

RESEARCH ARTICLE

Low LEF1 expression is a biomarker of early T-cell precursor, an aggressive subtype of T-cell lymphoblastic leukemia

Mei Wang^{1*}, Chi Zhang²

1 School of Life Science, Peking University, Beijing, People's Republic of China, **2** Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing, People's Republic of China

* mei_wang@outlook.com**OPEN ACCESS**

Citation: Wang M, Zhang C (2020) Low LEF1 expression is a biomarker of early T-cell precursor, an aggressive subtype of T-cell lymphoblastic leukemia. PLoS ONE 15(5): e0232520. <https://doi.org/10.1371/journal.pone.0232520>

Editor: Francesco Bertolini, European Institute of Oncology, ITALY

Received: January 14, 2020

Accepted: April 15, 2020

Published: May 13, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0232520>

Copyright: © 2020 Wang, Zhang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All the data in this study is available in public domains. The URL links of the public datasets (GSE42328 and TARGET T-ALL) are as follows: a. TARGET T-ALL: <https://ocg.cancer.gov/programs/target>, dbGaP ID:

Abstract

Early T-cell precursor (ETP) is the only subtype of acute T-cell lymphoblastic leukemia (T-ALL) listed in the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia. Patients with ETP tend to have worse disease outcomes. ETP is defined by a series of immune markers. The diagnosis of ETP status can be vague due to the limitation of the current measurement. In this study, we performed unsupervised clustering and supervised prediction to investigate whether a molecular biomarker can be used to identify the ETP status in order to stratify risk groups. We found that the ETP status can be predicted by the expression level of Lymphoid enhancer binding factor 1 (LEF1) with high accuracy (AUC of ROC = 0.957 and 0.933 in two T-ALL cohorts). The patients with ETP subtype have a lower level of LEF1 comparing to the those without ETP. We suggest that incorporating the biomarker LEF1 with traditional immune-phenotyping will improve the diagnosis of ETP.

Introduction

Early T-cell precursor (ETP) is the only subtype of acute T-cell lymphoblastic leukemia (T-ALL) listed in the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia [1]. It is defined by immune-phenotyping as lack of CD1a and CD8, weak expression of CD5 and positive for one or more of the myeloid/stem cell markers [2]. The unique phenotype represents the characteristics of immature T-cells which correspond to the early stage of normal T-cell development. Contrary to “matured” T-ALLs, patients with ETP have worse disease outcomes [2–5]. Nowadays, most of the T-ALLs can be treated. However, the relapse rate is observed as up to 40% [6]. Targeting the specific ETP sub-group which has a higher relapse rate is crucial for treating T-ALLs.

As one of the routine examinations, the expression of relevant cell surface markers was measured by the flow cytometry [7]. However, the diagnosis can be varied depending on different pathologists and/or different parameters, for instance, the voltage setting. In order to improve the reliability of the diagnosis, Khogeer et al. developed a scoring system based on 11 surface markers to re-define the ETP subgroup [7]. Still, it is solely based on the measurement

phs000464 b.GSE42328: <https://www.ncbi.nlm.nih.gov/geo/> c.TCGA PanCancer: UCSC Xena platform (<https://pancanatlas.xenahubs.net>) Additionally, the authors have provided the Scripts data as a Supporting Information file.

Funding: This research is supported by National Natural Science Foundation of China (81900155) to M.W., the 100 Young Talents Program of Chinese Academy of Sciences and the Strategic Priority Research Program of Chinese Academy of Sciences (XDB26000000) to C.Z.

Competing interests: None of the authors has a relevant conflict of interest.

of flow cytometry. In this study, we investigate whether a molecular biomarker can be used to identify the ETP status in order to stratify risk groups.

Results and discussion

Inconsistent classification of ETP status yield by different classifiers

Currently, the most well accepted classification of ETP status was defined by immune phenotype proposed by Coustan-Smith et al. in 2009 [2]. According to their study, the sub-group of patients was discovered by unsupervised clustering using 35 differentially expressed genes in mice thymic early T-cell precursors. Patients in that group manifested a unique molecular character. Since then, this immune phenotype has been accepted as the definition of ETP [8]. How the ETP subtype was discovered indicates that the classification by immune-phenotyping and clustering by gene expression profile should reach an agreement. To evaluate if and to what extent the previously reported clustering methods could identify ETP status, we performed unsupervised clustering in two T-ALL datasets (TARGET T-ALL [9] and GSE42328 [10]). The characteristics of the two datasets were described in Materials and Method below.

