
Complete Genome analysis: Structural variation detection

SVDetect tutorial

Alban LERMINE – Institut Curie, INSERM U900, Mines ParisTech

Workshop outlines

1) SVDetect tool presentation

Methodology & functionalities

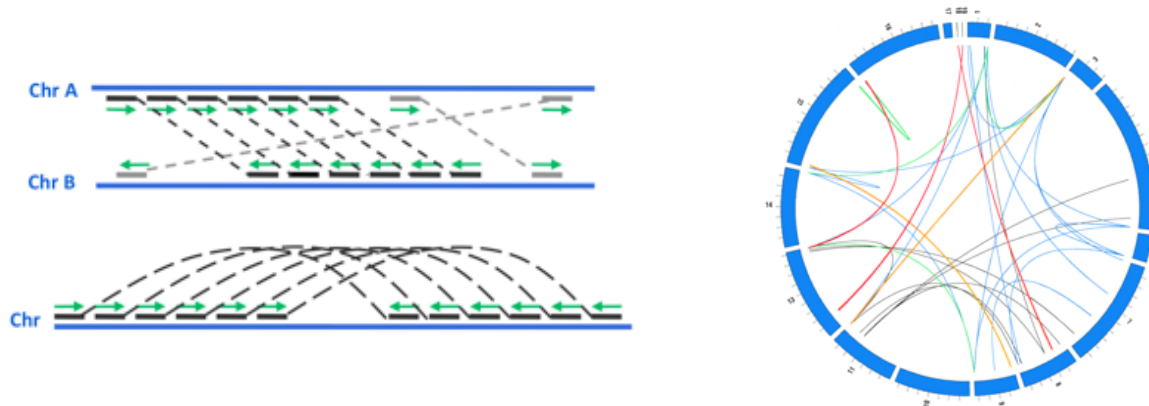
2) SVDetect tutorial on Galaxy

Hands on workshop

SVDetect

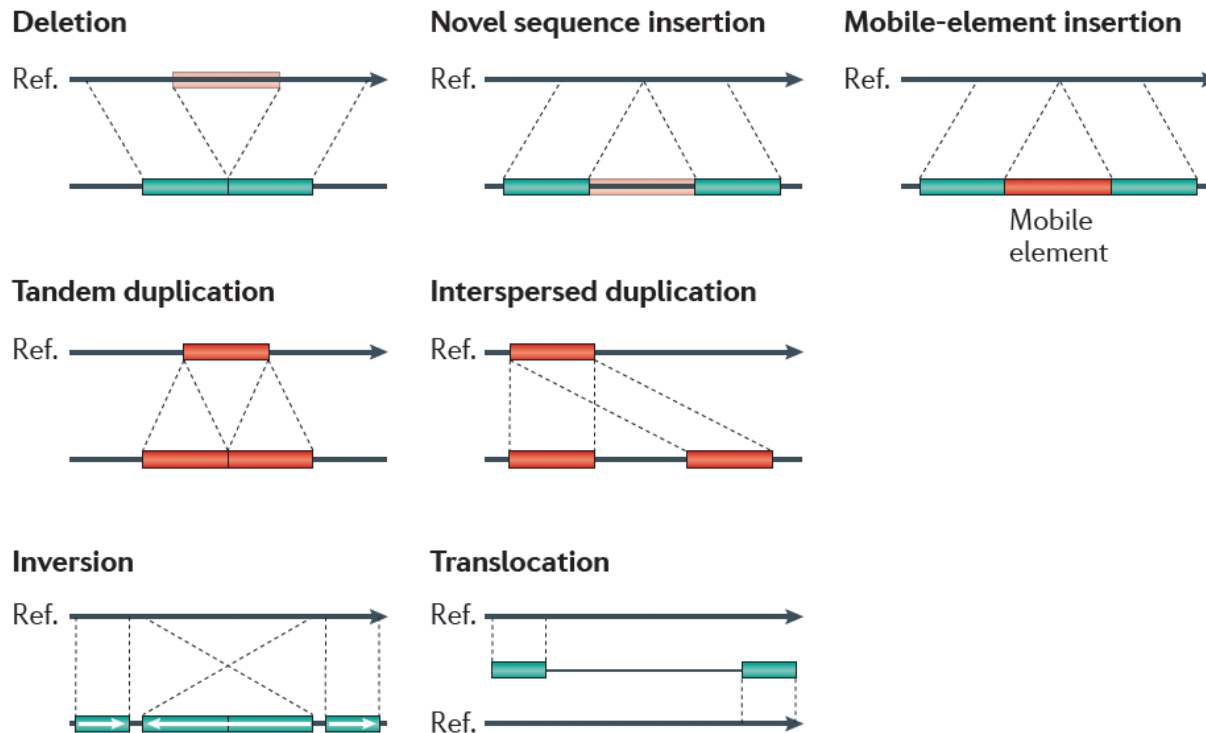
SVdetect is a tool dedicated to detection and prediction of inter/intra chromosomal rearrangements from paired-end/mate-pair sequencing data

