SOUTH AUSTRALIAN GENOMICS CENTRE
SAGC

SAGC BIOINFORMATICS WORKSHOP

RNAseq analysis using nf-core

A practical guide to RNAseq analysis using nextflow-core pipelines.
Recent development of computational biology tools and initiatives like nf-core have enhanced accessibility to analysis options for many genomics technologies including RNAseq. Discover how nextflow simplifies and streamlines RNAseq and other genomics analyses.

Who should attend?

It is intended to be approachable to new users of RNAseq. The focus will be on understanding analysis options for RNAseq, although familiarity with command line tools (unix & R) will be necessary to run nextflow pipelines.

Register Here

TOPICS

A hands-on walk-through of nf-core analysis pipelines for RNAseq
run nf-core bioinformatic pipelines using NGS data; QC metrics, read trimming, differential expression analysis, data visualisation, mRNA and small RNA.

Description of the key metrics and analyses for RNAseq analysis
Gain familiarity with key bioinformatic tools and , and how to decipher and use the outputs.

Location
'Flinders City Campus (Festival Tower), Rm: 505

Lead Trainer
Dr Daniel Thomson SAGC

Time and Date
10:00am - 4:00pm
(Registration opens at 9:30 am)
Thursday, 10th October 2024

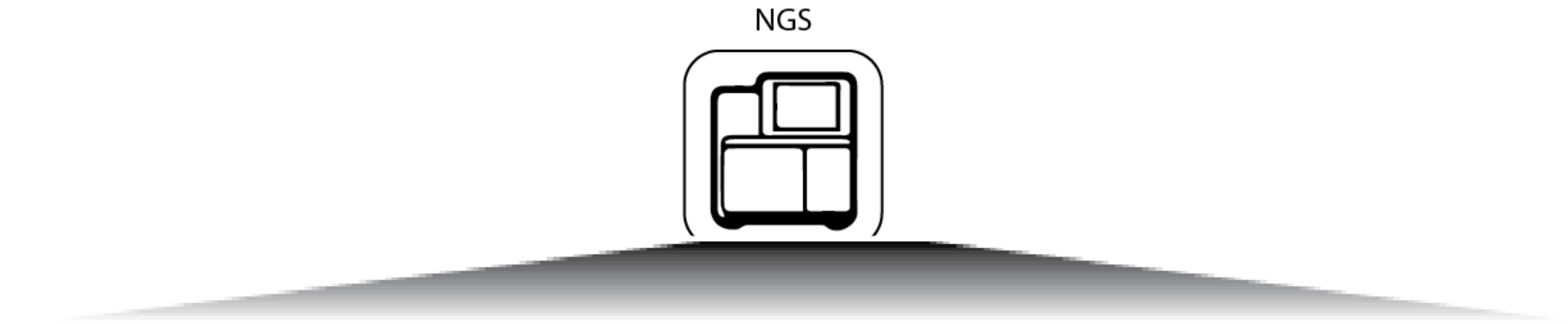
Cost
$100 Non-Student $50 Student

HEAPS GOOD GENOMICS!

