Role of Galaxy in a bioinformatic plant breeding platform

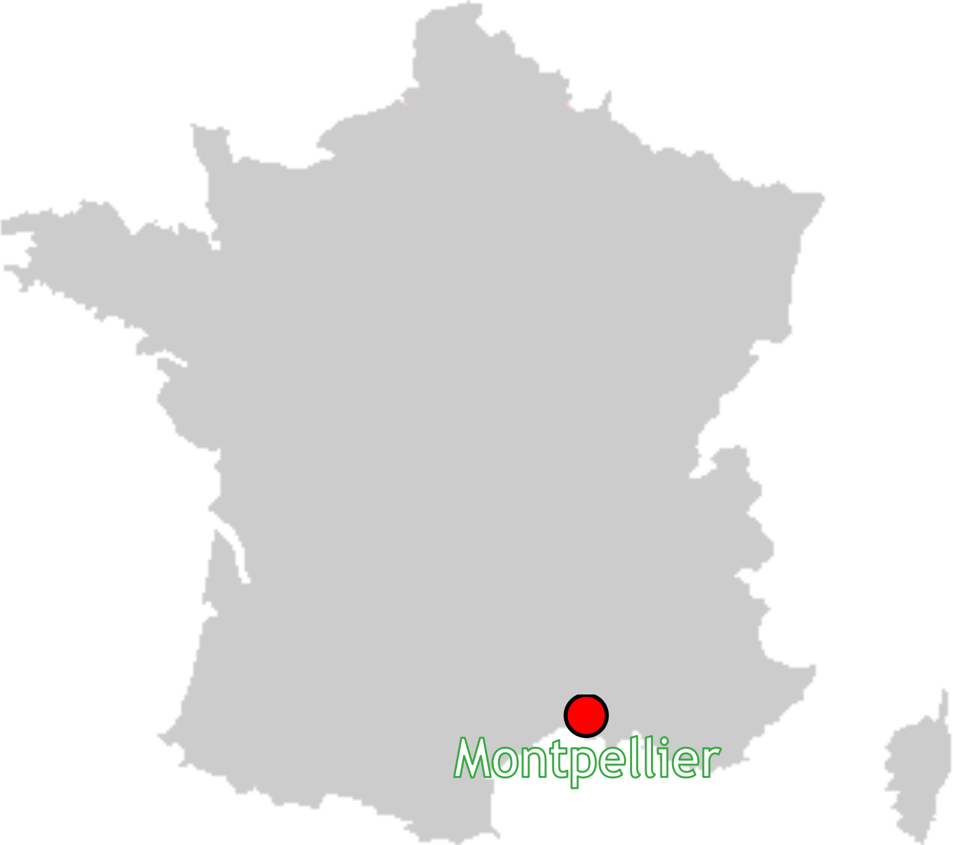
Vincent Maillol (INRA)
Jean-François Dufayard (CIRAD)

Vincent Maillol, Roberto Bacilieri, Stéphanie Bocs, Jean-Michel Boursiquot, Grégory Carrier, Alexis Dereeper, Gaétan Droc, Cécile Fleury, Pierre Larmande, Loïc Le Cunff, Jean-Pierre Péros, Bertrand Pitollat, Manuel Ruiz, Gautier Sarah, Guilhem Sempéré, Marilyne Summo, Patrice This, and Jean-Francois Dufayard

