Genomic tools

Libor Mořkovský, Václav Janoušek

https://ngs-course.readthedocs.io/en/praha-february-2019/

samtools

- Working with SAM/BAM files (i.e read alignment data)
- Manipulation with SAM/BAM (sorting, merging, subsetting)
- Summary statistics (read depth by position)
- Viewing read alignment in command line:

https://samtools.github.io

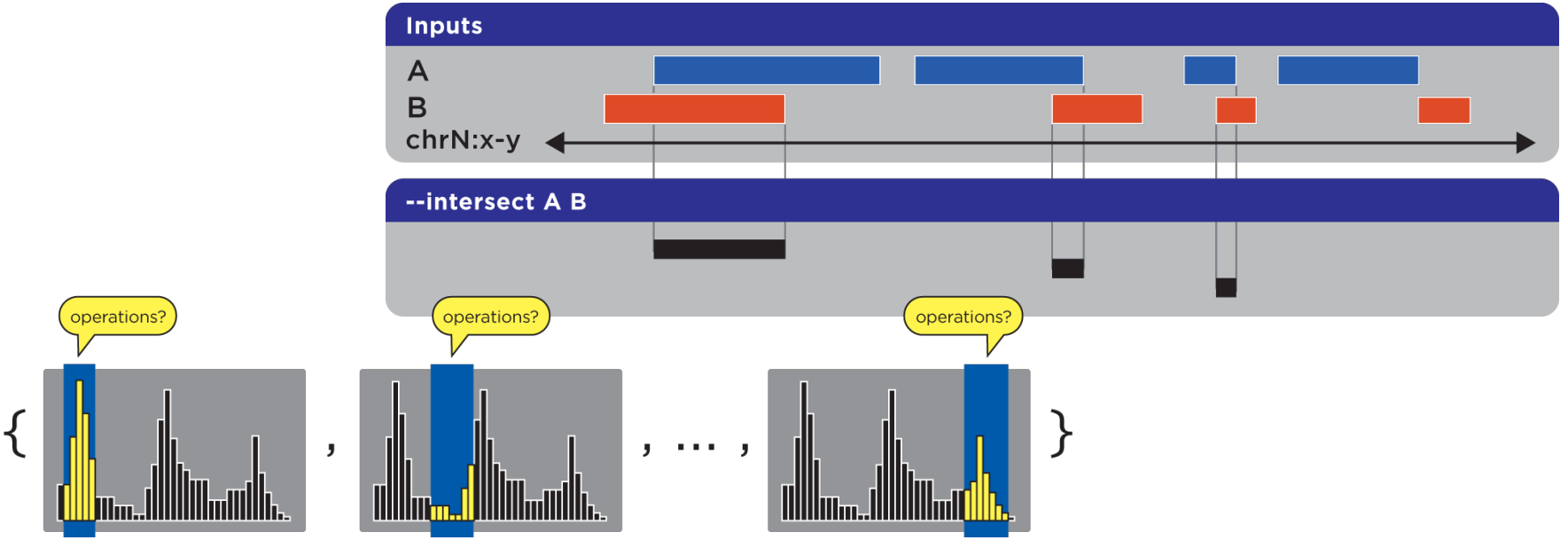
bcftools/vcftools

- variant call files (vcf/bcf)
- bcftools:
  - annotation, concatenation, merging, converting to different formats, filtering based on various criteria, variant calling
- vcftools:
  - mainly filtering/creating subsets
  - population genetics (allele frequency, Hardy-Weinberg, Fst, Pi, Tajima, linkage disequilibrium,...)

https://vcftools.github.io/index.html
https://samtools.github.io/bcftools/

bedtools/bedops

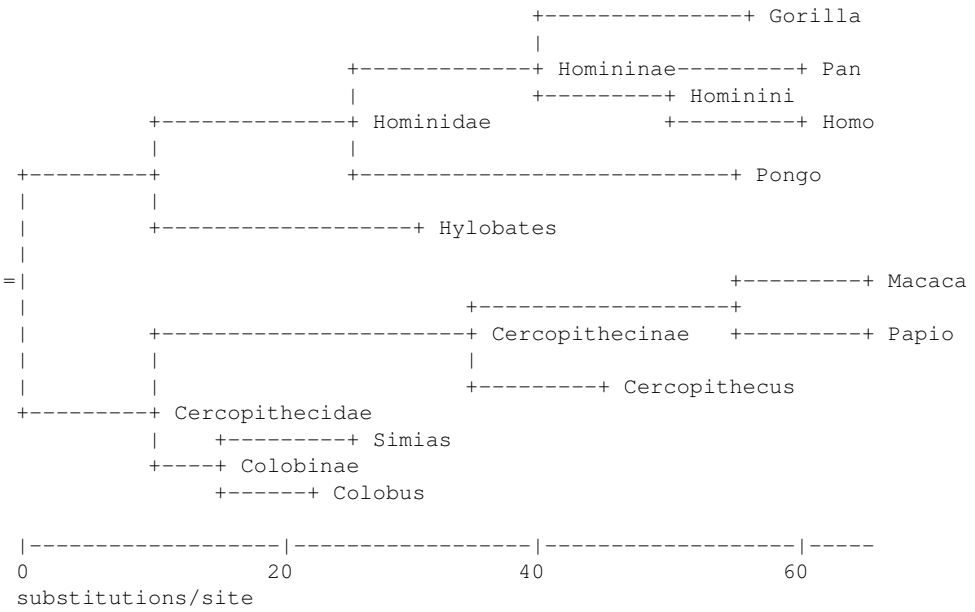
- Operations with genomic data based on their physical position in genome (chromosome, feature start, feature end, strand)
- Usually intersections, overlaps, summary by specific regions (e.g. coverage), sliding window analysis, randomization



https://bedtools.readthedocs.io/en/latest/
https://github.com/bedops/bedops

newick-utils

- newick tree format
- manipulation of phylogenetic trees:
  - re-rooting
  - extracting
  - subsetting
  - viewing
  - …



https://github.com/tjunier/newick_utils/wiki

BuddySuite

- python based but can be used in unix command line (pipes work)
- tools:
  - SeqBuddy − manipulation with FASTA/FASTQ files
  - AlignBuddy − manipulation with alignment files (phylip, clustal, etc.)
  - PhyloBuddy − manipulation with phylogenetic trees

https://github.com/biologyguy/BuddySuite