Comparative gene expression profiling of T-cell acute lymphoblastic leukemias (T-ALL) expressing the T-cell receptor $\alpha\beta$ or $\gamma\delta$ (T-ALL $\alpha\beta$ / T-ALL $\gamma\delta$)

Shahan Mamoor, MS

Abstract

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive form of leukemia with inferior treatment outcomes. The T-cell receptor (TCR) exists in two major forms: the $\alpha\beta$ TCR or the $\gamma\delta$ TCR, and 20-35% of T-ALL cases express either the $\alpha\beta$ TCR or the $\gamma\delta$ TCR (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$). Using a published dataset from a cohort of 14 TCR+ T-ALL patients, I found a series of genes that are differentially expressed among patients T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$. Any number of these differentially expressed genes may be a scientifically and/or clinically actionable target in TCR+ T-ALL.

Keywords: T-ALL, $\gamma\delta$ TCR, T $\gamma\delta$, δ T-cell, transcriptome, differentially expressed genes, T-ALL $\alpha\beta$, T-ALL $\gamma\delta$.

Introduction

Gene expression profiling in malignancies can be useful for patient stratification in diagnosis, prognosis, and treatment (1). T-ALL is an aggressive form of leukemia frequently featuring central nervous system infiltration that requires identification of novel therapeutic targets (2). The phenotype of lymphoblasts is TdT positive and variable for: CD1a, CD2, CD3, CD4, CD5, CD7 and CD8 (3,4). Expression of the $\alpha\beta$ TCR or the $\gamma\delta$ TCR occurs in 20-35% of T-ALL cases (3,4). To facilitate the discovery of novel therapeutic targets in TCR+ T-ALL, I performed a comparative gene expression analysis on a dataset from T-ALL patients with confirmed expression of $\alpha\beta$ TCR or $\gamma\delta$ TCR. These data can serve as a resource for clinicians and scientists working to increase quality of life and extend survival in patients with T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$.

Methods

Comparative Gene Expression Profiling

GEO (Gene Expression Omnibus) Dataset GSE37389 was utilized for this analysis (31). GEO2R identifies the most differentially expressed transcripts in a given comparison by ranking the log-fold change of RNA expression value between samples when provided with a dataset. In the first analysis, genes were identified blindly by their propensity to segregate based on expression level (e.g., “all up” versus “all down”)-in the second, pattern of expression (e.g., broad versus narrow, positive and negative versus negative only) was used to hand-pick computationally identified genes of interest.

Statistics

Statistical analyses comparing the mean RNA expression value of each differentially expressed transcript from leukemic blasts from the peripheral blood of 14 patients with T-ALL were performed using GraphPad Prism Version 8.0. All of the genes presented here are considered significant (p -value less than 0.05) by GEO2R with respect to differential expression between T-ALL types, but a second statistical analysis comparing was performed (Unpaired t test [with two-tailed p -value]) and shown alongside each graph. To maximize retrieval of transcripts with differential expression patterns, a second analysis was performed with log-transformation of mRNA expression, as well as adjustment of p -values by multiple methods (Benjamini and Hochberg, Benjamini and Yekutieli, Bonferroni, and Hommel). The “Submitter supplied” category of platform annotation was used to identify gene/RNA names. P -values for each comparison are denoted on each graph.

Results

Kinases and phosphatases represent ideal pharmacological therapeutic avenues because their enzymatic functions can be targeted with small molecules (5). Kinase inhibition in cancers has proved clinically efficacious (7, 8, 9, 10, 11, 12, 13, 14).

Ppm1j was expressed at higher levels in patients with T-ALL $\gamma\delta$ versus patients with T-ALL $\alpha\beta$ and this trended towards statistical significance (Figure 1; $p=0.0734$). Protein phosphatase regulatory subunit Ppp1r14a was expressed at higher levels in patients with T-ALL $\gamma\delta$ and this approached statistical significance (Figure 1; $p=0.0566$). Another protein phosphatase regulatory subunit, Ppp4r1 was expressed at higher levels in T-ALL $\gamma\delta$ but this was not statistically significant (Figure 1; $p=0.1200$).

