

Comprehensive & Precise Structural Variant Detection With a Single Assay

J. Zachary Sanborn¹, James Durbin¹, Mital Bhakta¹, Lisa Munding¹

¹Dovetail Genomics, Part of Cantata Bio, LLC. Scotts Valley, CA, USA

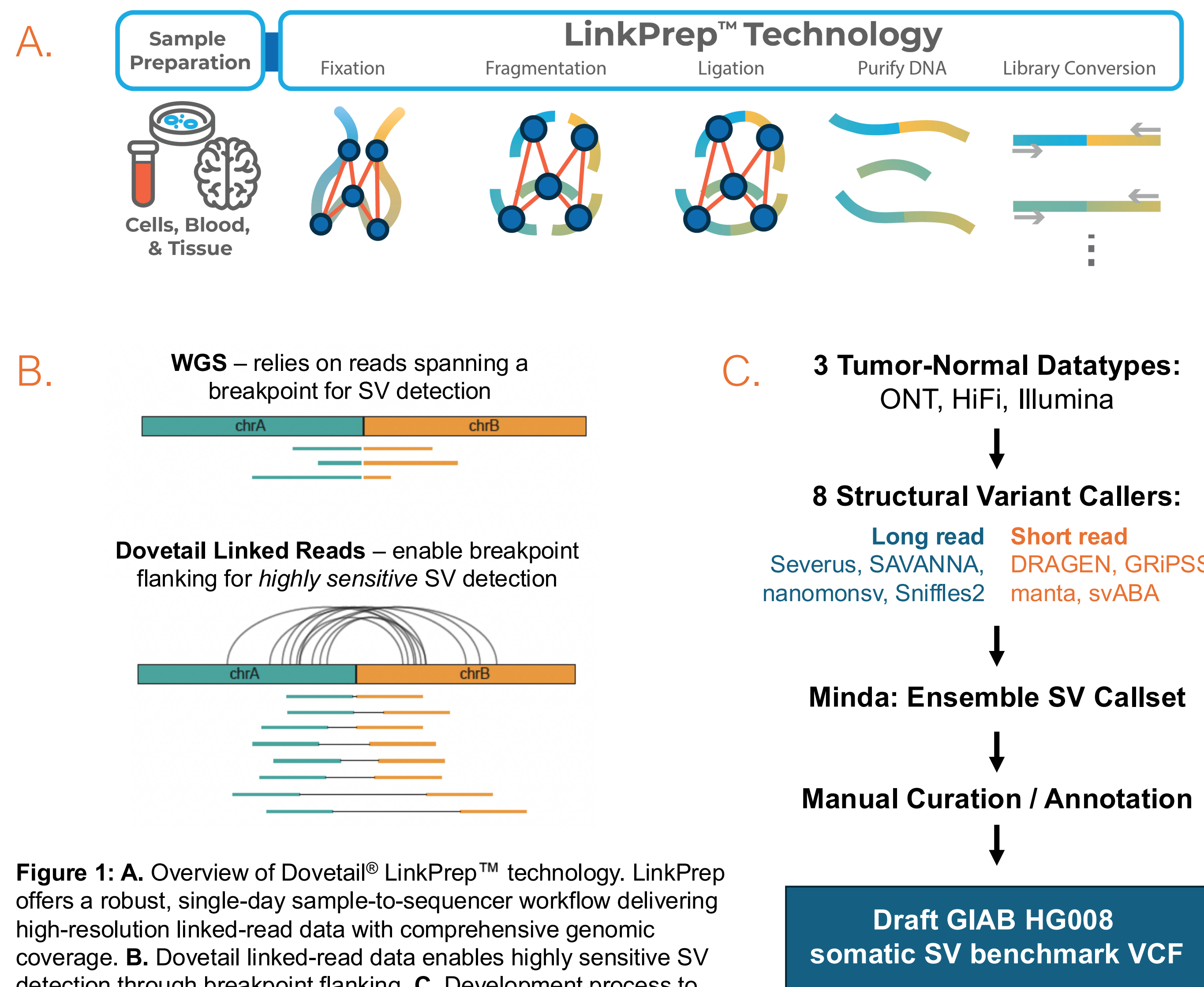


Introduction

Structural variants (SVs) are increasingly recognized as key drivers of tumorigenesis, yet current technologies struggle to detect and accurately characterize them due to resolution limitations and sequencing constraints. As a result, attempts to produce complete somatic SV truth sets often suffer due to the gaps in detection performance of the sequencing technology utilized, impacting the truth sets' usefulness for benchmarking alternative technologies and/or SV callers. To this end, the Genome in a Bottle Consortium (GIAB) developed a draft benchmark for somatic SV detection, aiming to fully characterize the structural changes of a pancreatic ductal carcinoma (PDAC, labeled HG008-T) by integrating 16 whole genome sequencing technologies and analyzing with multiple datatype-specific SV callers¹. After curating these SV calls, GIAB has made public a draft clonal Benchmark SV callset including deletions, duplications, and complex translocations and inversions. Here, we tested the capability of Dovetail® LinkPrep™ technology, a linked-read NGS-based assay enabling high-sensitivity SV detection, to reproduce the draft GIAB Benchmark callset using both open source and proprietary SV callers.

