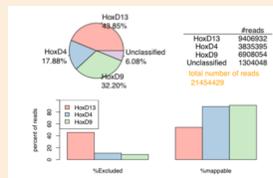
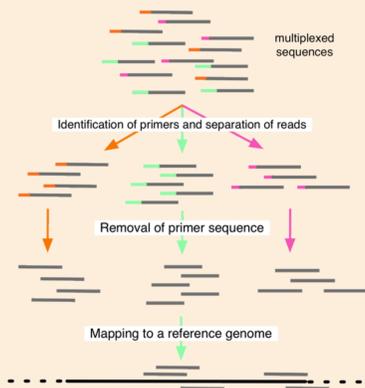


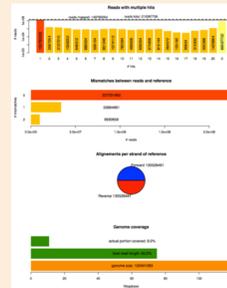
A web application for High Throughput Sequencing data analysis

Fabrice PA David *, Solenne Carat *, Julien Delafontaine *, Frederick Ross, Gregory Lefebvre, Lucas Sinclair, Yohan Jarosz, Marion Leleu, Jacques Rougemont
 Bioinformatics and Biostatistics Core Facility, School of life science, EPFL

Demultiplexing / Mapping



demultiplexing report



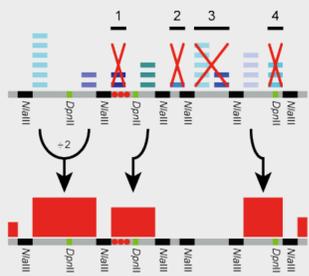
mapping report

Separate & map to a reference genome

- Separate reads, remove barcodes and filter undesired sequences
- Use bowtie to map reads to genome
- Include and compensate for multiple mappings in densities
- Remove multiple identical reads (PCR artifacts)

4C-seq

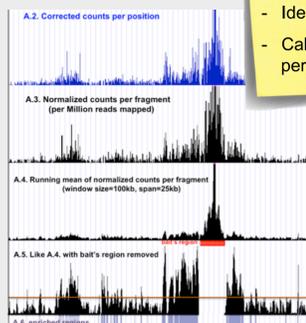
removal of non informative reads
 1: repeat sequences
 2: too short NlaIII fragment
 3: absence of DpnII site
 4: DpnII and NlaIII sites too close



translocation to interacting restriction fragment

identification of interacting regions

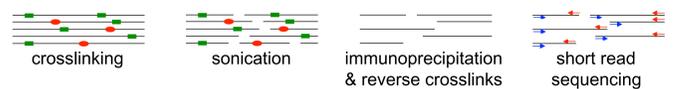
running mean algorithm
 Window size: 29 informatives fragments; FDR: 0%



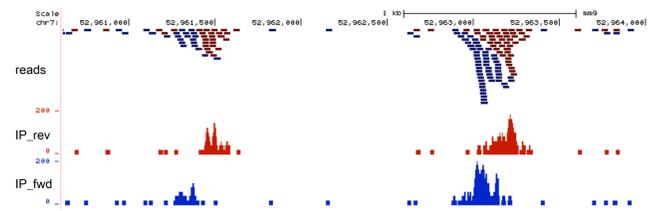
Identification of long-range 3D interaction partners

- Create library of fragments
- Identify informative fragments
- Calculate a normalized score per informative fragment

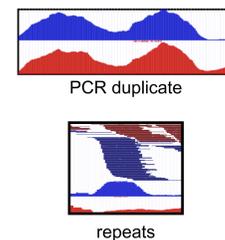
ChIP-seq



Mapping & compute densities



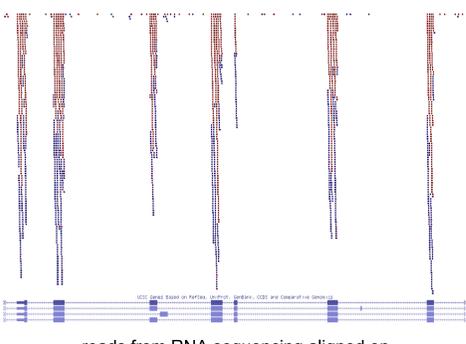
ARTIFACTS



Locate & quantify protein bound sites

- Merge strands
- Peaks detection with MACS
- Artifacts correction
- Peaks annotation & motif discovery