Clustering by the whole transcriptome and by the 35-gene ETP signature were conducted on the two T-ALL cohorts (TARGET and GSE42328). Clustering based on the whole transcriptome is aiming to investigate whether distinct subtypes with molecular characteristics exist, and clustering based on the 35 ETP signature genes is to identify the ETP group in patients [2]. The subjects were divided into two clusters, but inconsistent classifications were yield by different approaches (Fig 1). Approximately 10–20% of T-ALLs were classified as ETP ($n = 19$, 10.05% in TARGET, $n = 10$, 18.87% in GSE42328) based on the immune-phenotype. However, more than half of the samples were predicted as ETP ($n = 100$, 60.61% in TARGET; $n = 32$, 60.38% in GSE42328) when clustering using the whole transcriptome, while 62 patients (37.58%) were predicted as ETP in TARGET and 34 (64.15%) in GSE32428 when clustering using the 35-gene panel. Thus, the classification of ETP by immune-phenotyping is inconsistent with that by unsupervised clustering of gene expression profile, which deviates from the original finding [8]. It indicates that the ETP status cannot be correctly classified by unsupervised clustering approaches using either the whole transcriptome or the 35-gene ETP signature.

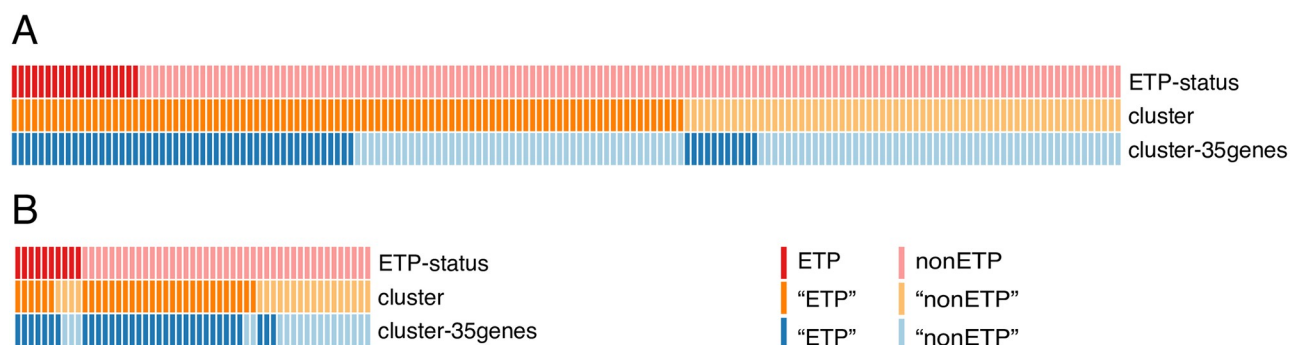


Fig 1. Classifications of ETP by immune-phenotype, transcriptome, and 35-gene ETP signature genes in TARGET T-ALL (A) and GSE42328 (B) datasets. Each column is a patient and each row represents a classification method. Patients were stratified to ETP (dark color) or nonETP (light color) groups based on immune phenotyping (first row), the whole transcriptome clustering (second row) or the 35-gene ETP signature clustering (third row), respectively.

<https://doi.org/10.1371/journal.pone.0232520.g001>

A potential biomarker of ETP status—LEF1

We then developed a lasso regression model to evaluate whether the ETP status could be predicted correctly based on transcriptomic data. The prediction accuracy of the model is first-rate that the area under the curve (AUC) of the receiver operating characteristic (ROC) achieved 0.957 (confidence interval (CI): 0.946–0.969, Fig 2A). It is concluded that the immune-phenotype defined ETP status can be predicted by gene expression profile.

To investigate which gene and to what extent it contributes to the prediction model, we looked into the features which were selected in the 100 rounds of outer cross-validation. The top 10 predictors were listed in Fig 2B. The coefficient of LEF1, CD5, HOXC10, and RSPO4 were high comparing to the other top predictors (Fig 2C). To predict the ETP status by expression level of individual genes, LEF1, CD5 and OGN achieved high accuracy in terms of AUC of ROC (Fig 2D).