Methods

LinkPrep libraries were generated for the HG008-T cell line (p21) and sequenced to ~55x coverage. Sequencing data was aligned to hg38 with bwa and SV calls generated by running a standard, HiC-based SV caller, HiC-Breakfinder² (HiC-B), and Dovetail Precise (DP), a new SV caller developed by Dovetail Genomics. These callsets were compared to the curated GIAB Benchmark callset to assess their performance based on SV type and size.



Results

Tool Overview and Comparison

	HiC-Breakfinder	Dovetail Precise
Description	- Contact map-based SV caller that looks for unexpected linkage within/between chromosomes. - Makes calls at various resolutions (bin sizes). - Breakpoint output is imprecise (bin-level resolution of breakpoint).	- Blend of multiple SV calling methods designed to take fuller advantage of the unique aspects of LinkPrep data. - Utilizes information from both long-range linkage density (kb resolution) as well as read-level sequencing data (bp resolution).
Advantages	High sensitivity for inter-chromosomal SVs calls	- Can provide precise location & orientation of breakpoints. - No limitation on SV size or orientation. - Reduced false positives.
Limitations	- Challenged at detecting smaller (< 1Mb) intra-chromosomal SVs. - Cannot detect SVs smaller than called resolution. - Error-prone in complex regions harboring many copy number, structural alterations.	Base pair breakpoint resolution requires sufficient coverage over breakpoint.

