Quality control and preprocessing tools

General QC tools

- **FastQC**
  - https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
  - https://github.com/s-andrews/FastQC

- **MultiQC**
  - Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047-3048. doi:https://doi.org/10.1093/bioinformatics/btw354
  - https://multiqc.info/
  - https://github.com/ewels/MultiQC

Trimming tools

- **Skewer**
  - Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*. 2014;15(1). doi:https://doi.org/10.1186/1471-2105-15-182
  - https://github.com/relipmoc/skewer

- **fastp**
  - Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890. doi:https://doi.org/10.1093/bioinformatics/bty560
  - https://github.com/OpenGene/fastp

Long-read QC

- **PycoQC**
  - Leger A, Leonardi T. pycoQC, interactive quality control for Oxford Nanopore Sequencing. *Journal of Open Source Software*. 2019;4(34):1236. doi:https://doi.org/10.21105/joss.01236
  - https://a-slide.github.io/pycoQC/
  - https://github.com/a-slide/pycoQC

- **Porechop**

    - https://github.com/rrwick/Porechop

- **NanoPack**
    - De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*. 2018;34(15):2666-2669. doi:https://doi.org/10.1093/bioinformatics/bty149
    - https://github.com/wdecoster/nanopack

- **Filtlong**
    - https://github.com/rrwick/Filtlong

- **LongQC**
    - Fukasawa Y, Ermini L, Wang H, Carty K, Cheung MS. LongQC: A Quality Control Tool for Third Generation Sequencing Long Read Data. *G3: Genes, Genomes, Genetics*. Published online February 10, 2020:g3.400864.2019. doi:https://doi.org/10.1534/g3.119.400864
    - https://github.com/yfukasawa/LongQC

Alignment and mapping tools

RNA-Seq

- **Kallisto**
    - Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*. 2016;34(5):525-527. doi:https://doi.org/10.1038/nbt.3519
    - https://github.com/pachterlab/kallisto
    - https://pachterlab.github.io/kallisto/about

- **STAR** (Spliced Transcripts Alignment to a Reference)
    - Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2012;29(1):15-21. doi:https://doi.org/10.1093/bioinformatics/bts635
    - https://github.com/alexdobin/STAR

- **BWA** (Burrows-Wheeler Aligner)
    - Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:https://doi.org/10.1093/bioinformatics/btp324
    - Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589-595. doi:https://doi.org/10.1093/bioinformatics/btp698
    - https://github.com/lh3/bwa

Short-read DNA-Seq

- **BWA-MEM**

- Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997v2 [q-bio.GN]*. 2013: https://doi.org/10.48550/arXiv.1303.3997
- https://github.com/bwa-mem2/bwa-mem2

- **BWA**
  - (same as above)

- **bowtie2**
  - Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-359. doi:https://doi.org/10.1038/nmeth.1923
  - https://github.com/BenLangmead/bowtie2

Long reads

- **Minimap2**
  - Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094-3100. doi:https://doi.org/10.1093/bioinformatics/bty191
  - https://github.com/lh3/minimap2

- **LRA (Long Read Aligner)**
  - Ren J, Chaisson MJP. lra: A long read aligner for sequences and contigs. *PLOS Computational Biology*. 2021;17(6):e1009078. doi:https://doi.org/10.1371/journal.pcbi.1009078
  - https://github.com/ChaissonLab/LRA

- **NGMLR (Next Generation Mapping and Long Read Mapping)**
  - Sedlazeck FJ, Rescheneder P, Smolka M, et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*. 2018;15(6):461-468. doi:https://doi.org/10.1038/s41592-018-0001-7
  - https://github.com/philres/ngmlr