High-Resolution Detection of Structural Variants Associated with Leukemia Using Electronic Genome Mapping for Molecular Diagnostics

Yang Zhang[1], Sangjun Lee[2], Shuk Shukor[3], Hanae Sugiura[3], Anna Lozar[1], Lily Nasanovsky[1], Pasko Lee[2], Erika Headrick[1], Kyoko Roberts[3], Eiichi Araki[3], Jefferson Y. Chan[2]
[1]Nabsys LLC, Providence, RI; [2]University of California, Irvine, Department of Pathology & Laboratory Medicine, Irvine, CA; [3]Hitachi High-Tech America, San Deigo, CA

Abstract

Structural variants (SVs) such as *TP53* deletion and *PML::RARA* gene fusion are clinically significant SVs in hematologic malignancies, and their accurate and timely detection is essential for risk stratification and guiding effective treatment. Detecting such SVs with precision remains a significant clinical challenge. Short-read sequencing (SRS) often misses large SVs due to read-length constraints, while long-read sequencing (LRS) and traditional cytogenetic techniques, chromosomal microarray analysis (CMA) and Fluorescence in situ hybridization (FISH), struggle to balance resolution, throughput, and cost. Here, we explored the use of electronic genome mapping (EGM), a novel genome mapping technology, for detecting SVs in leukemic cells.

The OhmX Platform, which uses electronic genome mapping (EGM) and solid-state, nanochannel detection, was used to detect *TP53* deletions and *PML::RARA* gene fusion in various human leukemic cell lines. DNA was isolated using the NEB Monarch HMW DNA extraction kit and processed with a dual-enzyme nicking, labeling, tagging and coating workflow. Data were collected using the OhmX Analyzer. SVs were detected using Human Chromosome Explorer's (HCE) *de novo* assembly pipeline and confirmed with SV-Verify (SVV), a targeted, alignment-based variant confirmation pipeline. Findings were further supported by orthogonal techniques.

HCE identified the breakpoints and sizes for *TP53* deletions ranging from 87 kb to 3.8 kb, as well as fusion sites of *PML::RARA* translocation. These SV calls were consistent with published data generated by orthogonal methods, demonstrating strong agreement in results.

EGM offers streamlined analysis and scalable, cost-effective SV detection, serving as a valuable complement to traditional cytogenetic and sequencing technologies. Its ability to detect clinically relevant variants position it as a strong candidate for integration into diagnostic workflows in hematopathology.

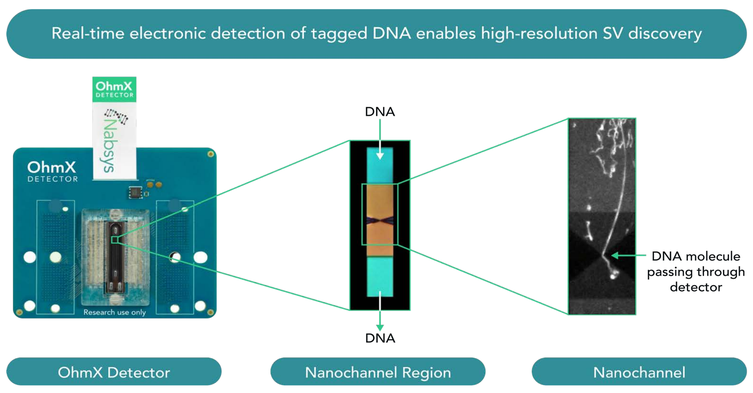
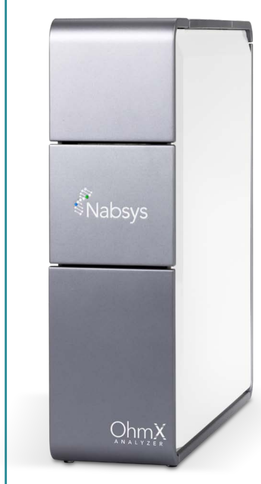


Figure 1: EGM detector. EGM is enabled by the novel, solid-state nanodetectors developed by Nabsys. The OhmX Detector houses 256 parallel nanochannels, each with its own electronic sensor. Nabsys has developed DNA tagging chemistry that provides a high signal-to-noise ratio allowing superior detection in EGM versus optical methods. The high signal and spatial resolution allows tags to be more closely spaced on the high molecular weight DNA molecules. This yields optimal tag density such that structural variations as small as 300 bp can be routinely identified and mapped genome-wide.



