



SEQ-ing the genetic constellation of acute lymphoblastic leukaemia

by Zhenhua Li and Allen Eng Juh Yeoh

Received: November 20, 2023.

Accepted: December 12, 2023.

Citation: Zhenhua Li and Allen Eng Juh Yeoh.

SEQ-ing the genetic constellation of acute lymphoblastic leukaemia.

Haematologica. 2023 Dec 21. doi: 10.3324/haematol.2023.284456 [Epub ahead of print]

Publisher's Disclaimer.

E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.

E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.

SEQ-ing the genetic constellation of acute lymphoblastic leukaemia

Zhenhua Li^{1,2} and Allen Eng Juh Yeoh^{2,3}

¹Department of Pharmacy and Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, ²Department of Pediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, ³VIVA-University Children's Cancer Centre, Khoo Teck Puat-National University Children's Medical Institute, National University Health System, Singapore

Correspondence

Allen Eng Juh Yeoh, MD

Department of Paediatrics, National University of Singapore, Level 12 NUHS Tower Block, 1E Kent Ridge Road, Singapore 119228, SINGAPORE

email: allen.yeoh@nus.edu.sg

Phone: (65) 6772 4406

Fax: (65) 6779 7486

Running title: SEQ-ing ALL subtypes

Disclosures

No conflicts of interest to disclose.

Contributions

Both authors wrote, reviewed and approved the manuscript.

Acute lymphoblastic leukaemia (ALL) is a heterogeneous cancer driven by a constellation of diverse recurrent genetic aberrations. The ease to sample bone marrow allows easy access to the cancer cells and enables deep exploration of the genetics that drive ALL. Naturally, with every new genetic tool, the genetic constellation of ALL is often the first frontier to be explored. These deep explorations result in a detailed map of the genetic constellation of ALL (**Figure 1**) which is the basis of WHO classification of tumours of haematopoietic and lymphoid tissues.

From the 1960s, when karyotyping and chromosomal banding were established, investigators embarked on this 60-year journey of discovery. This discovery started with abnormal whole chromosome copy numbers termed aneuploidy. Excess chromosomes >50, also known as hyperdiploidy, was the most common driver (**Figure 1**). Translocations, where bits of chromosomes were aberrantly fused, led to the discovery of Philadelphia (Ph) chromosome $t(9;22)/BCR::ABL1$ and $t(1;19)/TCF3::PBX1$. Translocations which do not change the banding patterns like $t(12;21)/ETV6::RUNX1$ took a while longer before yielding to discovery.

Paralleling this discovery is better treatment. With better treatment, investigators found that these genetic drivers are prognostic i.e., they predict the risk of relapse. This prognostic value of genetic subtypes gave birth to genetic risk stratification and eventually genetically driven treatment like addition of imatinib and dasatinib for Ph ALL. However, the difficulty to karyotype lymphoblasts and the need for many different diagnostic platforms, like multiple FISH probes, limited widespread use of genetic stratification.

In the 2000s, gene arrays enticingly promised a single platform to interrogate genetic drivers of ALL. Gene expression microarrays, which simultaneously measured the expression levels of tens of thousands genes, allowed the discovery of the “novel” subtype¹ (later found to be the *DUX4* subtype) and the Ph-like subtype². Using single nucleotide polymorphism (SNP) arrays which simultaneously genotyped hundreds of thousands of SNPs, deletion of a segment in pseudoautosomal region 1 (PAR1) next to *CRLF2* was identified³.

