Interpreting Structural Variations in Breast Cancer Cell Lines Using 3D Genome Data

Jingyu Hao[1], Yongyi Luo[2], Jiandong Shi[2], Depeng Wang[3], Shu Wang[4], Xiaodan Fan[2], and Weichuan Yu[1]

[1] Electronic and Computer Engineering, The Hong Kong University of Science and Technology, HK, China
[2] Statistics and Data Science, The Chinese University of Hong Kong, HK, China
[3] GrandOmics Inc, China
[4] Department of Breast Surgery, Peking University People's Hospital, China

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

WHY IT MATTERS

Structural variations (SVs) are a major source of genomic alterations in cancer, yet the majority of SVs occur in non-coding regions and remain difficult to interpret using linear genome annotations alone. 3D genome data provide additional spatial information for understanding potential regulatory impacts of SVs.

WHAT WE DID

01 Cross-cell-line survey
A unified comparison of linear and spatial SV context.

02 Same-donor comparison
A same-donor comparison reduces inter-individual genetic variation.

03 Pore-C spatial readout
Distal contacts and local boundary changes at candidate loci.

WHAT WE USED

Data

HCC1937
Long-read + Pore-C

HCC1954
Long-read + Hi-C

SK-BR-3
Long-read + Hi-C

Tools

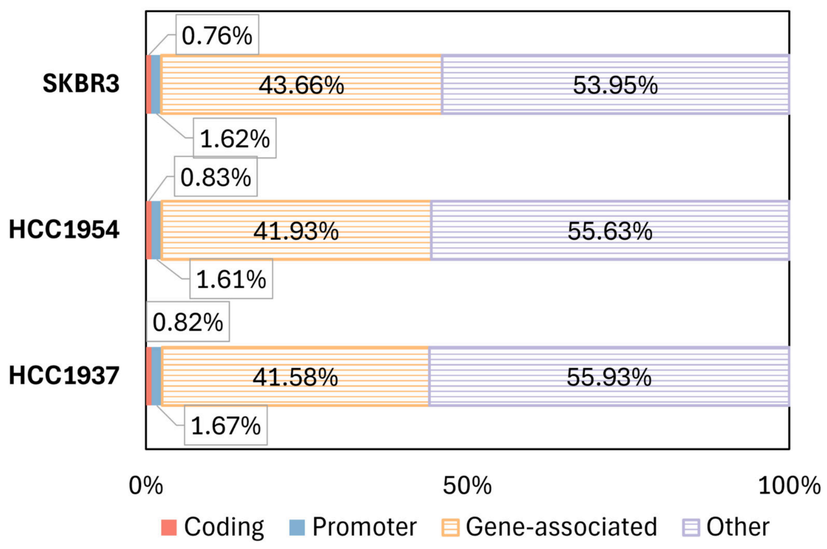
- Alignment: minimap2
- SV calling: cuteSV, DeBreak, gSV, SVision-pro, Sniffles
- Merging: SURVIVOR
- Annotation: ANNOVAR
- 3D genome analysis: HiCExplorer

