3D Genomics with Arima Hi-C Sequencing Enables Detection of Clinically Relevant Gene Fusions in Pediatric Cancer Samples

Kristin Sikkink[1], Derek Reid[1], Anthony Schmitt[1*], Midhat Farooqi[2*], Atif Ahmed[3*]

[1]Arima Genomics Inc, Carlsbad, CA; [2]Children's Mercy Hospital, Kansas City, MO, [3]Seattle Children's Hospital, Seattle, WA

*Equal contribution

Introduction

- Gene fusions as biomarkers have broad clinical utility in cancer patients including for accurate diagnosis, early detection, prognosis, and selection of optimal treatment regimens.
- Typically, these fusions are detected using low-resolution karyotyping, or low throughput and limited fluorescence in situ hybridization (FISH) assays. Alternatively, RNA sequencing approaches can be used but present challenges due to low transcript abundance or RNA degradation.
- Formalin-fixed paraffin-embedded (FFPE) tissues are a critical archival and clinical sample type, but typically perform poorly in molecular assays due to DNA damage and cross-linking from formalin.
- Gene fusions can be detected with Hi-C technology, which has been shown to have high accuracy for identifying inter- and intrachromosomal translocations and rearrangements[1,2].

Methods

To address these limitations, we developed a novel approach to identify structural variants and gene fusions from FFPE samples using the Arima-HiC platform and Illumina short-read sequencing. We then selected five archived pediatric alveolar rhabdomyosarcoma (ARMS) tumors (FFPE archival period range: 8-12 years)—known to be fusion-positive via prior clinical testing. All cases had undergone standard of care cytogenetic testing at the time of diagnosis, i.e., karyotyping and/or FISH).

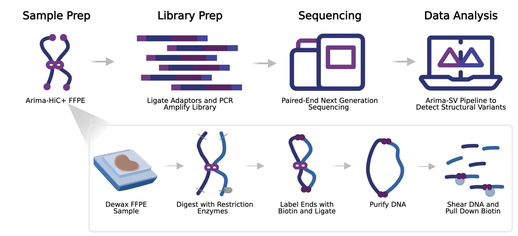

Figure 1. The Arima HiC+ FFPE Workflow[3]. Sample Prep: FFPE tissue scrolls were dewaxed, and the tissue rehydrated; samples then underwent chromatin digestion, end-labeling, and proximity ligation prior to DNA purification per the Arima-HiC FFPE protocol. Library Prep and Sequencing: Purified DNA was next prepared as a short-read sequencing library and sequenced on a HiSeq X.

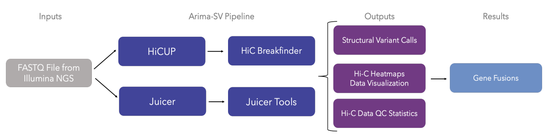

Figure 2. Data Analysis Workflow[4]. FASTQ files were inputted into the Arima-SV pipeline, which enabled the variant calling and produced Hi-C heatmaps for identification of gene fusions.

To expand our understanding of Arima Hi-C to detect gene fusions, we selected 2 patient samples with cryopreserved cells that did not have an identified causative genetic alteration via prior testing by FISH, karyotyping, microarray, and a targeted cancer NGS panel.

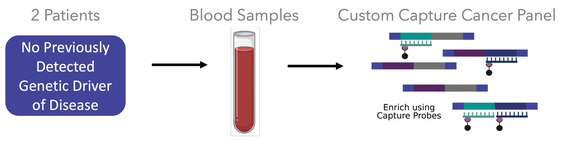

Figure 3. Detection of Gene Fusions in Blood Samples. The workflow used on 2 patient blood samples were subjected to Arima Capture HiC using a custom panel design for 884 known cancer-related genes.

Findings

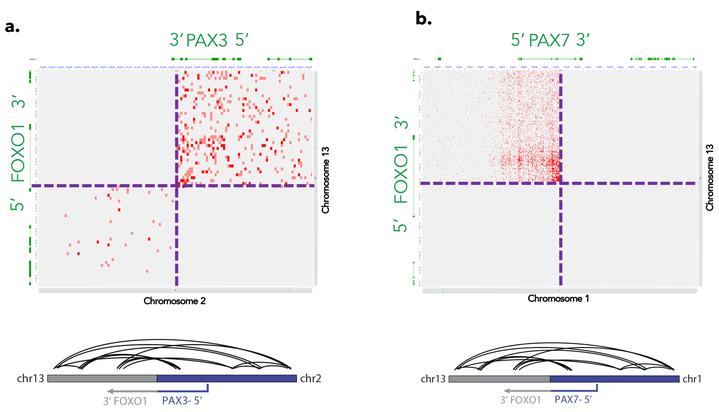

Figure 4. Leveraging 3D Genomics to Identify *FOXO1* Fusions in Alveolar Rhabdomyosarcoma. A Hi-C contact matrix and associated karyogram showing 3D interactions around a breakpoint creating a *PAX3-FOXO1* gene fusion. High 3D interaction signal is observed between the 5′ portion of *PAX3* and the 3′ portion of *FOXO1*. **b.** A Hi-C contact matrix and associated karyogram showing 3D interactions around a breakpoint creating a *PAX7-FOXO1* gene fusion. High 3D interaction signal is observed between the 5′ portion of *PAX7* and the 3′ portion of *FOXO1*.