RNAseq

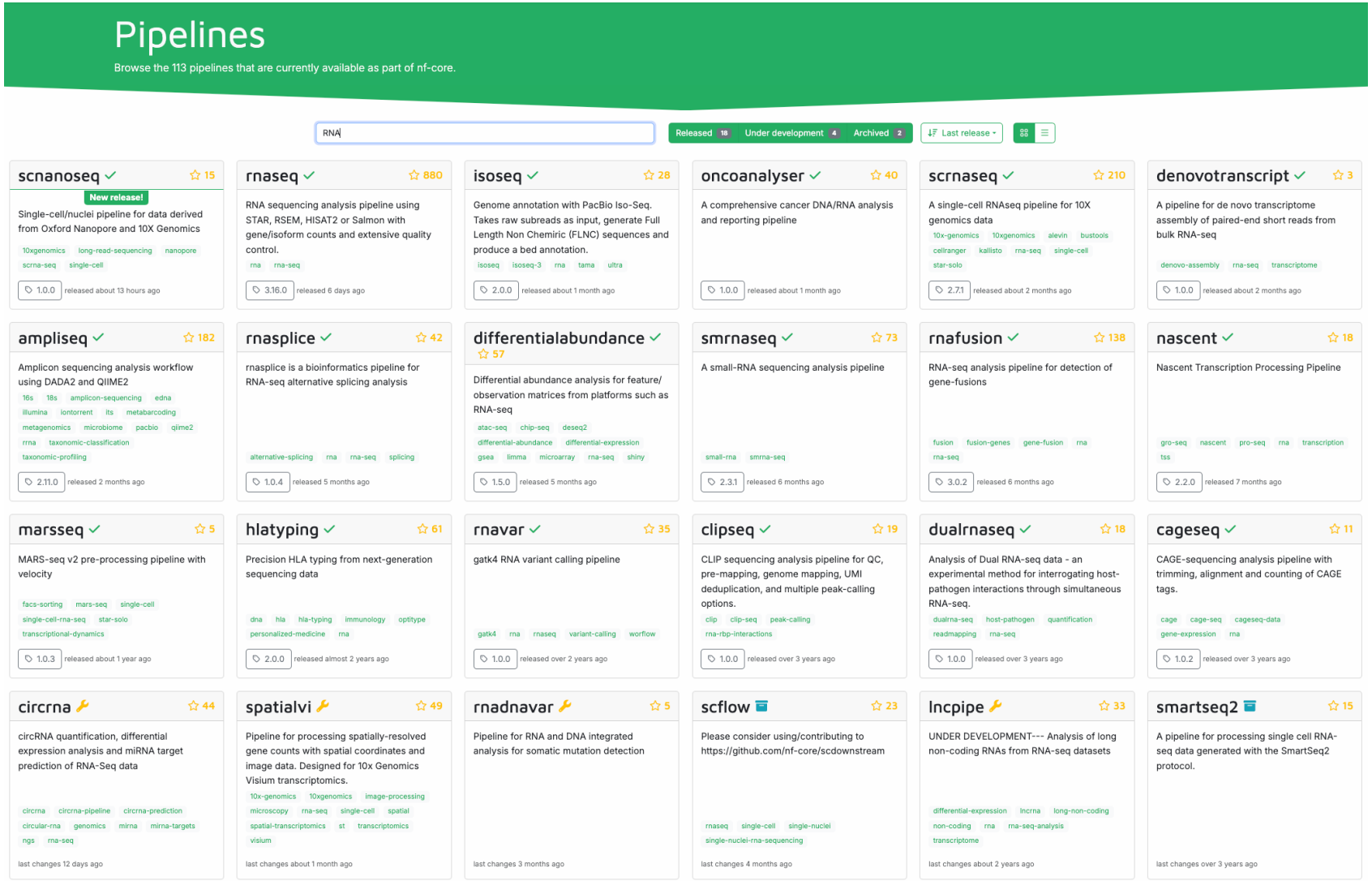
Quick background

RNAseq is a versatile tool



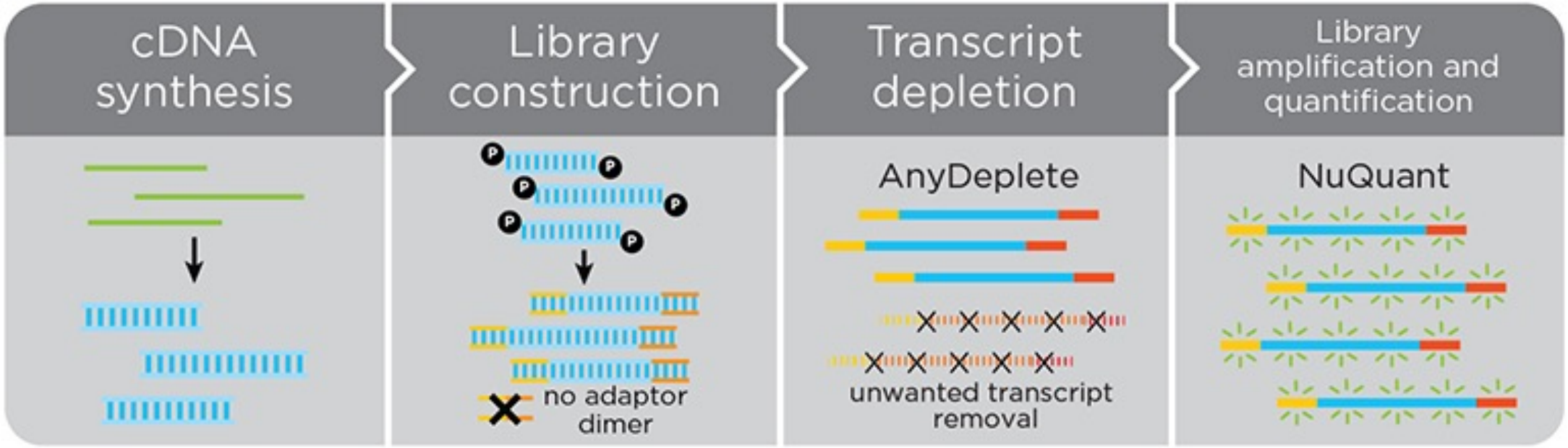
DNA	yes	yes*	no	no	no	yes
