

Long-read sequencing reveals multi-break rearrangements and aneuploidy patterns in cancer genomes

Ayşe Keskus¹, Tanveer Ahmad¹, Ataberk Donmez¹, Yi Xie², Isabel Rodriguez², Rose Milano², Nicole M. Rossi², Hong Lou³, Michael Dean², Mikhail Kolmogorov¹

¹Center for Cancer Research, NCI, Bethesda, MD, USA; ²Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA; ³Leidos Biomedical Research, Inc., National Laboratory for Cancer Research, Frederick, MD, USA;

ABSTRACT

Flowchart

Background

Structural variations (SV) including insertions, deletions, inversions, translocations and more complex events have demonstrated to contribute a large portion of population diversity as well as play a key role in a wide range of pathogenesis. Therefore comprehensive analysis of SVs is essential to decipher the genetic role players in population diversity, disease progression and tumorigenesis.

Method

We developed Severus to detect SVs from simpler indels to complex events with multiple breakpoints. Severus utilizes a framework that combines the ideas from long-read assembly and breakpoint graph methods. Severus uses abnormally mapped reads to build a breakpoint graph in which complex events with multiple breakpoints form connectivity clusters. Clusters are then classified based on the subgraph properties to characterize the structure of the derived cancer genome.

Severus also takes advantage of phased haplotypes and can incorporate multiple related datasets (such as in tumor-normal comparison or multi-site tumor sampling).