Linear annotation leaves a gap

High-confidence SV calls

Five SV callers · retained when supported by at least three tools

16,373 HCC1937 | **18,726** HCC1954 | **13,730** SK-BR-3



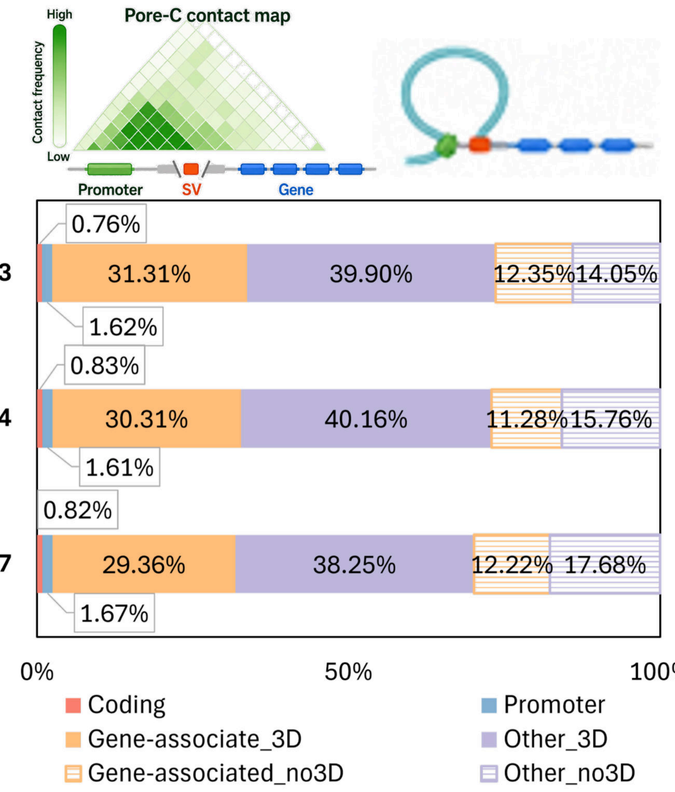
> 97 %
outside coding / promoter regions

3D contacts extend the annotation space

Differences in 3D chromatin architecture between HCC1937 and its matched control

35.7 %
A/B compartment switching

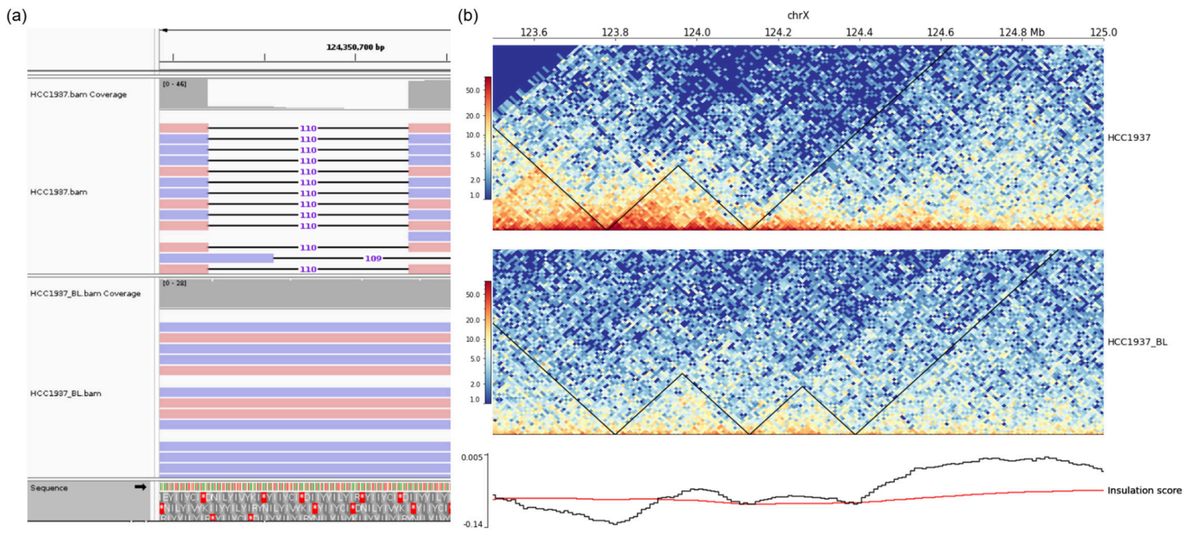
3D chromatin contacts provide additional functional annotation for SVs



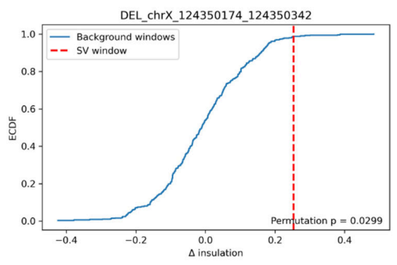
~ 70 %
SVs with high-contact interactions with distal promoters