RNA	yes	yes*	yes	yes	yes	yes

Why nextflow and nf-core?



RNAseq library preparation

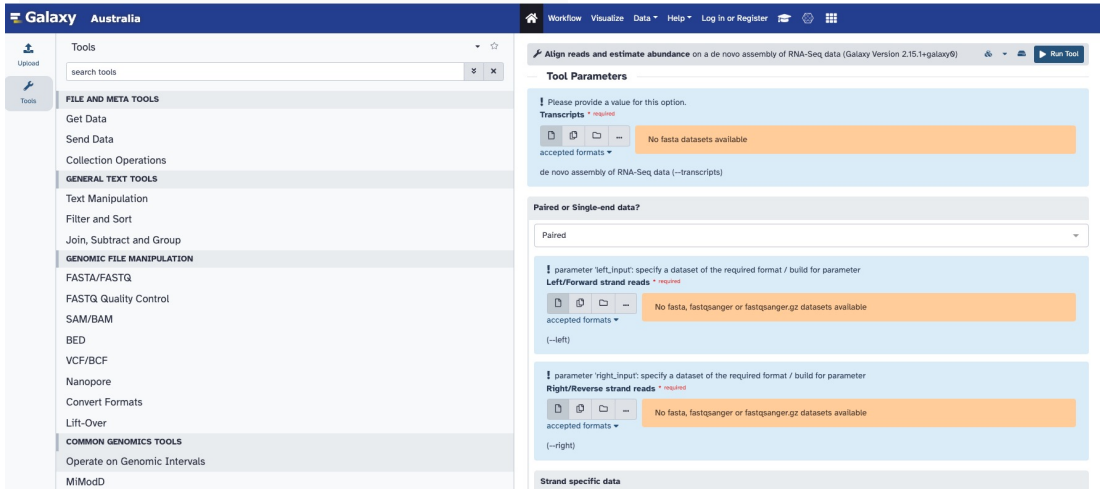
- Universal Plus Total RNA-Seq Library Preparation Kit