1. Alignment: minimap2



2. Phasing: PEPPER-Margin-DeepVariant



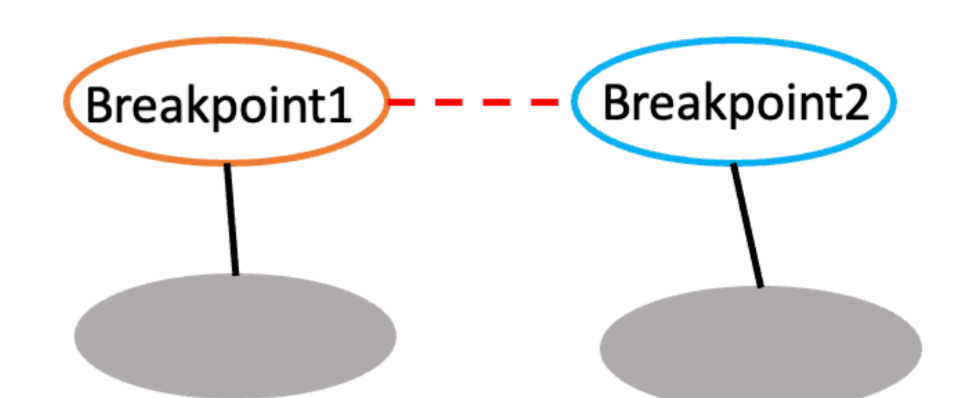
3. Extracting split reads



4. Breakpoint clustering



5. Build breakpoint graph



6. Cluster & output multi-breakpoint SVs

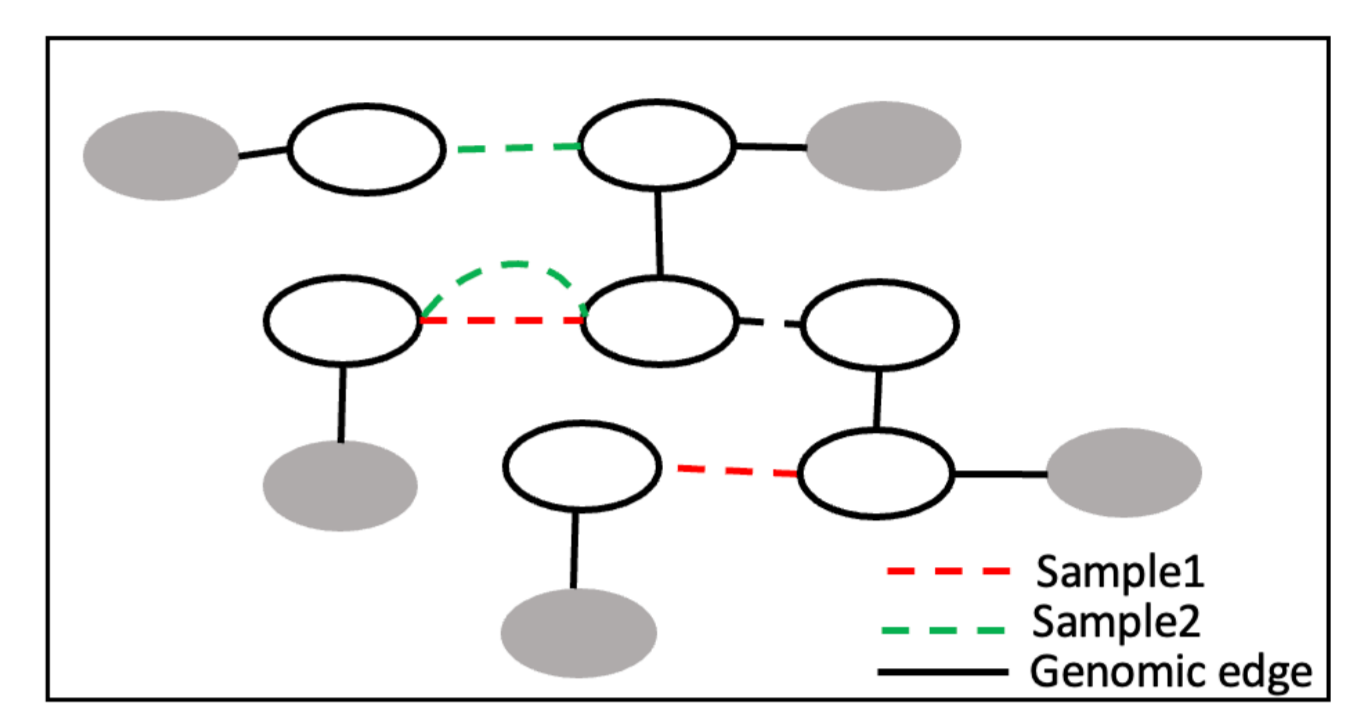


