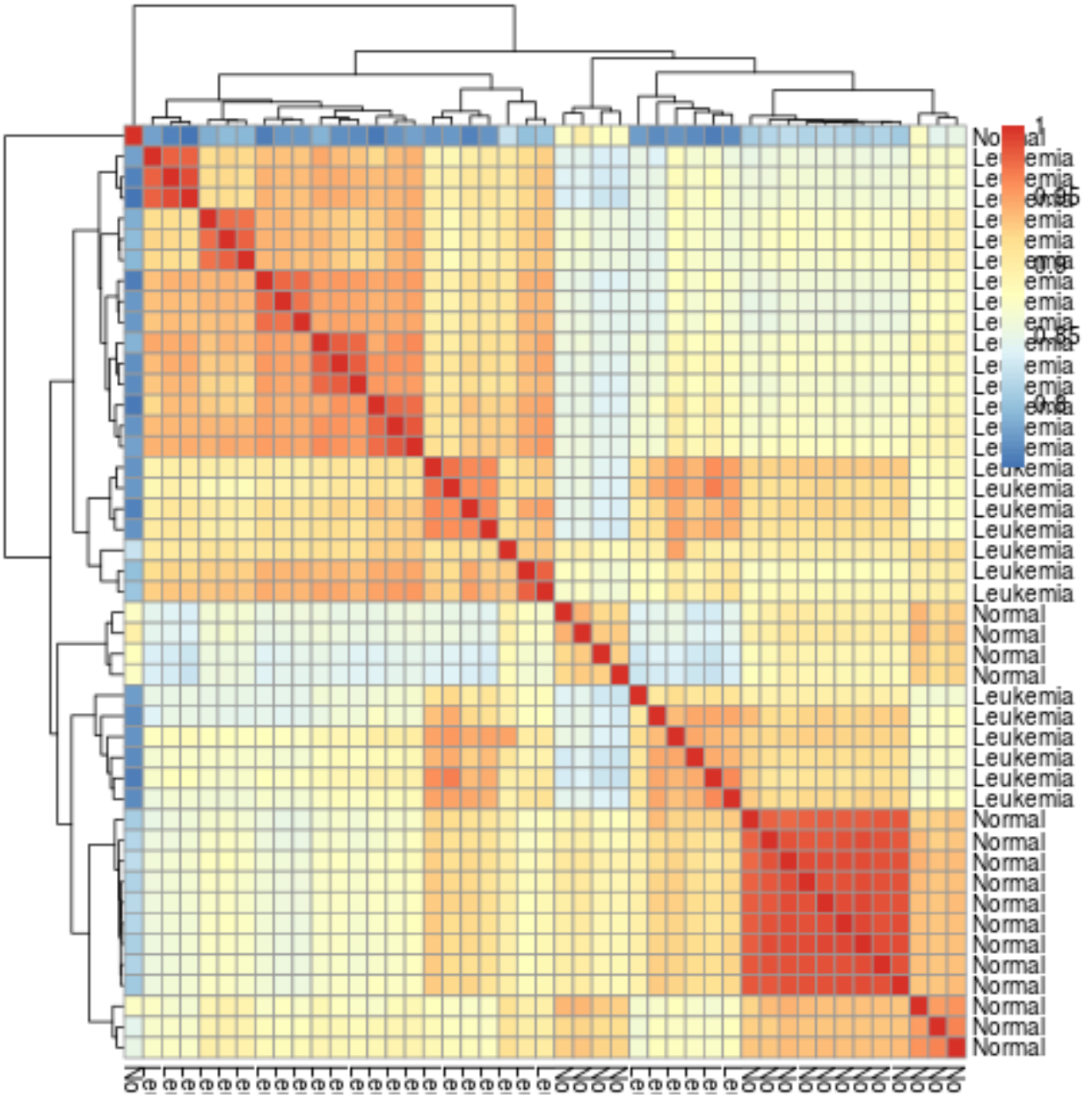

Figure 3: The heat map for T cells, showing relatively better clustering. Clustering score was not calculated due to the god damn deadline.

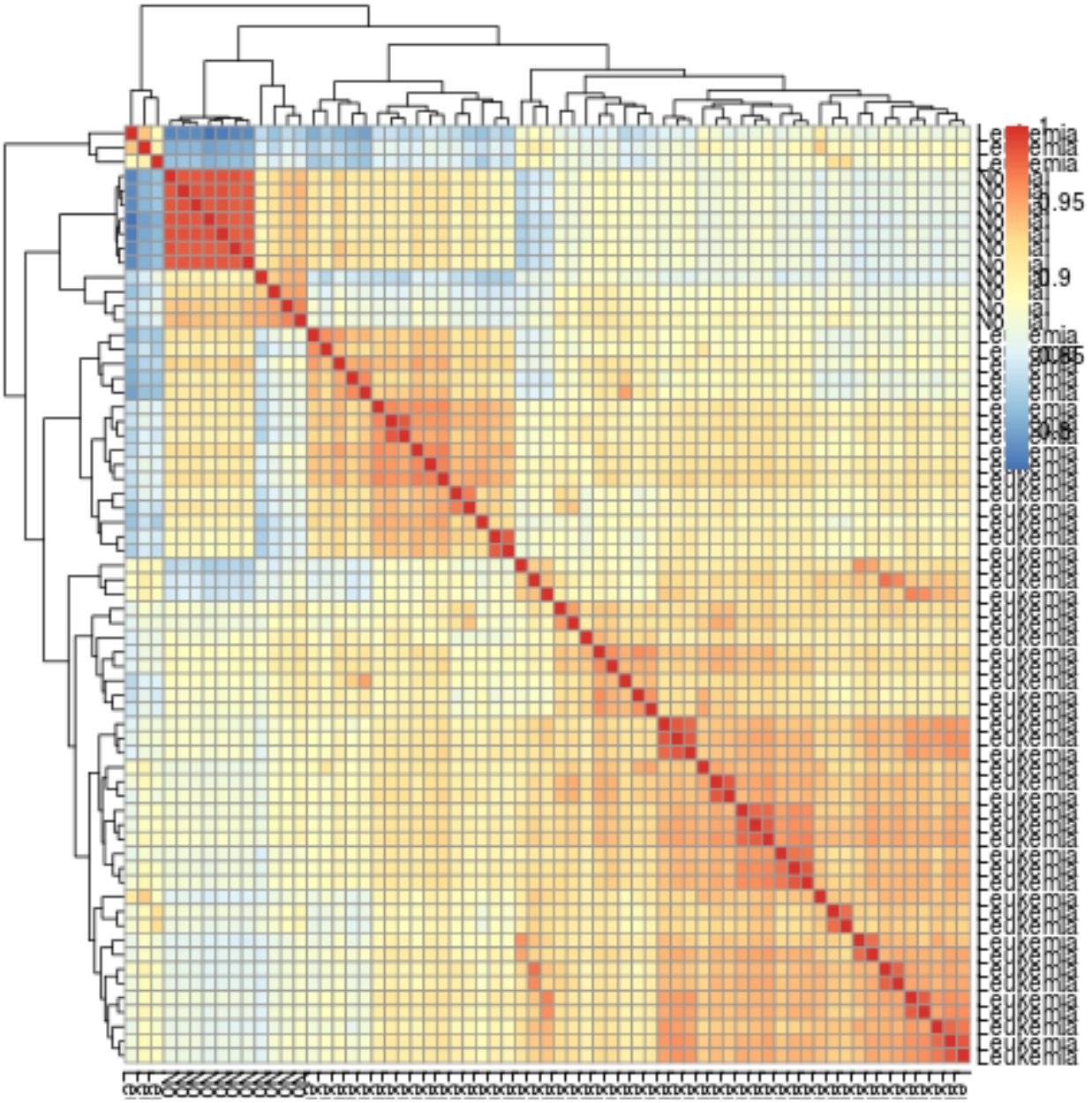

Figure 4: The heat map for the B cells, showing almost perfect clustering. Clustering score was not calculated due to the god damn deadline.