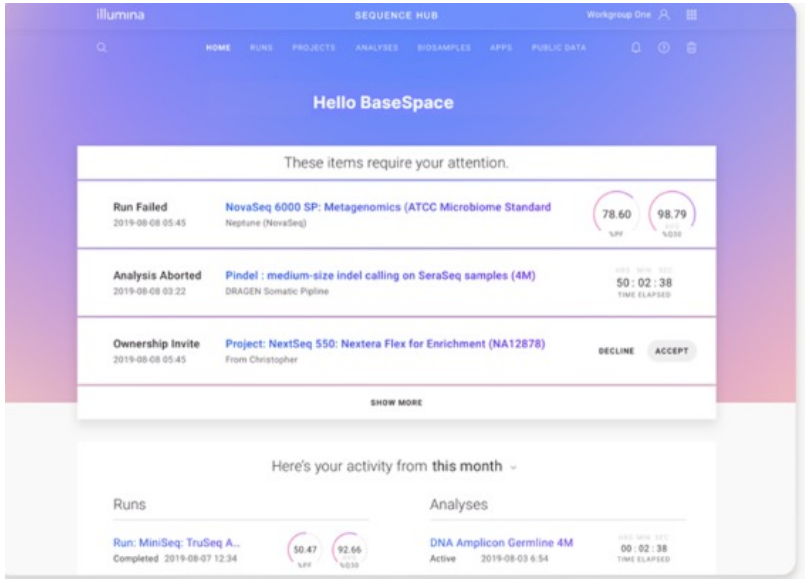


Other options in running Bioinformatics Pipelines

BaseSpace SEQUENCE HUB

Galaxy AUSTRALIA



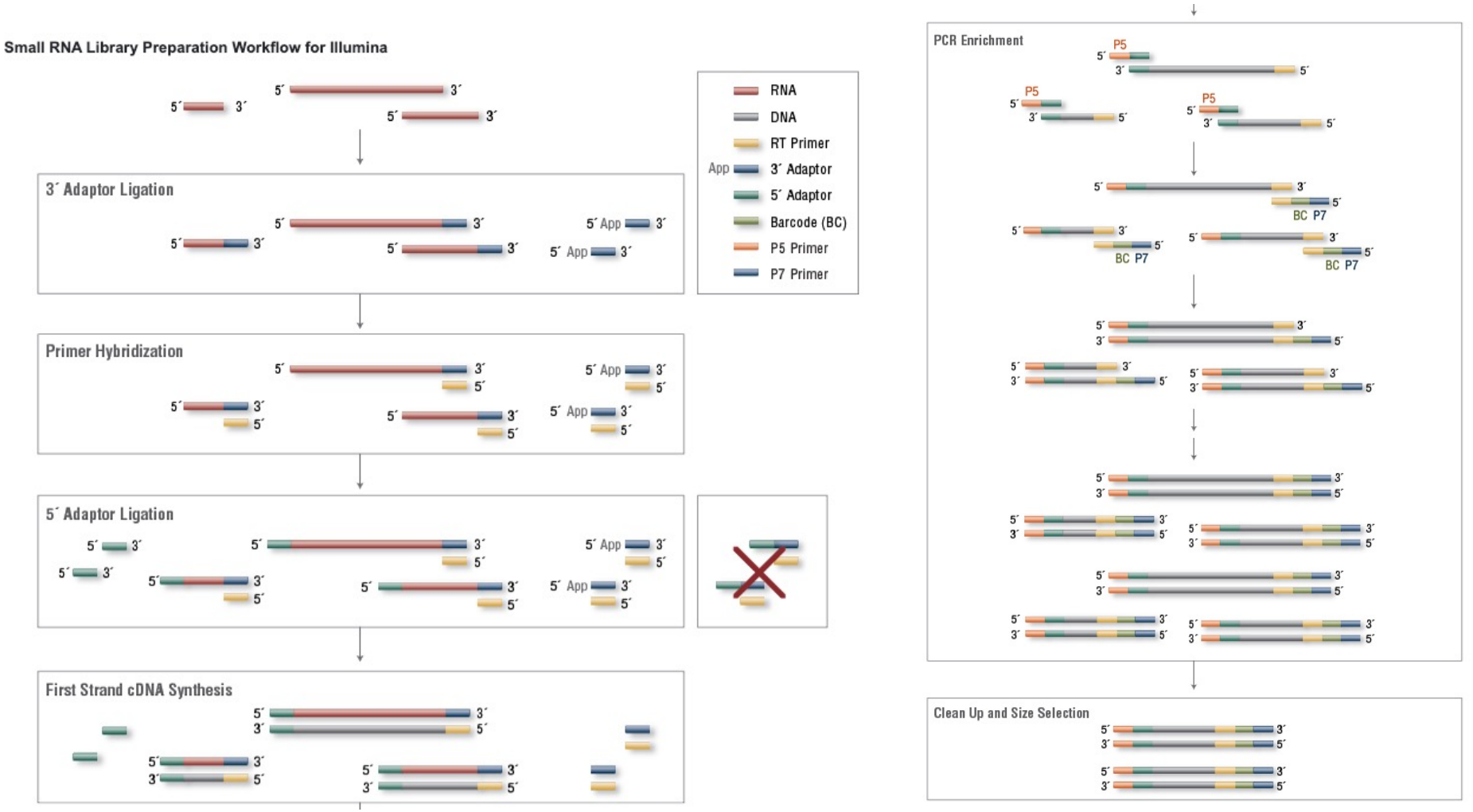


Sentieon

GitHub

Small RNAseq library preparation

- NEBNext Small RNA Library Prep Set for Illumina



