Investigating the Effectiveness of Nanopore Technology and Hi-C Analysis in Detecting Structural Variations in Human Soft-Tissue Sarcoma Samples

Kristy Mendoza Rangel[1], Kadir Caner Akdemir[2]

1 Columbia University **; 2** UT MD Anderson Cancer Center, Department of Neurosurgery

THE UNIVERSITY OF TEXAS MD Anderson Cancer Center

Making Cancer History®

Background

- Structural variations (SVs) are the most disruptive mutation type in human tumors, affecting a large fraction of the cancer genomes.
  - Challenging to detect in cancer genomes because of their complexity and technical limitations.
- Oxford Nanopore technology: long fragment of DNA or RNA passes through a nanopore that is embedded in an electrically resistant membrane[1].
  - Each base causes a characteristic current disruption which is then decoded using basecalling algorithms to determine the DNA sequence.
- Hi-C is a chromosome conformation capture method that involves cross-linking DNA within cells to create a snapshot of which genomic regions physically interact with each other[2].

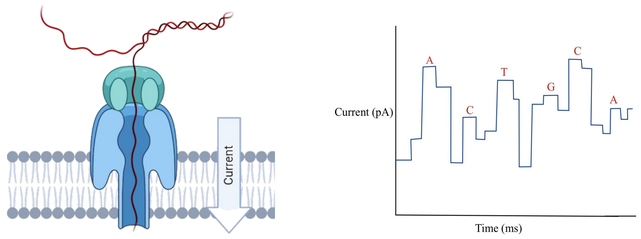

Fig. 1. Nanopore technology requires for a strand of DNA or RNA to enter the nanopore. The technology is able to sequence the fragment due to changes to the electrical current. The resulting signal is then decoded to provide the DNA or RNA sequence. A benefit of using Nanopore is the ability to sequence long fragments of DNA or RNA, however, it has a high sequencing error rate.

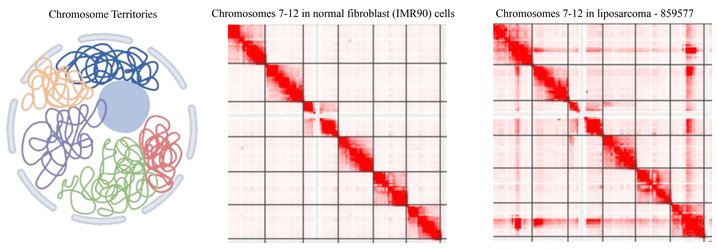

Fig. 2. Hi-C analyzes the way chromatin is organized and allows scientists to investigate chromatin interactions on a genome-wide scale. There are major differences in the Hi-C contact maps for normal fibroblast cells and tumor samples such as liposarcoma- S77. In normal fibroblast cells, the interchromosomal areas do not show any interaction while in liposarcoma- S77 there is high frequency interactions in these areas.

Methods

- Used Juicebox[3] as a data visualization tool to analyze Hi-C contact maps and SV annotations (Hi-C and Nanopore) of human tumor samples (Liposarcoma - S69, Liposarcoma - S77, and Osteosarcoma - S03).
- The SV calls were identified using HiCBreakFinder[4] algorithm for Hi-C and Sniffles[5] algorithm for Nanopore.

	Hi-C	Nanopore
Liposarcoma - S69	**X**	**X**
Liposarcoma - S77	**X**	**X**
Osteosarcoma - S03	**X**	**X**