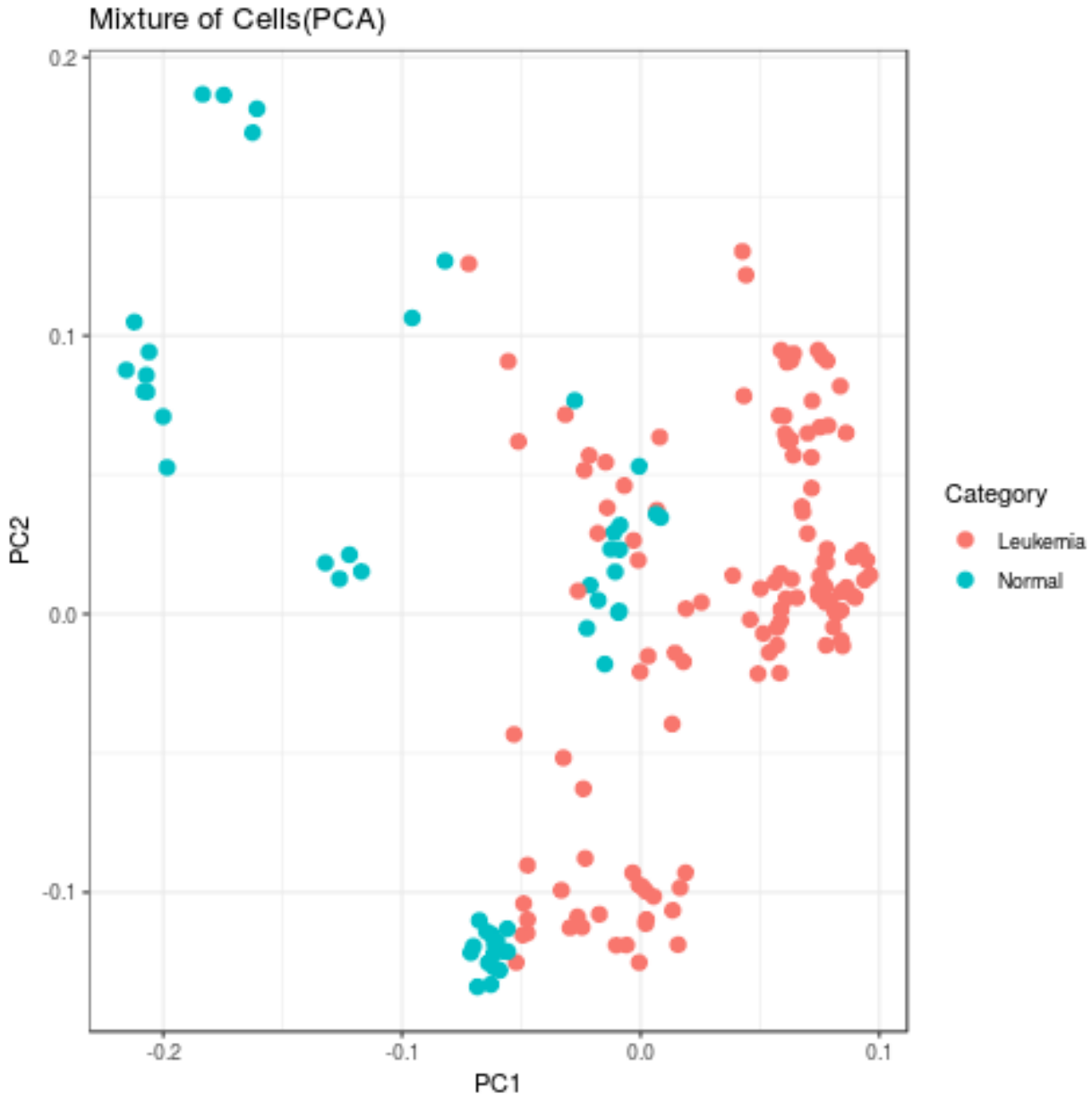

Figure 5: PCA indicates poor separability in cell mixture.

intuition about the comparability of the data. The data needed no normalization (figure1.) However, the heatmap and its interensic hierarchical clustering failed to assure the validity of the data.

extract some genes with significantly different expression levels in normal versus affected cells.

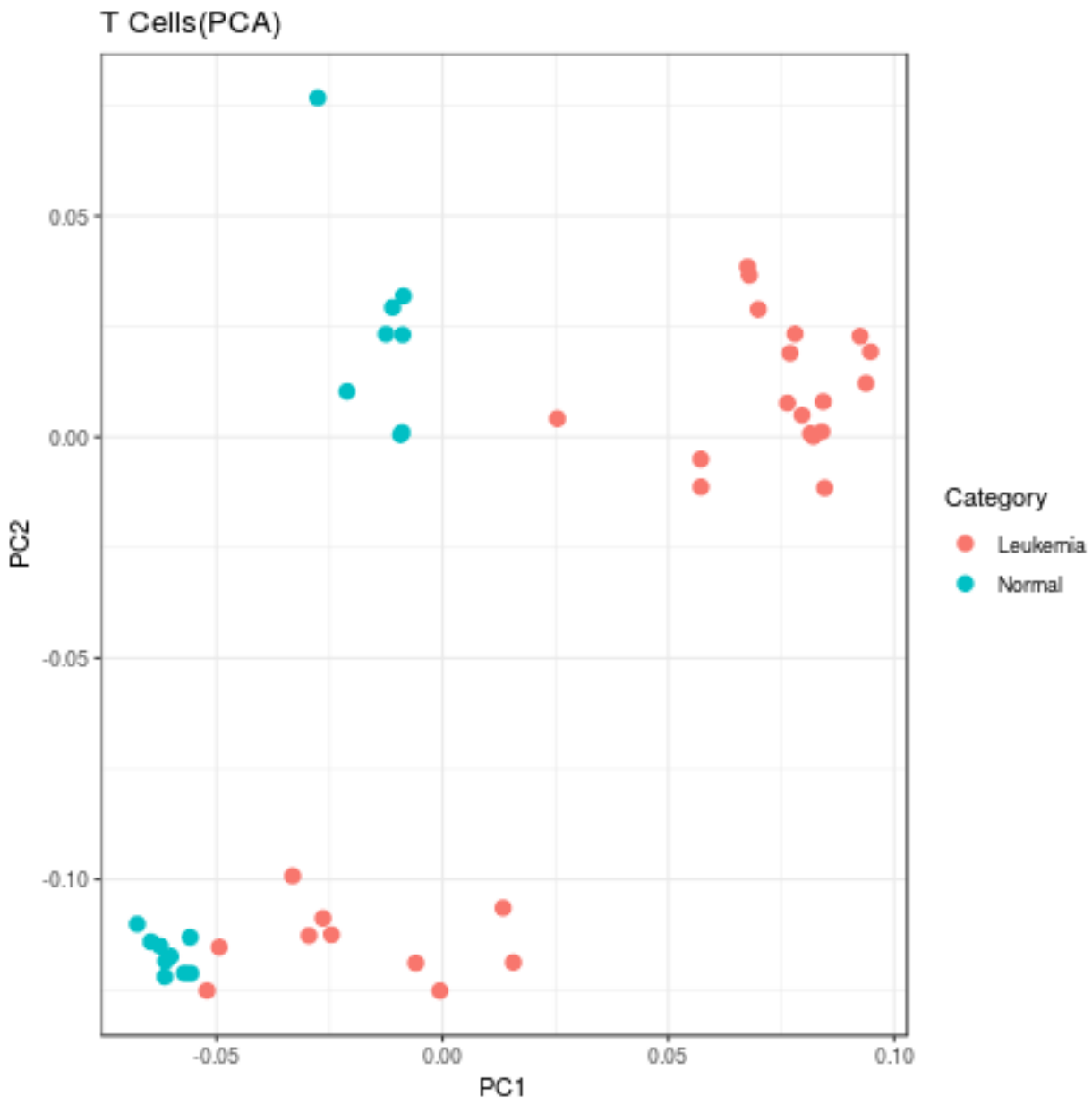

Figure 6: PCA indicates relatively better separability in T cells.

4. Gene Set Enrichment Analysis

As the gene sets extracted, they were passed to EnrichR to find possible causes and cures for AML. The GSEA results were truncated not to be represent ideas elusive to naked eyes. So only results with adjusted p-value less than .05 were selected. Moreover, at most 5 top cases extracted from each database were kept and the rest were eliminated. The results were sorted by combined score before the elimination process. The databases that were exploited for each purpose are listed in Table **??** . Appendix A would explain the reason of each