Range of Sequencing technologies

PACBIO®

Oxford NANOPORE™ Technologies

华大基因 BGI

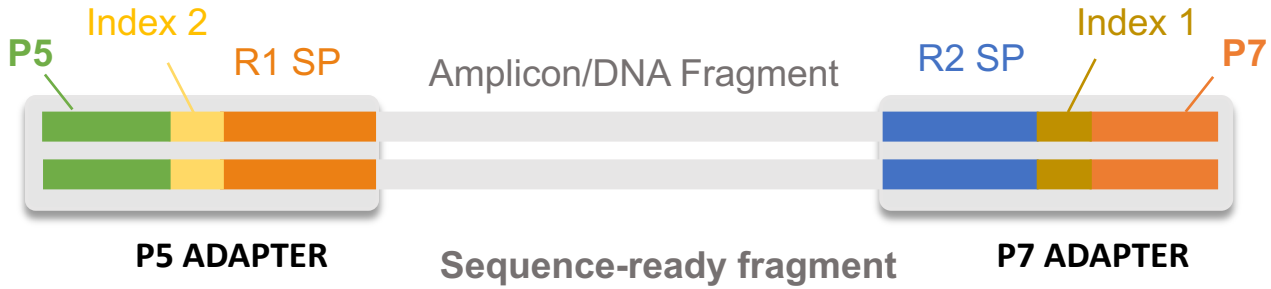
MGISEC-T7

ThermoFisher SCIENTIFIC

Ion Torrent™ Genexus™

Sequencing libraries

The aim of library prep is to obtain nucleic acid fragments with adapters attached on both ends