EGM is based on the principal of current blockade. When an object occupies a significant space in a current path, it creates a detectable drop in current. DNA in a nanochannel reduces current flow, with larger tagged DNA, causing a greater blockade. Measuring these current blockades provides information about the relative locations of tags. Transit times for these segments are converted to physical distances. The unique spacing between tags can be used to align and map individual molecules. The mapped or de novo assembled molecules can be compared to reference sequences for identification or confirmation of structural variants. No amplification or ligation is required, minimizing false positives with the very long DNA molecules (>100 kb).

Figure 2: OhmX Analyzer. The OhmX Analyzer is easy to operate with simple sample loading, flexible experiment run time capabilities, and automated protocols. With a footprint of 5" by 17" the instrument takes up minimal bench space and the entire OhmX Platform – including the OhmX Controller – takes up a footprint of 30" x 17".

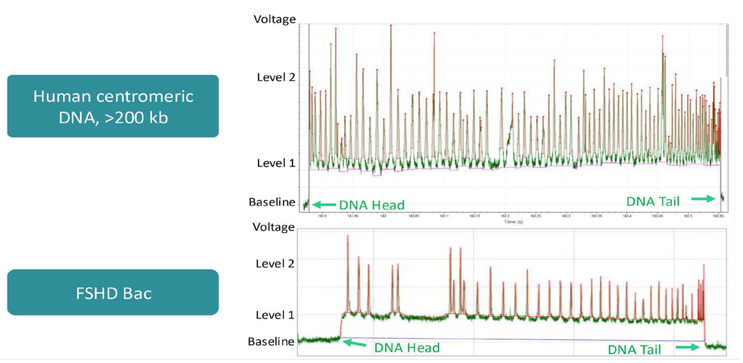


Figure 3: Electronic Profile for tagged DNA molecules over 150 kb. The baseline voltage is shown on the left. When the DNA molecules enter the nanochannel, a change in voltage is observed. Whenever a site-specific tag on the DNA is also present, the voltage changes again. The time between tags is converted to distance in base pairs.

TP53 Deletion in HL-60 and H1299

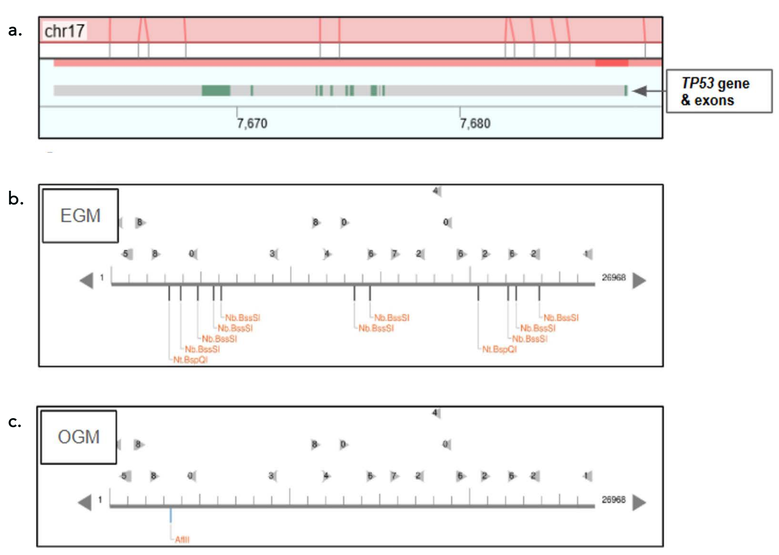


Figure 4: Comparison of label site density between EGM and OGM for the *TP53* gene a) *TP53* gene region (hg38) visualized in HCE, with the gene represented as grey and red rectangles and exons shown in green. b) In silico digestion of the *TP53* gene using EGM's dual-enzyme nicking (Nt.BspQI and Nb.BssSI) via NEBCutter 3.0. c) In silico digestion of the *TP53* gene region (hg38) using OGM's labeling chemistry (AfiII-like recognition sequence) via NEBCutter 3.0.