The neurotrophic receptor tyrosine kinase NTRK3 was differentially expressed in this cohort of fourteen T-ALL patients, at higher levels in T-ALL $\gamma\delta$ relative to T-ALL $\alpha\beta$ (Figure 1; $p=0.0596$).

On the contrary, another differentially expressed kinase salt-inducible kinase Sik3 was expressed at higher levels in patients with T-ALL $\alpha\beta$ (Figure 1; $p=0.0558$).

Together, these data show that enzymes involved in phosphorylation are among the most differentially expressed genes between patients with T-ALL $\alpha\beta$ compared to patients with T-ALL $\gamma\delta$.

Precise identification of the cell surface receptors that cancer cells express is critical for diagnostic purposes and can allow the stratification of patients into more refined prognostic groups. CD117 (Kit, or stem cell factor receptor) was among the most differentially expressed gene between cases of T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$ (Figure 2), and expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ ($p=0.0476$). CD1e and CD68 were expressed in a bimodal fashion in patients with TCR+ T-ALL, with high expression of CD1e and CD68 found in patients with T-ALL $\gamma\delta$ but not in patients with T-ALL $\alpha\beta$ (Figure 2; $p=0.1263$ and $p=.1119$). CD80 was expressed at significantly lower levels in patients with T-ALL $\alpha\beta$ (Figure 2; $p=0.0250$).

Differential expression of developmental transcription factors is a paradigm in hematopoietic cancers (23, 24). Eight transcription factors were differentially expressed in a group of fourteen patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$. The homeobox transcription factor HoxA5 and HoxA10 were among the genes most differentially expressed in patients with T-ALL $\gamma\delta$ when compared to patients with T-ALL $\alpha\beta$, and both expressed at higher levels in patients with T-ALL $\gamma\delta$. Higher expression of HoxA5 and HoxA10 in patients with T-ALL $\gamma\delta$ approached statistical significance (Figure 3; $p=0.0773$ and $p=0.0724$, respectively). I recently reported that HoxA10 is one of the most differentially expressed genes in healthy T $\gamma\delta$ cells when compared to CD4+ and CD8+ T $\alpha\beta$ counterparts.

Six other transcription factors were differentially expressed at significantly higher levels when comparing patients with T-ALL $\alpha\beta$ to patients with T-ALL $\gamma\delta$. Mybl1 (also known as A-Myb), Ets-like ternary complex factor Elk3, and SRY-Box 30 (Sox30) were expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 3; $p=0.0595$, $p=0.0689$ and $p=0.0202$, respectively). On the other hand, expression of Zinc finger and BTB-domain containing protein Zbtb18, Ets variant 5 (Etv5) and the putative transcription factor Znf525 was higher in patients with T-ALL $\alpha\beta$ (Figure 3; $p=0.0400$, $p=0.0844$, $p=0.0674$, and $p=0.0456$, respectively).

Cell adhesion molecules were also among the genes most differentially expressed when comparing the transcriptomes of patients with T-ALL $\alpha\beta$ to patients with T-ALL $\gamma\delta$.

Integrin alpha x (Itgax) or CD11c is an alpha integrin that heterodimerizes with CD18 to form the leukocyte-specific integrin complement component 3 receptor 4 subunit (CR4). Itgax is expressed on monocytes, macrophages, dendritic cells, and some B-cells (25). Integrin alpha 6, also known as CD49f or VLA-6, is expressed by embryonic, hematopoietic and neural stem cells (26). Itgax and Itga6 were both expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 4; $p=0.0935$ and $p=0.0720$, respectively).

Adhesion molecules ESAM, Icam1 and Mcam were all expressed at higher levels in T-ALL $\gamma\delta$ (Figure 4A; $p=0.0482$, $p=0.0702$, $p=0.0772$), respectively, while Tcam1p was expressed at higher levels in patients with T-ALL $\alpha\beta$ (Figure 4A; $p=0.1004$).

Interestingly, the Kell blood group precursor *Xkr* was among the top 10 genes most differentially expressed in patients with T-ALL $\gamma\delta$ versus patients with T-ALL $\alpha\beta$, expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 5; $p=0.0017$). Expression of another red cell-associated gene, blood group antigen *Xg*, was higher in patients with T-ALL $\alpha\beta$ (Figure 5; $p=0.0758$).