Representative somatic SVs associated with 3D chromatin remodeling

CASE A | chrX intronic DEL in SH2D1A



(c)

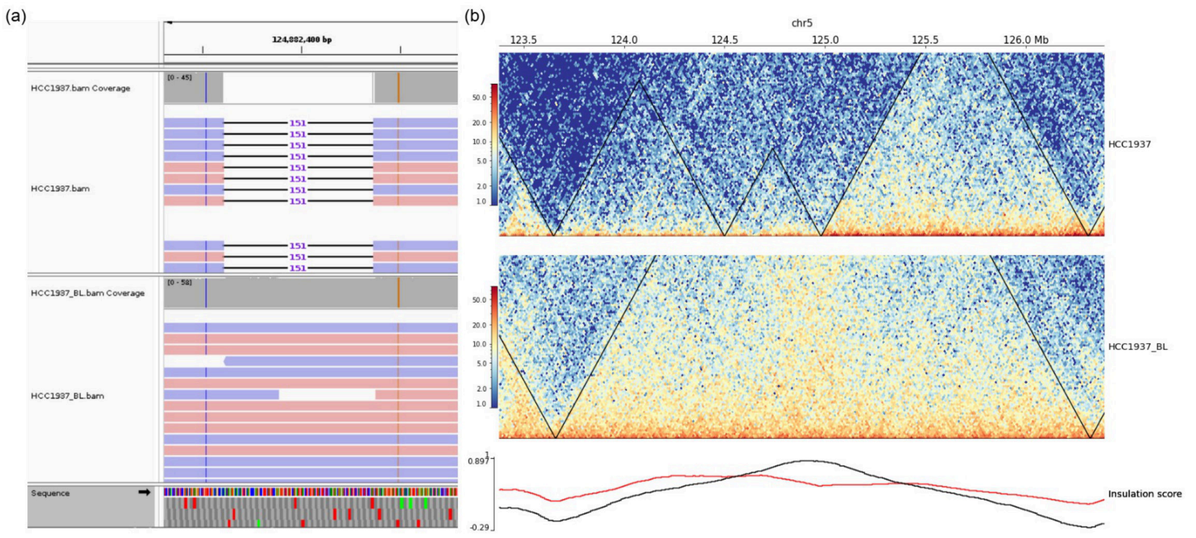


(d)

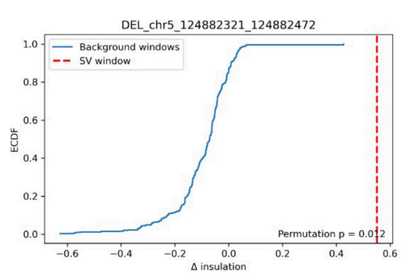
Chrom	Gene	Start	End	Chrom	Gene	Start	End
chr1	SLC2A1	42925354	42958868	chr2	FAP	162170683	162243445
chr7	HGF	81699011	81770047	chr8	CTSB	11842524	11868087
chr15	IQGAP1	90388241	90502239	chr15	NOX5	68930524	69062762
chrX	STAG2	123960559	124102655	chrX	NXF2B	102360394	102440008
chrX	XIAP	123859707	123913972	chrX	OCRL	129540258	129592553

Loss of a pre-existing TAD boundary

CASE B | chr5 intergenic DEL



(c)



(d)

Chrom	Gene	Start	End	Chrom	Gene	Start	End
chr5	SMAD5	136132844	136182734	chr5	CAVIN3	6318945	6320501
chr7	RHEB	151466011	151519895	chr5	GRIA1	153490523	153813873
chr8	EYA1	71197432	71548094	chr5	ZNF608	124636914	124748824
chr8	LOXL2	23296896	23404120	chr9	RAPGEF1	131576774	131740076
chr9	RAPGEF1	131576774	131740076	chr9	TLE4	79571772	79726881

Emergence of a new TAD boundary

Scan to learn more