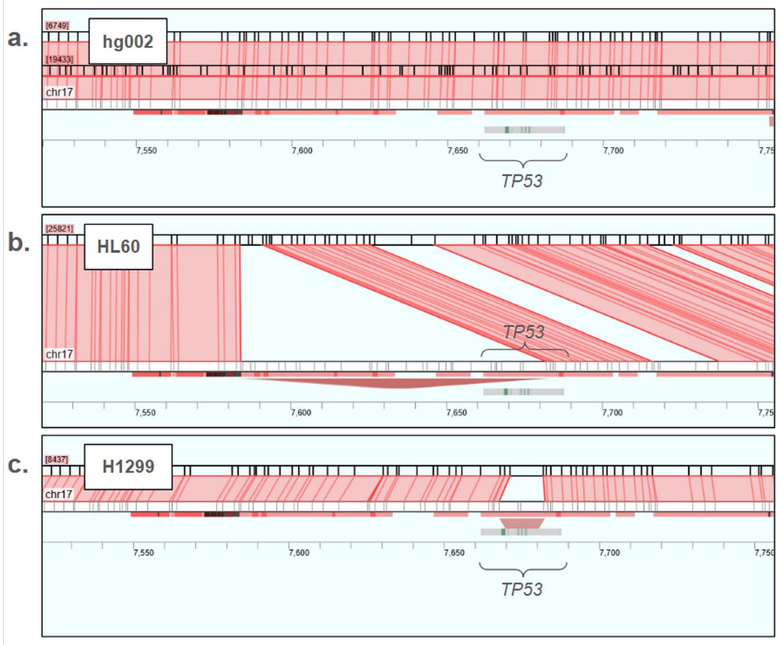


Figure 5: *TP53* deletions in healthy and cell line samples visualized in HCE *TP53* exon information is visualized using a custom BED file, with the gene represented as grey and red rectangles and exons shown in green. Deletions are represented as red triangles below the alignment. a) Negative control HG002 with WT *TP53* gene. b) Large, ~87 kb *TP53* deletion in the HL-60 cell line. c) Small, ~3.8 kb *TP53* deletion in the H1299 cell line.

Sample	Chr	Start	Length (bp)	Type	Zygosity
HL-60	17	7,583,898	90,663	Deletion	Homozygous
H1299	17	7,667,694	3,828	Deletion	Homozygous

Table 1: HCE summary statistics for *TP53* SV calls. *TP53* deletions were called in both the HL-60 and H1299 cell lines.

Sample	Chr	Start	Length (bp)	Type	Molecules Aligned to WT	Molecules Aligned to SV Hypothesis	Observed VAF SV Frequency
HL-60	17	7,589,540	87,389	Deletion	5	11	0.6875
H1299	17	7,667,694	3,828	Deletion	25.5	34	0.57

Table 2: SV-Verify summary statistics for *TP53* SV calls. *TP53* deletions were confirmed in both the HL-60 and H1299 cell lines. Approximately 69% and 57% of molecules aligned to the SV hypothesis in HL-60 and H1299, respectively.

Chr	hg38 Start Pos. (kb)	hg38 Start Pos. (kb)	Size (kb)
chr17	7,589,540	7,668,534	78.994
chr17	7,668,886	7,674,811	5.925
chr17	7,674,828	7,675,361	0.533
chr17	7,675,973	7,677,910	1.937
			87.389

Table 3: HL-60 *TP53* deletion Microarray data. The data indicating copy number loss events in the *TP53* region and the upstream genomic region in the HL-60 cell line