Figure 1: Workflow of structural variation detection using Severus

Structural Variations

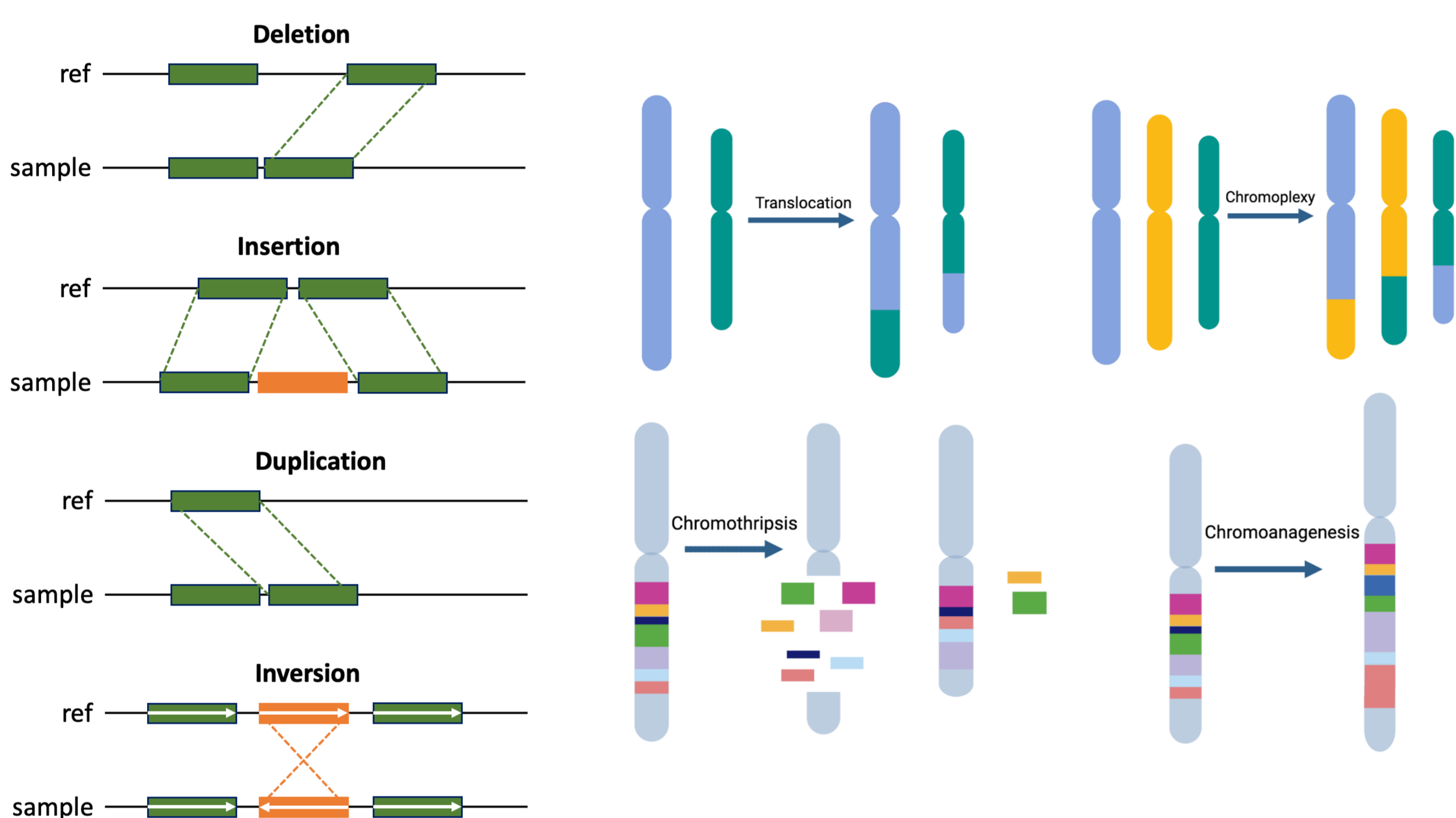


Figure 2: Examples of structural variation types from simple rearrangements (left) to more complex events; chromoplexy, chromothripsis, chromoanagenesis (right).