Lymphoid enhancer binding factor 1 (LEF1) is the top player, which was selected 94 times out of 100 (Fig 2B). LEF1 is one of the most contributed variables. The coefficient of LEF1 ranged from -0.826 to -0.012 within the 94 times (Fig 2C). This gene is an important transcription factor in T-cell development and malignancy [11, 12]. It is also one of the most differentially expressed genes (DEGs) between ETP and nonETP in TARGET T-ALL dataset (S1 Table). There were 8761 genes out of 33038 differentially expressed and LEF1 was ranked 12th in the DEGs. Patients with ETP have a lower expression of LEF1 comparing to patients with “matured” T-cells.

We further tested whether LEF1 could predict the ETP status in the GSE42328 dataset. The prediction of ETP status by LEF1 expression level reached 0.933 (CI: 0.858–1, Fig 2E). The expression level of LEF1 in patients with early immature subtype is significantly lower than the cortical/mature group (Fig 2F). LEF1 was also found differentially expressed between ETP and nonETP subtypes in two previous Microarray datasets (GSE28703 and GSE8879) [2, 13].

To quantify the prediction accuracy in terms of sensitivity and specificity, we considered the identification of ETP by immune-phenotyping as the golden standard. The sensitivity (0.9 in TARGET, 1 in GSE42328) and specificity (1 in TARGET and 0.8 in GSE42328) of the prediction by LEF1 expression level are higher than either by the whole transcriptome clustering or by the 35-gene signature clustering in both TARGET T-ALL and GSE42328 (Table 1).

It is known that LEF1 is essential for T- and B-cell differentiation and lineage determination [14–16]. Yu et al. demonstrated that Tcf1 and Lef1 transcription factors are intrinsic required in leukemic stem cells (LSCs) self-renewal [17]. It indicates that LEF1 plays a key role in T-ALL tumorigenesis and provides an important biological evidence to support our hypothesis that expression level of LEF1 could predict the ETP status. In this study, the high-risk subtype of T-ALL showed a lower expression of LEF1. Jia et al. also found that LEF1 expression is a favorable prognostic factor in a pediatric cohort with 94 B-ALLs and 28 T-ALLs [18]. To the contrast, the expression of LEF1 in solid tumors is higher than the corresponding normal tissue. The overexpression of LEF1 was observed in most of the solid tumors (15 out of 24 types in the TCGA Pancancer database, S2 Table). Moreover, overexpression of LEF1 was associated with worse overall survival in adrenocortical cancer, kidney clear cell carcinoma, kidney papillary cell carcinoma, rectum adenocarcinoma, and uveal melanoma (S3 Table). However, in thymoma, a negative association with the LEF1 expression level and overall survival rate was observed. It is interesting that thymus is the organ where the development of T- and B-cell takes place (S3 Table). Giambra et al. reported that the activity of leukemia stem cells is dependent on Lef1 in the Wnt-active subpopulation of T-ALL [19]. In addition, it has also been found that low expression of LEF1 is associated with worse survival in acute myeloid leukemia