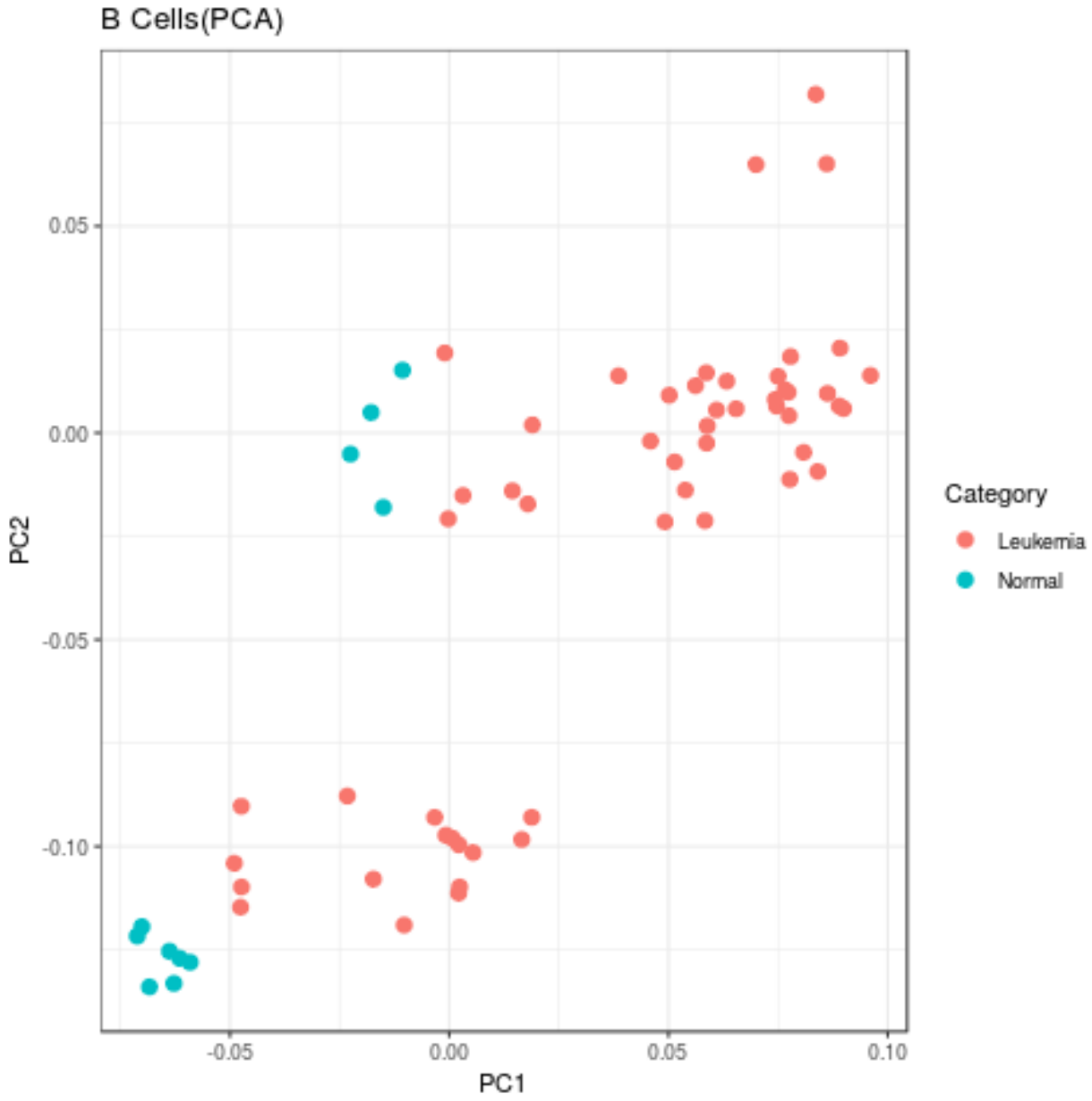

Figure 7: PCA indicates almost perfect separability in B cells.

database choice.

The comprehensive genes outputed by GSEA can be viewed inAppendix B. Amongth the extracted genes, pathways, etc, some were more attractive and listed here with some details. The headers are represented in the format of ([Source_Change], Name, Adjusted P-value, Z-score, Combined Score, database.)

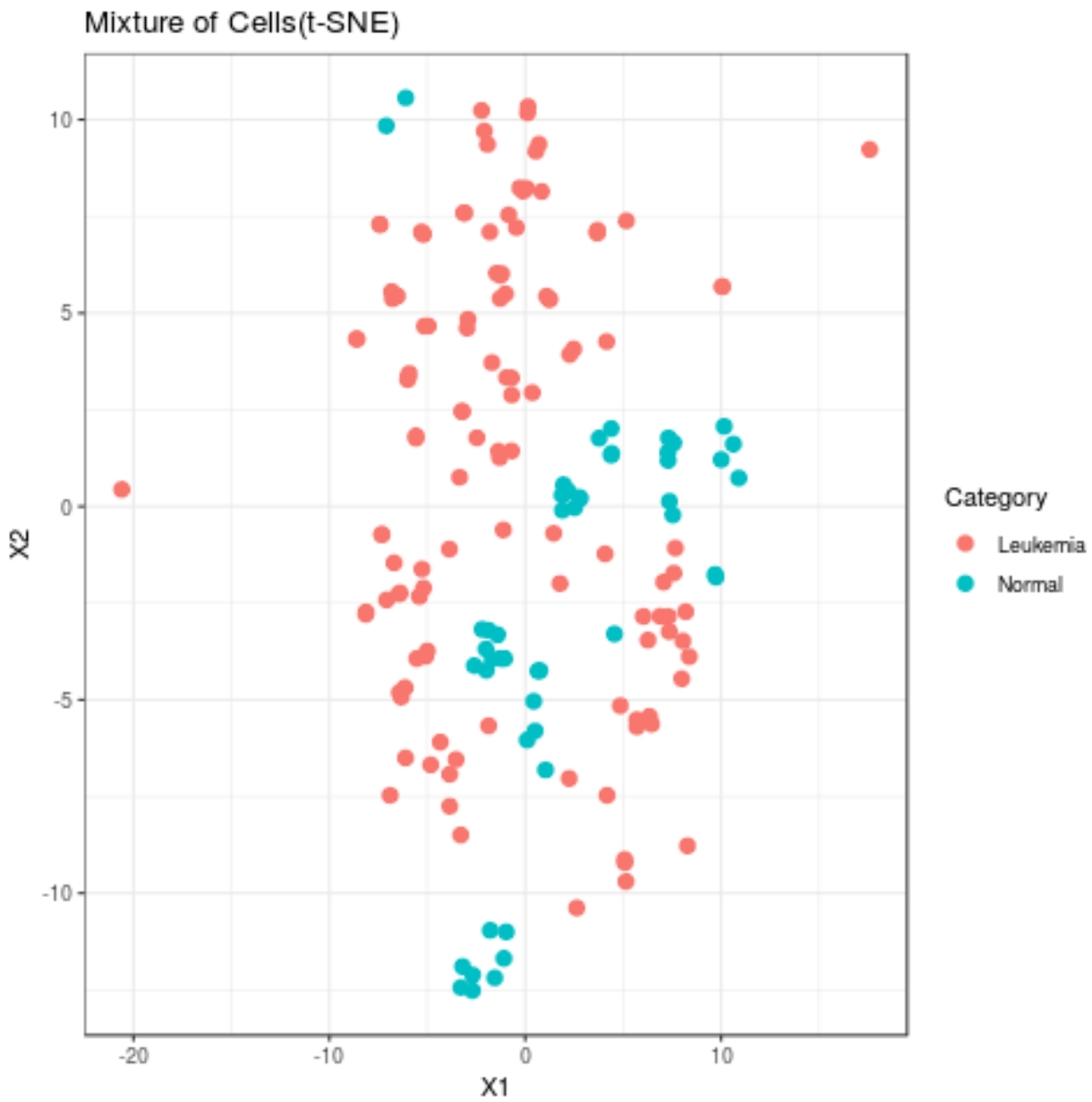

Figure 8: T-SNE indicates poor separability in cell mixture.

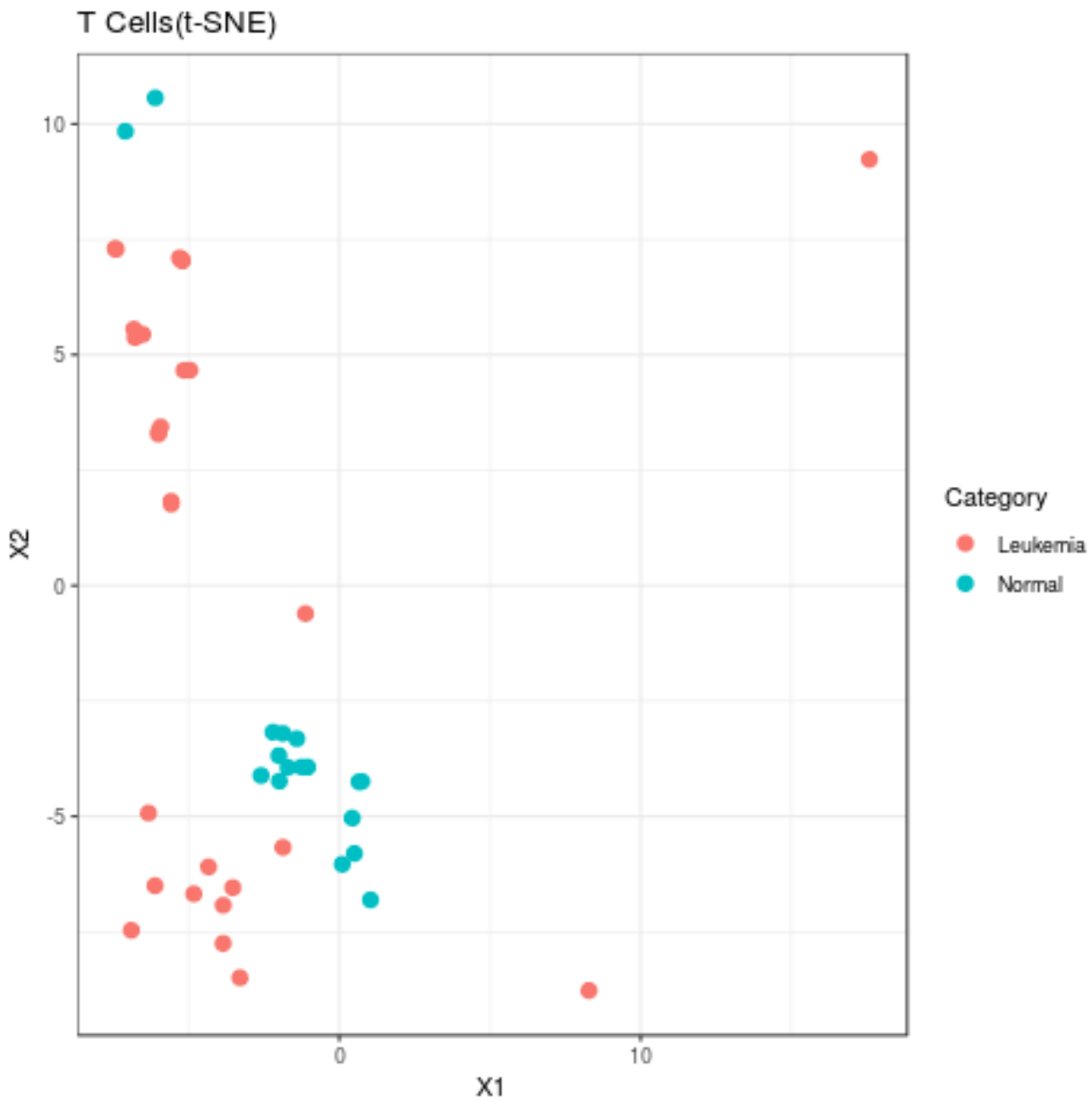

Figure 9: T-SNE indicates relatively better separability in T cells.