Dovetail Precise Improves Recall of GIAB SVs

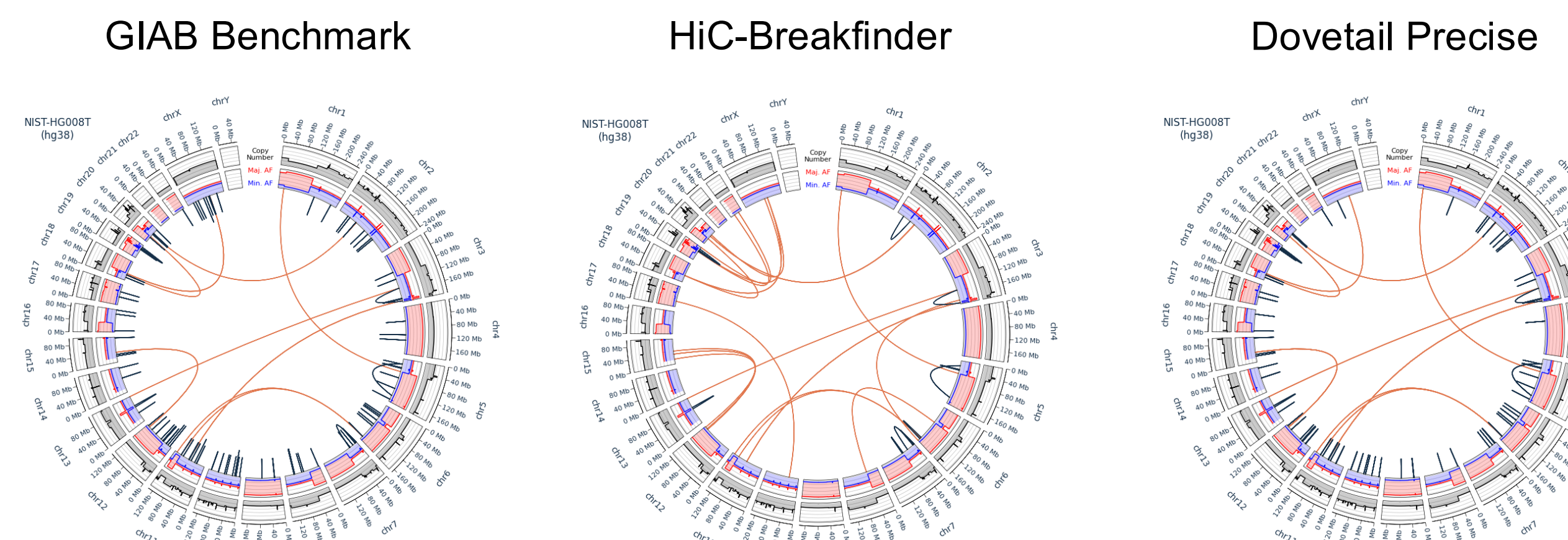


Figure 2. Circos plot overview of the GIAB SV benchmark callset (truth) against calls made by two tools (HiC-Breakfinder and Dovetail Precise) on 55x HG008-T LinkPrep Data.

Category	Type	GIAB	LinkPrep + HiC-Breakfinder	LinkPrep + Dovetail Precise
Inter-chrom	TLOC	12	8 – 11 (67 – 92%)	12 (100%)
Large Intra (> 2 Mb)	INV	5	5 (100%)	5 (100%)
	DEL	3	3 (100%)	3 (100%)
	DUP*	2	2 (100%)	2 (100%)
Medium Intra (100 kb – 2 Mb)	DUP	36	0 (0%)	35 (97%)
	INV	1	0 (0%)	1 (100%)
	DEL	6	0 (0%)	4 (67%)
Small-Medium Intra (5 – 100 kb)	DUP	7	0 (0%)	5 (71%)
	INV	8	0 (0%)	8 (100%)
Small Intra (100 bp – 5 kb)	DEL	17	0 (0%)	13 (76%)
	DUP	10	0 (0%)	9 (90%)
Tiny Intra (< 100 bp)	DEL	10	0 (0%)	9 (90%)
	ALL	108	19 – 21 (18 – 20%)	97 (90%)

Table 1. Recall performance comparison between HiC-Breakfinder and Dovetail Precise on 55x HG008-T LinkPrep data against the GIAB SV benchmark. Hi-C Breakfinder incorrectly calls orientation for 25% (3 of 12) inter-chromosomal truthset SVs and fails to call truthset SVs < 2 Mb. Dovetail Precise recalls 90% of truthset SVs across all SV size ranges. NOTE: Mobile insertions are not included in this analysis.

Dovetail Precise Achieves Base-Level Precision on SV Breakpoints

	Breakend Δ from GIAB HiC-Breakfinder	Breakend Δ from GIAB Dovetail Precise
Inter-chrom	M = 7,887 bp (3,363 - 27,196 bp)	M = 1 bp (0 - 1,363 bp)
Large Intra (> 2 Mb)	M = 4,117 bp (1,636 - 188,989 bp)	M = 1 bp (0 - 669 bp)
Other Intra (1 bp – 2 Mb)	N/A	M = 0 bp (0 - 16 bp)

Table 2. Breakend accuracy comparison between HiC-Breakfinder and Dovetail Precise on 55X HG008-T LinkPrep data. To determine the difference in breakend assignment by HiC-Breakfinder or Dovetail Precise the GIAB breakend call +/- CI was compared against the breakend assigned by the respective tool. Hi-C Breakfinder achieves bin-level (imprecise) resolution of the breakpoint position. Dovetail Precise achieves precise calls closely matching the GIAB benchmark.

Dovetail Precise Calls Small SVs that HiC-Breakfinder Cannot Detect, Enabling Accurate Local Reconstruction