Results

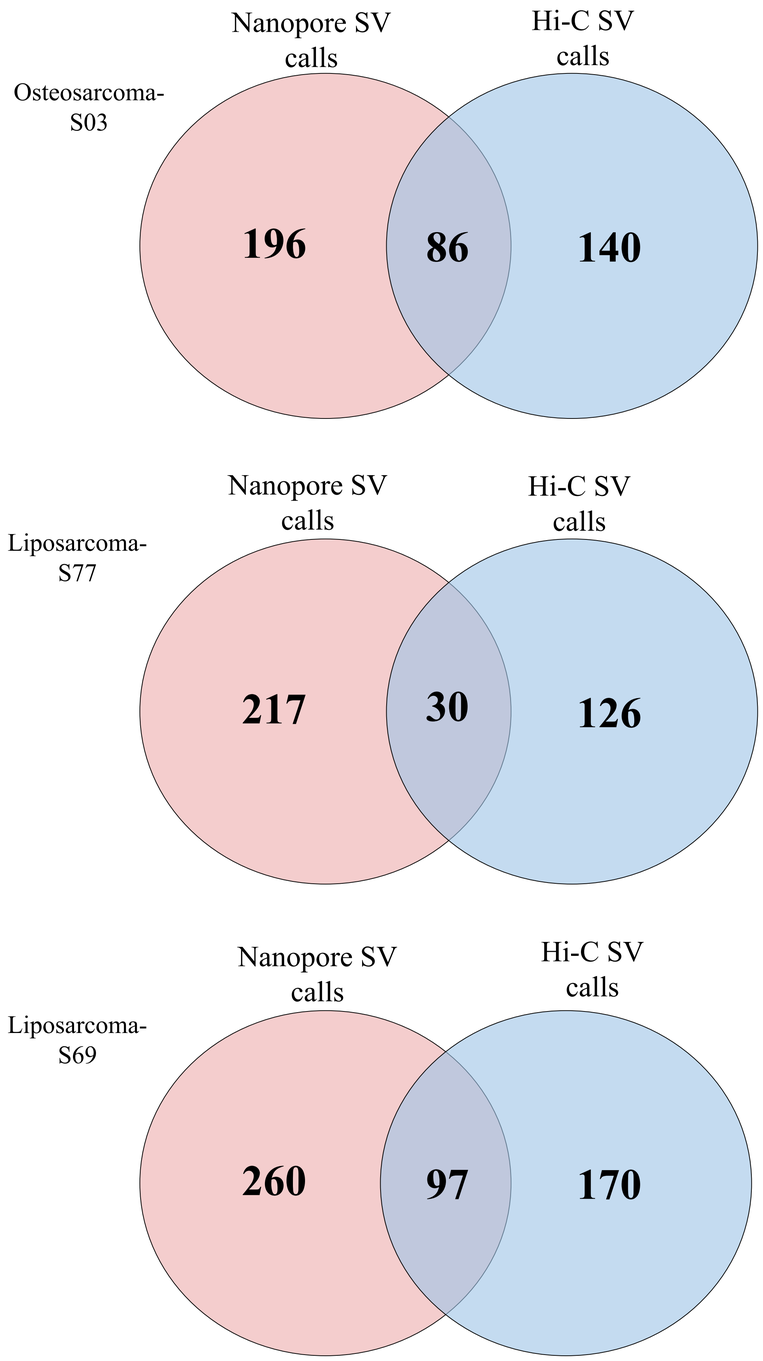

Fig. 3. The venn diagrams show how many SV calls were identified by only Hi-C or only Nanopore as well as the calls that they both identified in Liposarcoma - S69, Liposarcoma - S77, and Osteosarcoma - S03.

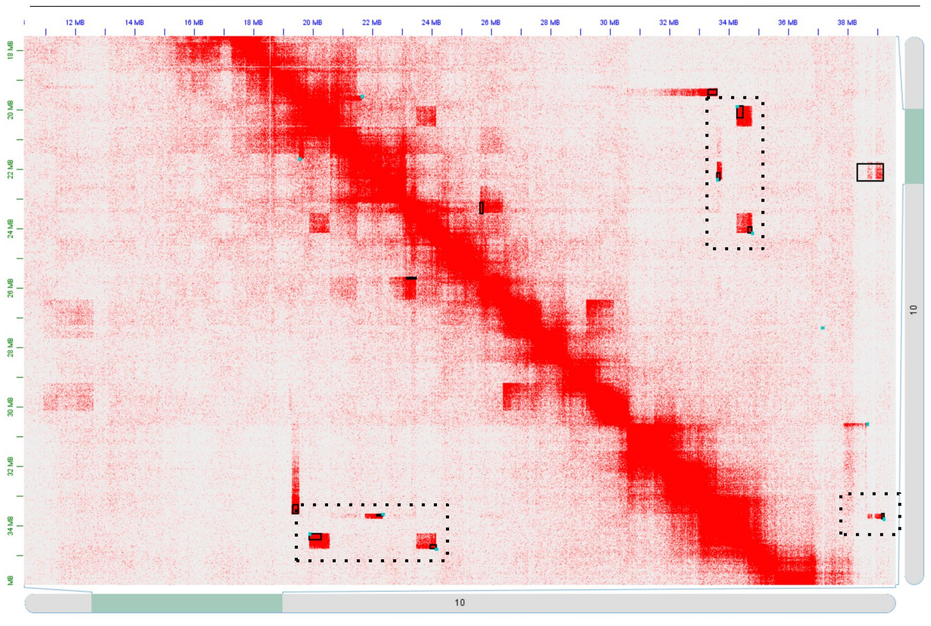

Fig. 4. The intrachromosomal locus in chromosome 10 of osteosarcoma - S03, there are several SVs seen in the Hi-C contact map, however, only 7 SVs were identified by Hi-C and Nanopore.