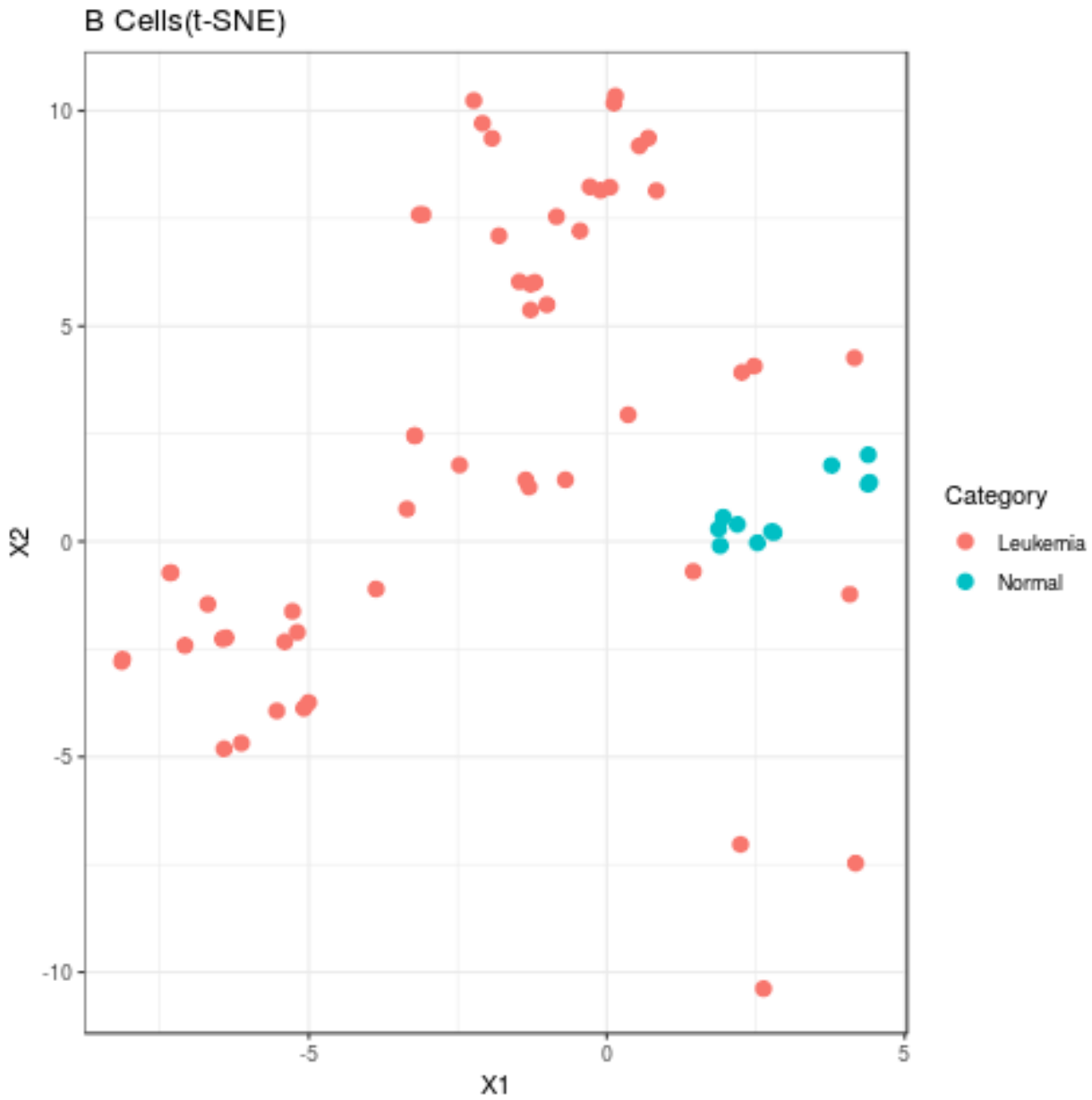

Figure 10: T-SNE indicates almost perfect separability in B cells.

Table 1: The samples were studied dependant and independant of their source.

Source Dependancy	Normal Sources	Leukemia Sources
Independent	B Cells, CD34+HSPC, Granulocytes, Monocytes, T Cells	AML Cell Line, AML Patient, B ALL Cell Line, B ALL Patient, T ALL Cell Line, T ALL Patient
T cells	T Cells	T ALL Cell Line, T ALL Patient
B cells	B Cells	B ALL Cell Line, B ALL Patient

Table 2: Databases in Use for GSEA

Analysis	Databases
Transcription Factors	TRANSFAC and JASPAR PWMs, miRTarBase 2017
Pathways	KEGG2019 Human, NCI-Nature 2016, Reactome 2016
Gene Ontology	GO Biological Process 2018, GO Molecular Function 2018, GO Cellular Component 2018
Drugs	ARCHS4 IDG Coexp

4.1. Transcription Factors (TFs)

This section has been widely influenced by proteinatlas and KEGG.

4.1.1. ETV4

Transcription regulator, e.g is unfavorable to liver cancer while favorable to thyroid cancer.

4.1.2. E2F4

Transcription facator, significant role in cell cycle regulation.

4.1.3. E2F1

Transcription factor of the E2F TF family.

4.1.4. NRF1

Nuclear respiratory factor 1, TF regulating key metabolic genes.

4.1.5. NFIC

TF in cells and replication factor for adenovirus.

4.1.6. NFYA

TF associated with a subunit of a high specified and affined DNA binding complex.

4.1.7. POU2F1

Of the same group of oct4 i.e octamer transcription factors.

4.2. Pathways (PWs)

4.2.1. Glycosaminoglycan degradation

Glycosaminoglycans are polar molecules found in extracellular matrix and interact with growth factors (Ernst et al. (1995)).

4.2.2. Glycosphingolipid biosynthesis

Glycosphingolipids form a major part of lipid rafts in cell membrane, in which some receptors lay (RL and T. (2017)).

4.2.3. RIG-I-like receptor signaling pathway

RIG-I-like receptors have accepted the role of pathogen sensing of RNA viruses (Loo and Gale (2011)).

4.2.4. Autoimmune thyroid disease

A disease in which thyroid cells are defected and secrete antigens consequently, attracting B and T cells toward themselves, therefore, driven to necrosis/apoptosis. (**?**)

4.2.5. Graft-versus-host disease

Allogeneic hematopoietic transplantations causes some disorders invoking donor cells against the host ones. (Ferrara et al. (2009), **?**)

4.2.6. Endosomal/Vacuolar pathway

Antigenic peptides are generated and cross presented, resulting in loaded MHC-I cells through this pathway. (**?**)

4.2.7. Interferon alpha/beta signaling

The Interferon alpha/beta signaling pathway results in some IFN-stimulated and -induced gene expressions and site bindings. (**?**)

4.2.8. Antigen Presentation: Folding, assembly and peptide loading of class I MHC

Antigen peptides are loaded to MHC I molecules and cytotoxic T cells are invoked by them. This pathway deals with MHC since Folding in Golgi up to the peptide loading of it.

4.2.9. Immune System

Self explanatory.

4.2.10. Interferon gamma signaling

Interferon Gamma Receptors induce phosphorylation upon binding to IFNG and consequently activate genes containing gamma-interferon activation sequence (GAS) in their promoter along this pathway. (**?**)

4.2.11. DNA replication

Self explanatory.

4.2.12. Homologous recombination

A pathway in which double strand breaks are fixed using the homologous DNA sequence, which is a necessary error correction mechanism for the DNA replication. (Alberts B)

4.2.13. Oocyte meiosis

The process of forming mature female gametocyte.

4.2.14. Fanconi anemia pathway

A pathway in which some replication barriers are removed. (Rodríguez and D'Andrea (2017))

4.2.15. E2F transcription factor network

E2F factors are of the most important S-phase entry in the cell cycle, although their roles are not limited to the aforementioned one. (Dimova and Dyson (2005))

4.2.16. PLK1 signaling events

Yet another pathway to ensure the mitosis is done correctly. (Lera et al. (2016))

4.2.17. FOXM1 transcription factor network

FOXM1 TFs have been found to be essential for mitotic progression and some gene encodings as well. (Wang et al. (2005))

4.2.18. Aurora B signaling

Aurora B is assigned the duty of chromosome separation during cell divisions by the evolution. (Krenn and Musacchio (2015)) It is alleged that some aneuploidies are the consecuence of its malfunctioning.

4.2.19. ATR signaling pathway

DNA damage response is a mechanism in cell that maintains DNA integrity. This mechanism is mainly regulated by ATR signaling pathway. (Nam and Cortez (2011))

4.2.20. Cell Cycle

A process including but not limited to DNA replication and organelle division in order to form daughter cells.

4.2.21. M Phase

A phase of cell cycle in which nuclear and cytoplasm are divided to form two daughter cells.

4.2.22. Cell Cycle Checkpoints

In order to have the cell cycle under control, evolution has defined some checkpoints between each two consecutive phases of cell cycle to ensure whether it is necessary to continue the cycle.