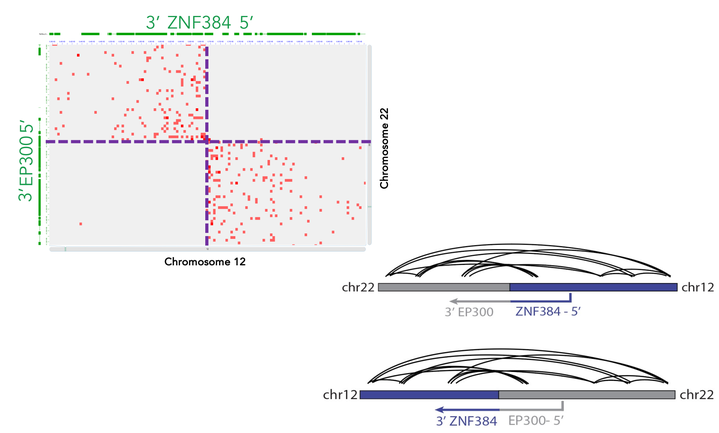

Figure 5. 3D Genomics Identifies a Previously Undetected Structural Variant of Prognostic Significance in a Pediatric ALL Case with No Previously Detected Genetic Driver. Hi-C contact matrix showing 3D interactions around a breakpoint creating a *ZNF384-EP300* gene fusion not detected by cytogenetics or NGS testing (targeted, DNA-based, gene panel. The observation of a *ZNF384-EP300* fits the diagnosis of ALL, and published reports suggest a favorable prognosis[5].

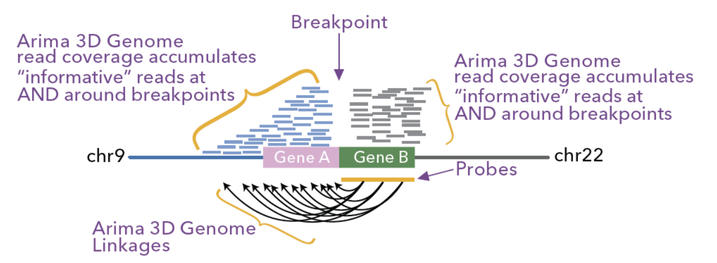

Figure 6. Schematic Representation of how Arima Capture HiC Enables Detection of Breakpoints. We created a cancer gene panel to further interrogate genes of interest. Using probes for designed for Gene B, gene fusions can be detected by assessing 3D genome linkages.