Nucleic acid polymerases were represented in the genes most differentially expressed in patients with T-ALL $\alpha\beta$ versus patients with T-ALL $\gamma\delta$. RNA Polymerase *Polr3c* was expressed at higher levels in T-ALL $\alpha\beta$ (Figure 6; $p=0.0504$), and DNA Polymerase *Pol2rd* trended towards higher expression in T-ALL $\alpha\beta$ (Figure 6; $p=0.1562$).

Two TATA-binding associated factors were also among the genes most differentially expressed in the transcriptome of patients with $\gamma\delta$ TCR+ T-ALL compared to the transcriptome of patients with $\alpha\beta$ TCR+ T-ALL. *Taf5l* and *Taf4b* were expressed in an opposing fashion in patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$. *Taf5l* expression was significantly higher in patients with T-ALL $\alpha\beta$ (Figure 6; $p=0.0305$), while *Taf4b* was expressed at significantly higher levels in T-ALL $\gamma\delta$ (Figure 6; $p=0.0564$).

Arid5b was expressed at higher levels in patients with T-ALL $\alpha\beta$ ($p=0.0704$) while *Zcchc4*, a methyltransferase involved in post-transcriptional modification of ribosomal RNAs with N⁶-methyladenosine was also expressed at higher levels in patients with T-ALL $\alpha\beta$ ($p=0.0659$) (27).

In addition to machinery involved in transcription, specific non-coding RNA were among the most differentially expressed transcripts when comparing the global gene expression profiles of patients with T-ALL $\gamma\delta$ versus patients with T-ALL $\alpha\beta$.

The non-coding RNA *Linc00852* was expressed at significantly higher levels in patients with T-ALL $\alpha\beta$ than in patients with T-ALL $\gamma\delta$ (Figure 7; $p=0.0237$).

Another non-coding RNA, deleted in leukemia (*Dleu1*), was expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 7; $p=0.0061$) (28, 29, 30).

Anti-sense transcripts were differentially transcribed in TCR+ T-ALL patients expressing either the $\alpha\beta$ TCR or the $\gamma\delta$ TCR. Antisense transcript for *B4galt1* (*B4galt1-AS1*) and the antisense transcript for *Rap2c* (*Rap2c-AS1*) were among the most differentially expressed genes when comparing the transcriptomes of patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$. *B4galt1-AS1* was expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 7; $p=0.0343$), while *Rap2c-AS1* was expressed at significantly higher levels in patients with T-ALL $\alpha\beta$ (Figure 7; $p=0.0406$).

In addition, retinitis pigmentosa pseudogene *Rp9p* was also among the most differentially expressed genes when comparing the global gene expression profiles of patients with T-ALL $\alpha\beta$ and patients with T-ALL $\gamma\delta$, expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 7, $p=0.0143$).

A number of receptors at the cell surface were represented in the genes that were most differentially expressed in patients with either $\gamma\delta$ TCR+ or $\alpha\beta$ TCR+ T-ALL. The prokineticin receptor, *PROKR* is a G-protein coupled receptor that has pro-angiogenic effects and displays broad expression, including the hematopoietic system (6). *PROKR* was one of the genes most differentially expressed among patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$ (Figure 8; $p=0.0407$).

Lilrb1, a leukocyte immunoglobulin-type receptor was among the most differentially expressed genes between T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$, with (Figure 9; $p=0.0673$).

The Wnt pathway receptor Frizzled3 was also among the most differentially expressed genes in patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$, and expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 10; $p=0.0167$).

C-type lectin receptor Clec2d was expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 11; $p=0.0760$).

Together these data show that a number of signaling receptors at the plasma membrane are among the most differentially expressed genes when comparing the gene expression profiles of patients with T-ALL $\alpha\beta$ and patients with T-ALL $\gamma\delta$.

A number of molecules which did not readily fit into clusters of genes were also among the most differentially expressed genes when comparing the transcriptomes of patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$.

Carbohydrate sulfotransferases Chst1 and Chst7 were both expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 12; $p=0.0100$ and $p=0.0777$, respectively).

Interestingly, a taste receptor Tas2r45 was expressed at significantly lower levels in patients with T-ALL $\gamma\delta$ (Figure 13; $p=0.0189$).