4.2.23. G2/M Checkpoints

Before entering mitosis phase, this checkpoint would check the DNA replication fidelity. (Stark and Taylor (2004))

4.2.24. Th17 cell differentiation

Th17 cells are a subset of CD4+ T cells. (**?**)

4.2.25. Allograft rejection

The process of recognizing donor cells as intruders and attacking them by T cells. (Martinu et al. (2011))

4.2.26. TNF receptor signaling pathway

Regulation of cell proliferation , differentiation, and apoptosis are processes in which TNF involves in. (**?**)

4.2.27. TRAIL signaling pathway

TRIAL is a death ligand which binds to death receptors and causes apoptosis. Metastasis suppression is of its main roles. (Thorburn (2007))

4.2.28. IL12-mediated signaling events

IL12 is a protein defending cell against intracellular viral infections attracting imune system against the corrupted cell. (Liu et al. (2005))

4.2.29. Downstream signaling in naive CD8+ T cells

4.2.30. Cytokine Signaling in Immune system

"Cytokines are small proteins that regulate and mediate immunity, inflammation, and hematopoiesis. They are secreted in response to immune stimuli, and usually act briefly, locally, at very low concentrations." (**?**)

4.2.31. p73 transcription factor network

p73 is a TF with low frequency of mutation in cancer despite of overlapping gene targets with p53 (Yu et al. (2006))

4.2.32. Mitotic Prometaphase

A phase of mitosis in which the nuclear membrane breaks apart.

4.2.33. Resolution of Sister Chromatid Cohesion

Eliminating cohesin complexes from chromosomes having them still connected at centromeres (**?**)

4.2.34. Intestinal immune network for IgA production

A mechanism of generating and differentiating noninflammatory immunoglobulin A originated from mucosal B cells. (**?**)

4.2.35. TNF signaling pathway

TNFs play wide range of roles varying from apoptosis to survival reactions. (**?**)

4.2.36. Asthma

A pathway in which allergens invoke T cells against themselves causing inflammation. (**?**)

4.2.37. Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell

There are some receptors and surface attaching molecules which adjust the response of lymphoid cells to the environmental condition. (**?**)

4.2.38. Adaptive Immune System

A major strategy of the immune system is to produce and devote a specific kind of cells to each pathogen. This is the responsibility of Adaptive Immune System.

4.2.39. CD22 mediated BCR regulation

Out of control BCR expression, which causes autoimmune disease, is avoided a set of mechanisms including CD22 receptors signaling. (**?**)

4.2.40. TNF receptor superfamily (TNFSF) members mediating non-canonical NF-kB pathway

4.2.41. Human T-cell leukemia virus 1 infection

A retroviral infection mostly nonlethal and even asymptomatic. (TCLv1)

4.2.42. Regulation of retinoblastoma protein

Of the main G0/G1 and G1/S regulators, Rb proteins can be made. They regulates target genes of E2F TFs. (Hume and Kalejta (2009))

4.2.43. Mitotic G1-G1/S phases

Respectively growth focused and transition states in which cell's commitment to the proliferation is ensured. (Bertoli et al. (2013))

It can be obviously observed that the listed pathways are mainly concerning cell cycle and immune system. It can be guessed that cell cycle regulation errors are the causes of AML, while immune system deficiencies are the consequences of the cancer. The former is widely studied in Saito et al. (2010), Stiehl et al. (2015), Banker et al. (1998), Schnerch et al. (2012), and a handful of other articles. This article would follow the later to compile pieces of information about immune system deficiencies and AML relation.

4.3. Gene Ontology

Table up to Table would cover extracted Gene Ontologies which are in simple English and, therefore, intelligible using only high school biology knowledge. Skimming through tables entries, it was observed that highly scored ontologies are strongly linked to cell cycle regulation and fidelity mechanisms, and cellular components. While low scored ontologies, which are usually observed among genes which have been expressed less in AML affected cells, seems related to immune system regulation mechanisms and actions. Exploiting existing knowledge about AML, it could be inferred that due to failure of immune cells to differentiate and become mature, they lack the ability to accomplish their normal responsibility. Hence, the idea of immune deficiencies be side effects of AML is supported.

4.4. Consensus

As retroviral infections are not reported as significant causes of AML, it is not a strong suggestion to regard immune system weakness a major cause of AML. However, Ramadan et al. (2012) has indicated that autoimmune deficiencies can surge the risk of AML and supported this idea with a number of statistical tests. These tests approve the claim in some cases while stay wordless against others. Moreover, the immune system can be considered as a major therapy fails in AML. Bone marrow stem cell transplant is among the most well known therapies for AML while as indicated by Horowitz et al. (1990) it is suppressed by body immune system and donor cells are identified as intruders by the immune system. Previously reported pathways in this article has pointed this out. Lamble and Lind (2018) suggests immunotherapy methods to bypass the aforementioned problem.

4.5. Drugs

Usually cancer drugs target cell cycle phases to regulate the cell cycle, hoping to outperform mutational adaption of cancer cells (Shapiro and Harper (1999), Bai et al. (2017). A list of less studied drugs can be found in (TODO appx) , but due to the god damn deadline, any investigation about docking capability and other related drug assessments are ignored.

4.6. Future Works

During the study, two main challenges were faced. The first one was the absence of a clustering and separability metric with a reasonable biological interpretation which had left the researcher unsure whether to ignore some samples for the sake of accuracy. A workaround was to define a metric based on the variance in each cluster using only its data label. The second challenge was the absence of an automated tool to relate the enrichment analysis results to each other. A proposed idea which is supported by intuition but suffering major drawbacks is to feed the enrichment process again using the extracted TFs. This idea was tested and provided weak statistical results.

Acknowledgments

Special thanks to the teaching team and author's patient classmates, who were a great source of inspiration and knowledge.

Appendix A. GSEA Database Selection

TRANSFAC and JASPAR PWMs, and miRTarBase 2017 were chosen for TF enrichment because they were curated and validated experimentally. KEGG and Reactome was selected for the same reason, however due to its vast resources and manually curated essence, NCI-Nature was used as well. Once again, being manually curated, was an advantage of GO against Jensen group. ARCHS4 IDG Coexp was interesting for the research because of the article's approach of digging into less studied areas has been satisfied by it. In fact, ARCHS4 IDG Coexp covers numerous under-studied drugs.

Appendix B. GSEA Full Selected Genes

Table B.3: TF Analysis of genes under-expressed in AML (cell mixture), TRANSFAC and JASPAR PWMs

Name	Adj. P-value	Z-score	Combined Score
ETV4 (human)	0.043	-1.74	15.24

Table B.4: TF Analysis of genes over-expressed in AML (cell mixture), TRANSFAC and JASPAR PWMs

Name	Adj. P-value	Z-score	Combined Score
E2F4 (human)	6.248e-9	-1.97	48.40
E2F1 (human)	0.00002605	-1.55	23.43
NRF1 (human)	0.0001446	-1.71	22.44
NFIC (human)	0.001465	-1.69	18.00
NFYA (human)	0.002248	-1.60	16.03

Table B.5: TF Analysis of genes over-expressed in AML (cell mixture), miRTarBase 2017

Name	Adj. P-value	Z-score	Combined Score
hsa-miR-193b-3p	8.760050e-64	-6.364176	973.9777
hsa-miR-192-5p	2.501043e-31	-6.765765	522.3863
hsa-miR-215-5p	2.718386e-34	-5.184328	437.7653
hsa-miR-34a-5p	9.159129e-21	-5.117180	269.1568
hsa-miR-92a-3p	1.886866e-08	-9.381637	223.6528

Table B.6: TF Analysis of genes over-expressed in AML (T Cells), TRANSFAC and JASPAR PWMs

Name	Adj. P-value	Z-score	Combined Score
E2F4 (human)	0.006586	-1.95	18.87
E2F1 (human)	0.002198	-1.57	18.41

Table B.7: TF Analysis of genes over-expressed in AML (T Cells), miRTarBase 2017

Name	Adj. P-value	Z-score	Combined Score
hsa-miR-193b-3p	1.574e-35	-6.36	558.85
hsa-miR-192-5p	4.991e-29	-6.77	485.40
hsa-miR-215-5p	6.399e-30	-5.18	384.69
hsa-miR-34a-5p	1.639e-7	-5.12	112.13
hsa-let-7b-5p	0.008490	-7.90	85.63

Table B.8: TF Analysis of genes under-expressed in AML (B Cells), TRANSFAC and JASPAR PWMs

Name	Adj. P-value	Z-score	Combined Score
POU2F1 (human)	0.03246	-1.74	15.68

Table B.9: TF Analysis of genes over-expressed in AML (B Cells), TRANSFAC and JASPAR PWMs

Name	Adj. P-value	Z-score	Combined Score
E2F4 (human)	3.861e-8	-1.97	44.58
E2F1 (human)	0.005012	-1.55	15.12
NRF1 (human)	0.01787	-1.71	14.02

Table B.10: TF Analysis of genes over-expressed in AML (B Cells), miRTarBase 2017

Name	Adj. P-value	Z-score	Combined Score
hsa-miR-193b-3p	1.275e-39	-6.36	618.57
hsa-miR-192-5p	4.036e-18	-6.77	315.22
hsa-miR-215-5p	2.494e-19	-5.18	258.08
hsa-miR-34a-5p	5.211e-14	-5.12	188.50
hsa-let-7b-5p	0.00006949	-7.90	121.79

Table B.11: PW Analysis of genes under-expressed in AML (cell mixture), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
Glycosaminoglycan degradation	0.02246	-613.78	3210.85
Glycosphingolipid biosynthesis	0.4322	-900.86	1321.91
RIG-I-like receptor signaling pathway	0.004278	-97.67	704.97
Autoimmune thyroid disease	0.002948	-78.88	653.40
Graft-versus-host disease	0.001461	-68.48	636.28