South Green
bioinformatics platform

Southgreen platform, context

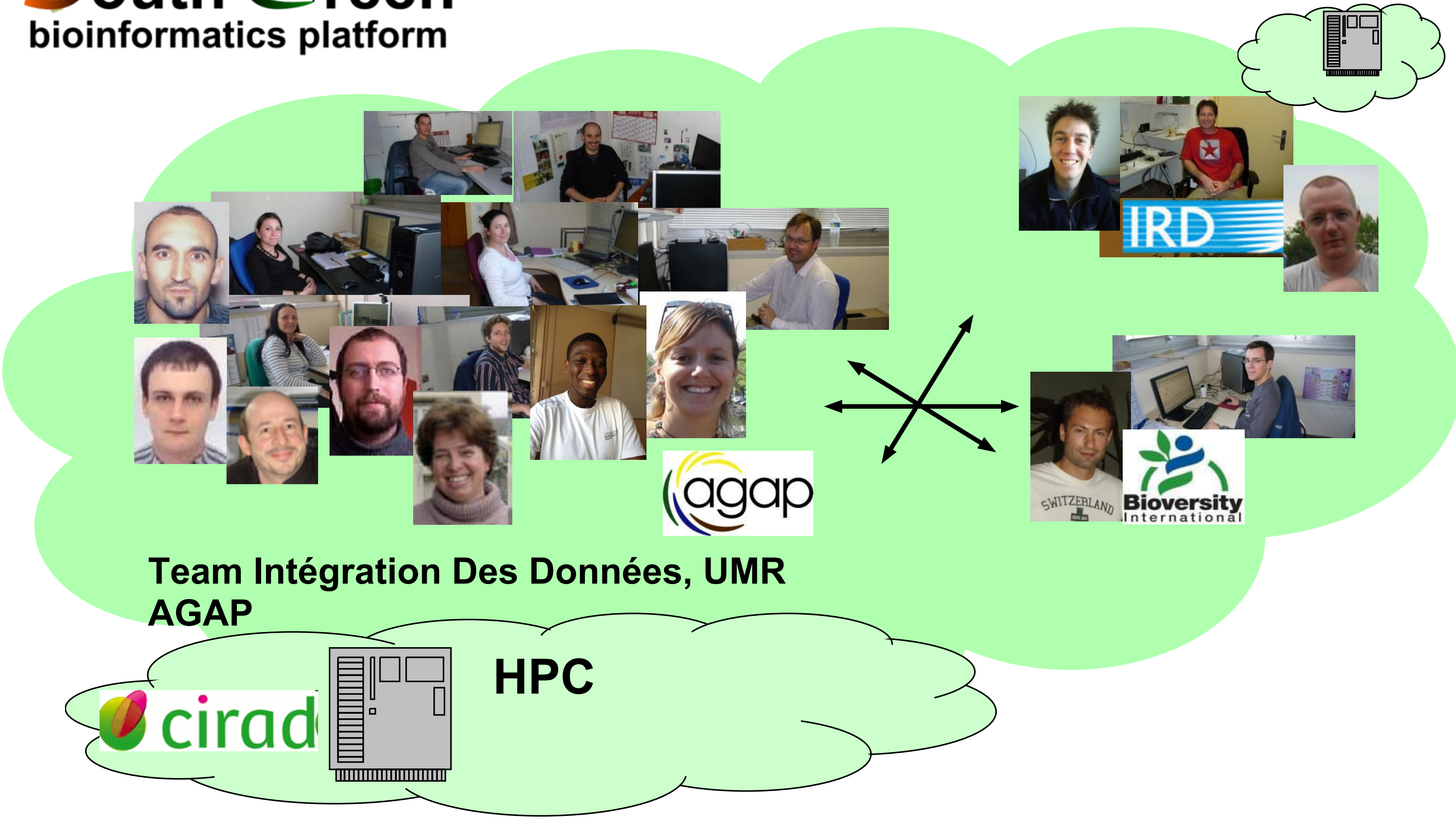
Southgreen partnerships



South Green ©
bioinformatics platform

Southgreen partnerships

South Green©
bioinformatics platform

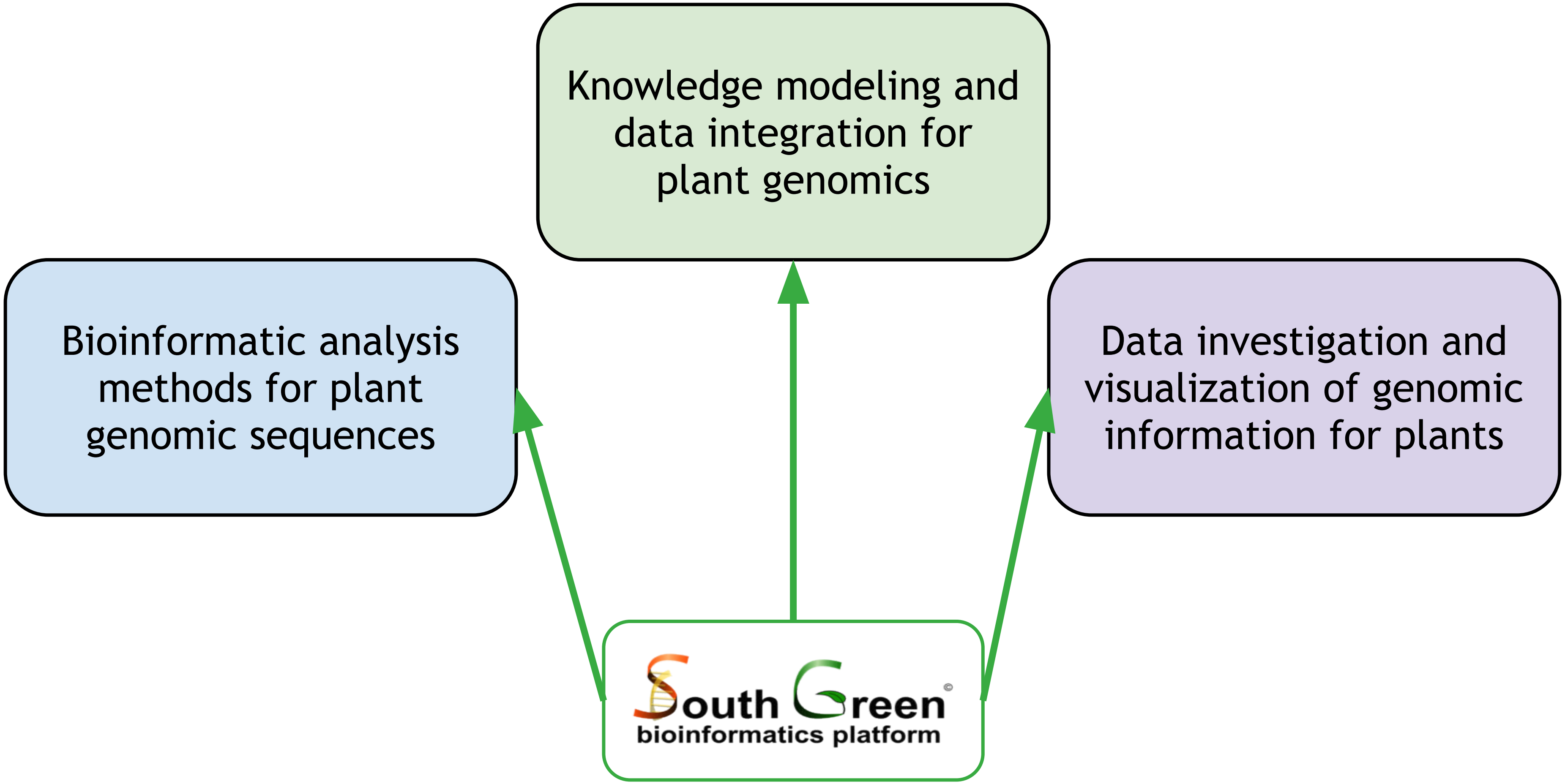


HPC

Team Intégration Des Données, UMR AGAP

HPC

Research team: "Data integration"



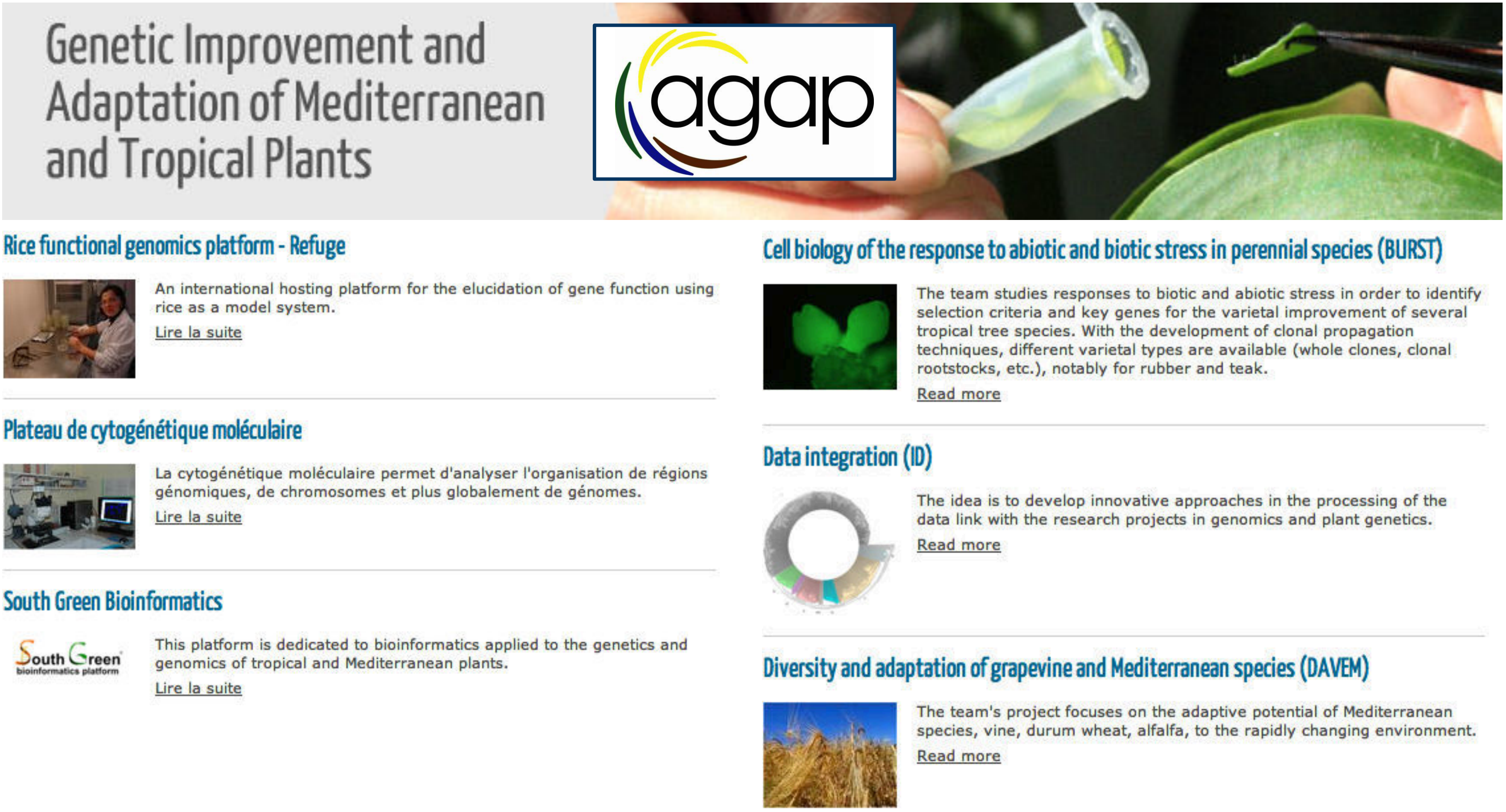
Knowledge modeling and data integration for plant genomics

Bioinformatic analysis methods for plant genomic sequences

Data investigation and visualization of genomic information for plants

South Green bioinformatics platform

UMR AGAP (Joined Research Unit)



Genetic Improvement and Adaptation of Mediterranean and Tropical Plants

agap

Rice functional genomics platform - Refuge

An international hosting platform for the elucidation of gene function using rice as a model system.

Lire la suite

Plateau de cytogénétique moléculaire

La cytogénétique moléculaire permet d'analyser l'organisation de régions génomiques, de chromosomes et plus globalement de génomes.

Lire la suite

South Green Bioinformatics

South Green bioinformatics platform

This platform is dedicated to bioinformatics applied to the genetics and genomics of tropical and Mediterranean plants.

Lire la suite

Cell biology of the response to abiotic and biotic stress in perennial species (BURST)

The team studies responses to biotic and abiotic stress in order to identify selection criteria and key genes for the varietal improvement of several tropical tree species. With the development of clonal propagation techniques, different varietal types are available (whole clones, clonal rootstocks, etc.), notably for rubber and teak.

Read more

Data integration (ID)

The idea is to develop innovative approaches in the processing of the data link with the research projects in genomics and plant genetics.