Insertion and Deletion- COLO829

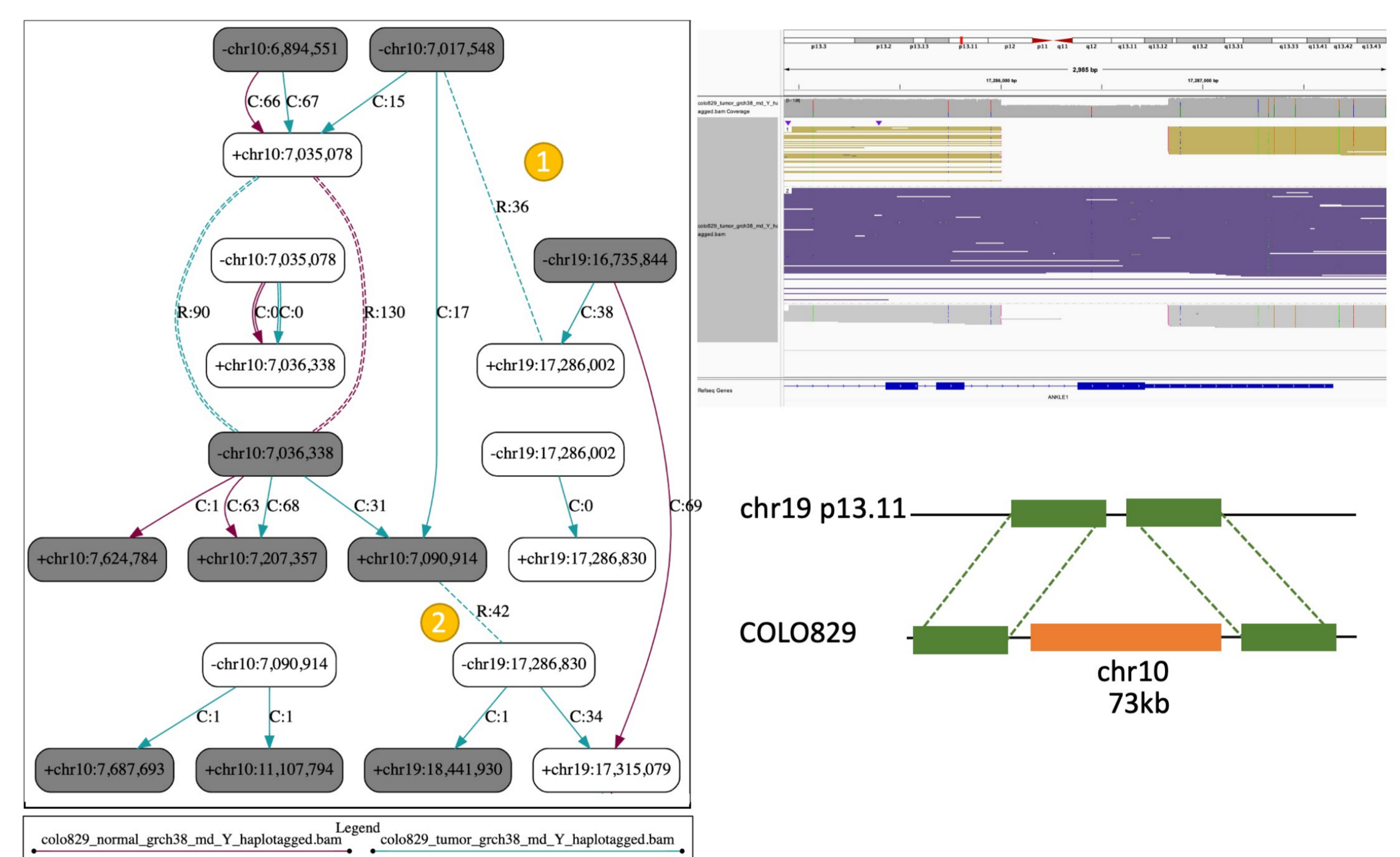


Figure 3: Haplotype-specific, somatic intrachromosomal insertion (chr10) and deletion (chr19) event in ANKLE1 gene functional in DNA damage repair. Breakpoint graph for the event (left). The breakpoints were labelled in genomic order. IGV representation of the breakpoints in chr19 (top). Schematic representation of the SV (bottom).