Table B.12: PW Analysis of genes under-expressed in AML (cell mixture), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Endosomal/Vacuolar pathway_Homo sapiens_R-HSA-1236977	0.000001320	-1.98	37.94
Interferon alpha/beta signaling_Homo sapiens_R-HSA-909733	0.00001904	-1.86	29.37
Antigen Presentation: Folding, assembly and peptide loading of class I MHC_Homo sapiens_R-HSA-983170	0.00002778	-1.93	28.99
Immune System_Homo sapiens_R-HSA-168256	0.0001276	-2.20	28.61
Interferon gamma signaling_Homo sapiens_R-HSA-877300	0.00008257	-1.76	24.06

Table B.13: PW Analysis of genes over-expressed in AML (cell mixture), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
DNA replication 6.096e-17	-49.69	1855.26	
Homologous recombination	5.327e-9	-49.65	1100.04
Oocyte meiosis	1.312e-10	-26.82	704.47
Fanconi anemia pathway	0.000001224	-41.92	687.64
Small cell lung cancer	0.0006399	-63.94	602.21

Table B.14: PW Analysis of genes over-expressed in AML (cell mixture), NCI-Nature

Name	Adj. P-value	Z-score	Combined Score
E2F transcription factor network_Homo sapiens_bb4d0fd3-6191-11e5-8ac5-06603eb7f303	1.616e-20	-1.73	85.49
PLK1 signaling events_Homo sapiens_e5e87977-6194-11e5-8ac5-06603eb7f303	1.127e-20	-1.66	83.87
FOXM1 transcription factor network_Homo sapiens_c51cda49-6192-11e5-8ac5-06603eb7f303	2.769e-16	-1.51	58.67
Aurora B signaling_Homo sapiens_304a75af-618c-11e5-8ac5-06603eb7f303	5.779e-18	-1.26	54.28
ATR signaling pathway_Homo sapiens_8991cbac-618b-11e5-8ac5-06603eb7f303	4.751e-15	-1.18	42.36

Table B.15: PW Analysis of genes over-expressed in AML (cell mixture), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Cell Cycle_Homo sapiens_R-HSA-1640170	1.426e-117	-2.46	677.98
Cell Cycle, Mitotic_Homo sapiens_R-HSA-69278	1.227e-106	-2.47	616.03
M Phase_Homo sapiens_R-HSA-68886	2.552e-50	-2.43	290.68
Cell Cycle Checkpoints_Homo sapiens_R-HSA-69620	6.168e-44	-2.34	244.94
G2/M Checkpoints_Homo sapiens_R-HSA-69481	2.432e-40	-2.32	222.45

Table B.16: PW Analysis of genes under-expressed in AML (T Cells), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
Glycosaminoglycan degradation	0.1424	-585.89	2091.24
Autoimmune thyroid disease	0.00003361	-80.82	1151.46
Th17 cell differentiation	0.00002734	-69.21	1041.32
Graft-versus-host disease	0.00004041	-68.48	923.26
Allograft rejection	0.00003361	-54.70	767.09

Table B.17: PW Analysis of genes under-expressed in AML (T Cells), NCI-Nature

Name	Adj. P-value	Z-score	Combined Score
TNF receptor signaling pathway_Homo sapiens_316be05e-6196-11e5-8ac5-06603eb7f303	0.0001737	-1.53	19.39
TRAIL signaling pathway_Homo sapiens_3a79fddf-6196-11e5-8ac5-06603eb7f303	0.0001737	-1.41	18.49
IL12-mediated signaling events_Homo sapiens_7acdea19-6193-11e5-8ac5-06603eb7f303	0.0008804	-1.54	16.39
Downstream signaling in naive CD8+ T cells_Homo sapiens_92180cef-6191-11e5-8ac5-06603eb7f303	0.0009973	-1.43	14.59
IL2-mediated signaling events_Homo sapiens_a2a1883c-6193-11e5-8ac5-06603eb7f303	0.002484	-1.55	14.07

Table B.18: PW Analysis of genes under-expressed in AML (T Cells), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Immune System_Homo sapiens_R-HSA-168256	0.00001975	-2.23	37.84
Cytokine Signaling in Immune system_Homo sapiens_R-HSA-1280215	0.00003972	-2.40	37.28
Endosomal/Vacuolar pathway_Homo sapiens_R-HSA-1236977	0.00004649	-1.90	28.15
Interferon alpha/beta signaling_Homo sapiens_R-HSA-909733	0.00004649	-1.84	27.00
Interferon gamma signaling_Homo sapiens_R-HSA-877300	0.00006090	-1.75	24.90

Table B.19: PW Analysis of genes over-expressed in AML (T Cells), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
Glyoxylate and dicarboxylate metabolism	0.5917	-701.48	867.99
Glycosaminoglycan biosynthesis	0.3410	-406.31	856.52
Homologous recombination	0.00001242	-50.16	744.50
Small cell lung cancer	0.001191	-67.62	628.70
DNA replication	0.00006099	-49.02	624.51

Table B.20: PW Analysis of genes over-expressed in AML (T Cells), NCI-Nature

Name	Adj. P-value	Z-score	Combined Score
PLK1 signaling events_Homo sapiens_e5e87977-6194-11e5-8ac5-06603eb7f303	1.853e-13	-1.66	55.78
FOXM1 transcription factor network_Homo sapiens_c51cda49-6192-11e5-8ac5-06603eb7f303	1.887e-11	-1.65	46.63
E2F transcription factor network_Homo sapiens_bb4d0fd3-6191-11e5-8ac5-06603eb7f303	3.319e-11	-1.68	45.91
p73 transcription factor network_Homo sapiens_a88c505e-6194-11e5-8ac5-06603eb7f303	8.821e-10	-1.38	32.22
ATR signaling pathway_Homo sapiens_8991cbac-618b-11e5-8ac5-06603eb7f303	1.914e-10	-1.24	31.11

Table B.21: PW Analysis of genes over-expressed in AML (T Cells), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Cell Cycle_Homo sapiens_R-HSA-1640170	3.743e-54	-2.46	317.68
Cell Cycle, Mitotic_Homo sapiens_R-HSA-69278	5.027e-44	-2.47	259.31
Mitotic Prometaphase_Homo sapiens_R-HSA-68877	1.905e-22	-2.03	111.73
M Phase_Homo sapiens_R-HSA-68886	2.687e-18	-2.41	107.58
Resolution of Sister Chromatid Cohesion_Homo sapiens_R-HSA-2500257	1.644e-20	-2.06	103.70
Intestinal immune network for IgA production	0.000003880	-99.15	1739.13

Table B.22: PW Analysis of genes under-expressed in AML (B Cells), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
Glycosphingolipid biosynthesis	0.8842	-781.54	574.04
TNF signaling pathway	0.06974	-74.45	391.55
Autoimmune thyroid disease	0.08780	-77.59	383.98
Asthma	0.09760	-81.53	76.60

Table B.23: PW Analysis of genes under-expressed in AML (B Cells), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Immune System_Homo sapiens_R-HSA-168256	0.0003718	-2.23	31.24
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell_Homo sapiens_R-HSA-198933	0.0006682	-2.00	25.45
Adaptive Immune System_Homo sapiens_R-HSA-1280218	0.006699	2.27	22.67
CD22 mediated BCR regulation_Homo sapiens_R-HSA-5690714	0.009758	-1.94	17.65
TNF receptor superfamily (TNFSF) members mediating non-canonical NF-kB pathway_Homo sapiens_R-HSA-5676594	0.009595	-1.88	17.61

Table B.24: PW Analysis of genes over-expressed in AML (B Cells), KEGG2019 Human

Name	Adj. P-value	Z-score	Combined Score
Small cell lung cancer	0.002421	-64.35	525.51
DNA replication	0.0004052	-48.35	515.36
Homologous recombination	0.0005634	-48.79	488.48
Fanconi anemia pathway	0.001818	-40.98	355.37
Human T-cell leukemia virus 1 infection	0.00003497	-18.10	242.43

Table B.25: PW Analysis of genes over-expressed in AML (B Cells), NCI-Nature

Name	Adj. P-value	Z-score	Combined Score
E2F transcription factor network_Homo sapiens_bb4d0fd3-6191-11e5-8ac5-06603eb7f303	6.635e-13	-1.78	57.97
FOXM1 transcription factor network_Homo sapiens_c51cda49-6192-11e5-8ac5-06603eb7f303	6.102e-9	-1.65	37.63
PLK1 signaling events_Homo sapiens_e5e87977-6194-11e5-8ac5-06603eb7f303	2.976e-7	-1.47	26.78
Aurora B signaling_Homo sapiens_304a75af-618c-11e5-8ac5-06603eb7f303	9.910e-8	-1.26	24.67
Regulation of retinoblastoma protein_Homo sapiens_407a3468-6195-11e5-8ac5-06603eb7f303	0.000003482	-1.52	23.55

Table B.26: PW Analysis of genes over-expressed in AML (B Cells), Reactome 2016

Name	Adj. P-value	Z-score	Combined Score
Cell Cycle_Homo sapiens_R-HSA-1640170	1.430e-40	-2.46	241.02
Cell Cycle, Mitotic_Homo sapiens_R-HSA-69278	7.756e-36	-2.47	213.08
Mitotic G1-G1/S phases_Homo sapiens_R-HSA-453279	1.077e-18	-2.11	98.22
G1/S Transition_Homo sapiens_R-HSA-69206	2.942e-15	-2.11	80.21
Mitotic Prometaphase_Homo sapiens_R-HSA-68877	1.569e-15	-2.02	78.71