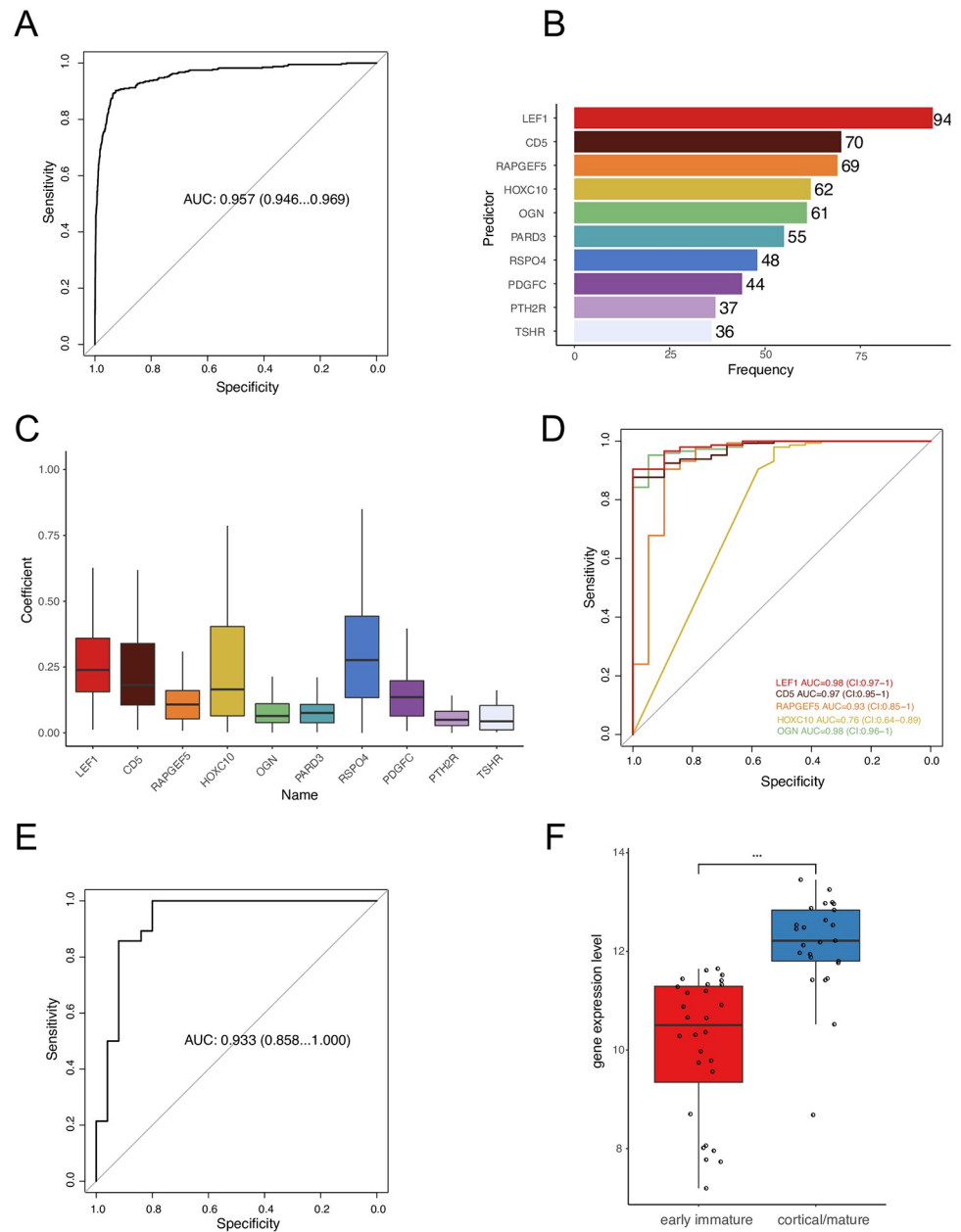


Fig 2. (A) AUC of ROC, predictions based on transcriptome against ETP status in the TARGET T-ALL dataset; **(B)** The top 10 predictors with non-zero coefficients in the 100 rounds of outer CV; **(C)** The coefficients of the top 10 predictors in the 100 rounds of outer CV. The y-axis is the absolute value of coefficient; **(D)** AUC of ROC, individual-gene expression against ETP status in the TARGET T-ALL dataset; **(E)** AUC of ROC, LEF1 expression against T-ALL subtypes (early immature, cortical/mature) in the GSE42328 dataset; **(F)** the expression level of LEF1 in T-ALL subtypes in the GSE42328 dataset. AUC, the area under the curve; ROC, the receiver operating characteristic; ETP, early T-cell precursor.

<https://doi.org/10.1371/journal.pone.0232520.g002>

Table 1. The accuracy (sensitivity and specificity) of predicting the ETP status (defined by immune markers).

		Sensitivity	Specificity
TARGET	cluster	1	0.45
	35-gene cluster	1	0.71
	LEF1	0.9	1
GSE42328	cluster	0.6	0.4
	35-gene cluster	0.3	0.6
	LEF1	1	0.8

<https://doi.org/10.1371/journal.pone.0232520.t001>

[20]. Overall, it suggests that LEF1 plays different roles in solid tumorigenesis and lymphoid malignancies. The underlying mechanism needs further investigation.