This tool detect structural variation (SV's) by using sliding windows and clustering strategies, and also allows you to visualize them at genome scale



Different classes of SV's

Lot of Tumors have chromosomal mutations in their genome. This can affect the number of chromosomes, and also its structure, implicating large events such as deletions, insertions, duplications, inverisions,translocations, etc.



SV's discovery from sequencing data

4 different methods:

- Based on coverage depth differences (Control- FREEC)
- Splitted reads method, local reads realignment around breaking point (Pindel)
- De novo assembly, new sequences - insertions -
- Paired-end/mate-pair sequences constraints (Breakdancer, PEMer, SVDetect)

Reviews on thema:

Computational methods for discovering structural variation with next-generation sequencing

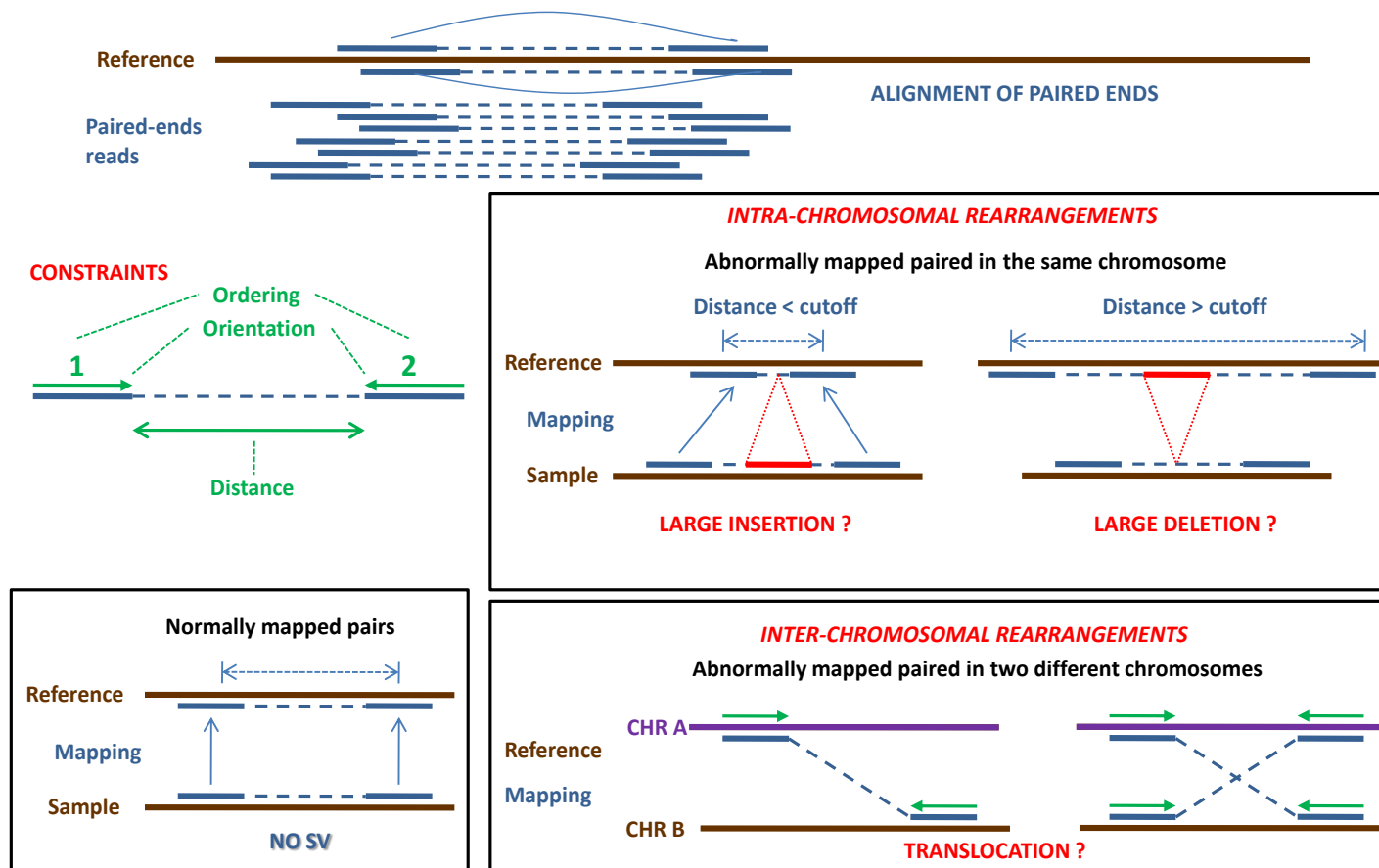
Paul Medvedev et al, Nature Methods 6, S13 - S20 (2009)