Read more

Diversity and adaptation of grapevine and Mediterranean species (DAVEM)

The team's project focuses on the adaptive potential of Mediterranean species, vine, durum wheat, alfalfa, to the rapidly changing environment.

Read more

...
6 service platforms

...
13 research teams

300+ permanent staff
(researchers, engineers, technicians)

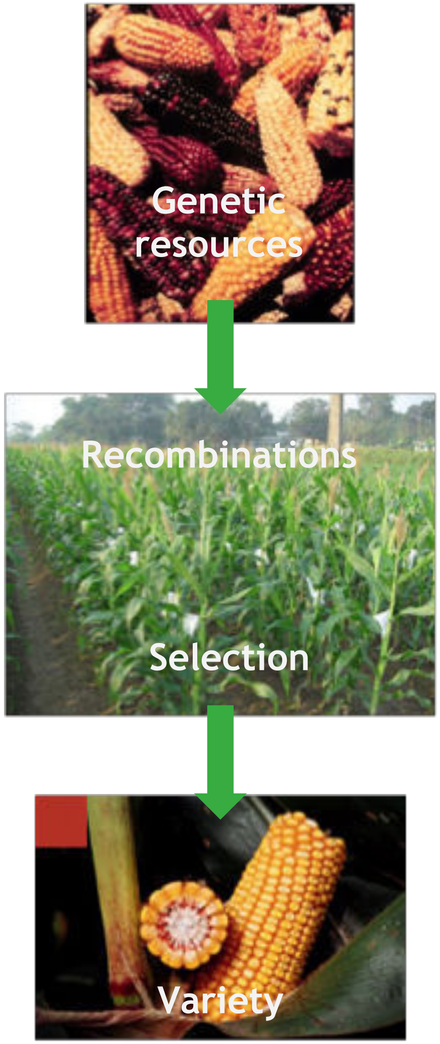
South Green©
bioinformatics platform

Southgreen platform, activities

Species of interest

...
A large variety of tropical and mediterranean species

Plant breeding



Characterize biodiversity	More markers...
Understand crop plants	Sequencing, annotation, expression, epigenetic ...
Detect polymorphisms of interest for agronomy	Markers, QTL, association genetic
Facilitate new allelic combinations	Assisted selection

Public systems and databases

- **Annotation and comparative genomics**

1. GNPAnnot
2. GreenPhyl
3. Analysis of genome sequences
4. Comparative population genomics

…

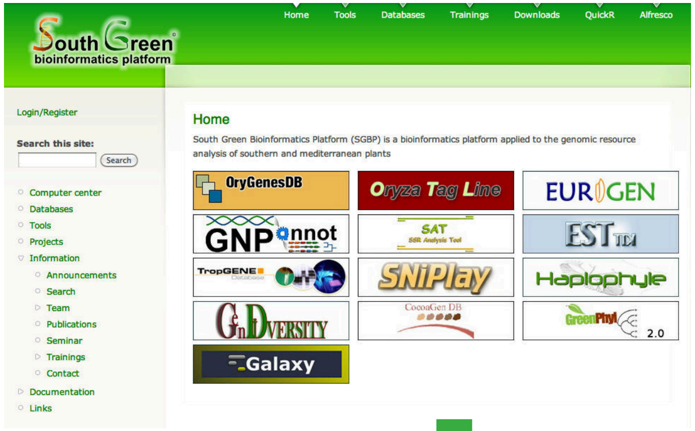
- **Information systems**

1. TropGene
2. Integrated rice functional genomics

- **Integrated workflows**

1. ESTtik
2. SNiPlay

Platform entry points and Galaxy



Home Tools Databases Trainings Downloads QuickR Alfresco

South Green bioinformatics platform

Login/Register

Search this site:

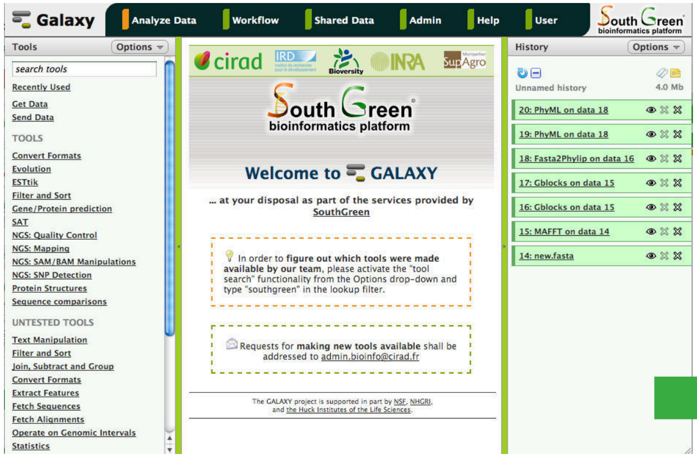
Search

- Computer center
- Databases
- Tools
- Projects
- Information
  - Announcements
  - Search
  - Team
  - Publications
  - Seminar
  - Trainings
  - Contact
- Documentation
- Links

Home

South Green Bioinformatics Platform (SGBP) is a bioinformatics platform applied to the genomic resource analysis of southern and mediterranean plants

OryGenesDB | Oryza Tag Line | EUROGEN
GNPAnnot | SAT SSR Analysis Tool | ESTtik
TropGENE Database | SNiPlay | Haplophyle
GeDiversity | CocoaGen DB | GreenPhyl 2.0
Galaxy



Galaxy | Analyze Data | Workflow | Shared Data | Admin | Help | User | South Green bioinformatics platform

Tools Options

search tools

Recently Used

Get Data

Send Data

TOOLS

Convert Formats
Evolution
ESTtik
Filter and Sort
Gene/Protein prediction
SAT
NGS: Quality Control
NGS: Mapping
NGS: SAM/BAM Manipulations
NGS: SNP Detection
Protein Structures
Sequence comparisons

UNTESTED TOOLS

Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Fetch Alignments
Operate on Genomic Intervals
Statistics

cirad IRD Bioversity INRA SupAgro

South Green bioinformatics platform

Welcome to GALAXY

… at your disposal as part of the services provided by SouthGreen

In order to **figure out which tools were made available by our team**, please activate the "tool search" functionality from the Options drop-down and type "southgreen" in the lookup filter.

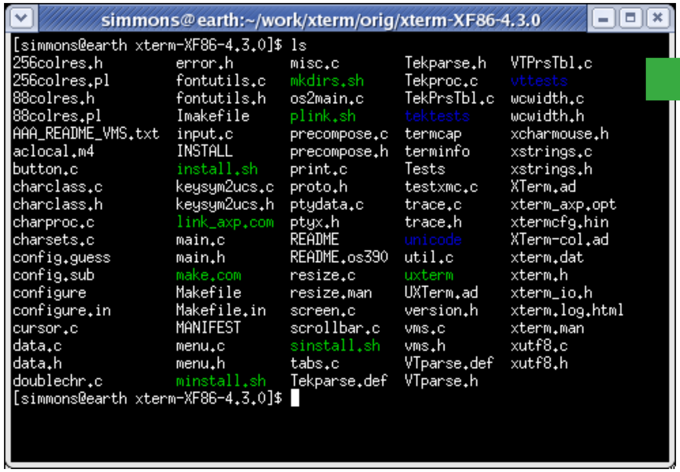
Requests for **making new tools available** shall be addressed to admin.bioinfo@cirad.fr

The GALAXY project is supported in part by NSF, NHGRI, and the Huck Institutes of the Life Sciences.