In late 2010s, transcriptomic sequencing (RNA-Seq) promised another revolution. With RNA-Seq, we can study both genetic expression profiles and the sequences of mRNA. Together, RNA-Seq allowed us to identify gene rearrangements, karyotype, gene expression patterns as well as sequence mutations. Using RNA-Seq of leukaemic blasts, Gu et al elegantly showed that >90% of ALL patients can be assigned to a specific genetic subtype⁴. We and others have tried to implement RNA-Seq in clinical practice^{5,6}. With standardization of RNA-Seq library preparations and affordable sequencing services, perhaps the most significant obstacle remaining was bioinformatics analysis. In this issue of Haematologica, Hu et al⁷ shared the Molecular Diagnosis of ALL (MD-ALL), an integrated analysis software for ALL subtype classification using RNA-Seq. Using published RNA-Seq data, they carefully selected the feature genes responsible for each subtype distinction, constructed machine learning models to perform gene expression analysis, and combined gene expression and genomic alterations to classify ALL subtypes. MD-ALL advanced the bioinformatics analysis for RNA-Seq-based ALL classification by addressing 3 key areas:

- 1) a reliable reference dataset. Hu et al. assembled an RNA-Seq dataset with 2,955 ALL cases around the world, representing more than 20 subtypes from both children and adult patients.
- 2) standardization of gene expression analysis. With different analysis method or features used, gene expression defined subtypes can be variable. For example, the *BCR::ABL1*-like subtype defined by European researchers have minor variations compared to the Ph-like subtype defined by St. Jude investigators^{2,3}. Hu et al. tested the different feature selection methods and streamlined gene expression analysis using multiple machine learning methods. This enhances reproducibility and robustness for clinical use.
- 3) integration of multiple types of information into a final call. Though majority of cases can be uniquely assigned to a subtype, multiple genetic events may appear together. For example, high hyperdiploidy can occur with *BCR::ABL1* fusion, and low

hypodiploidy with *TP53* mutations. A decision-making workflow is implemented in MD-ALL.

The recent International Consensus Classification of acute lymphoblastic leukaemia/lymphoma included nearly 30 subtypes. Efforts like MD-ALL is an important tool for clinical use of the newly discovered subtypes, particularly in resource-constrained settings.

ALL subtypes have distinct sensitivity patterns to commonly used chemotherapy agents⁸, targeted therapy, and even to immune therapy⁹. How to integrate these subtypes into risk stratification or treatment protocols need further investigations. For example, the *DUX4* subtype, despite poorer end of induction minimal residual disease (hence treated with intense treatment), have excellent outcomes. Yet, de-intensification for this favourable subtype needs to be done cautiously. On the other hand, intensifying therapy or use of novel treatment for newly discovered unfavourable subtypes, such as *TCF3::HLF* and *MEF2D*, is necessary. In addition, targeted or immune therapy could be used for certain subtypes, e.g. *ABL1* inhibitors and blinatumomab for Ph ALL creating a chemotherapy free regimen is exciting¹⁰.

We are in the cusps of a brave new world of ALL: better understanding of the biological basis of each genetic subtype and better ways to treat them. With better and more ways to treat ALL, exploration of the genetic constellation of ALL is no longer an academic exercise, it transforms care.

References

1. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*. 2002;1(2):133-143.
2. Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol*. 2009;10(2):125-134.
3. Mullighan CG, Collins-Underwood JR, Phillips LA, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet*. 2009;41(11):1243-1246.
4. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet*. 2019;51(2):296-307.
5. Chin WHN, Li Z, Jiang N, et al. Practical Considerations for Using RNA Sequencing in Management of B-Lymphoblastic Leukemia: Malaysia-Singapore Acute Lymphoblastic Leukemia 2020 Implementation Strategy. *J Mol Diagn*. 2021;23(10):1359-1372.
6. Yu CH, Wu G, Chang CC, et al. Sequential Approach to Improve the Molecular Classification of Childhood Acute Lymphoblastic Leukemia. *J Mol Diagn*. 2022;24(11):1195-1206.
7. Hu Z, Jia Z, Liu J, et al. MD-ALL: an Integrative Platform for Molecular Diagnosis of B-cell Acute Lymphoblastic Leukemia. *Haematologica*. xxx
8. Lee SHR, Yang W, Gocho Y, et al. Pharmacotypes across the genomic landscape of pediatric acute lymphoblastic leukemia and impact on treatment response. *Nat Med*. 2023;29(1):170-179.
9. Li Y, Moriyama T, Yoshimura, et al. PAX5 epigenetically orchestrates CD58 transcription and modulates blinatumomab response in acute lymphoblastic leukemia. *Sci Adv*. 2022;8(50):eadd640.
10. Foà R, Bassan R, Vitale A, et al. Dasatinib–blinatumomab for Ph-positive acute lymphoblastic leukemia in adults. *N Engl J Med*. 2020;383(17):1613-1623.