Deletion and Insertion - HCC1937

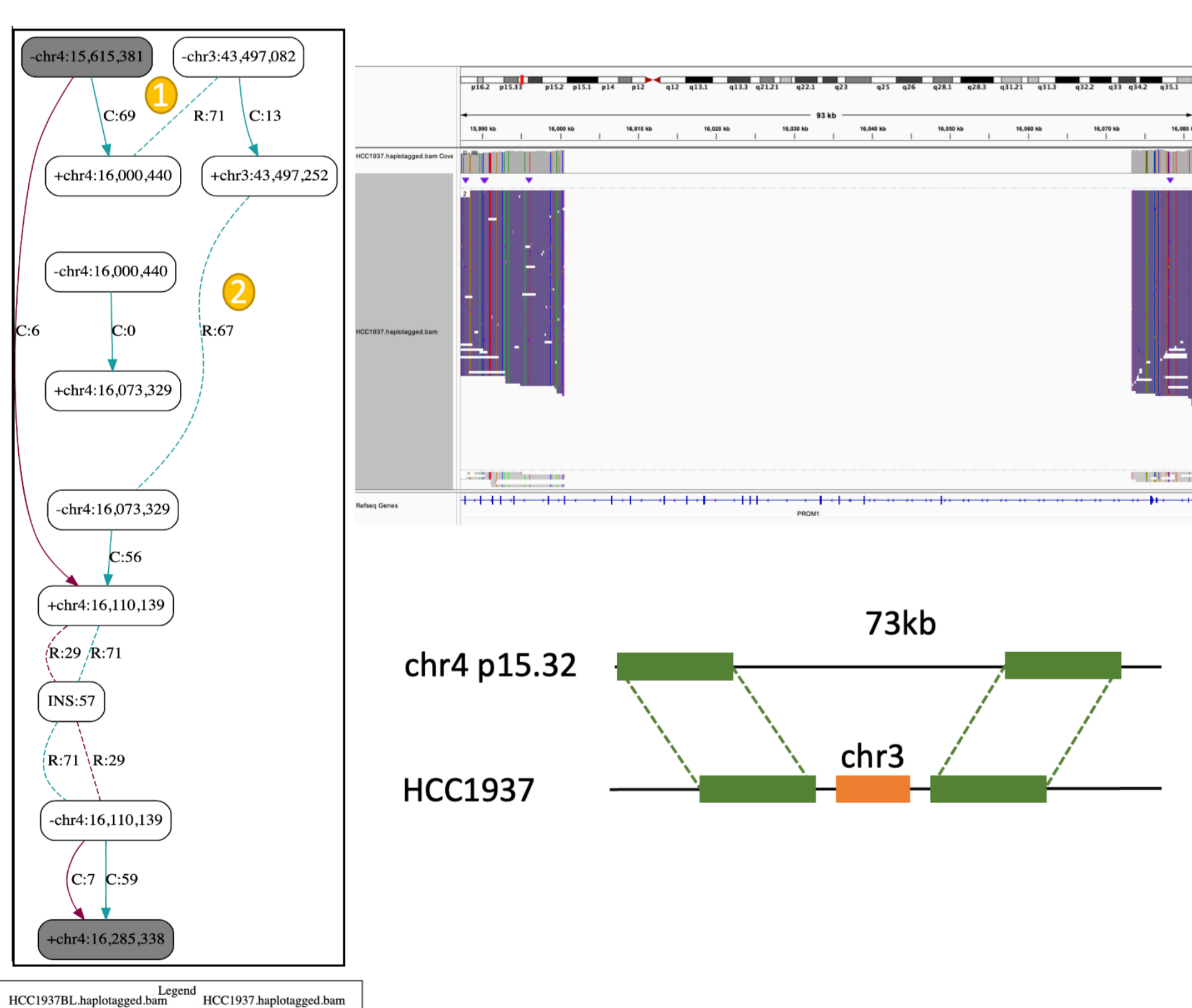


Figure 4: Haplotype-specific, somatic deletion (chr4) and intrachromosomal insertion (chr3) event in PROM1 gene; known cancer stem cell marker. Breakpoint graph for the event (left). IGV representation of the breakpoints in chr4 (top). Schematic representation of the SV (bottom).