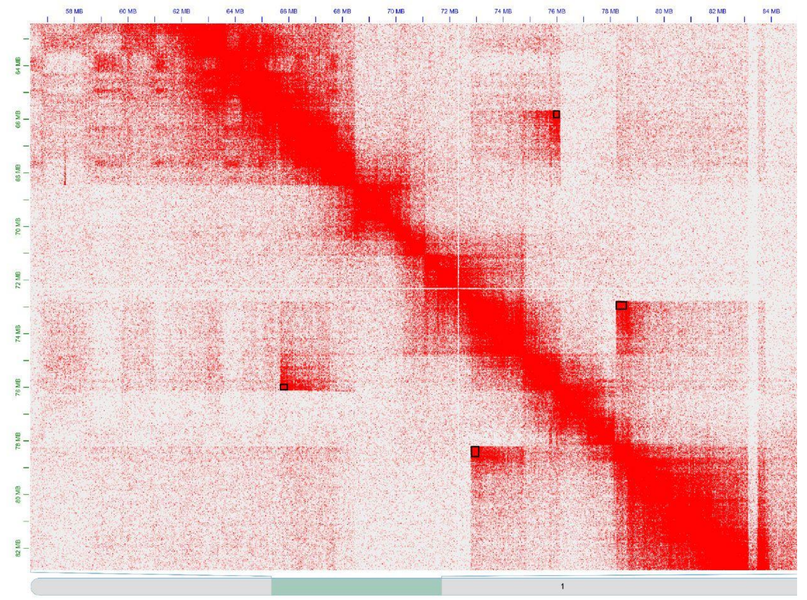

Fig. 5. This intrachromosomal locus of chromosome 1 in liposarcoma- S77 shows an instance where only Hi-C identifies SVs. The 4 structural variations identified are outside of TADs and show very high interaction frequency. There are other SVs in the locus however the have a relatively low interaction frequency. Hi-C does not tend to identify SVs with low interaction frequency. .

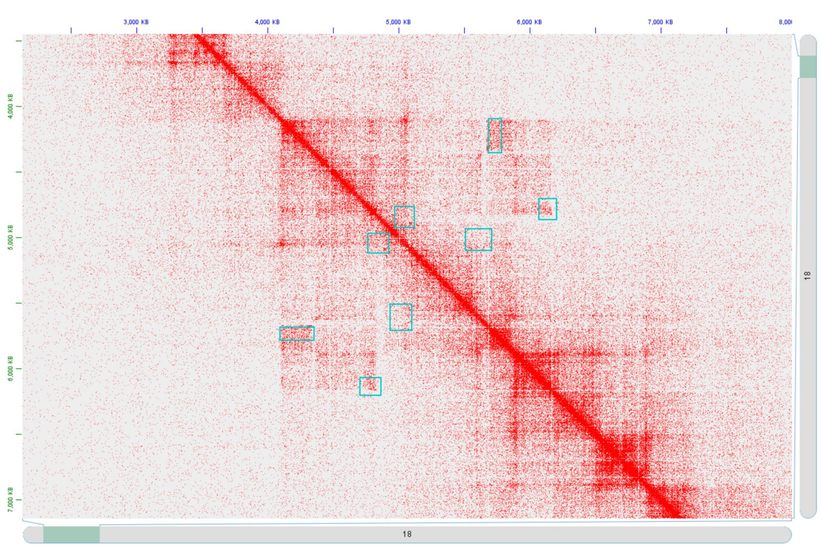

Fig. 6. In Liposarcoma - S69, the intrachromosomal locus of chromosome 18 shows several SVs inside of a TAD and they were only identified by Nanopore. These 8 SVs show relatively low interaction frequency compared to those SVs that Hi-C tends to identify.

Conclusions

- Revealed that detecting SVs with a single method might have certain limitations.
  - Nanopore was better at identifying SVs inside TADs and in interchromosomal areas.
  - Hi-C did better at identifying SVs in intrachromosomal areas compared to Nanopore.
- Therefore, a combination of orthogonal methods is needed for precise SV detections in human tumor samples.
- Further investigation will reveal other potential improvements that can be made to Nanopore and Hi-C and more samples can increase the validity of the data.

References

1) Oxford Nanopore Technologies. Oxford, UK.
2) Kim et al. Seminars in Cell & Developmental Biology 2022; 121
3) Durand et al. Cell Systems 2016; 3
4) Dixon et al. Nat Genetics 2018; 50
5) Sedlazeck et al. Nat Methods 2018;15