E3 Ligase Trim71 was expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 14; $p=0.1104$).

MHC Class II molecule HLA-DQB1 was also expressed at higher levels in patients with T-ALL $\gamma\delta$ (Figure 15; $p=0.0843$).

G protein-coupled receptor Gpr65 was expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ (Figure 16; $p=0.0385$).

Finally, two proteins involved in the nuclear pore complex, Nup43 and Nup62CL were expressed at higher levels in patients with T-ALL $\alpha\beta$ ($p=0.0551$ and $p=0.1066$, respectively).

Conclusions

I used a publicly available transcriptome dataset from 14 patients with acute lymphoblastic T-cell leukemia (T-ALL) (31) to identify the most differentially expressed genes when comparing patients with T-ALL expressing the $\gamma\delta$ TCR (T-ALL $\gamma\delta$) to patients with T-ALL expressing the $\alpha\beta$ TCR (T-ALL $\alpha\beta$).

ETV6 has been described in relation to cancers before. Fusion of NTRK3 to ETV6 is a defining event in congenital fibrosarcoma and in secretory breast cancers (17, 18). Evidence of clinically efficacious targeting of NTRK3 in cancers includes one case of ETV6-NTRK3 infantile fibrosarcoma that was successfully treated with LOXO-1, a tropomyosin related kinase inhibitor (19). Additionally, a pan-TRK kinase inhibitor expressing ETV6-NTRK3 has displayed anti-tumor activity in acute myeloid leukemia cell lines and animal models (37).

In this analysis, the stem cell factor receptor Kit/CD117 was among the most differentially expressed genes when comparing the transcriptomes of patients with T-ALL $\gamma\delta$ to that of patients with T-ALL $\alpha\beta$. It is not known whether the lymphoblasts of patients with T-ALL $\gamma\delta$ are more stem-like, or have more self-renewal capacity than the lymphoblasts of patients with T-ALL $\alpha\beta$. This could be experimentally tested in vivo by comparing the level of engraftment in secondary transplants of nod/scid IL2R γ $-/-$ (non-obese diabetic severe combined immunodeficiency IL2 receptor γ null) mice transplanted with the bone marrow or purified lymphoblasts of patients with T-ALL $\gamma\delta$ versus that of patients with T-ALL $\alpha\beta$ (32, 33, 34).

Several molecules involved in cell adhesion were also significantly differentially transcribed between patients with T-ALL $\gamma\delta$ and patients with T-ALL $\alpha\beta$. Itga6 has previously been reported to be over-expressed in T-ALL (by the authors of the study from which this data was obtained, [31]). It may play a role in the maintenance of the stem cell state as it is expressed in higher amounts in human embryonic stem cells as opposed to mesenchymal stem cells and this higher expression is decreased upon hES cell differentiation (38). Oct4 and Sox2 directly occupy the promoter of Itga6, and knockdown of Itga6 leads to up-regulation of p53 (38).

I recently reported that Wnt pathway receptors Frizzled2 and Frizzled5 were among the receptors that distinguish the transcriptional landscape of T $\gamma\delta$ from that of CD4 $^{+}$ and CD8 $^{+}$ T $\alpha\beta$ (39). The finding here that Frizzled3 is over-expressed in T-ALL $\gamma\delta$ suggests that selective targeting of the Wnt pathway by inhibition of Frizzled3 may represent a novel therapeutic avenue in T-ALL $\gamma\delta$ and merits further investigation.

I also recently reported that Lilrb1 was among the receptors most differentially expressed by T $\gamma\delta$ cells when compared to T $\alpha\beta$ CD4 $^{+}$ and CD8 $^{+}$ cells (39). The finding here that Lilrb1 is expressed at significantly higher levels in patients with T-ALL $\gamma\delta$ when compared to patients with T-ALL $\alpha\beta$ suggests that differential expression of Lilrb1 is linked to the $\gamma\delta$ T-cell state and not some other property of the transformed T-ALL $\gamma\delta$ cell.

One recent study compared microarray data from 18 lung adenocarcinoma samples to that of 18 spinal metastases from patients with lung adenocarcinoma and found that Linc00852 was associated with spinal metastases in patients with lung adenocarcinoma (35). It is not clear what relation (if any) this bears to the differential and higher expression of Linc00852 in patients with T-ALL $\alpha\beta$ compared to patients with T-ALL $\gamma\delta$.