Deletion and Inversion- H2009

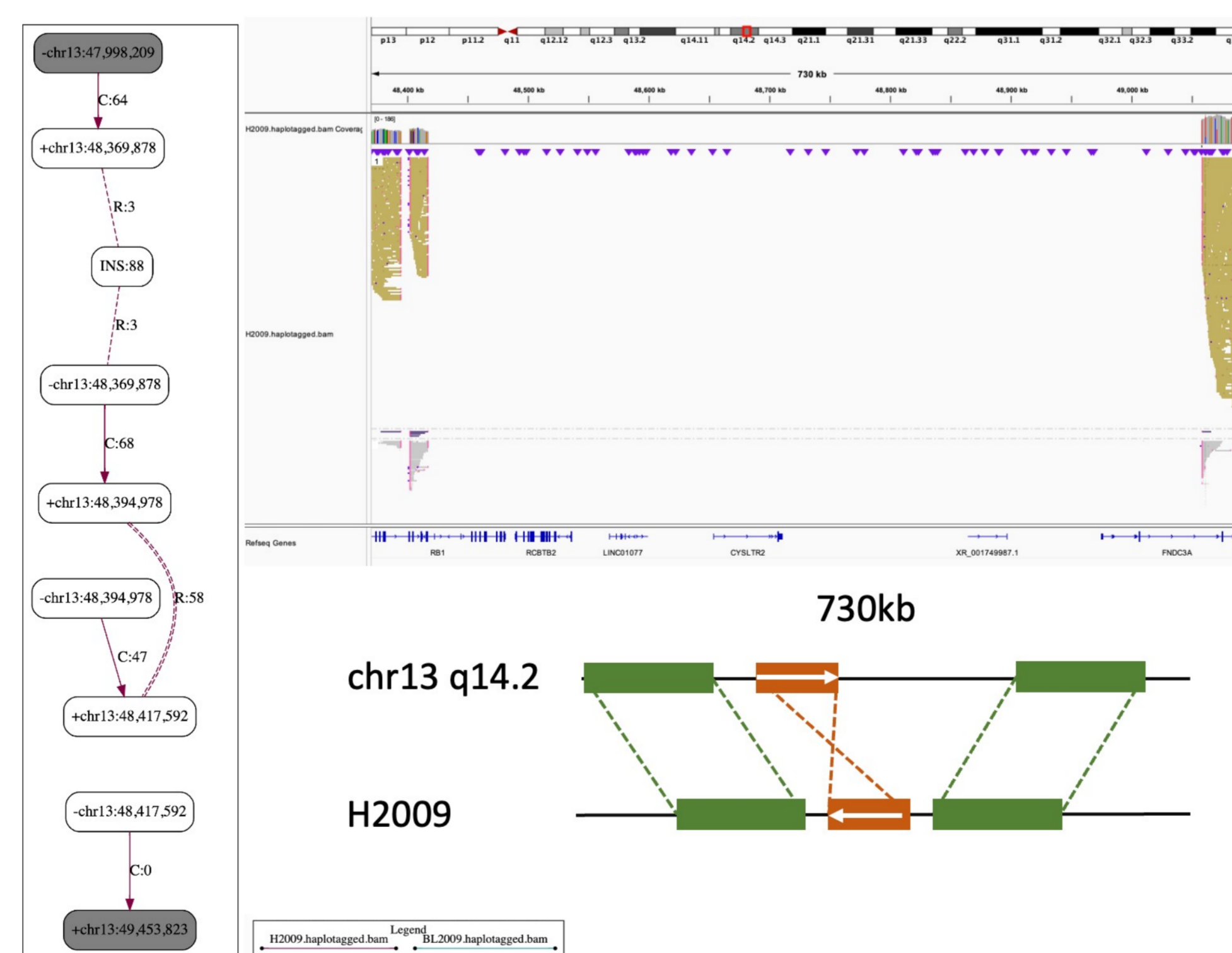


Figure 5: Breakpoint graph for somatic homozygous deletion and inversion event in H2009 cell line (A). IGV view of the region of interest (top). Schematic representation of the SV (B).