P5 and P7 regions are complementary to the oligos bound to the flow cell surface

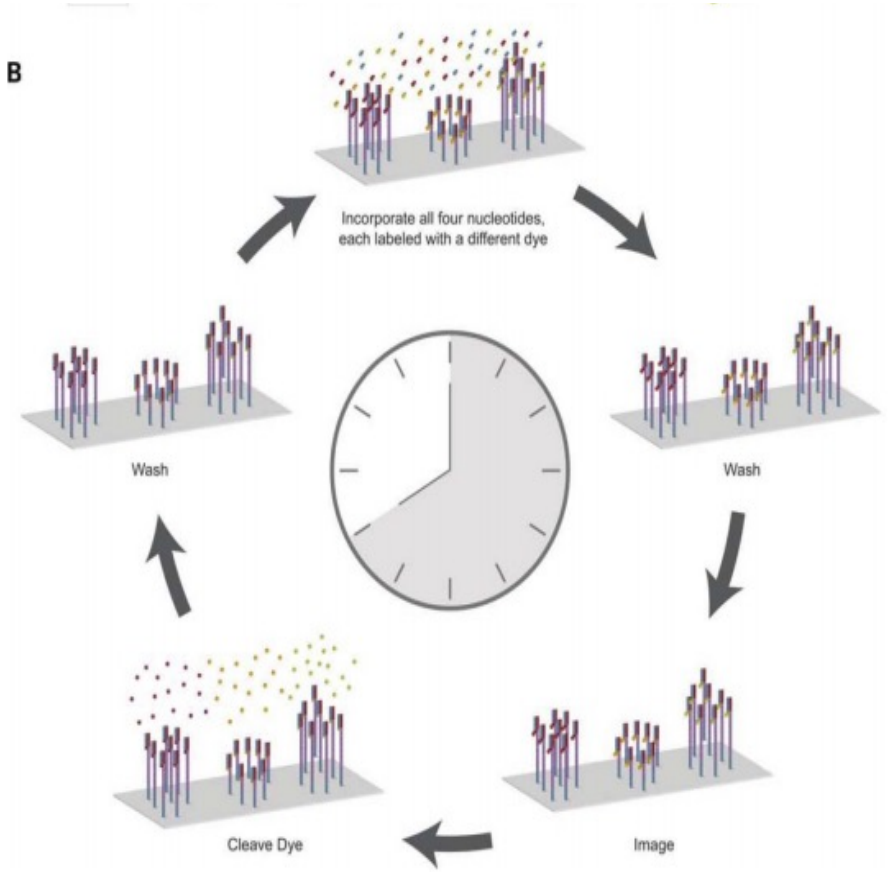
Index sequences are used to tag individual samples to allow for pooling