History Options

Unnamed history 4.0 Mb

20: PhyML on data 18
19: PhyML on data 18
18: Fasta2Phylip on data 16
17: Gblocks on data 15
16: Gblocks on data 15
15: MAFFT on data 14
14: new.fasta



simmons@earth:~/work/xterm/orig/xterm-XF86-4.3.0

```
[simmons@earth xterm-XF86-4.3.0]$ ls
256colres.h         error.h       misc.c        Tekparse.h   VTPrsTbl.c
256colres.pl        fontutils.c   mkdirs.sh     Tekproc.c    vttests
88colres.h          fontutils.h   os2main.c     TekPrsTbl.c  wcwidth.c
88colres.pl         Imakefile     plink.sh      tektests     wcwidth.h
AAA_README_VMS.txt  input.c       precompose.c  termcap      xcharmouse.h
aclocal.m4          INSTALL       precompose.h  terminfo     xstrings.c
button.c            install.sh    print.c       Tests        xstrings.h
charclass.c         keysym2ucs.c  proto.h       testxmc.c    XTerm.ad
charclass.h         keysym2ucs.h  ptydata.c     trace.c      xterm_axp.opt
charproc.c          link_axp.com  ptyx.h        trace.h      xtermcfg.hin
charsets.c          main.c        README        unicode      XTerm-col.ad
config.guess        main.h        README.os390  util.c       xterm.dat
config.sub          make.com      resize.c      uxterm       xterm.h
configure           Makefile      resize.man    UXTerm.ad    xterm_io.h
configure.in        Makefile.in   screen.c      version.h    xterm.log.html
cursor.c            MANIFEST      scrollbar.c   vms.c        xterm.man
data.c              menu.c        sinstall.sh   vms.h        xutf8.c
data.h              menu.h        tabs.c        VTparse.def  xutf8.h
doublechr.c         minstall.sh   Tekparse.def  VTparse.h
[simmons@earth xterm-XF86-4.3.0]$
```

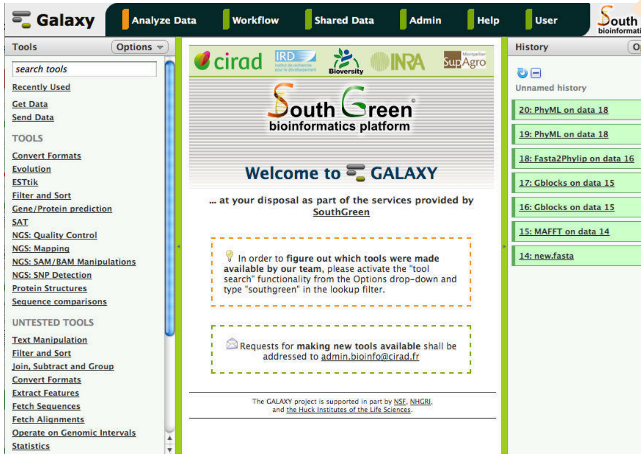
Cluster (208 cores, 50 TO storage)

Public databases

http://southgreen.cirad.fr/

South Green bioinformatics platform

Platform entry points and Galaxy



Recently Used

Get Data
Send Data

TOOLS

Convert Formats
Evolution
ESTtik
Filter and Sort
Gene/Protein prediction
SAT
NGS: Quality Control
NGS: Mapping
NGS: SAM/BAM Manipulations
NGS: SNP Detection
Protein Structures
Sequence comparisons

UNTESTED TOOLS

Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Fetch Alignments
Operate on Genomic Intervals
Statistics

Welcome to GALAXY

... at your disposal as part of the services provided by SouthGreen

In order to **figure out which tools were made available by our team**, please activate the "tool search" functionality from the Options drop-down and type "southgreen" in the lookup filter.

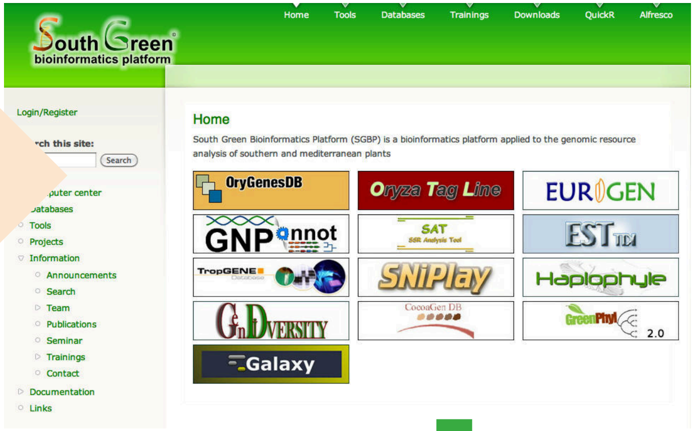
Requests for **making new tools available** shall be addressed to admin.bioinfo@cirad.fr

The GALAXY project is supported in part by NSF, NHGRI, and the Huck Institutes of the Life Sciences.

Unnamed history 4.0 Mb

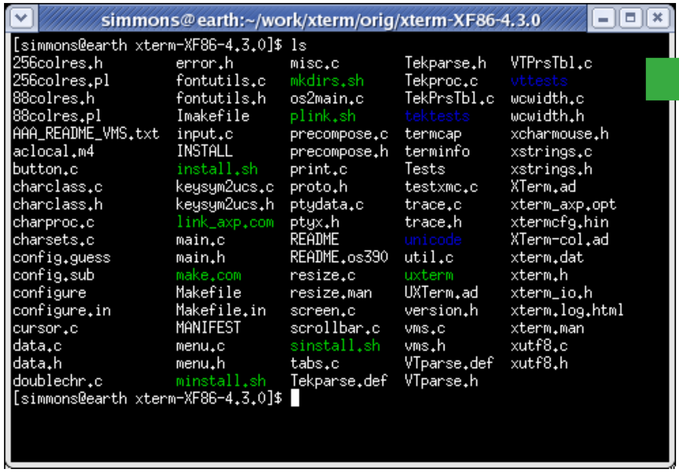
20: PhyML on data 18
19: PhyML on data 18
18: Fasta2Phylip on data 16
17: Gblocks on data 15
16: Gblocks on data 15
15: MAFFT on data 14
14: new.fasta

Reproducibility



Home

South Green Bioinformatics Platform (SGBP) is a bioinformatics platform applied to the genomic resource analysis of southern and mediterranean plants



```
[simmons@earth xterm-XF86-4.3.0]$ ls
256colres.h      error.h      misc.c         Tekparse.h    VTPrsTbl.c
256colres.pl     fontutils.c  mkdirs.sh      Tekproc.c     vttests
88colres.h       fontutils.h  os2main.c      TekPrsTbl.c   wcwidth.c
88colres.pl      Imakefile    plink.sh       tektests      wcwidth.h
AAA_README_VMS.txt input.c    precompose.c   termcap       xcharmouse.h
aclocal.m4       INSTALL      precompose.h   terminfo      xstrings.c
button.c         install.sh   print.c        Tests         xstrings.h
charclass.c      keysym2ucs.c proto.h        testxmc.c     XTerm.ad
charclass.h      keysym2ucs.h ptydata.c      trace.c       xterm_axp.opt
charproc.c       link_axp.com ptyx.h         trace.h       xtermcfg.hin
charsets.c       main.c       README         unicode       XTerm-col.ad
config.guess     main.h       README.os390   util.c        xterm.dat
config.sub       make.com     resize.c       uxterm        xterm.h
configure        Makefile     resize.man     UXTerm.ad     xterm_io.h
configure.in     Makefile.in  screen.c       version.h     xterm.log.html
cursor.c         MANIFEST     scrollbar.c    vms.c         xterm.man
data.c           menu.c       sinstall.sh    vms.h         xutf8.c
data.h           menu.h       tabs.c         VTparse.def   xutf8.h
doublechr.c      minstall.sh  Tekparse.def   VTparse.h
[simmons@earth xterm-XF86-4.3.0]$ █
```

Cluster (208 cores, 50 TO storage)

Public databases

http://southgreen.cirad.fr/

South Green bioinformatics platform

Animation and maintenance

Agile programming for maintenance and developments

1 session of 4 hours, every 2 weeks

6-10 developers, researchers (bioinformatic, biology), using pair programming

Integration of new softwares

Code maintenance, refactoring, documentation.

Galaxy, bioinformatic training platform for biologists

Galaxy is widely used during Southgreen trainings, as a complete replacement of command line system.