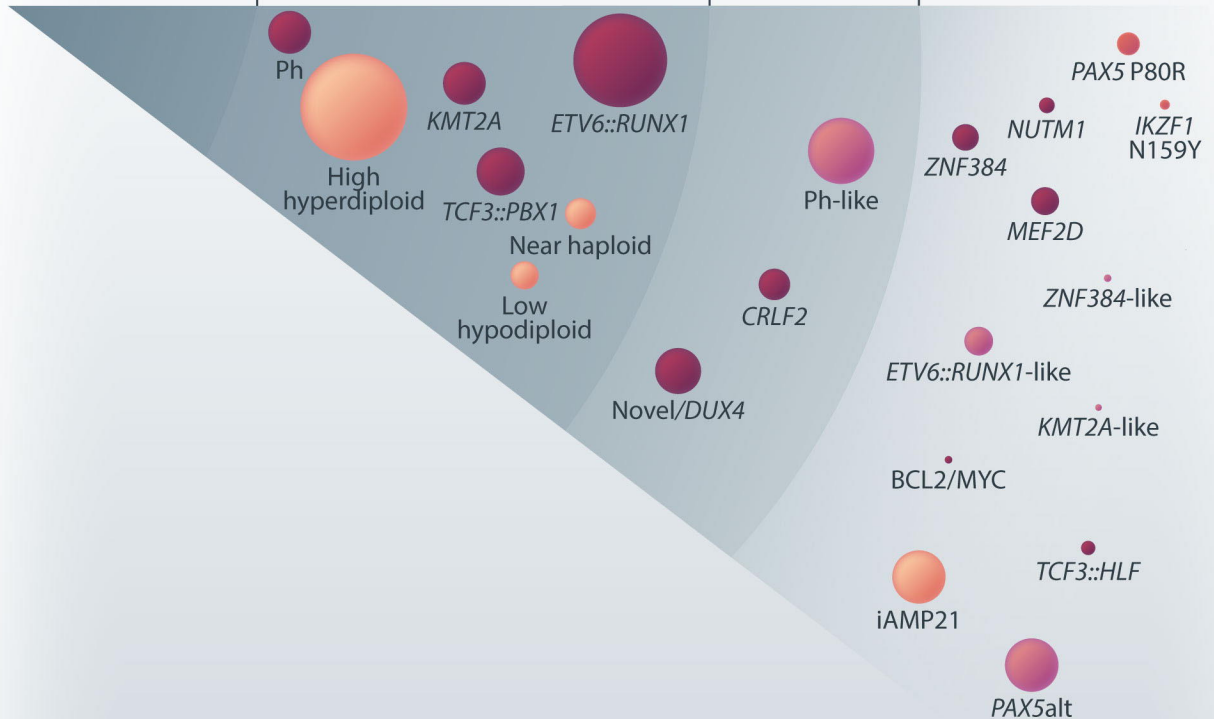
Figure Legend

Figure 1. A brief history of ALL subtype classification. This figure summarizes the main technologies available and the subtypes discovered during different time periods. Sizes of the circles indicate approximate relative frequencies of ALL subtypes in children.

1960
cytogenetics

2000
microarray

2015
RNA sequencing



● Aneuploidy

● Gene rearrangement

● Gene expression

● Mutation