Read 1 & Read 2 Sequencing Primers are used to initiate sequencing

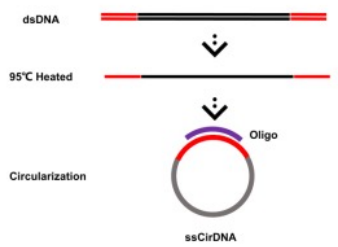
“short read” sequencing

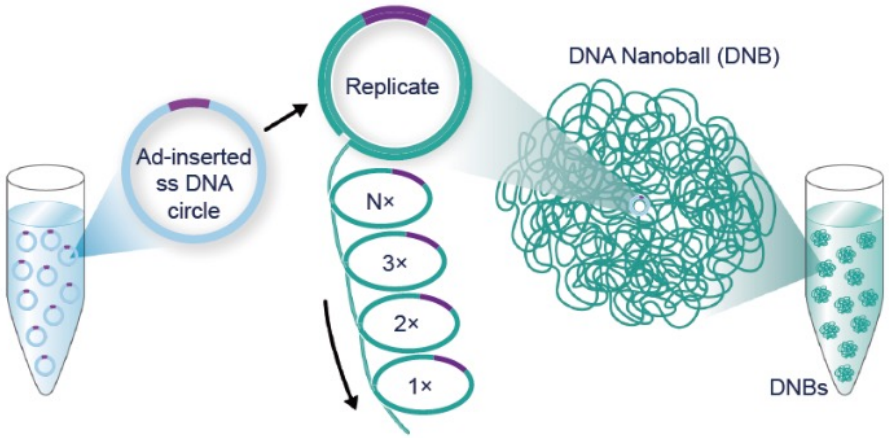
illumina®

Sequencing by Synthesis



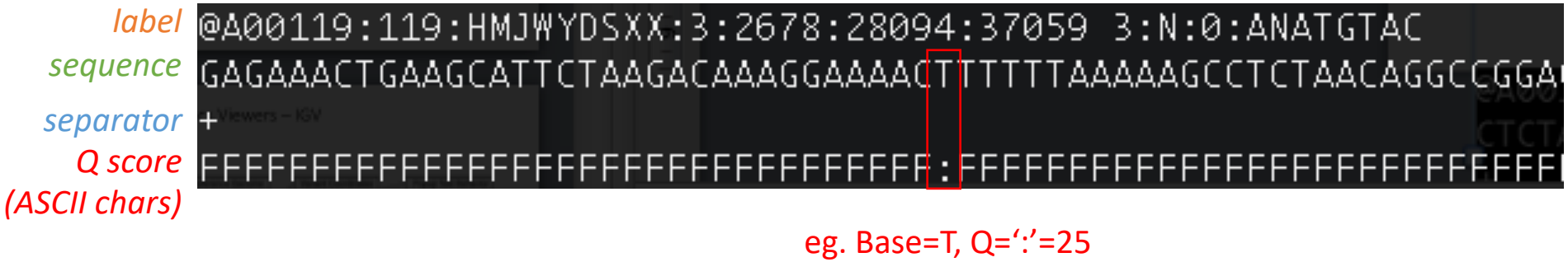
MGI





DNBSEQ™

Fastq files – raw sequence data



- Read information for each sample stored as FASTQ files
  - Unmapped raw sequence file
  - Text format (.fastq.gz)
  - Has a 4-line entry for each read
  - Contains base call information and quality scores for each base
  - Contains reference information for each read

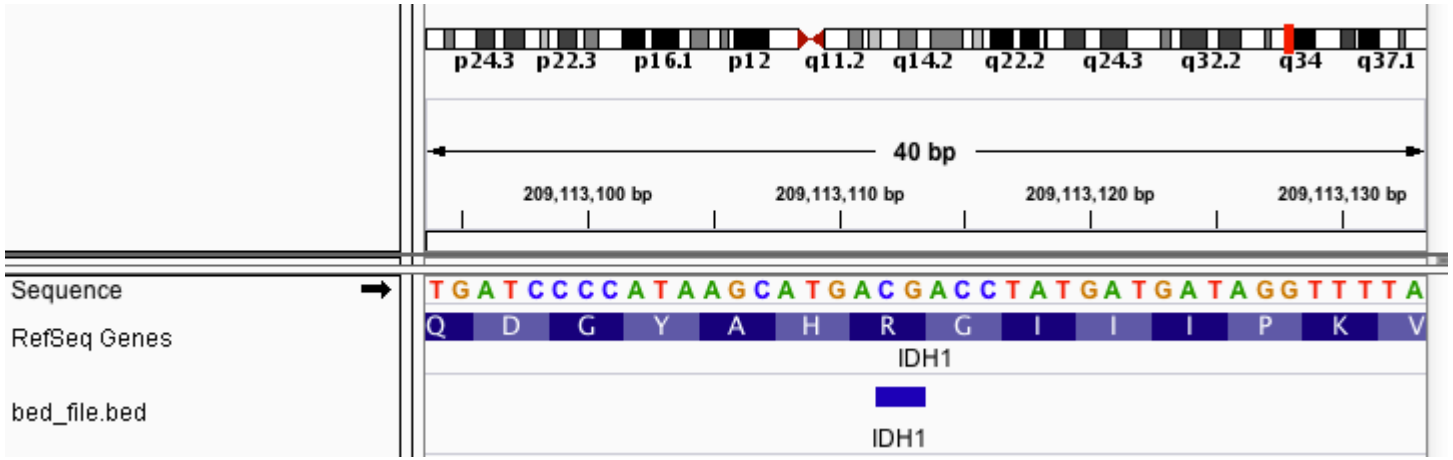
BED file

- Browser Extensible Data (BED) file
- 4 -12 columns

chromosome	start	stop	name
chr21	37518705	37518706	CBR3
chr2	209113111	209113113	IDH1
chr3	37053566	37053567	MLH1