Seven transcription factors (HoxA5, HoxA10, Sox30, Yes1, Etv5, Elk3 and Zbtb18) were among the most differentially expressed genes when comparing the transcriptomes of patients with T-ALL $\gamma\delta$ to that of patients with T-ALL $\alpha\beta$. HoxA5, HoxA10, Sox30 and Elk3 were expressed at higher levels in patients with T-ALL $\gamma\delta$ while Yes1, Etv5 and Zbtb18 were expressed at higher levels in patients with T-ALL $\alpha\beta$.

Approximately 30% of patients with T-cell acute lymphoblastic leukemia express a T-cell receptor of either $\gamma\delta$ or $\alpha\beta$ clonotype. In this study, expression of the $\gamma\delta$ TCR was associated with significant gene expression differences and some of these differences may be of relevance to the way TCR $^{+}$ T-ALL is diagnosed and treated.

References

1. Mirji, G., Bhat, J., Kode, J., Banavali, S., Sengar, M., Khadke, P., Sait, O. and Chiplunkar, S., 2016. Risk stratification of T-cell Acute Lymphoblastic Leukemia patients based on gene expression, mutations and copy number variation. *Leukemia research*, 45, pp.33-39.
2. Van Vlierberghe, P. and Ferrando, A., 2012. The molecular basis of T cell acute lymphoblastic leukemia. *The Journal of clinical investigation*, 122(10), pp.3398-3406.
3. Buccheri, V., Matutes, E., Dyer, M.J. and Catovsky, D., 1993. Lineage commitment in biphenotypic acute leukemia. *Leukemia*, 7(6), pp.919-927.
4. Flohr, T., Schrauder, A., Cazzaniga, G., Panzer-Grümayer, R., Van der Velden, V., Fischer, S., Stanulla, M., Basso, G., Niggli, F.K., Schäfer, B.W. and Sutton, R., 2008. Minimal residual disease-directed risk stratification using real-time quantitative PCR analysis of

- immunoglobulin and T-cell receptor gene rearrangements in the international multicenter trial AIEOP-BFM ALL 2000 for childhood acute lymphoblastic leukemia. *Leukemia*, 22(4), p. 771.
5. Zhang, J., Yang, P.L. and Gray, N.S., 2009. Targeting cancer with small molecule kinase inhibitors. *Nature reviews cancer*, 9(1), p.28.
6. Lin, D.C.H., Bullock, C.M., Ehlert, F.J., Chen, J.L., Tian, H. and Zhou, Q.Y., 2002. Identification and molecular characterization of two closely related G protein-coupled receptors activated by prokineticins/endocrine gland vascular endothelial growth factor. *Journal of Biological Chemistry*, 277(22), pp.19276-19280.
7. Llovet, J.M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.F., De Oliveira, A.C., Santoro, A., Raoul, J.L., Forner, A. and Schwartz, M., 2008. Sorafenib in advanced hepatocellular carcinoma. *New England journal of medicine*, 359(4), pp.378-390.
8. Motzer, R.J., Hutson, T.E., Tomczak, P., Michaelson, M.D., Bukowski, R.M., Rixe, O., Oudard, S., Negrier, S., Szczyluk, C., Kim, S.T. and Chen, I., 2007. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *New England Journal of Medicine*, 356(2), pp.115-124.
9. Grothey, A., Van Cutsem, E., Sobrero, A., Siena, S., Falcone, A., Ychou, M., Humblet, Y., Bouché, O., Mineur, L., Barone, C. and Adenis, A., 2013. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *The Lancet*, 381(9863), pp.303-312.
10. Saglio, G., Kim, D.W., Issaragrisil, S., Le Coutre, P., Etienne, G., Lobo, C., Pasquini, R., Clark, R.E., Hochhaus, A., Hughes, T.P. and Gallagher, N., 2010. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *New England Journal of Medicine*, 362(24), pp.2251-2259.
11. Kantarjian, H., Giles, F., Wunderle, L., Bhalha, K., O'Brien, S., Wassmann, B., Tanaka, C., Manley, P., Rae, P., Mietlowski, W. and Bochinski, K., 2006. Nilotinib in imatinib-resistant CML and Philadelphia chromosome–positive ALL. *New England Journal of Medicine*, 354(24), pp.2542-2551.
12. Demetri, G.D., Von Mehren, M., Blanke, C.D., Van den Abbeele, A.D., Eisenberg, B., Roberts, P.J., Heinrich, M.C., Tuveson, D.A., Singer, S., Janicek, M. and Fletcher, J.A., 2002. Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *New England Journal of Medicine*, 347(7), pp.472-480.
13. Druker, B.J., Guilhot, F., O'Brien, S.G., Gathmann, I., Kantarjian, H., Gattermann, N., Deininger, M.W., Silver, R.T., Goldman, J.M., Stone, R.M. and Cervantes, F., 2006. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine*, 355(23), pp.2408-2417.
14. Talpaz, M., Shah, N.P., Kantarjian, H., Donato, N., Nicoll, J., Paquette, R., Cortes, J., O'Brien, S., Nicaise, C., Bleickardt, E. and Blackwood-Chirchir, M.A., 2006. Dasatinib in imatinib-resistant Philadelphia chromosome–positive leukemias. *New England Journal of Medicine*, 354(24), pp.2531-2541.
15. Lu, Z., Wan, G., Guo, H., Zhang, X. and Lu, X., 2013. Protein phosphatase 1 inhibits p53 signaling by dephosphorylating and stabilizing Mdmx. *Cellular signalling*, 25(4), pp.796-804.

