



Structural variant calling - long read data

Anticipated workshop duration when delivered to a group of participants is **4 hours**.

For queries relating to this workshop, contact Melbourne Bioinformatics (bioinformatics-training@unimelb.edu.au).

Overview

Topic

- ☒ Genomics
- ☐ Transcriptomics
- ☐ Proteomics
- ☐ Metabolomics
- ☐ Statistics and visualisation
- ☐ Structural Modelling
- ☐ Basic skills

Skill level

- ☐ Beginner
- ☒ Intermediate
- ☐ Advanced

This workshop is designed for participants with no command line knowledge. A web-based platform called Galaxy will be used to run our analysis. We *will* be using **1 line awk programs** to process text output, but these will be supplied and explained.

Description

Long reads have turbo-charged structural variant detection - be part of the renaissance!

This tutorial uses **sniffles** to implement a structural variant calling pipeline.

(short summary) Structural Variants (SVs)...

Data: Nanopore reads: bacterial & human (FASTQ), genomic feature annotations (GFF), human reference genome hg38

Pipeline: Read summaries & QC, alignment, SV calling, text processing, visualisation

Tools: NanoPlot, Filtlong, minimap2, CalMD, SortSam, sniffles, VCFsort, VCFannotate, awk, Circos, IGV

Section 1 covers bacterial SV calling and benchmarking of bioinformatics tools.

Section 2 will demonstrate SV calling on a human sample to diagnose a patient condition.

Learning Objectives

At the end of this introductory workshop, you will :

- Be able to perform SV calling in model *and* non-model organisms
 - Be familiar with the current field of SV calling
 - Gain an understanding of *why* and *when* SV calling is an appropriate analysis to perform.
-

Requirements and preparation

#should this section be omitted since it is zoom?

!!! attention "Important" **Attendees are required to bring their own laptop computers.**

At least one week before the workshop, participants should install the software and data files below. This should provide sufficient time for participants to liaise with their own IT support should they encounter any IT problems.

Preparing your laptop prior to starting this workshop

1. Go to: <https://www>.
2. Download and install the [required software](#) (free) on your laptop.
3. Download and install the [required data](#) on your laptop.
4. Check that the software and data are correctly installed by executing this test: ...

Required Software

- No additional software needs to be installed for this workshop.

Required Data

- No additional data needs to be downloaded for this workshop.
-

Author Information

Written by: Grace Hall

Melbourne Bioinformatics, University of Melbourne

Created/Reviewed: March 2021

Background

What is structural variation?

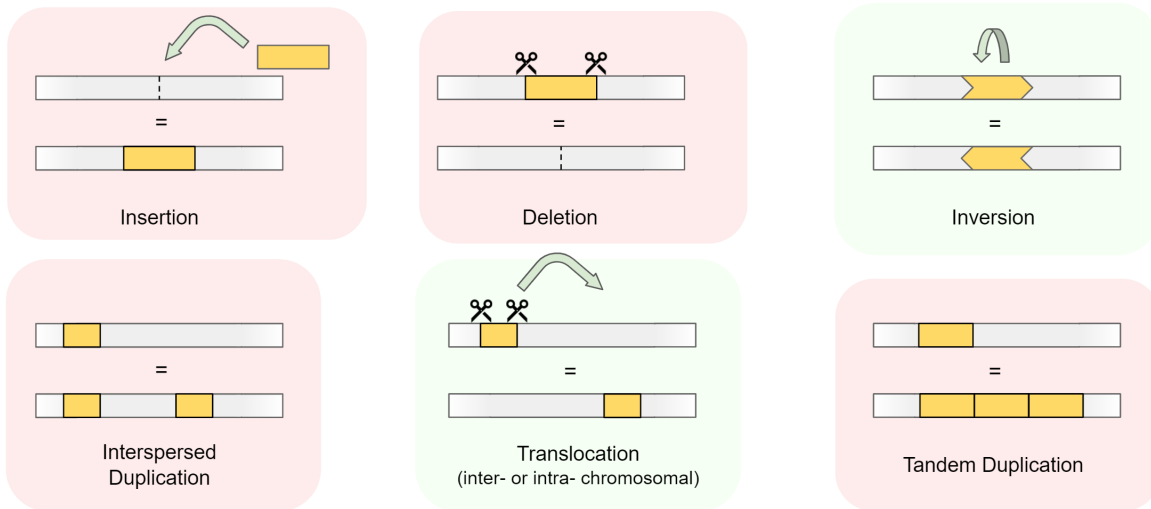
General info (cover the common terms section of slides at least)

SV types

SV Types

Destructive (non-balanced)

Non-destructive (balanced)



biological mechanism/motivation - human and bacteria

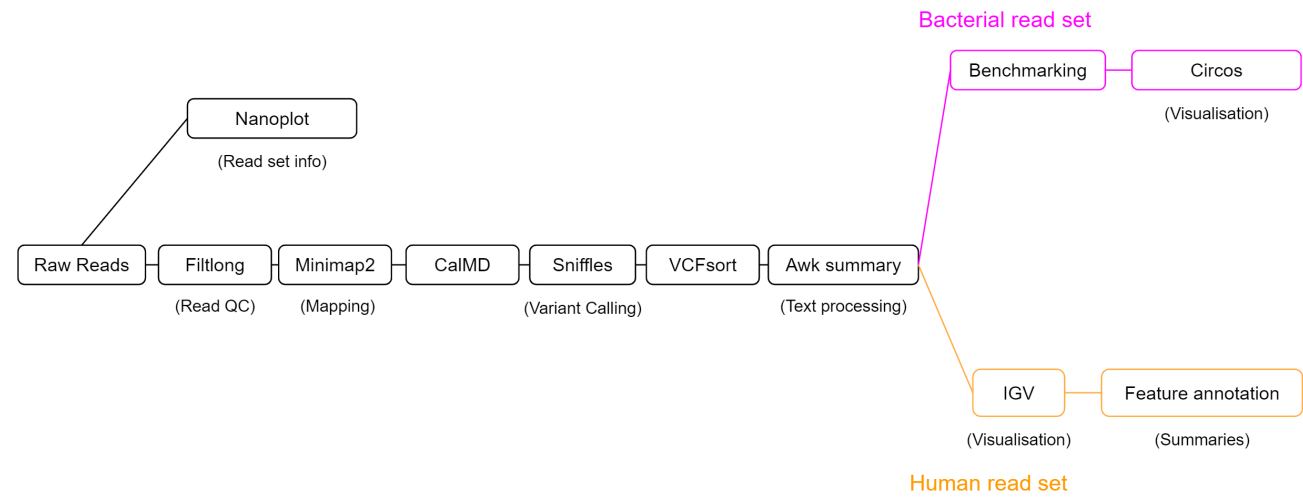
Implications of structural variation

- human - disease
- plant - agriculture
- bacteria - HGT/evolution, virulence / antibiotic resistance

SV calling pipeline

SV calling consists of 5 key steps:

- Read QC
- Alignment
- SV calling
- Feature annotation
- Visualisation



Section 1: Title of section 1

In this section you will (bacterial dataset)...

- Read QC
- alignment
- SV calling
- awk text processing
-

Section 2: Title of section 2

In this section we will ...

Additional reading

Links to additional recommended reading and suggestions for related tutorials.