Conclusions

In this study, we investigated the molecular characteristics of the risky subtype of T-ALL-ETP. The current routine diagnosis of ETP is vague. We found that the ETP status can be predicted by the expression level of LEF1 with high accuracy. We propose that incorporating the biomarker LEF1 with traditional immune-phenotyping will improve the diagnosis of ETP.

Materials and methods

Study subjects

The TARGET T-ALL (n = 165, children and adolescent) dataset was used as the training set to develop the prediction model. Gene expression and clinical data were retrieved from Therapeutically Applicable Research To Generate Effective Treatments (TARGET, <https://ocg.cancer.gov/programs/target>, dbGaP ID: phs000464). The complete data of TARGET T-ALL in the original study were composed of 264 children and young adults diagnosed with T-ALL [9]. 190 cases had immune phenotypes in terms of ETP status, 25 out of which were labelled as "nearETP" and excluded due to the ambiguous definition, resulting in 165 cases used in this study. We compared the demographic and clinical characteristics of the selected samples to those in the original publication (Table 2). There are 33,038 genes totally in our dataset. The gene expression data were normalized using Trimmed Mean of M-values (TMM) [21] and log₂ transformed.

The GSE42328 adult T-ALL (n = 53) was the validation dataset. The GSE42328 dataset had 53 adult T-ALLs. The clinical characteristics of the samples were well described in their

Table 2. The demographic and basic clinical characteristics of the samples selected from TARGET T-ALL.

	Original study (n = 265)	Used this study (n = 165)
Age (Median, range)	9 (1–29)	9 (1–22)
Gender (male, n, %)	202 (76.23%)	122 (73.94%)
WBC (mean)	167.54	163.3
Bone marrow blasts at diagnosis (mean)	91.56	91.51
ETP status		
ETP (n)	19	19
nonETP (n)	146	146
nearETP (n)	25	NA
Unknown (n)	75	NA

<https://doi.org/10.1371/journal.pone.0232520.t002>

original study [10]. Data was downloaded from Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). In total, 47,212 probes were measured representing 31324 genes. The value of gene expression provided were quantile normalized and log₂ transformed.

The TCGA pan-cancer dataset was used to investigate the association of the expression level of LEF1 with survival in 32 types of tumor. Data were downloaded from Pan-Cancer Atlas Hub at UCSC Xena platform (<https://pancanatlas.xenahubs.net>) [22]. In total, 9110 samples were selected with sample type labeled as primary tumor. The RNA sequencing data were batch-effects normalized and log₂ transformed.

All the data in this study are available in public domains. Ethical approval is not applicable for this study.

Unsupervised clustering

Clustering by the whole transcriptome and by the 35-gene ETP signature were conducted respectively on the two T-ALL cohorts (TARGET and GSE42328). For the clustering by the whole 75 transcriptome, the top 5,000 most variable genes were selected [23]. For the clustering by 35-gene ETP signature, only the 35 genes (or the corresponding probes) were selected [2]. Hierarchy clustering with Pearson's correlation coefficient was performed 10,000 times to achieve a consensus [24]. The procedure was conducted by R package ConsensusClusterPlus v1.48 [23].

Supervised prediction

A regression model that uses L1-regularization (lasso)[25] was applied to fit the data and to predict the ETP status by transcriptomic gene expression data (RNA sequencing). Nested cross validation (CV) was conducted to evaluate the model performance. For outer CV, 80% of the patients (132 out of 165) were randomly assigned to the training set and the rest 20% (33 out of 165) to the test set. The outer CV was run 100 times. For inner CV, 5-fold CV was performed to optimize the tuning parameter lambda. The TARGET T-ALL (n = 165) dataset was used to develop the prediction model and GSE42328 adult T-ALL (n = 53) was the validation dataset. The prediction was conducted by R package glmnet v2.0 [26].

Receiver operating characteristic (ROC) curves of prediction against true ETP status were constructed. The area under the ROC curve (AUC) and 95% confidence interval (CI) were generated to compare model performance. The AUCs of ROC curves were compared by the DeLong test [27]. The decision boundary was determined at the point closest to the top-left part of the ROC curve using the R package pROC v1.15 [28]. Individuals were predicted as ETP if their predicted probability was larger than or equal to the cut-off point; otherwise, they were classified as nonETP.