16. Kashiwaba, M., Katsura, K., Ohnishi, M., Sasaki, M., Tanaka, H., Nishimune, Y., Kobayashi, T. and Tamura, S., 2003. A novel protein phosphatase 2C family member (PP2C ζ) is able to associate with ubiquitin conjugating enzyme 9. *FEBS letters*, 538(1-3), pp.197-202.
17. Knezevich, S.R., McFadden, D.E., Tao, W., Lim, J.F. and Sorensen, P.H., 1998. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nature genetics*, 18(2), p.184.
18. Tognon, C., Knezevich, S.R., Huntsman, D., Roskelley, C.D., Melnyk, N., Mathers, J.A., Becker, L., Carneiro, F., MacPherson, N., Horsman, D. and Poremba, C., 2002. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer cell*, 2(5), pp.367-376.
19. Nagasubramanian, R., Wei, J., Gordon, P., Rastatter, J.C., Cox, M.C. and Pappo, A., 2016. Infantile fibrosarcoma with NTRK3–ETV6 fusion successfully treated with the tropomyosin-related kinase inhibitor LOXO-101. *Pediatric blood & cancer*, 63(8), pp.1468-1470.
20. Brechmann, M., 2010. Identification and characterization of protein phosphatase 4 regulatory subunit 1 (PP4R1) as a suppressor of NF-kappaB in T lymphocytes and T cell lymphomas (Doctoral dissertation).
21. Kraus, T.S., Sillings, C.N., Saxe, D.F., Li, S. and Jaye, D.L., 2010. The role of CD11c expression in the diagnosis of mantle cell lymphoma. *American journal of clinical pathology*, 134(2), pp.271-277.
22. Yu, K.R., Yang, S.R., Jung, J.W., Kim, H., Ko, K., Han, D.W., Park, S.B., Choi, S.W., Kang, S.K., Schöler, H. and Kang, K.S., 2012. CD49f enhances multipotency and maintains stemness through the direct regulation of OCT4 and SOX2. *Stem cells*, 30(5), pp.876-887.
23. Zheng, R. and Blobel, G.A., 2010. GATA transcription factors and cancer. *Genes & cancer*, 1(12), pp.1178-1188.
24. Darnell Jr, J.E., 2002. Transcription factors as targets for cancer therapy. *Nature Reviews Cancer*, 2(10), p.740.
25. Bradford, B.M., Sester, D.P., Hume, D.A. and Mabbott, N.A., 2011. Defining the anatomical localisation of subsets of the murine mononuclear phagocyte system using integrin alpha X (Itgax, CD11c) and colony stimulating factor 1 receptor (Csf1r, CD115) expression fails to discriminate dendritic cells from macrophages. *Immunobiology*, 216(11), pp.1228-1237.
26. Yang, R., Zheng, Y., Burrows, M., Liu, S., Wei, Z., Nace, A., Guo, W., Kumar, S., Cotsarelis, G. and Xu, X., 2014. Generation of folliculogenic human epithelial stem cells from induced pluripotent stem cells. *Nature communications*, 5, p.3071.
27. Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S.K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y.G. and Lan, F., 2019. N 6-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. *Nature chemical biology*, 15(1), p.88.
28. Wang, L.L., Sun, K.X., Wu, D.D., Xiu, Y.L., Chen, X., Chen, S., Zong, Z.H., Sang, X.B., Liu, Y. and Zhao, Y., 2017. DLEU 1 contributes to ovarian carcinoma tumorigenesis and development by interacting with miR-490-3p and altering CDK 1 expression. *Journal of cellular and molecular medicine*, 21(11), pp.3055-3065.