Table B.27: GO Analysis of genes under-expressed in AML (cell mixture), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent (GO:0002480)	0.0001851	-2.33	35.76
type I interferon signaling pathway (GO:0060337)	0.0002905	-2.32	30.11
retrograde transport, vesicle recycling within Golgi (GO:0000301)	0.02414	-4.03	29.32
regulation of necrotic cell death (GO:0010939)	0.003712	-2.48	23.91
regulation of necroptotic process (GO:0060544)	0.0002905	-1.62	20.96

Table B.28: GO Analysis of genes under-expressed in AML (cell mixture), GO Cellular Component 2018

Name	Adj. P-value	Z-score	Combined Score
death-inducing signaling complex (GO:0031264)	0.01076	-3.30	22.26
recycling endosome membrane (GO:0055038)	0.00003093	-1.46	21.80
integral component of lumenal side of endoplasmic reticulum membrane (GO:0071556)	0.001205	-1.95	20.60
COPII-coated ER to Golgi transport vesicle (GO:0030134)	0.001228	-2.06	19.47
MHC protein complex (GO:0042611)	0.001980	-2.05	17.74

Table B.29: GO Analysis of genes over-expressed in AML (cell mixture), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
DNA metabolic process (GO:0006259)	1.366e-44	-1.36	147.85
DNA replication (GO:0006260)	1.493e-35	-1.56	135.24
DNA repair (GO:0006281)	3.977e-26	-1.71	109.74
mitotic cell cycle phase transition (GO:0044772)	2.146e-32	-1.17	92.67
chromatin remodeling at centromere (GO:0031055)	1.246e-15	-2.28	88.85

Table B.30: GO Analysis of genes over-expressed in AML (cell mixture), GO Molecular Function 2018

Name	Adj. P-value	Z-score	Combined Score
DNA helicase activity (GO:0003678)	7.421e-10	-2.66	66.98
ATPase activity (GO:0016887)	4.563e-9	-2.31	53.57
DNA-dependent ATPase activity (GO:0008094)	3.454e-13	-1.53	51.82
DNA binding (GO:0003677)	1.223e-16	-1.20	50.97
3'-5' DNA helicase activity (GO:0043138)	0.000003135	-2.82	44.63

Table B.31: GO Analysis of genes over-expressed in AML (cell mixture), GO Cellular Component 2018

Name	Adj. P-value	Z-score	Combined Score
nuclear chromosome part (GO:0044454)	4.133e-37	-1.25	111.11
chromosome, centromeric region (GO:0000775)	5.429e-20	-2.25	108.81
microtubule organizing center (GO:0005815)	1.986e-18	-2.02	88.76
spindle (GO:0005819)	1.837e-29	-1.18	83.06
chromatin (GO:0000785)	1.158e-19	-1.66	78.42

Table B.32: GO Analysis of genes under-expressed in AML (T Cells), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
retrograde transport, vesicle recycling within Golgi (GO:0000301)	0.009160	-4.03	37.61
cytokine-mediated signaling pathway (GO:0019221)	6.018e-9	-1.35	35.45
type I interferon signaling pathway (GO:0060337)	0.001496	-2.32	29.10
regulation of lymphocyte activation (GO:0051249)	0.005443	-2.70	27.84
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent (GO:0002480)	0.001988	-2.33	27.55

Table B.33: GO Analysis of genes under-expressed in AML (T Cells), GO Molecular Function 2018

Name	Adj. P-value	Z-score	Combined Score
cytokine receptor activity (GO:0004896)	0.0002861	-1.84	25.68
tumor necrosis factor-activated receptor activity (GO:0005031)	0.04099	-2.56	20.28

Table B.34: GO Analysis of genes under-expressed in AML (T Cells), GO Cellular Component 2018

Name	Adj. P-value	Z-score	Combined Score
clathrin vesicle coat (GO:0030125)	0.03999	-3.52	19.54
integral component of lumenal side of endoplasmic reticulum membrane (GO:0071556)	0.002593	-1.94	19.08
MHC protein complex (GO:0042611)	0.003483	-2.08	18.73
T cell receptor complex (GO:0042101)	0.02267	-2.55	16.46
clathrin coat of trans-Golgi network vesicle (GO:0030130)	0.05913	-3.12	15.73

Table B.35: GO Analysis of genes over-expressed in AML (T Cells), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
strand displacement (GO:0000732)	1.727e-11	-2.58	75.45
kinetochore organization (GO:0051383)	2.163e-11	-2.56	73.81
mitotic nuclear division (GO:0140014)	1.157e-11	-2.22	65.92
DNA biosynthetic process (GO:0071897)	1.520e-12	-1.97	62.81
DNA replication (GO:0006260)	2.687e-15	-1.55	61.43

Table B.36: GO Analysis of genes over-expressed in AML (T Cells), GO Molecular Function 2018

Name	Adj. P-value	Z-score	Combined Score
DNA helicase activity (GO:0003678)	0.000009750	-2.62	38.96
microtubule motor activity (GO:0003777)	1.938e-7	-1.85	37.43
ATPase activity (GO:0016887)	0.000003774	-2.31	37.41
motor activity (GO:0003774)	2.992e-7	-1.58	30.62
DNA-dependent ATPase activity (GO:0008094)	4.744e-7	-1.52	28.12

Table B.37: GO Analysis of genes over-expressed in AML (T Cells), GO Cellular Component 2018

Name	Adj. P-value	Z-score	Combined Score
chromosome, centromeric region (GO:0000775)	4.185e-16	-2.26	89.52
condensed chromosome, centromeric region (GO:0000779)	5.224e-12	-2.09	61.06
spindle microtubule (GO:0005876)	5.953e-12	-1.95	56.22
condensed chromosome kinetochore (GO:0000777)	3.066e-10	-2.22	54.12
spindle (GO:0005819)	1.758e-16	-1.18	48.64

Table B.38: GO Analysis of genes under-expressed in AML (B Cells), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
renal filtration (GO:0097205)	0.002162	-3.50	38.03
phagocytosis, engulfment (GO:0006911)	0.0001266	-2.39	35.59
regulation of immune effector process (GO:0002697)	0.001193	-2.89	34.48
glomerular filtration (GO:0003094)	0.001895	-3.00	33.73
positive regulation of lymphocyte activation (GO:0051251	0.0001266	-2.21	32.93

Table B.39: GO Analysis of genes under-expressed in AML (B Cells), GO Molecular Function 2018

Name	Adj. P-value	Z-score	Combined Score
immunoglobulin receptor binding (GO:0034987)	0.00008611	-2.41	36.35

Table B.40: GO Analysis of genes over-expressed in AML (B Cells), GO Biological Process 2018

Name	Adj. P-value	Z-score	Combined Score
DNA metabolic process (GO:0006259)	1.976e-20	-1.36	71.65
microtubule cytoskeleton organization involved in mitosis (GO:1902850)	4.370e-12	-1.65	53.05
sister chromatid segregation (GO:0000819)	3.095e-9	-2.04	50.33
mitotic spindle organization (GO:0007052)	3.027e-13	-1.39	49.10
DNA replication (GO:0006260)	1.211e-11	-1.55	47.77

Table B.41: GO Analysis of genes over-expressed in AML (B Cells), GO Molecular Function 2018

Name	Adj. P-value	Z-score	Combined Score
histone kinase activity (GO:0035173)	0.0001513	-3.09	41.13
DNA helicase activity (GO:0003678)	0.0001706	-2.67	34.36
DNA polymerase binding (GO:0070182)	0.0004986	-2.31	25.27
3'-5' DNA helicase activity (GO:0043138)	0.004400	-2.80	22.77
DNA-dependent ATPase activity (GO:0008094)	0.00006878	-1.53	22.22

Table B.42: GO Analysis of genes over-expressed in AML (B Cells), GO Cellular Component 2018

Name	Adj. P-value	Z-score	Combined Score
microtubule organizing center (GO:0005815)	1.411e-10	-2.07	54.44
chromosome, centromeric region (GO:0000775)	2.555e-8	-2.21	45.55
nuclear chromosome part (GO:0044454)	1.751e-13	-1.25	42.91
centrosome (GO:0005813), 1.411e-10	-1.61	42.58	
condensed chromosome kinetochore (GO:0000777)	5.286e-7	-2.27	39.74

Table B.43: Drug Enrichment of genes under-expressed in AML (cell mixture), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
STK17A_IDG_kinase_ARCHS4_coexpression	5.174e-28	-1.62	110.43
P2RY10_IDG_GPCR_ARCHS4_coexpression	1.710e-25	-1.55	95.41
MAP3K14_IDG_kinase_ARCHS4_coexpression	9.806e-18	-1.54	66.84
GPR25_IDG_GPCR_ARCHS4_coexpression	1.197e-15	-1.77	66.61
GPR174_IDG_GPCR_ARCHS4_coexpression	8.520e-17	-1.58	63.96

Table B.44: Drug Enrichment of genes over-expressed in AML (cell mixture), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
UCK2_IDG_kinase_ARCHS4_coexpression	9.315e-75	-1.61	283.39
PKMYT1_IDG_kinase_ARCHS4_coexpression	2.303e-65	-1.53	235.44
CHRNA9_IDG_ionchannel_ARCHS4_coexpression	2.957e-46	-1.54	168.06
RIOK1_IDG_kinase_ARCHS4_coexpression	7.258e-36	-1.60	136.09
PKN3_IDG_kinase_ARCHS4_coexpression	7.575e-35	-1.55	127.31