Chromoplexy in CaSki

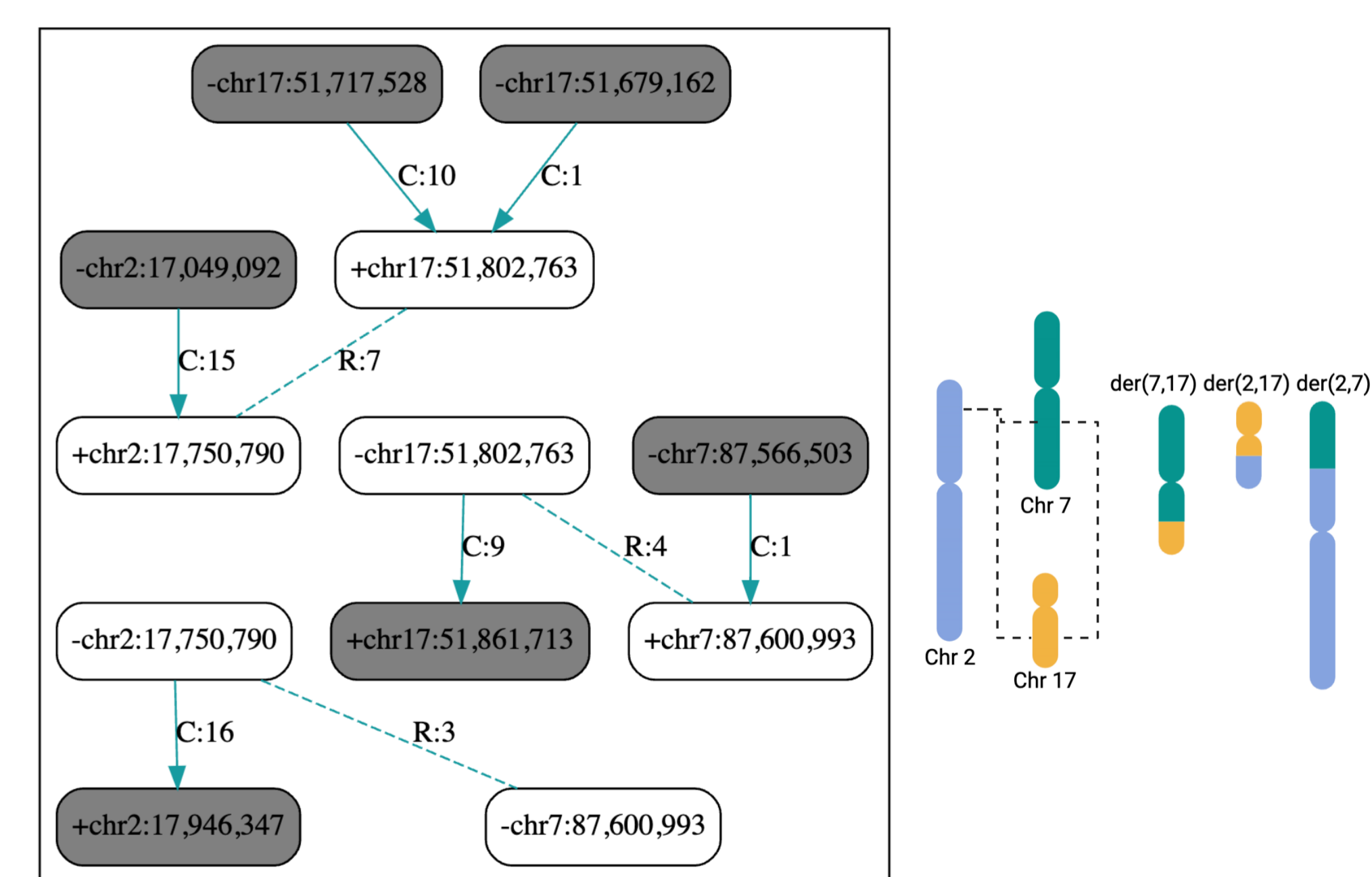


Figure 6: Breakpoint graph for chromoplexy events between chr2, chr7 and chr17 in CaSki cell line (A). Schematic representation of the SV (B).