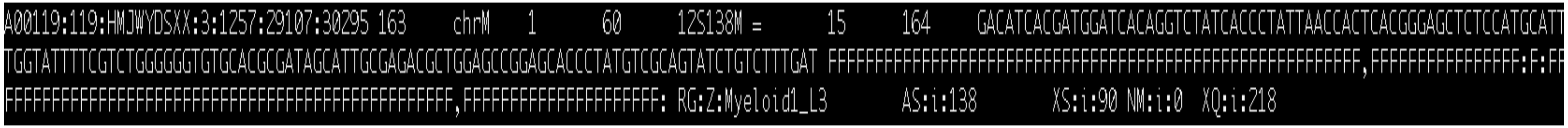
tab - delimited

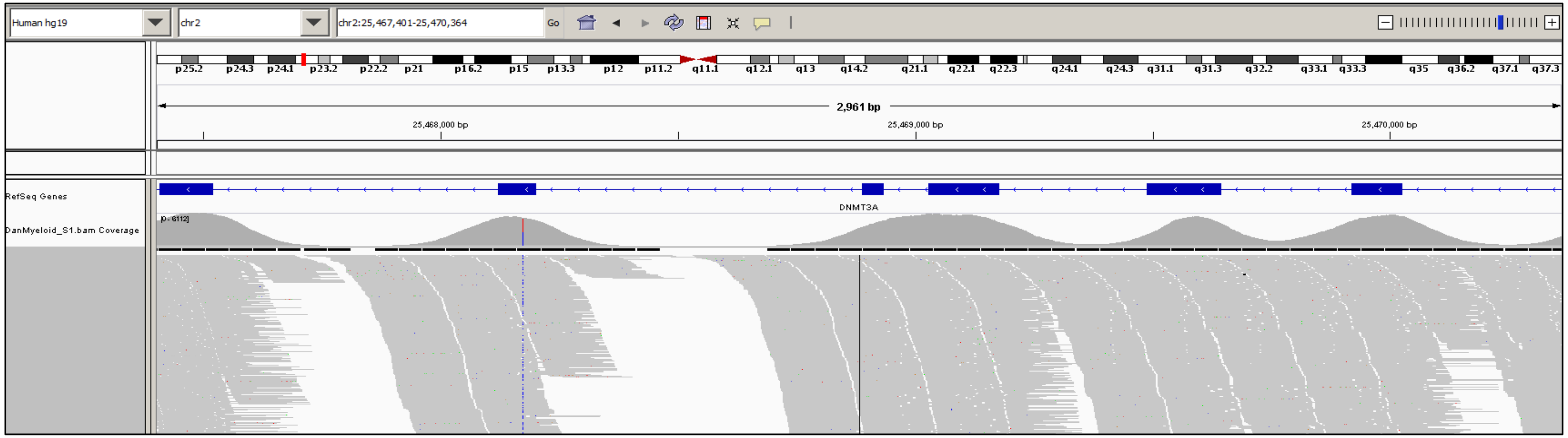
IGV – genomics viewer



Bam file

- Binary sequence allignment / map
  - Contains sequence information, mapping information and quality metrics
  - Can be viewed on genomics viewer - each line is one mapped sequence





BIOINFORMATICS PIPELINES

The SAGC provides a suite of analysis pipelines developed both externally and in-house, based on community best practises.

Workflows designed for SAGC sequenced libraries with set endpoints for quick turnaround.

Whole-Genome (WGS)

Whole-Exome (WES)

RNA seq

DNA Methylation, or Whole-Genome Bisulfite (WGBS)

Single Cell RNA (scRNA-seq)

Spatial Transcriptomics

Chromatin Immunoprecipitation Sequencing (ChIP-Seq)

Small RNA-seq

Assay of Transposase Accessible Chromatin (ATAC-Seq)

Metagenomics

Microbial Profiling (Microbiome)

Amplicons