29. Lee, S., Luo, W., Shah, T., Yin, C., O'Connell, T., Chung, T.H., Perkins, S.L., Miles, R.R., Ayello, J., Morris, E. and Harrison, L., 2017. The effects of DLEU1 gene expression in Burkitt lymphoma (BL): potential mechanism of chemoimmunotherapy resistance in BL. *Oncotarget*, 8(17), p.27839.
30. Shao, W., Li, Y., Chen, F., Jia, H., Jia, J. and Fu, Y., 2018. Long non-coding RNA DLEU1 contributes to the development of endometrial cancer by sponging miR-490 to regulate SP1 expression. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, 73(7), pp. 379-385.
31. Mirji, G., Bhat, J., Kode, J., Banavali, S., Sengar, M., Khadke, P., Sait, O. and Chiplunkar, S., 2016. Risk stratification of T-cell Acute Lymphoblastic Leukemia patients based on gene expression, mutations and copy number variation. *Leukemia research*, 45, pp.33-39.
32. Cox, C.V., Martin, H.M., Kearns, P.R., Virgo, P., Evely, R.S. and Blair, A., 2007. Characterization of a progenitor cell population in childhood T-cell acute lymphoblastic leukemia. *Blood*, 109(2), pp.674-682.
33. Chiu, P.P., Jiang, H. and Dick, J.E., 2010. Leukemia-initiating cells in human T-lymphoblastic leukemia exhibit glucocorticoid resistance. *Blood*, 116(24), pp.5268-5279.
34. Ma, F., Manabe, A., Wang, D., Ito, M., Kikuchi, A., Wada, M., Ohara, A., Hosoya, R., Asano, S. and Tsuji, K., 2002. Growth of human T cell acute lymphoblastic leukemia lymphoblasts in NOD/SCID mouse fetal thymus organ culture. *Leukemia*, 16(8), p.1541.
35. Liu, P., Wang, H., Liang, Y., Hu, A., Xing, R., Jiang, L., Yi, L. and Dong, J., 2018. LINC00852 Promotes Lung Adenocarcinoma Spinal Metastasis by Targeting S100A9. *Journal of Cancer*, 9(22), p.4139.
36. Fortunel, N.O., Otu, H.H., Ng, H.H., Chen, J., Mu, X., Chevassut, T., Li, X., Joseph, M., Bailey, C., Hatzfeld, J.A. and Hatzfeld, A., 2003. Comment on'"Stemness': transcriptional profiling of embryonic and adult stem cells" and" a stem cell molecular signature"(I). *Science*, 302(5644), pp.393-393.
37. Smith, K.M., Fagan, P.C., Pomari, E., Germano, G., Frasson, C., Walsh, C., Silverman, I., Bonvini, P. and Li, G., 2018. Antitumor Activity of Entrectinib, a Pan-TRK, ROS1, and ALK Inhibitor, in ETV6-NTRK3–Positive Acute Myeloid Leukemia. *Molecular cancer therapeutics*, 17(2), pp.455-463.
38. Yu, K.R., Yang, S.R., Jung, J.W., Kim, H., Ko, K., Han, D.W., Park, S.B., Choi, S.W., Kang, S.K., Schöler, H. and Kang, K.S., 2012. CD49f enhances multipotency and maintains stemness through the direct regulation of OCT4 and SOX2. *Stem cells*, 30(5), pp.876-887.
39. Mamoor, S., 2019. A new family of markers to characterize the human $\gamma\delta$ T-cell lineage.