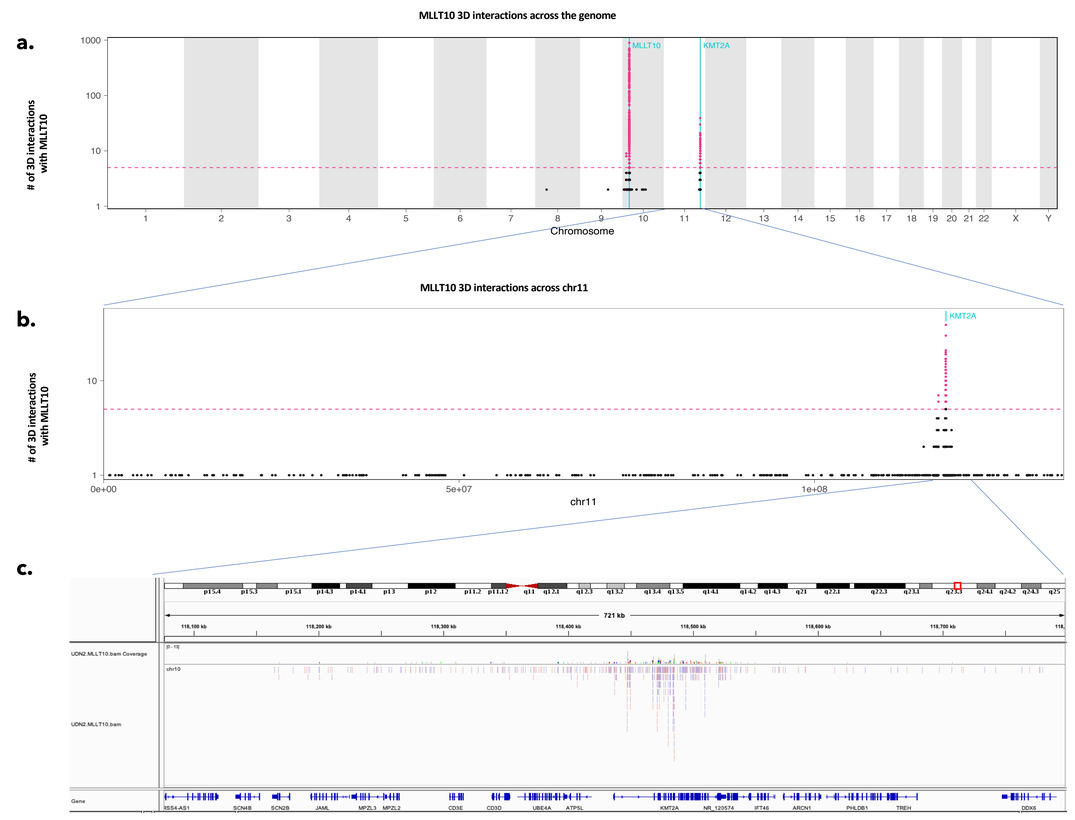

Figure 7. Targeted 3D Genomics Identifies a Previously Undetected Structural Variant of Prognostic Significance in a Pediatric AML Case with No Previously Detected Genetic Driver. a. A genome-scan plot of all 3D interactions between *MLLT10* and the rest of the genome obtained from a targeted Hi-C assay using a cancer gene panel for enrichment. Each genomic 1kb bin is plotted along the x-axis, and the y-axis corresponds to the observed 3D interaction counts between a given 1kb bin and *MLLT10*. **b.** A zoomed-in view of the genome-scan analysis on chr11, where a significant "hit" was found at *KMT2A*, indicating a *MLLT10-KMT2A* gene fusion. **c.** IGV browser view of all reads where one read end maps to *MLLT10*, and the other maps to the locus shown around *KMT2A*. The greatest abundance of 3D interaction reads are observed within *KMT2A*, supporting the *MLLT10-KMT2A* fusion call. The observation of an *MLLT10-KMT2A* fusion carries an unfavorable prognosis[6].

Gene Fusions Detected in FFPE Alveolar Rhabdomyosarcoma Tumor Samples.

Tumor ID	Archival Period (yrs)	Input DNA (ng)	Gene Fusions Detected (Primary Result)	Gene Fusions Detected (Orthogonal Result)	Orthogonal Assay
AR1	12	2.9	PAX3-FOXO1	FOXO1	FISH
AR6	12	97.6	PAX3-FOXO1	FOXO1	Karyotype
AR8	11	6.2	PAX7-FOXO1	FOXO1	Karyotype
AR19	8	157.5	PAX7-FOXO1	PAX7-FOXO1	FISH
AR3	unknown	12.1	PAX3-FOXO1	FOXO1	FISH

Table 1. Summary of Gene Fusions Detected in FFPE Alveolar Rhabdomyosarcoma Tumor Samples. Gene fusions were detected using Arima-HiC for FFPE and Arima-SV bioinformatics analyses (primary result) and compared to historical results obtained using FISH or karyotype (orthogonal result).

Conclusions

- In summary, this study demonstrates how Arima-HiC sequencing provides molecular diagnostic value in archived pediatric solid and liquid tumor specimens via the identification of clinically relevant gene fusions.
- This workflow can provide improved access to critical genomic information from FFPE blocks for the identification of clinically relevant gene fusion events and other structural variants across tumor types.
- Targeted 3D genomic approaches in the form of cancer gene panels can be used to identify structural variants in cases with no previously detected genetic driver.
- Information gained from 3D genomic interrogation of FFPE samples can provide diagnostic, prognostic, and therapeutic insights for pediatric cancer.

Acknowledgements

Drs. Farooqi and Ahmed would like to thank the Department of Pathology & Laboratory Medicine at Children's Mercy Hospital, as well as the Black & Veatch Foundation, Big Slick, and Braden's Hope for Childhood Cancer for their generous financial support.

References

1. Dixon, J. R., et al. (2018). Integrative detection and analysis of structural variation in cancer genomes. Nature Genetics, 50(10), 1388–1398.
2. Harewood, L., et al. (2017). Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. Genome Biology, 18(1), 125.
3. Product Flyer: Arima-HiC FFPE. Arima Genomics Literature.
4. Bioinformatics User Guide: Arima Structural Variant Pipeline. Arima Genomics.
5. Shago, M., et al. (2016). Frequency and outcome of pediatric acute lymphoblastic leukemia with ZNF384 gene rearrangements including a novel translocation resulting in an ARID1B/ZNF384 gene fusion. *Pediatric Blood & Cancer*, *63*(11), 1915–1921.
6. Forgione, M. O., et al. (2020). MLLT10 rearranged acute leukemia: Incidence, prognosis, and possible therapeutic strategies. *Genes, Chromosomes and Cancer*, *59*(12), 709–721.

@ArimaGenomics

Scan for a digital copy