South Green bioinformatics platform

South Green
bioinformatics platform

Bacchus pipeline

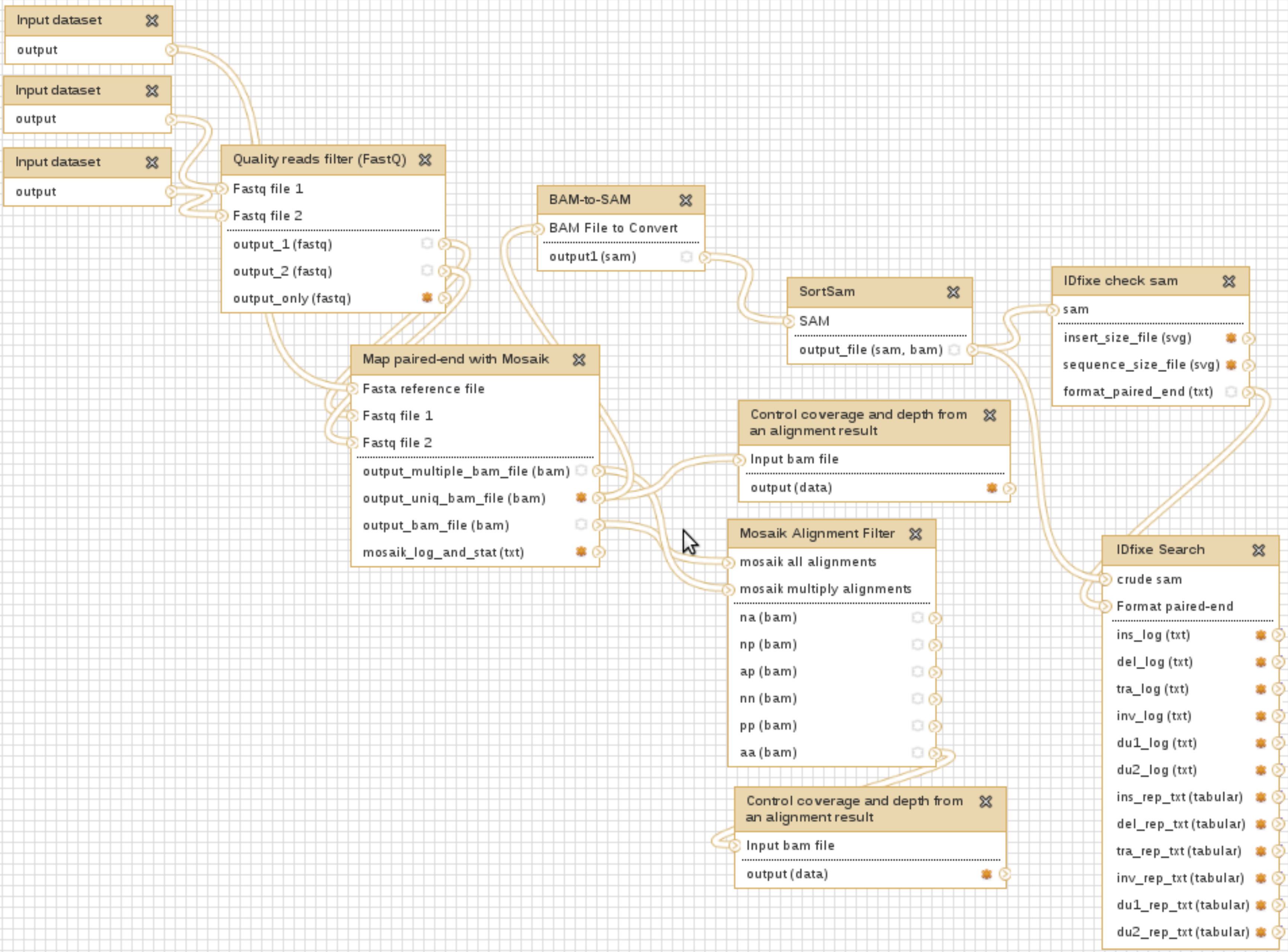
Vincent Maillol (INRA)

Bacchus pipeline

Designed to analyze NGS sequenced grape vines

And search polymorphism between genotype

Bacchus pipeline



Input dataset
output

Input dataset
output

Input dataset
output

Quality reads filter (FastQ)
Fastq file 1
Fastq file 2
output_1 (fastq)
output_2 (fastq)
output_only (fastq)

BAM-to-SAM
BAM File to Convert
output1 (sam)

Map paired-end with Mosaik
Fasta reference file
Fastq file 1
Fastq file 2
output_multiple_bam_file (bam)
output_uniq_bam_file (bam)
output_bam_file (bam)
mosaik_log_and_stat (txt)

SortSam
SAM
output_file (sam, bam)

IDfixe check sam
sam
insert_size_file (svg)
sequence_size_file (svg)
format_paired_end (txt)

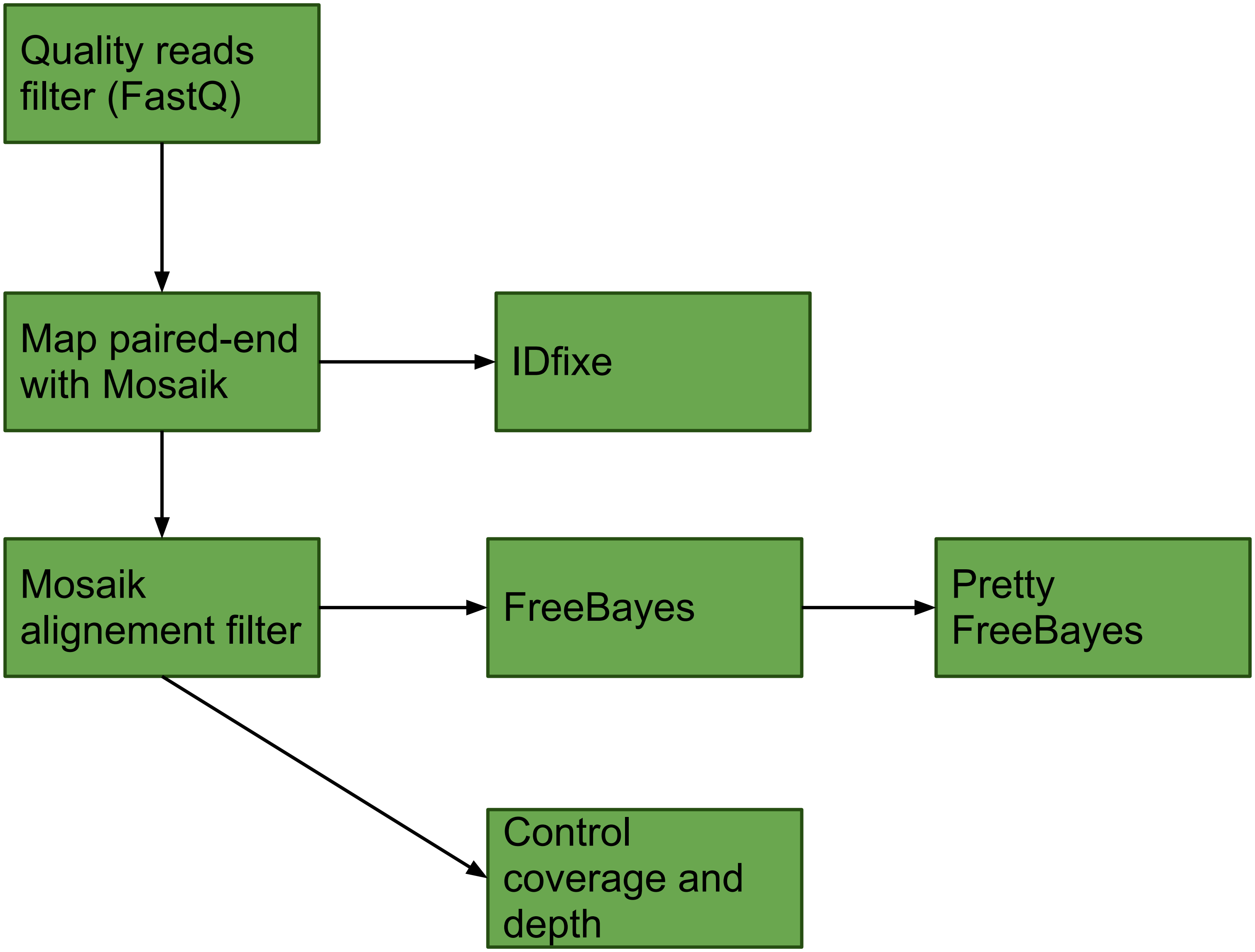
Control coverage and depth from an alignment result
Input bam file
output (data)