Figure 1: Kinases and phosphatases are differentially expressed in patients with T-ALL $\gamma\delta$ versus patients with T-ALL $\alpha\beta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 2: Differential expression of cluster of differentiation markers in T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 3: Eight transcription factors are differentially expressed in a group of fourteen patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 4: Differential expression of cell adhesion molecules in T-ALL $\alpha\beta$ / T-ALL $\gamma\delta$

Figure 4A: Integrins Itgax and Itga6 are expressed at higher levels in patients with T-ALL $\gamma\delta$ compared to patients with T-ALL $\alpha\beta$

Figure 4B: ESAM, Icam1 and Mcam are expressed at significantly higher levels in T-ALL $\gamma\delta$, while Tcam1p is differentially expressed in patients with T-ALL $\alpha\beta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 5: The Kell blood group precursor Xkr is among the Top 10 genes most differentially expressed in T-ALLγδ, while higher expression of blood group antigen Xg is unique to T-ALLαβ.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALLαβ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALLγδ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALLαβ or T-ALLγδ) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 6: The transcriptional apparatus is differentially expressed among patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$

Figure 6A: TATA-binding associated factors Taf5l and Taf4b are expressed in an opposing fashion in T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$, respectively.

Figure 6B: RNA Polymerase Polr3c is expressed at higher levels in T-ALL $\alpha\beta$, and DNA Polymerase Pol2rd trends towards higher expression in T-ALL $\alpha\beta$

Figure 6C: Arid5b and Zcchc4 are expressed at higher levels in patients with T-ALL $\alpha\beta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 7: Differential expression of five non-coding RNAs in patients with T-ALL $\alpha\beta$ compared to patients with T-ALL $\gamma\delta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 8: PROKR1 is expressed at significantly higher levels in patients with T-ALL $\gamma\delta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 9: Leukocyte immunoglobulin-type receptor LILRB1 is expressed at higher levels in patients with T-ALL $\gamma\delta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 10: Wnt receptor Frizzled3 is among the top ten most differentially expressed genes in patients with T-ALL $\alpha\beta$ and T-ALL $\gamma\delta$, and expressed at significantly higher levels in patients with T-ALL $\gamma\delta$.

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 11: C-type lectin receptor Clec2d is expressed at higher levels by patients with T-ALLγδ when compared to patients with T-ALLαβ

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALLαβ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALLγδ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALLαβ or T-ALLγδ) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 12: Chst1 and Chst7 are expressed at higher levels in patients with T-ALL $\gamma\delta$ when compared to patients with T-ALL $\alpha\beta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome. Figure 12: Chst1 and Chst7 are expressed at a broader range in patients with T-ALL $\gamma\delta$ when compared to patients with T-ALL $\alpha\beta$

Figure 13: Taste receptor Tas2r45 is expressed at significantly lower levels in patients with T-ALL $\gamma\delta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 14: Trim71 is expressed at higher levels in patients with T-ALL $\gamma\delta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 15: HLA-DQB1 is expressed at higher levels in patients with T-ALL $\gamma\delta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 16: Gpr65 is expressed at significantly higher levels in patients with T-ALL $\gamma\delta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.

Figure 17: Two components of the nuclear pore complex are among the genes most differentially expressed in patients with T-ALL $\gamma\delta$ and T-ALL $\alpha\beta$

The RNA expression level for each gene is displayed in a box and whisker plot per cohort of 14 patients, with each blue dot representing the RNA expression level of an individual T-ALL $\alpha\beta$ patient ($n=9$) and each red square representing the RNA expression level of an individual T-ALL $\gamma\delta$ patient ($n=5$ patients). The mean RNA expression level for either group of patients (T-ALL $\alpha\beta$ or T-ALL $\gamma\delta$) is marked with a (+) sign. An unpaired t-test (two-tailed) was performed for each comparison, and the exact p -value is denoted on each graph. All results shown here are statistically significant with respect to the magnitude of differential expression between groups; the unpaired t-test is an additional test evaluating the significance of the difference between the RNA expression levels independent of the rest of the transcriptome.