Genome structural variation discovery and genotyping.

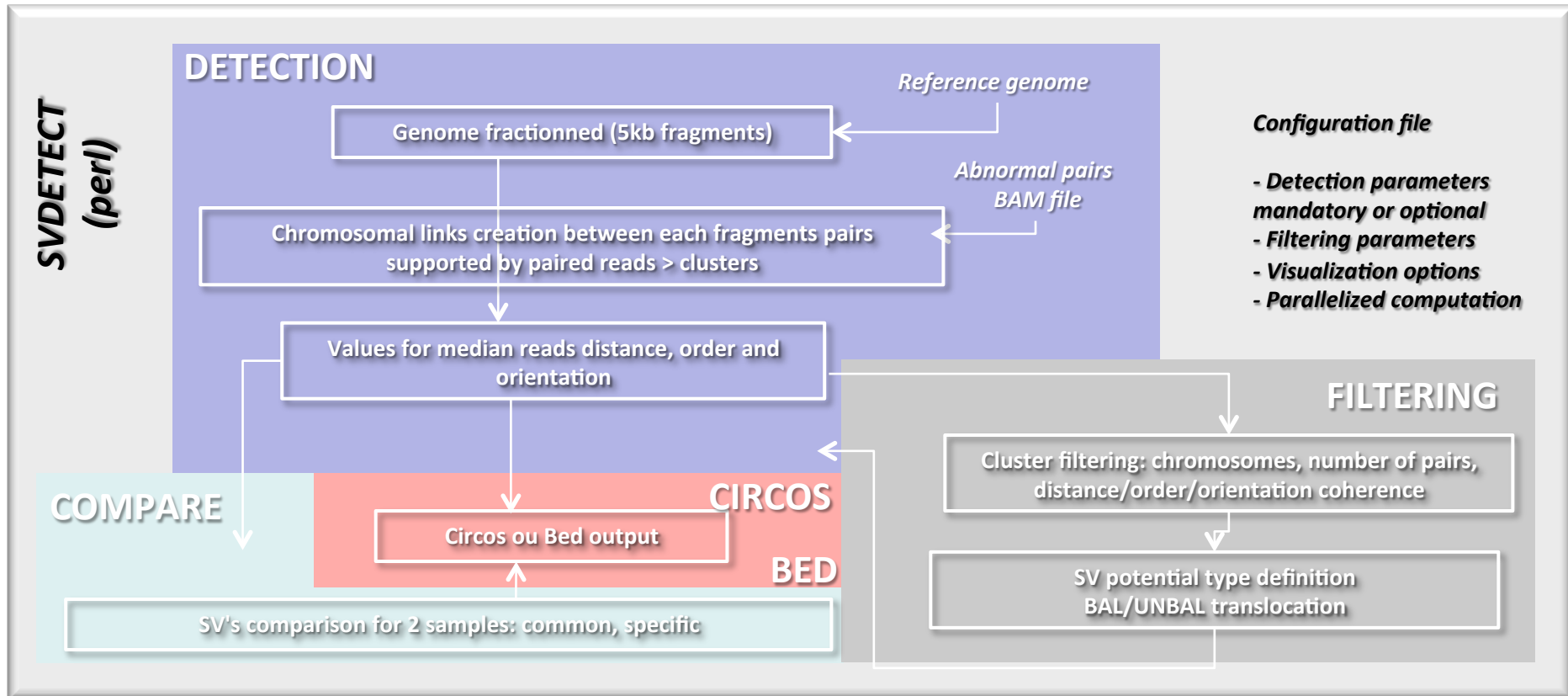
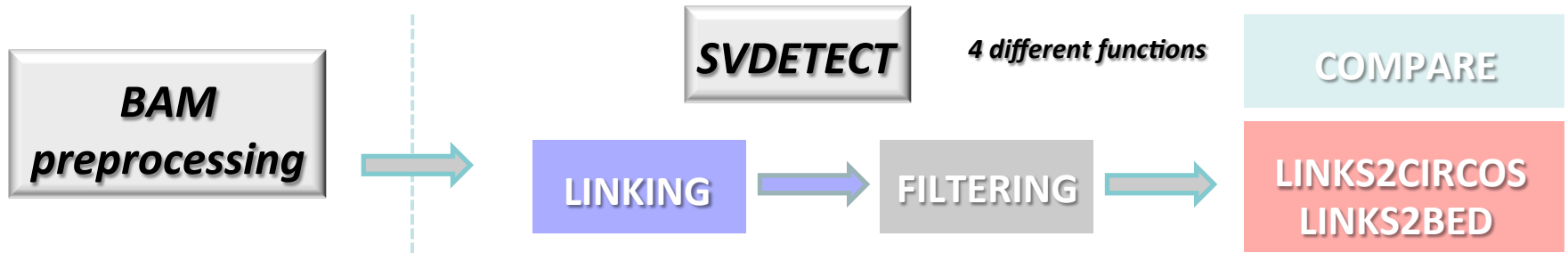
Alkan C et al, Nat Rev Genet. (2012)