Mosaik Alignment Filter
mosaik all alignments
mosaik multiply alignments
na (bam)
np (bam)
ap (bam)
nn (bam)
pp (bam)
aa (bam)

Control coverage and depth from an alignment result
Input bam file
output (data)

IDfixe Search
crude sam
Format paired-end
ins_log (txt)
del_log (txt)
tra_log (txt)
inv_log (txt)
du1_log (txt)
du2_log (txt)
ins_rep_txt (tabular)
del_rep_txt (tabular)
tra_rep_txt (tabular)
inv_rep_txt (tabular)
du1_rep_txt (tabular)
du2_rep_txt (tabular)

Bacchus pipeline



Quality reads filter (FastQ)

Map paired-end with Mosaik

IDfixe

Mosaik alignement filter

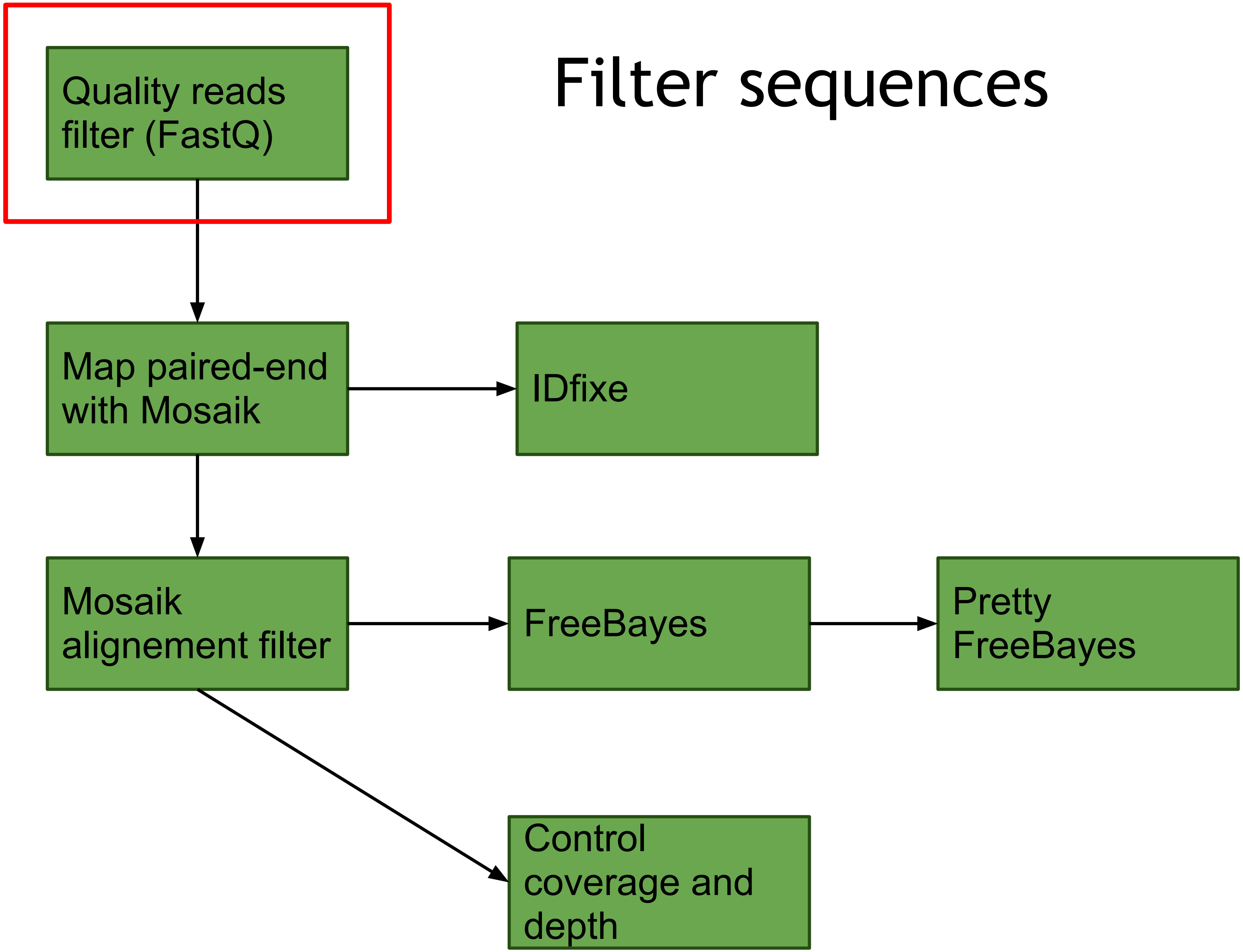
FreeBayes

Pretty FreeBayes

Control coverage and depth

Bacchus pipeline

Filter sequences



Quality reads filter (FastQ)