Other statistics. Gene expression data (RNA sequencing) were compared between with ETP and without ETP to determine differentially expressed genes (DEGs). Differential expression analysis was conducted by R package DESeq2 v1.24 [29]. To test if LEF1 was abnormally expressed in tumor, nonparametric Wilcoxon two-group comparison was conducted in each tumor type in the TCGA Pan-Cancer dataset. Cox regression [30] was used to test the association between LEF1's expression level and overall survival in each tumor type.

To quantify the prediction accuracy in terms of sensitivity and specificity, we considered the identification of ETP by immune-phenotyping as the golden standard. Sensitivity is calculated as the proportion of patients with immune positive that are correctly identified by other methods (clustering or LEF1 expression level). Specificity is calculated as the proportion of patients with immune negative that are correctly identified as such.

Supporting information

S1 Table. Differential expression analysis on ETP status in TARGET T-ALL.

(XLSX)

S2 Table. Association analysis between LEF1 expression level and cancer in TCGA Pan-cancer dataset.

(XLSX)

S3 Table. Survival analysis on LEF1 expression level and patients' overall survival in TCGA Pan-cancer dataset.

(XLSX)

S1 File.

(ZIP)

Author Contributions

Conceptualization: Mei Wang, Chi Zhang.

Data curation: Mei Wang.

Formal analysis: Mei Wang.

Funding acquisition: Mei Wang, Chi Zhang.

Investigation: Mei Wang.

Methodology: Mei Wang.

Project administration: Mei Wang.

Visualization: Chi Zhang.

Writing – original draft: Mei Wang.

Writing – review & editing: Mei Wang, Chi Zhang.

References

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016; 127(20):2391–405. Epub 2016/04/14. <https://doi.org/10.1182/blood-2016-03-643544> PMID: 27069254.
2. Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, Pei D, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol*. 2009; 10(2):147–56. [https://doi.org/10.1016/S1470-2045\(08\)70314-0](https://doi.org/10.1016/S1470-2045(08)70314-0) PMID: 19147408
3. Ma M, Wang X, Tang J, Xue H, Chen J, Pan C, et al. Early T-cell precursor leukemia: a subtype of high risk childhood acute lymphoblastic leukemia. *Front Med*. 2012; 6(4):416–20. Epub 2012/10/16. <https://doi.org/10.1007/s11684-012-0224-4> PMID: 23065427.
4. Bond J, Graux C, Lhermitte L, Lara D, Cluzeau T, Leguay T, et al. Early Response-Based Therapy Stratification Improves Survival in Adult Early Thymic Precursor Acute Lymphoblastic Leukemia: A Group for Research on Adult Acute Lymphoblastic Leukemia Study. *J Clin Oncol*. 2017; 35(23):2683–91. Epub 2017/06/13. <https://doi.org/10.1200/JCO.2016.71.8585> PMID: 28605290.
5. Jain N, Lamb AV, O'Brien S, Ravandi F, Konopleva M, Jabbour E, et al. Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype. *Blood*. 2016; 127(15):1863–9. <https://doi.org/10.1182/blood-2015-08-661702> PMID: 26747249
6. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet*. 2008; 371(9617):1030–43. Epub 2008/03/25. [https://doi.org/10.1016/S0140-6736\(08\)60457-2](https://doi.org/10.1016/S0140-6736(08)60457-2) PMID: 18358930.
7. Khogeer H, Rahman H, Jain N, Angelova EA, Yang H, Quesada A, et al. Early T precursor acute lymphoblastic leukaemia/lymphoma shows differential immunophenotypic characteristics including