PML::RARA Translocation in NB-4

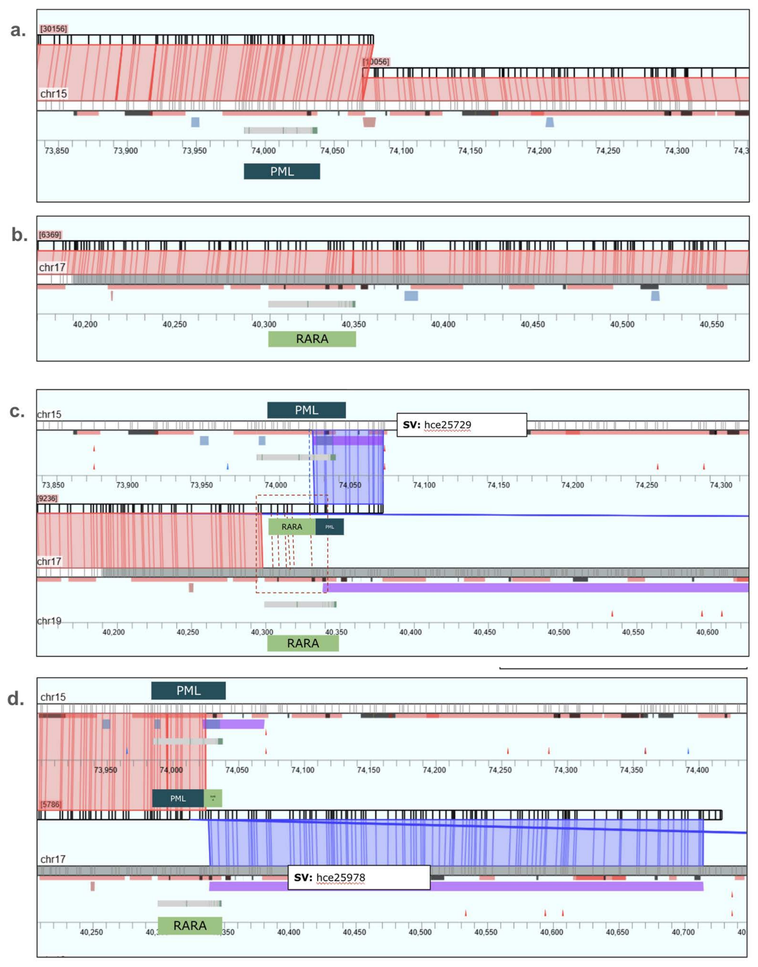


Figure 6: PML::RARA fusion in healthy and leukemia cell line samples visualized in HCE. *PML* exon information is visualized using a custom BED file, with the gene represented as grey and red rectangles. Translocations are represented by purple rectangles. a) Negative control HG002 with WT *PML* gene. b) Negative control HG002 with WT *RARA* gene. c) First fusion breakpoint showing a small region of *PML* fused to *RARA*. d) Second fusion breakpoint showing a small region of *RARA* fused to *PML*.

Sample	Chr	Start	Length (bp)	Chr (To)	Start (To)	Length (To)	Type
NB41A3	15	74,032,264	1,358	17	40,307,908	42,932,215	Transloc
NB41A3	17	40,358,576	1,168	15	74,034,747	27,942,760	Transloc

Table 4: HCE summary statistics for *PML::RARA* SV calls. A balanced reciprocal *PML::RARA* translocation was called in the NB41A3 cell line. Zygosity could not be determined using HCE.

Sample	Chr	Start	Chr (To)	Start (To)	Type	Molecules Aligned to WT	Molecules Aligned to SV Hypothesis	Observed VAF
NB41A3	15	74,034,032	17	40,345,923	Breakpoint Fusion	60.5	37	0.3795

Table 5: SV-Verify summary statistics for *PML::RARA* SV calls. A *PML::RARA* breakpoint fusion was confirmed in the NB41A3 cell line.

Conclusion

This study demonstrates the analytical utility of EGM and the OhmX Platform for high-resolution SV detection in leukemia models. EGM enabled high-confidence mapping of *TP53* deletions, as well as a reciprocal *PML::RARA* translocation, using a dense, sequence-specific tagging strategy and an electronic detection system with approximately 300 bp resolution.

Variants were detected across a range of sizes with support from raw molecule data, contig-level assemblies in HCE, and independent hypothesis-based confirmation using SV-Verify. The platform's ability to resolve complex genomic alterations that are difficult to detect by traditional cytogenetic or sequencing-based approaches highlights its value in translational cancer research.

EGM data visualization and de novo assembly were conducted within HCE, while structural variant confirmation was carried out using the separate SV-Verify analysis pipeline. Together, these tools support flexible and scalable workflows for SV discovery and confirmation. Future efforts will focus on expanding the use of EGM for more diverse clinical and research applications and integrating and enhancing analytical outputs.

References

- Wolf, D. & Rotter, V. Major deletions in the gene encoding the p53 tumor antigen cause lack of p53 expression in HL-60 cells. Proc. Natl. Acad. Sci. 82, 790–794 (1985).
- Mitsudomi, T. *et al.* p53 gene mutations in non-small-cell lung cancer cell lines and their correlation with the presence of ras mutations and clinical features. Oncogene 7, 171–180 (1992).
- H de Thé *et al.* The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature 347, 558–561 (1990).
- Middlezong, W. *et al.* Rapid Detection of *PML::RARA* Fusions in Acute Promyelocytic Leukemia: CRISPR/Cas9 Nanopore Sequencing with Adaptive Sampling. Biomolecules, 14(12), 1595 (2024).
- Oliver, J. S. *et al.* High-Definition Electronic Genome Maps from Single Molecule Data. 139840 Preprint at https://doi.org/10.1101/139840 (2017).

Learn more at Nabsys.com