Map paired-end with Mosaik

IDfixe

Mosaik alignement filter

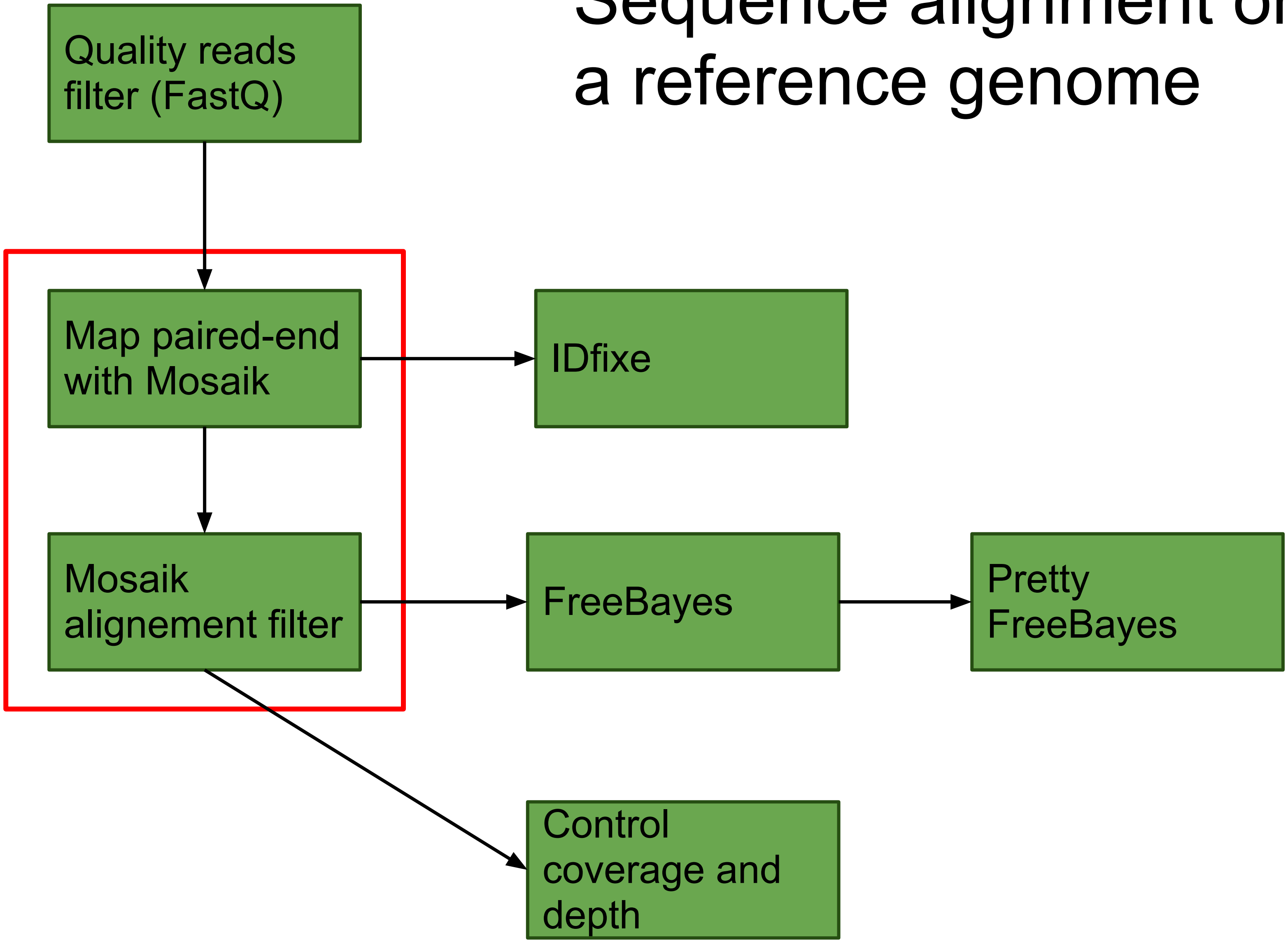
FreeBayes

Pretty FreeBayes

Control coverage and depth

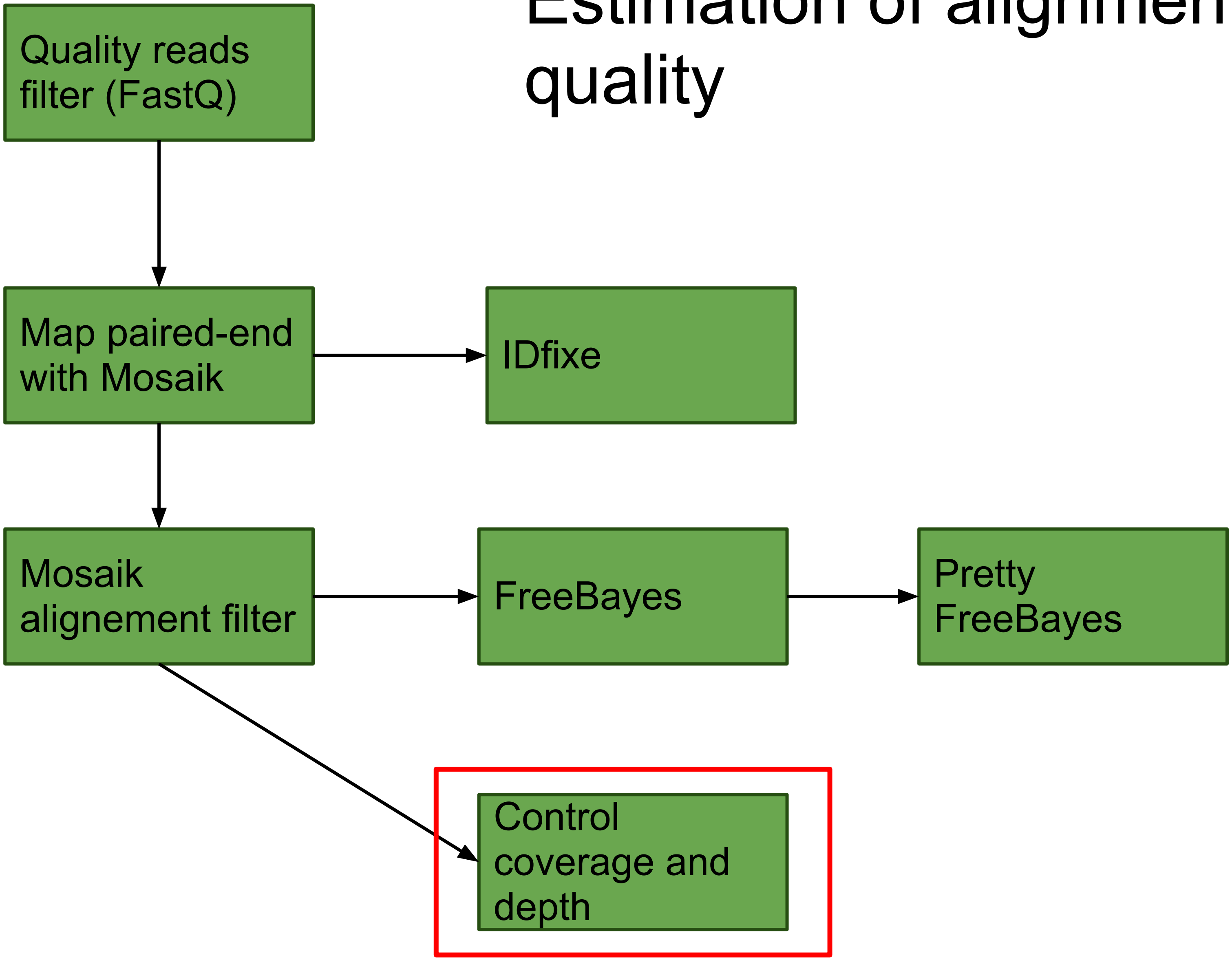
Bacchus pipeline

Sequence alignment on a reference genome

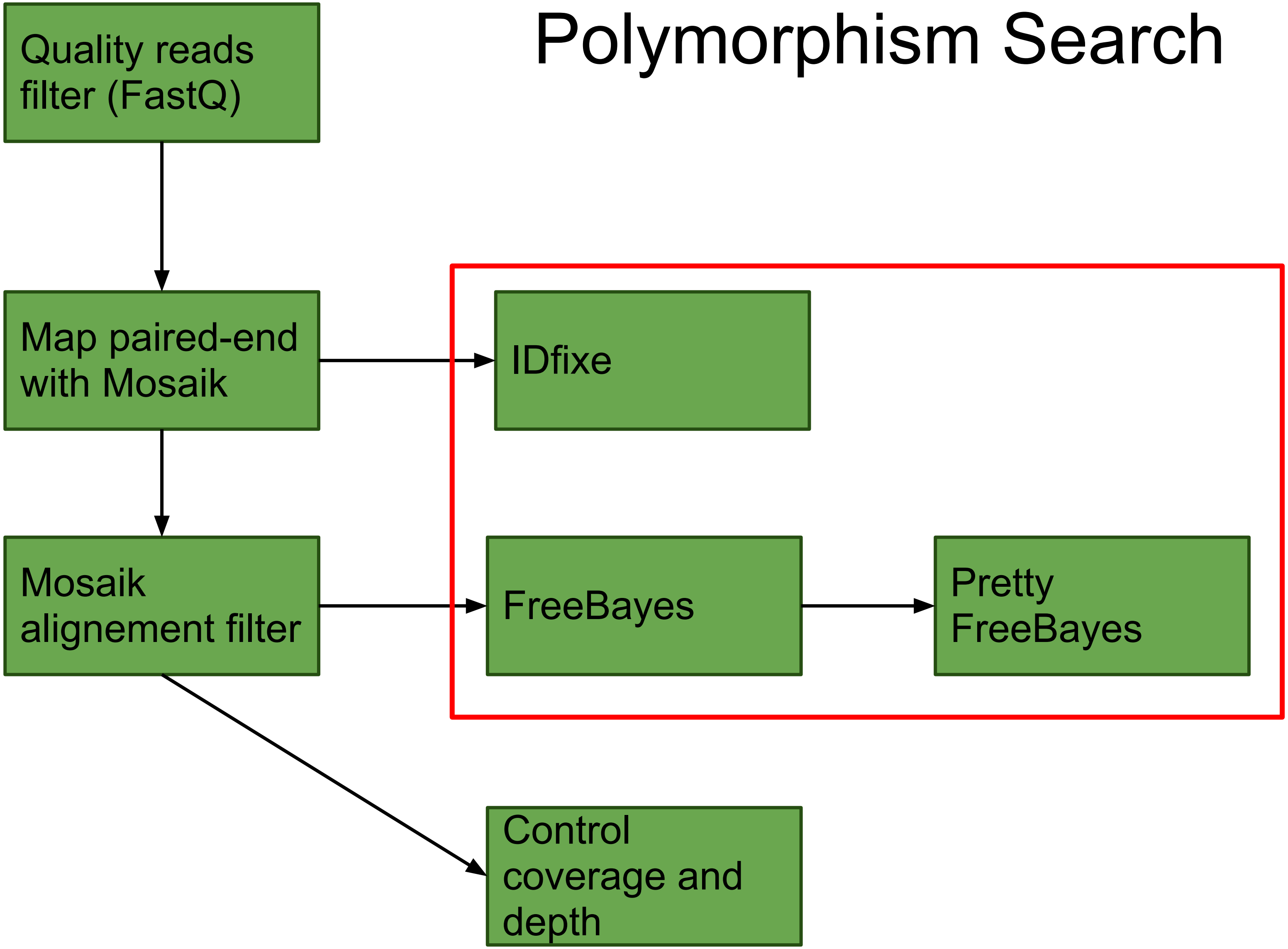

Quality reads filter (FastQ)
Map paired-end with Mosaik
IDfixe
Mosaik alignement filter
FreeBayes
Pretty FreeBayes
Control coverage and depth