SVDetect - strategy

Paired-end/mate-pair based methods, such as SVDetect, evaluate reads distance, order, and orientation after alignment on a reference genome. Abnormal or discordants pairs are clustered to define their type of SV



SVDetect - implementation



Testing data

Biological data coming from neuroblastoma tumor

Two samples: Tumoral vs Control (Constitutional DNA)

Illumina GAI mate-pair sequencing, 50x50 bp, insert size of 3kb

Alignments files in binary format: BAM

SVDetect need 2 types of files as input:

- Bam file containing all appariated and aligned reads
 - > « sample_mates.bam » & « reference_mates.bam »
- LEN file giving reference chromosome size
 - > « hs18_chr5_chr11.len »

1	chr5	180857966
2	chr11	134452384

Let's start hands on SVDetect tutorial!

Galaxy Analyze Data Workflow Shared Data Visualization Admin Help User Using 665.6 Gb

Tools Options ▾


- Multiple regression
- Multivariate Analysis
- Evolution
- Motif Tools
- Multiple Alignments
- Metagenomic analyses
- FASTA manipulation
- NCBI BLAST+
- NGS TOOLBOX BETA
- NGS: QC and manipulation
- NGS: Picard (beta)
- NGS: Mapping
- NGS: Indel Analysis
- NGS: RNA Analysis
- NGS: Target Analysis
- NGS: SAM Tools
- NGS: BED Tools
- NGS: GATK Tools (old beta)
- NGS: GATK Tools (beta)
- NGS: Peak Calling
- NGS: Peak Annotation
- NGS: Simulation
- NGS: SVDetect
 - DETECTION
 - Import data BAM, chromosome info or sv files
 - BAM preprocessing to get abnormal pairs
 - Detect clusters of anomalously mapped pairs and identify structural variants
 - Compare structural variants between two samples
 - VISUALIZATION
 - Circos plots
- NGS: GFAP
- NGS: GFAP r1.0
- RGENETICS
 - SNP/WGA: Data: Filters
 - SNP/WGA: QC; LD; Plots
 - SNP/WGA: Statistical Models
 - Human Genome Variation
 - Genome Diversity
 - VCF Tools
 - PacBio/Illumina Assembly

History Options ▾

Unnamed history 0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

WWFSMD?
grow noodly appendages...



usegalaxy.org

This project is supported in part by NSF, NHGRI, and the Huck Institutes of the Life Sciences.

NGS: SVDetect

DETECTION

- Import data BAM, chromosome info or sv files
- BAM preprocessing to get abnormal pairs
- Detect clusters of anomalously mapped pairs and identify structural variants
- Compare structural variants between two samples

VISUALIZATION

- Circos plots