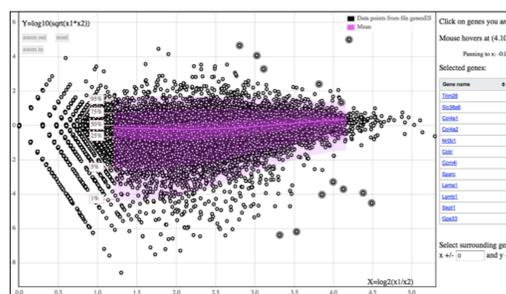
RNA-seq



reads from RNA sequencing aligned on annotated exome

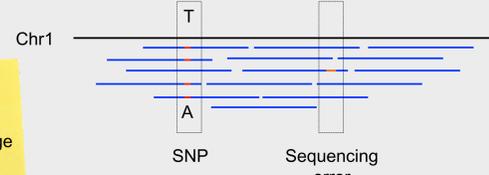
Transcriptome & exome profiling

- Calculate counts and RPKM per transcript, gene and exon from a single alignment
- "Optimal" estimation of transcripts expression
- Interactive MA-plot of differential expression between samples



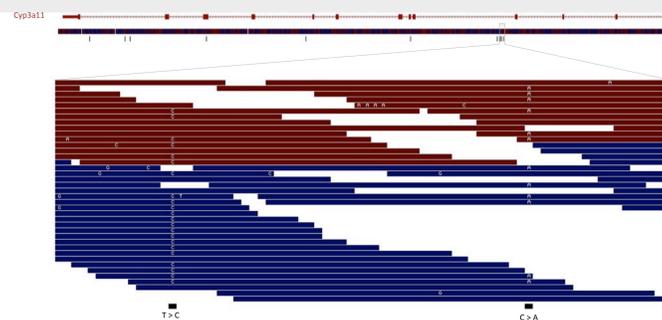
Interactive MA-plot to look for differential gene/transcript expression

SNP calling



Detect & annotate SNPs

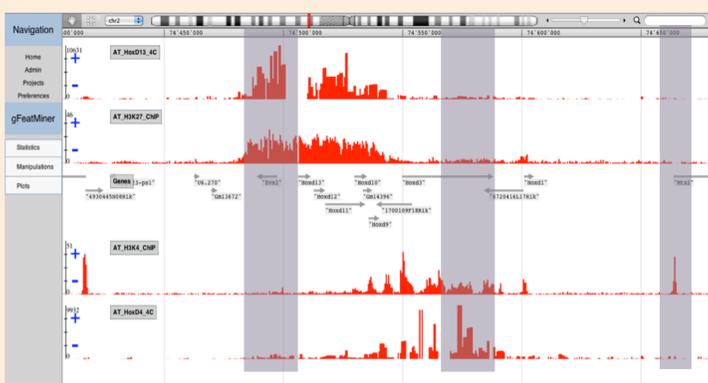
- Detect SNPs from reads coverage and frequencies
- Locate SNPs (exon, promoter, ...)
- Indicate amino acid changes



Chr position reference	SNP	Codon
1 4657	T A	synonymous
1 12654	C 90%/T / 10%/C	P->L
...		

SNP analysis report

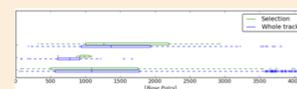
Visualisation with GDV



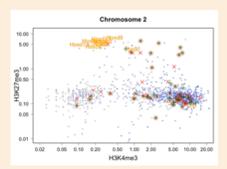
Genome browser

- Fast, responsive and interactive visualisation
- Manage, export and share data
- Generate sub-dataset by selection
- Link with gene documentation

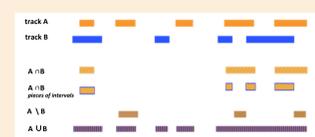
Data manipulations with GDV



descriptive statistics



advanced plots



Tracks manipulation

How to use HTSstation?

Standard analysis can be easily launched from the web interface: <http://htsstation.epfl.ch>. For advanced users, each module can be customised with configuration files.

Developpers can re-use all scripts available on github (<http://github.com/bbcf>) and integrate them with their own programs.