Amplification and Insertion in H2009

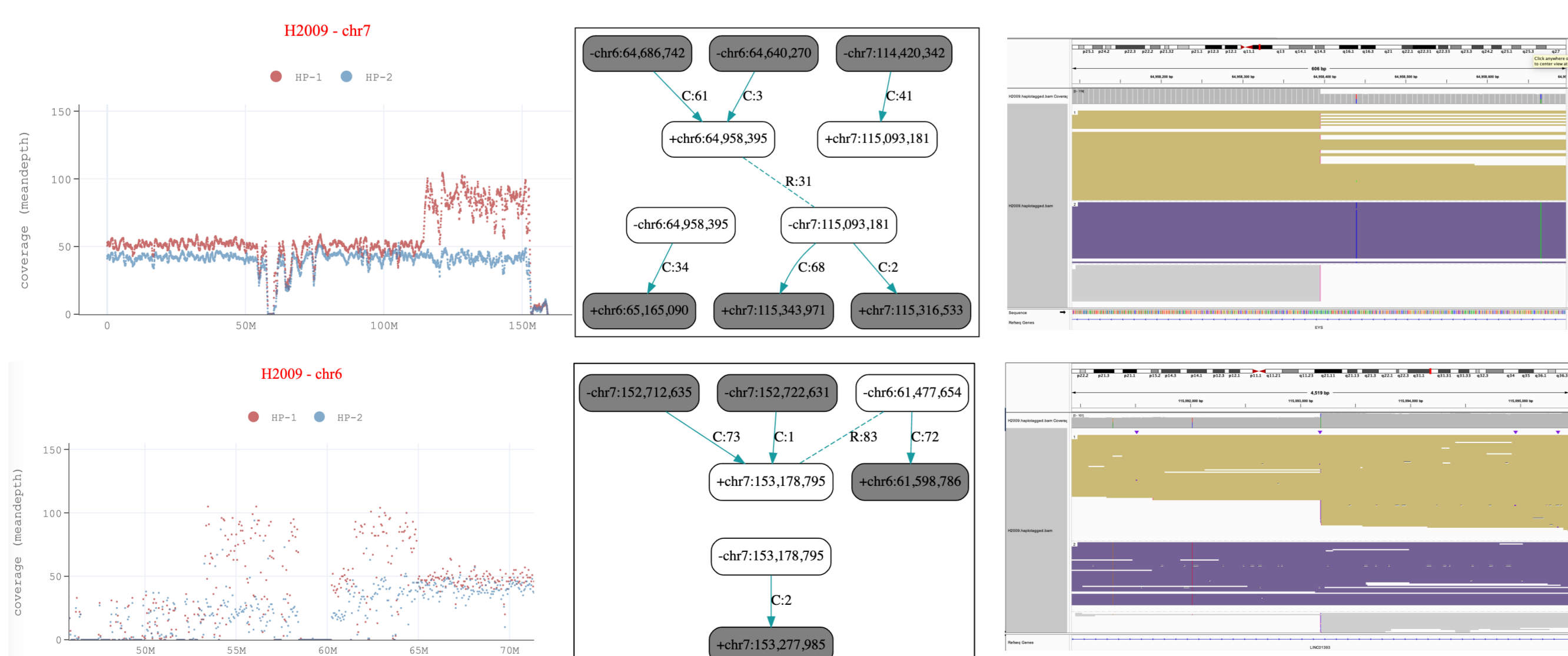


Figure 7: Intrachromosomal insertion and amplification between chr6 and chr7 in H2009 cell line

Breakage-Fusion-Bridge

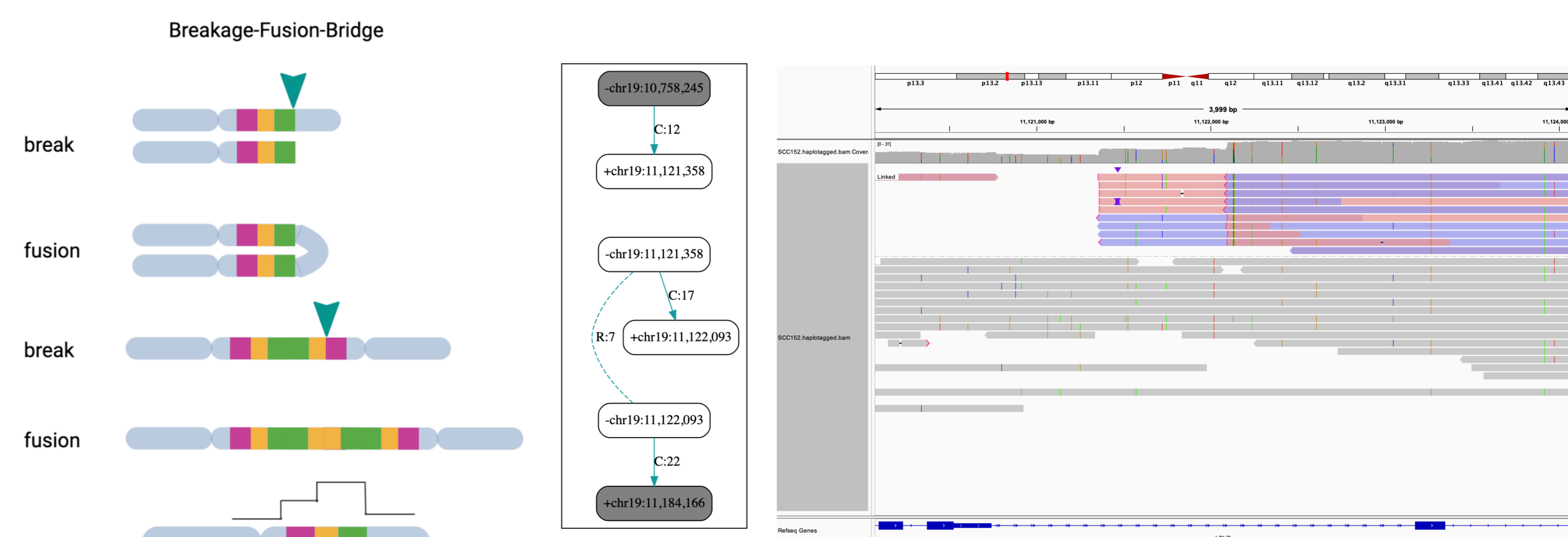


Figure 8: Schematic representation of breakpoint graph of breakage-fusion-bridge (BFB) cycle event. BFB in chr19 in SCC152 (A).