

PB1709 A COMPLETE DIGITAL KARYOTYPE OF THE B-CELL LEUKEMIA REH CELL LINE RESOLVED BY LONG-READ SEQUENCING

Topic: 1. Acute lymphoblastic leukemia – Biology & Translational Research

Mariya Lysenkova Wiklander^{*1}, Gustav Arvidsson¹, Ignas Bunikis^{2,3}, Anders Lundmark¹, Amanda Raine^{1,3}, Yanara Marincevic-Zuniga^{1,3}, Henrik Gezelius^{1,3}, Anna Bremer^{2,4}, Lars Feuk^{2,3}, Adam Ameur^{2,3}, Jessica Nordlund^{1,3}

¹Department Of Medical Sciences And Science For Life Laboratory, Uppsala University, Uppsala, Sweden; ²Department Of Immunology, Genetics And Pathology And Science For Life Laboratory, Uppsala University, Uppsala, Sweden; ³National Genomics Infrastructure Hosted By Science For Life Laboratory, Uppsala, Sweden; ⁴Department Of Clinical Genetics, Uppsala University Hospital, Uppsala, Sweden

Background:

The B-cell acute lymphoblastic leukemia (ALL) cell line REH, with the t(12;21) *ETV6-RUNX1* translocation, is known to have a complex karyotype defined by a series of large-scale chromosomal rearrangements. Taken from a 15-year-old at relapse, the cell line offers a valuable model for the study of high-risk pediatric B-ALL patients. However, traditional karyotyping techniques and short-read next-generation sequencing alone have been unable to fully resolve the aberrations in this widely used genome.

Aims:

We explore the use of long-read PacBio and Oxford Nanopore whole genome sequencing (WGS) and IsoSeq RNA-sequencing alongside short-read sequencing to create a comprehensive digital karyotype of the REH cell line.

Methods:

DNA and RNA were extracted from the human pre B-cell ALL cell line REH and subjected to long- and short-read sequencing alongside traditional g-banding. In total, we generated three WGS datasets (Illumina PCR-free on HiSeq X; Oxford Nanopore Ultralong on PromethION24; PacBio HiFi on Sequel II) and three RNA-seq datasets (TruSeq stranded total RNA on NovaSeq 6000; 2x PacBio IsoSeq RNA-seq on Sequel II, with standard- and full-length transcripts). The short-read datasets were processed using nf-core pipelines (sarek with TIDBIT; rnafusion). The long-read datasets were mapped to GRCh38 using Minimap2, with structural variant (SV) calling using Sniffles2 and fusion gene calling using cDNA_Cupcake and JAFFAL. The SV and fusion gene candidates were programmatically filtered on quality, read support and size criteria (SVs > 100 kbp), and high-confidence candidates were confirmed manually in IGV and Ribbon.

Results:

WGS refined the breakpoints of known and expected aberrations in the REH cell line and clarified the molecular traits of disrupted ALL-associated genes *BTG1* and *TBL1XR1*, as well as the glucocorticoid receptor *NR3C1*. Several previously underreported structural variants were also uncovered, including deletions affecting the ALL-associated genes *VPREB1* and *NFATC1*. In total, the consensus callset (intersection of all three WGS callsets) found 2100 SVs, and 23 SVs > 100 kbp were manually confirmed. Meanwhile, transcriptome sequencing identified seven fusion genes within the genomic breakpoints, including two in-frame fusions (*ETV6-RUNX1* and *RUNX1-PRDM7*), two fusions resulting from the balanced translocation t(5;12), and the GC receptor-disrupting *NR3C1-ARHGAP26*.

Summary/Conclusion:

Copyright Information: (Online) ISSN: 2572-9241

© 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access Abstract Book distributed under the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) which allows third parties to download the articles and share them with others as long as they credit the author and the Abstract Book, but they cannot change the content in any way or use them commercially.

Abstract Book Citations: Authors, Title, HemaSphere, 2023;7(S3):pages. The individual abstract DOIs can be found at <https://journals.lww.com/hemasphere/pages/default.aspx>.

Disclaimer: Articles published in the journal HemaSphere exclusively reflect the opinions of the authors. The authors are responsible for all content in their abstracts including accuracy of the facts, statements, citing resources, etc.

Our findings highlight the need for comprehensive genomic analysis of commonly used and perturbed cell lines. Further, we underscored the importance of the increased informational context of long reads and how they can identify critical biologically important SVs. Finally, we have cataloged multiple aberrations in REH that are frequently observed in high-risk ALL patients, making this cell line a highly relevant model for researchers investigating the biological underpinnings of relapse and GC resistance. Together, our extensive whole-genome investigation makes high-quality open-source data available to the leukemia genomics community.

Copyright Information: (Online) ISSN: 2572-9241

© 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access Abstract Book distributed under the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) which allows third parties to download the articles and share them with others as long as they credit the author and the Abstract Book, but they cannot change the content in any way or use them commercially.

Abstract Book Citations: Authors, Title, HemaSphere, 2023;7(S3):pages. The individual abstract DOIs can be found at <https://journals.lww.com/hemasphere/pages/default.aspx>.

Disclaimer: Articles published in the journal HemaSphere exclusively reflect the opinions of the authors. The authors are responsible for all content in their abstracts including accuracy of the facts, statements, citing resources, etc.