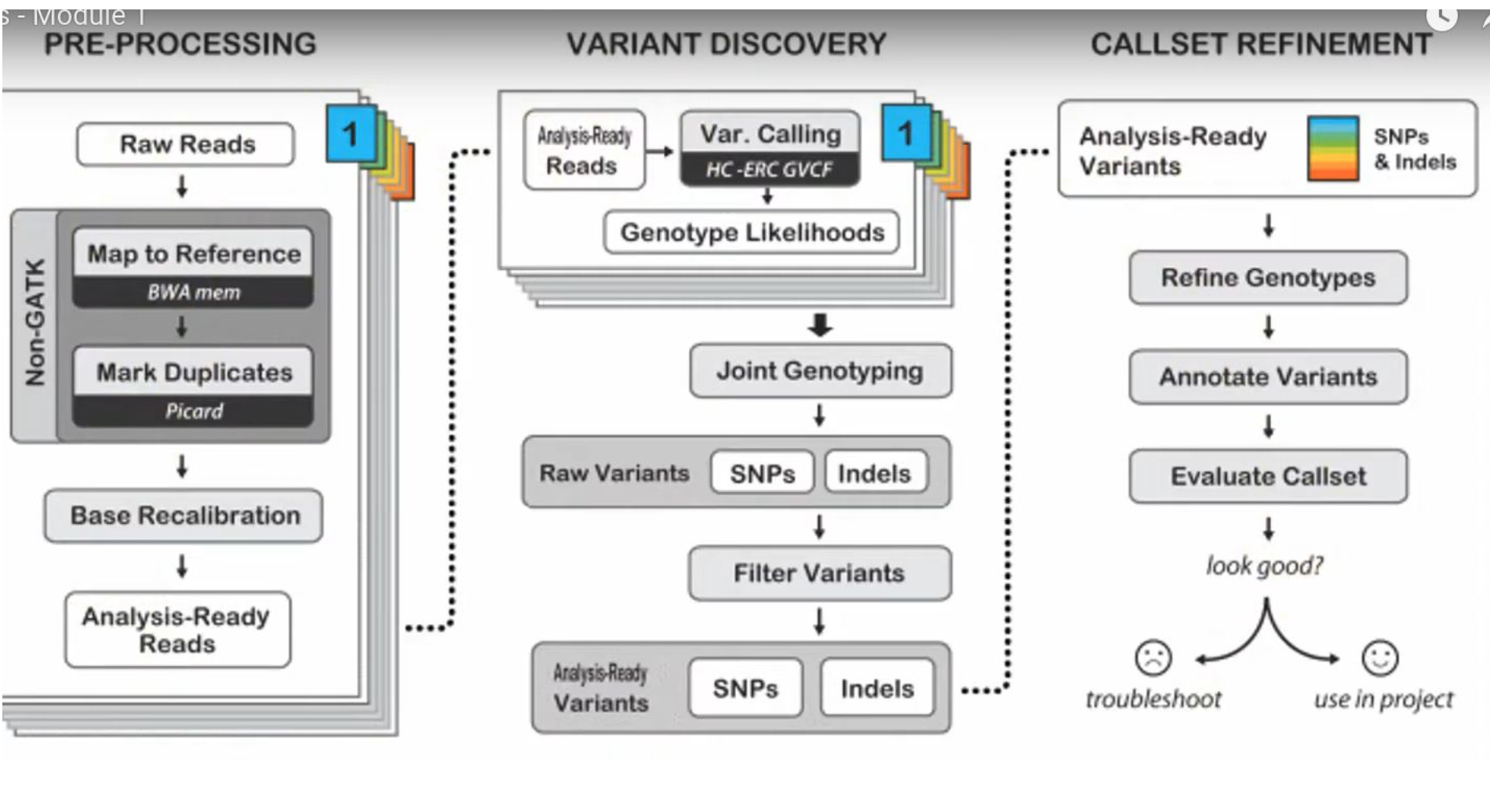
**BWA**

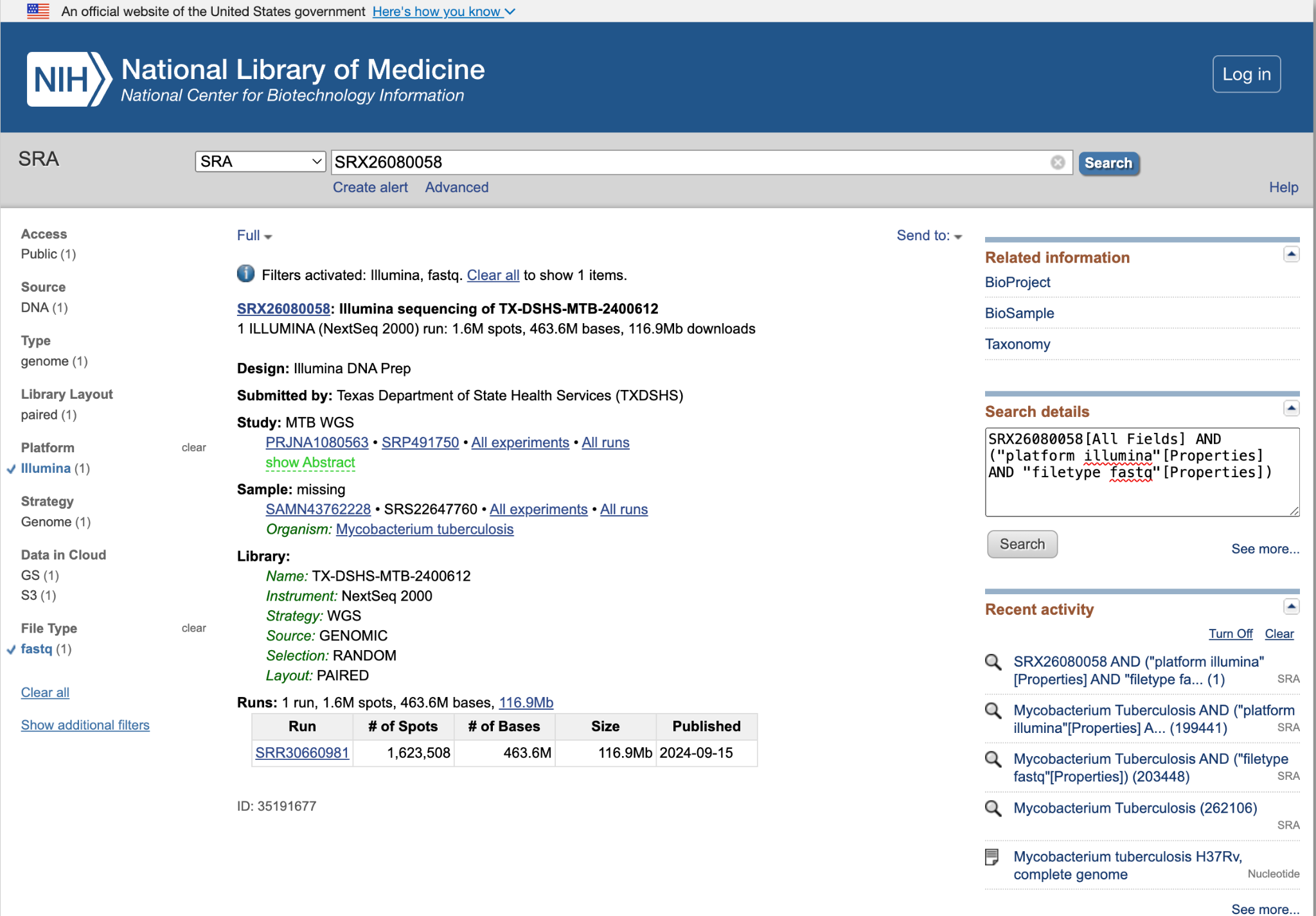


Resources:

1. <https://omics.sbs/blog/bwa/bwa.html>
2. <https://github.com/Jwomers/burrows_wheeler_alignment/blob/master/BWA.py>
3. <https://mtbgenomicsworkshop.readthedocs.io/en/latest/material/day3/mapping.html>
4. <https://chatgpt.com/c/66ed8896-78e4-8001-b655-127570892867>
5. <https://chatgpt.com/c/66eebbf8-9334-8001-807f-270d6e5652ac>

Dataset: [https://www.ncbi.nlm.nih.gov/sra/SRX26101592[accn](https://www.ncbi.nlm.nih.gov/sra/SRX26101592%5Baccn)]

Initial run dataset:



Doubts:

1. In the reading How the reading of the reference is done, because we dot know the exact sequence.
2. Why 2 observations two times?

**Picard**

Resources:

1. <https://chatgpt.com/c/66f0f309-81dc-8001-b85c-3605b55c3009>

Picard is a set of command-line tools designed for manipulating and analyzing high-throughput sequencing (HTS) data, particularly in the BAM (Binary Alignment/Map) and SAM (Sequence Alignment/Map) formats. It is widely used in bioinformatics workflows for tasks such as:

1. **Sorting and Merging BAM/SAM Files**: Picard can sort and merge alignment files, which is a common step before further analysis or visualization.
2. **Marking Duplicates**: It identifies duplicate reads, which can arise during PCR amplification. These duplicates are marked so they can be ignored in downstream analysis to avoid bias.
3. **Collecting Quality Metrics**: It can calculate various quality metrics like insert size distribution, alignment summary statistics, and more.
4. **File Conversion**: Picard can convert between different file formats, such as BAM to SAM and vice versa.
5. **Manipulating Headers**: It provides tools for modifying the headers of BAM/SAM files, such as adding, removing, or modifying read groups.

**Base Recalibration**

Resources:

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8743552/>