- frequent CD33 expression and in vitro response to targeted CD33 therapy. *British Journal of Haematology*. 2019; 0(0). <https://doi.org/10.1111/bjh.15960> PMID: 31115909
8. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009; 114(5):937–51. Epub 2009/04/10. <https://doi.org/10.1182/blood-2009-03-209262> PMID: 19357394.
 9. Liu Y, Easton J, Shao Y, Maciaszek J, Wang Z, Wilkinson MR, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat Genet*. 2017; 49(8):1211–8. Epub 2017/07/04. <https://doi.org/10.1038/ng.3909> PMID: 28671688
 10. Van Vlierberghe P, Ambesi-Impiombato A, De Keersmaecker K, Hadler M, Paietta E, Tallman MS, et al. Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. *Blood*. 2013; 122(1):74–82. Epub 2013/05/21. <https://doi.org/10.1182/blood-2013-03-491092> PMID: 23687089
 11. Yu S, Zhou X, Steinke Farrah C, Liu C, Chen S-C, Zagorodna O, et al. The TCF-1 and LEF-1 Transcription Factors Have Cooperative and Opposing Roles in T Cell Development and Malignancy. *Immunity*. 2012; 37(5):813–26. <https://doi.org/10.1016/j.immuni.2012.08.009> PMID: 23103132
 12. Staal FJT, M. Sen J. The canonical Wnt signaling pathway plays an important role in lymphopoiesis and hematopoiesis. *European Journal of Immunology*. 2008; 38(7):1788–94. <https://doi.org/10.1002/eji.200738118> PMID: 18581335
 13. Gutierrez A, Kentsis A, Sanda T, Holmfeldt L, Chen S-C, Zhang J, et al. The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood, The Journal of the American Society of Hematology*. 2011; 118(15):4169–73.
 14. Santiago L, Daniels G, Wang D, Deng FM, Lee P. Wnt signaling pathway protein LEF1 in cancer, as a biomarker for prognosis and a target for treatment. *Am J Cancer Res*. 2017; 7(6):1389–406. Epub 2017/07/04. PMID: 28670499
 15. Reya T, O’Riordan M, Okamura R, Devaney E, Willert K, Nusse R, et al. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity*. 2000; 13(1):15–24. [https://doi.org/10.1016/s1074-7613\(00\)00004-2](https://doi.org/10.1016/s1074-7613(00)00004-2) PMID: 10933391
 16. Weber BN, Chi AW-S, Chavez A, Yashiro-Ohtani Y, Yang Q, Shestova O, et al. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature*. 2011; 476(7358):63–8. <https://doi.org/10.1038/nature10279> PMID: 21814277
 17. Yu S, Li F, Xing S, Zhao T, Peng W, Xue H-H. Hematopoietic and leukemic stem cells have distinct dependence on Tcf1 and Lef1 transcription factors. *Journal of Biological Chemistry*. 2016; 291(21):11148–60. <https://doi.org/10.1074/jbc.M116.717801> PMID: 27044748
 18. Jia M, Zhao H-Z, Shen H-P, Cheng Y-P, Luo Z-B, Li S-S, et al. Overexpression of lymphoid enhancer-binding factor-1 (LEF1) is a novel favorable prognostic factor in childhood acute lymphoblastic leukemia. *International journal of laboratory hematology*. 2015; 37(5):631–40. <https://doi.org/10.1111/ijlh.12375> PMID: 25955539
 19. Giambra V, Gusscott S, Gracias D, Song R, Weng AP. Lef1 Is a Critical Mediator of Wnt/ β -Catenin Signaling in T-Cell Acute Lymphoblastic Leukemia (T-ALL). *Blood*. 2016; 128(22):5083.
 20. Fu Y, Zhu H, Wu W, Xu J, Chen T, Xu B, et al. Clinical significance of lymphoid enhancer-binding factor 1 expression in acute myeloid leukemia. *Leukemia & Lymphoma*. 2014; 55(2):371–7. <https://doi.org/10.3109/10428194.2013.805759> PMID: 23713453
 21. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010; 11(3):R25. <https://doi.org/10.1186/gb-2010-11-3-r25> PMID: 20196867
 22. Goldman M, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. *bioRxiv*. 2019:326470. <https://doi.org/10.1101/326470>
 23. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010; 26(12):1572–3. <https://doi.org/10.1093/bioinformatics/btq170> PMID: 20427518
 24. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn*. 2003; 52(1–2):91–118.
 25. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996; 58(1):267–88.
 26. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010; 33(1):1–22. Epub 2010/09/03. PMID: 20808728
 27. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988; 44(3):837–45. Epub 1988/09/01. PMID: 3203132.

28. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011; 12:77. <https://doi.org/10.1186/1471-2105-12-77> PMID: 21414208
29. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
30. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1972; 34(2):187–202.