Table B.45: Drug Enrichment of genes under-expressed in AML (T Cells), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
STK17A_IDG_kinase_ARCHS4_coexpression	6.729e-22	-1.62	88.38
GPR174_IDG_GPCR_ARCHS4_coexpression	2.546e-21	-1.62	84.65
P2RY10_IDG_GPCR_ARCHS4_coexpression	2.546e-21	-1.54	80.07
DYRK2_IDG_kinase_ARCHS4_coexpression	8.547e-12	-1.60	47.72
GPR171_IDG_GPCR_ARCHS4_coexpression	3.522e-10	-1.72	44.50

Table B.46: Drug Enrichment of genes over-expressed in AML (T Cells), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
UCK2_IDG_kinase_ARCHS4_coexpression	1.384e-31	-1.60	121.31
PKMYT1_IDG_kinase_ARCHS4_coexpression	1.384e-31	-1.54	117.12
CHRNA9_IDG_ionchannel_ARCHS4_coexpression	1.934e-23	-1.54	87.25
PKN3_IDG_kinase_ARCHS4_coexpression	3.962e-15	-1.55	57.22
CSNK2A2_IDG_kinase_ARCHS4_coexpression	3.962e-15	-1.47	54.49

Table B.47: Drug Enrichment of genes under-expressed in AML (B Cells), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
P2RY10_IDG_GPCR_ARCHS4_coexpression	2.375e-20	-1.56	79.30
MAP3K14_IDG_kinase_ARCHS4_coexpression	3.895e-7	-1.55	30.67
STK38L_IDG_kinase_ARCHS4_coexpression	0.000007491	-1.61	26.45
GPR174_IDG_GPCR_ARCHS4_coexpression	0.00002773	-1.60	23.70
GPR152_IDG_GPCR_ARCHS4_coexpression	0.003520	-1.98	17.89

Table B.48: Drug Enrichment of genes over-expressed in AML (B Cells), ARCHS4 IDG Coexp

Name	Adj. P-value	Z-score	Combined Score
PKMYT1_IDG_kinase_ARCHS4_coexpression	8.882e-32	-1.54	118.82
UCK2_IDG_kinase_ARCHS4_coexpression	1.443e-24	-1.60	95.40
CHRNA9_IDG_ionchannel_ARCHS4_coexpression	9.312e-17	-1.54	63.51
PKN3_IDG_kinase_ARCHS4_coexpression	9.496e-14	-1.56	52.99
CSNK2A2_IDG_kinase_ARCHS4_coexpression	6.583e-11	-1.46	39.83

Alberts B, Johnson A, L. J. e. a. Molecular biology of the cell. 4th edition.

Bai, J., Li, Y., and Zhang, G. (2017). Cell cycle regulation and anticancer drug discovery. *Cancer biology & medicine*, 14(4):348–362. 29372101[pmid].

Banker, D. E., Groudine, M., Willman, C. L., Norwood, T., and Appelbaum, F. R. (1998). Cell cycle perturbations in acute myeloid leukemia samples following in vitro exposures to therapeutic agents. *Leukemia Research*, 22(3):221 – 239.

Bertoli, C., Skotheim, J. M., and de Bruin, R. A. M. (2013). Control of cell cycle transcription during g1 and s phases. *Nature reviews. Molecular cell biology*, 14(8):518–528. 23877564[pmid].

Cramer-Morales, K., Nieborowska-Skorska, M., Scheibner, K., Padget, M., Irvine, D. A., Sliwinski, T., Haas, K., Lee, J., Geng, H., Roy, D., Slupianek, A., Rassool, F. V., Wasik, M. A., Childers, W., Copland, M., Müschen, M., Civin, C. I., and Skorski, T. (2013). Personalized synthetic lethality induced by targeting rad52 in leukemias identified by gene mutation and expression profile. *Blood*, 122(7):1293–1304.

Dimova, D. K. and Dyson, N. J. (2005). The e2f transcriptional network: old acquaintances with new faces. *Oncogene*, 24(17):2810–2826.

Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995). Enzymatic degradation of glycosaminogiycans. *Critical Reviews in Biochemistry and Molecular Biology*, 30(5):387–444.

Ferrara, J. L. M., Levine, J. E., Reddy, P., and Holler, E. (2009). Graft-versus-host disease. *Lancet (London, England)*, 373(9674):1550–1561. 19282026[pmid].

Horowitz, M., Gale, R., Sondel, P., Goldman, J., Kersey, J., Kolb, H., Rimm, A., Ringden, O., Rozman, C., and Speck, B. (1990). Graft-versus-leukemia reactions after bone marrow transplantation. *Blood*, 75(3):555–562.

Hume, A. J. and Kalejta, R. F. (2009). Regulation of the retinoblastoma proteins by the human herpesviruses. *Cell division*, 4:1–1. 19146698[pmid].

Krenn, V. and Musacchio, A. (2015). The aurora b kinase in chromosome bi-orientation and spindle checkpoint signaling. *Frontiers in oncology*, 5:225–225. 26528436[pmid].

Lamble, A. J. and Lind, E. F. (2018). Targeting the immune microenvironment in acute myeloid leukemia: A focus on t cell immunity. *Frontiers in oncology*, 8:213–213. 29951373[pmid].

Lera, R. F., Potts, G. K., Suzuki, A., Johnson, J. M., Salmon, E. D., Coon, J. J., and Burkard, M. E. (2016). Decoding polo-like kinase 1 signaling along the kinetochore-centromere axis. *Nature chemical biology*, 12(6):411–418. 27043190[pmid].

Liu, J., Cao, S., Kim, S., Chung, E. Y., Homma, Y., Guan, X., Jimenez, V., and Ma, X. (2005). Interleukin-12: an update on its immunological activities, signaling and regulation of gene expression. *Current immunology reviews*, 1(2):119–137. 21037949[pmid].

Loo, Y.-M. and Gale, M. J. (2011). Immune signaling by rig-i-like receptors. *Immunity*, 34(5):680–692. 21616437[pmid].

Martinu, T., Pavlisko, E. N., Chen, D.-F., and Palmer, S. M. (2011). Acute allograft rejection: cellular and humoral processes. *Clinics in chest medicine*, 32(2):295–310. 21511091[pmid].

Nam, E. A. and Cortez, D. (2011). Atr signalling: more than meeting at the fork. *The Biochemical journal*, 436(3):527–536. 21615334[pmid].

Parkin, D. M., Bray, F., Ferlay, J., and Pisani, P. Global cancer statistics, 2002. *CA: A Cancer Journal for Clinicians*, 55(2):74–108.

Ramadan, S. M., Fouad, T. M., Summa, V., Hasan, S. K., and Lo-Coco, F. (2012). Acute myeloid leukemia developing in patients with autoimmune diseases. *Haematologica*, 97(6):805–817. 22180424[pmid].

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7):e47–e47.

RL, S. and T., K. (2017). *Essentials of Glycobiology [Internet], chapter 11*. Cold Spring Harbor (NY), Reading, Massachusetts.

Rodríguez, A. and D'Andrea, A. (2017). Fanconi anemia pathway. *Current Biology*, 27(18):R986–R988.

Saito, Y., Uchida, N., Tanaka, S., Suzuki, N., Tomizawa-Murasawa, M., Sone, A., Najima, Y., Takagi, S., Aoki, Y., Wake, A., Taniguchi, S., Shultz, L. D., and Ishikawa, F. (2010). Induction of cell cycle entry

eliminates human leukemia stem cells in a mouse model of aml. *Nature Biotechnology*, 28:275 EP –.

Schnerch, D., Yalcintepe, J., Schmidts, A., Becker, H., Follo, M., Engelhardt, M., and Wäsch, R. (2012). Cell cycle control in acute myeloid leukemia. *American journal of cancer research*, 2(5):508–528. 22957304[pmid].

Shapiro, G. I. and Harper, J. W. (1999). Anticancer drug targets: cell cycle and checkpoint control. *The Journal of clinical investigation*, 104(12):1645–1653. 10606615[pmid].

Stark, G. R. and Taylor, W. R. (2004). *Checkpoint Controls and Cancer: Volume 1: Reviews and Model Systems*, chapter Analyzing the G2/M Checkpoint, pages 51–82. Humana Press, Totowa, NJ.

Stiehl, T., Baran, N., Ho, A. D., and Marciniak-Czochra, A. (2015). Cell division patterns in acute myeloid leukemia stem-like cells determine clinical course: A model to predict patient survival. *Cancer Research*, 75(6):940–949.

TCLv1.

Thorburn, A. (2007). Tumor necrosis factor-related apoptosis-inducing ligand (trail) pathway signaling. *Journal of Thoracic Oncology*, 2(6):461–465.

Wang, I.-C., Chen, Y.-J., Hughes, D., Petrovic, V., Major, M. L., Park, H. J., Tan, Y., Ackerson, T., and Costa, R. H. (2005). Forkhead box m1 regulates the transcriptional network of genes essential for mitotic progression and genes encoding the scf (skp2-cks1) ubiquitin ligase. *Molecular and cellular biology*, 25(24):10875–10894. 16314512[pmid].

Yu, J., Baron, V., Mercola, D., Mustelin, T., and Adamson, E. D. (2006). A network of p73, p53 and egr1 is required for efficient apoptosis in tumor cells. *Cell Death And Differentiation*, 14:436 EP –. Original Paper.

Author biography

Reza Asakereh A good guy with bad fruits