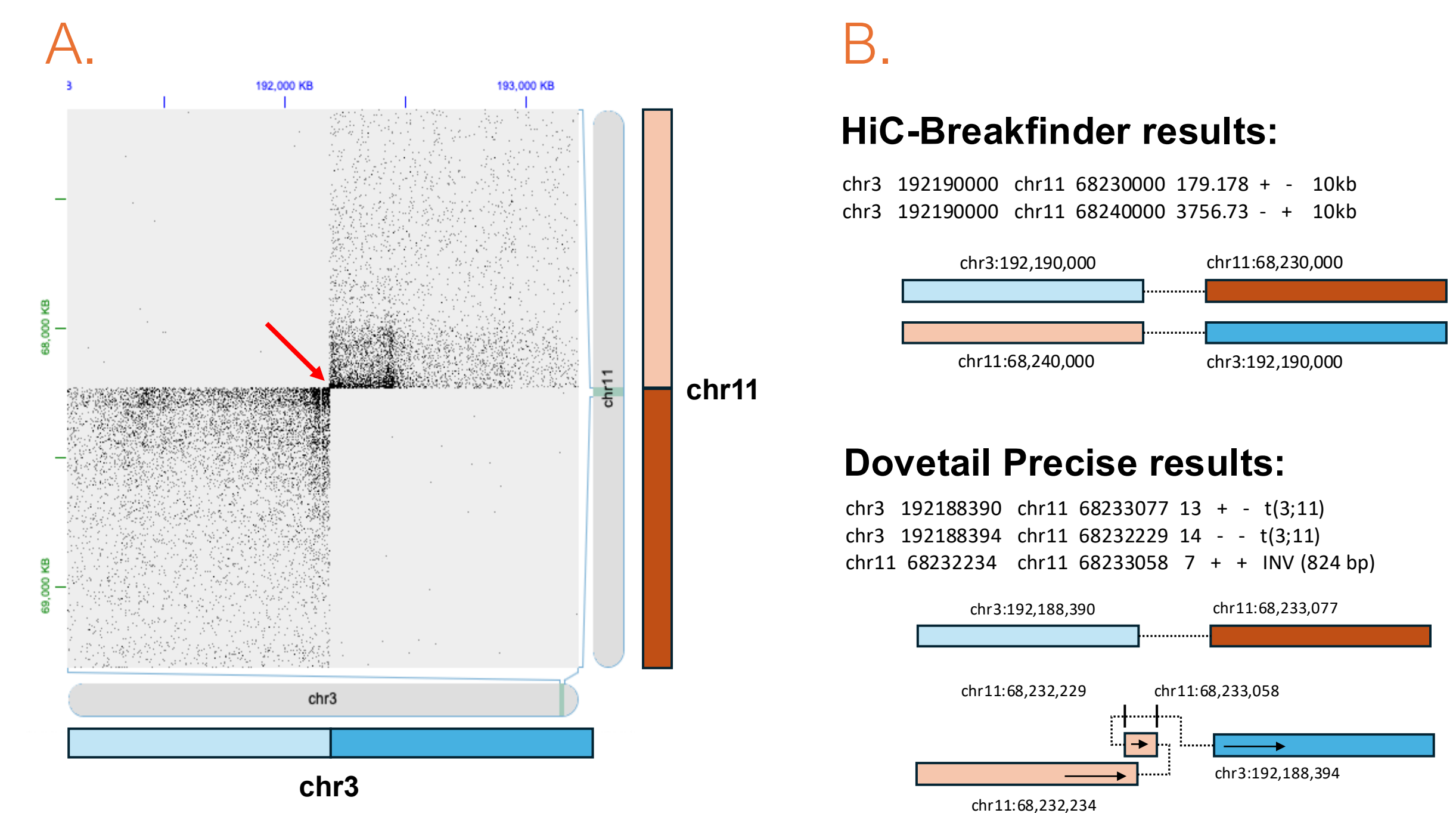


Figure 3. A. Link-density plot of t(3;11) balanced translocation. **B.** Hi-C Breakfinder misses a small (824 bp) inversion involved in the t(3;11) event resulting in its incorrect (local) call orientation. Dovetail Precise accurately calls all breakends and orientations involved in t(3;11) event to enable accurate local reconstruction.

Summary

Dovetail® LinkPrep™ technology is a linked-read NGS assay with comprehensive genome coverage. LinkPrep delivers highly sensitive structural variant detection while maintaining strong performance for SNVs, Indels, and CNVs – enabling more comprehensive genomic variant characterization with a single sequencing approach.

SV recall performance was benchmarked using 55x LinkPrep HG008-T data and two callers: Hi-C Breakfinder (open-source tool) and Dovetail Precise (proprietary tool in development). Comparisons were made against the GIAB SV benchmark v0.3.

LinkPrep + HiC-Breakfinder's ability to detect SVs was limited to imprecise calls on translocations and large intra-chromosomal SVs (> 2 Mb) achieving only 18% recall of the SV truthset.

LinkPrep + Dovetail Precise achieved 90% recall of the SV truthset and accurate base-level precision on SV breakpoints.

LinkPrep + Dovetail Precise offer promise of enabling more accurate diagnostics and novel therapeutic target development.

References

- McDaniel, J.H. *et al.* Development and extensive sequencing of a broadly-consented Genome in a Bottle matched tumor-normal pair. *Scientific Data* (2025).
- Dixon, J.R. *et al.* Integrative detection and analysis of structural variation in cancer genomes. *Nature Genetics* (2018).