Bacchus pipeline

Estimation of alignment quality


Quality reads filter (FastQ)
Map paired-end with Mosaik
IDfixe
Mosaik alignement filter
FreeBayes
Pretty FreeBayes
Control coverage and depth

Bacchus pipeline



Polymorphism Search

Quality reads filter (FastQ)

Map paired-end with Mosaik

Mosaik alignement filter

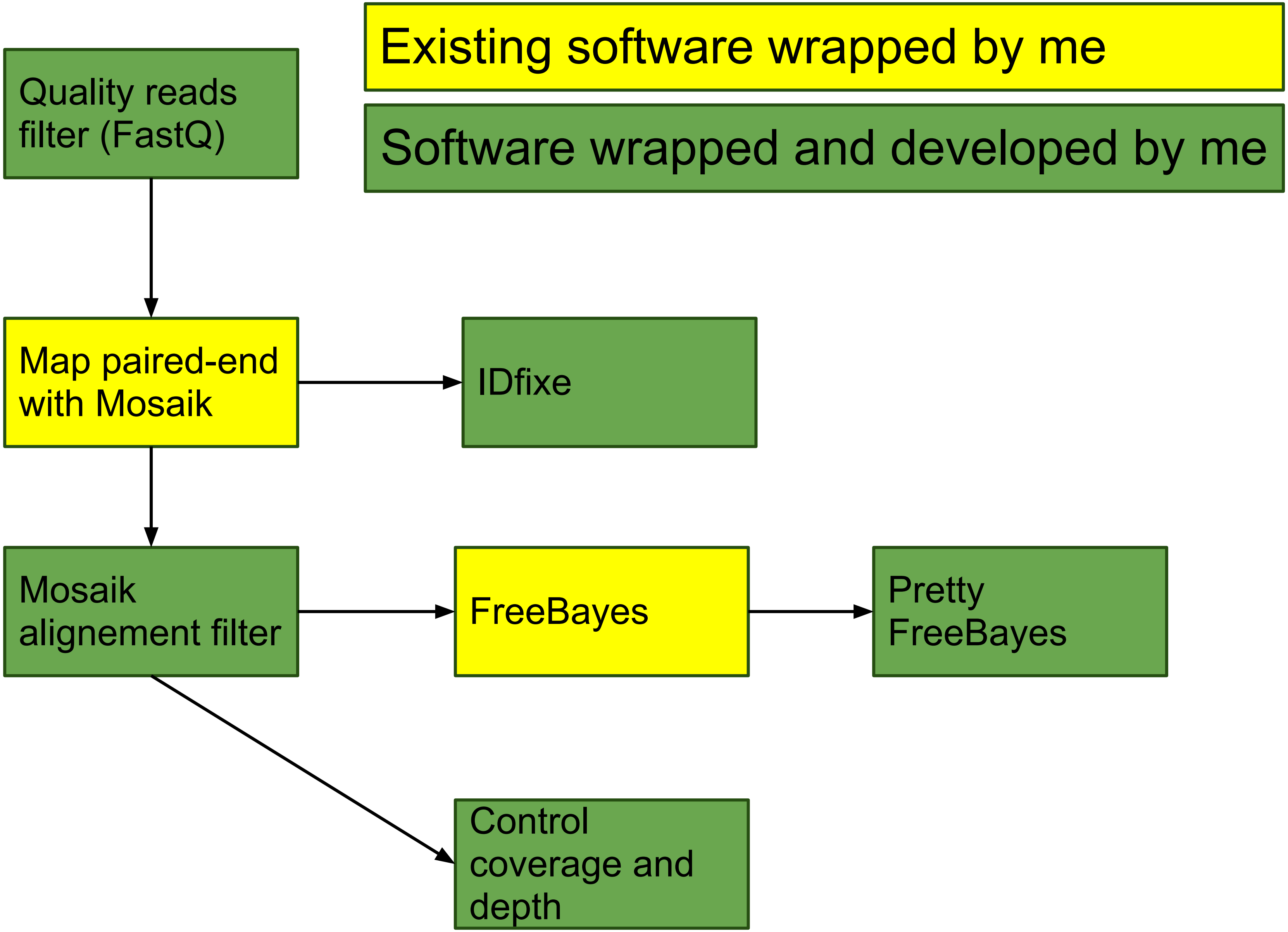
IDfixe

FreeBayes

Pretty FreeBayes

Control coverage and depth

Bacchus pipeline


Existing software wrapped by me

Software wrapped and developed by me

Quality reads filter (FastQ)

Map paired-end with Mosaik

IDfixe

Mosaik alignement filter

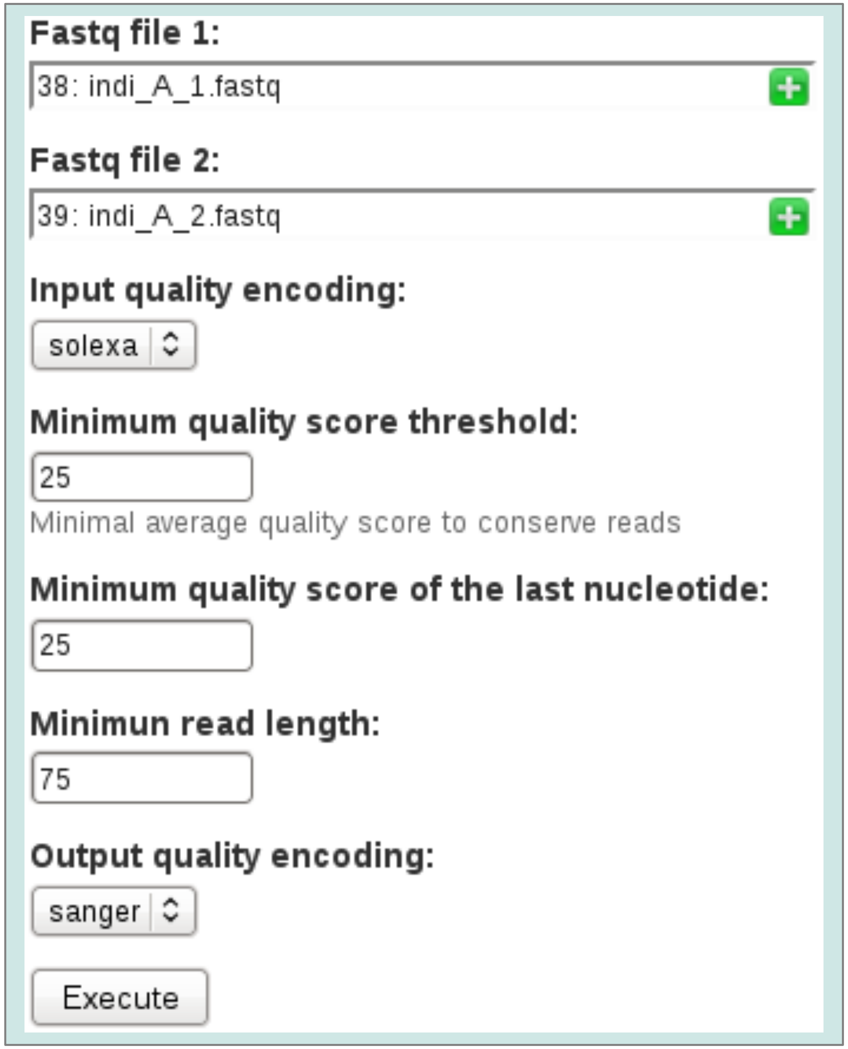
FreeBayes

Pretty FreeBayes

Control coverage and depth

Read quality filter (FastQ)

Quality reads filter (FastQ) | Map paired-end with Mosaik | Mosaik Align. filter | coverage and depth | Freebayes | IDfixe | Pretty Freebayes



Fastq file 1:
38: indi_A_1.fastq

Fastq file 2:
39: indi_A_2.fastq

Input quality encoding:
solexa

Minimum quality score threshold:
25
Minimal average quality score to conserve reads

Minimum quality score of the last nucleotide:
25

Minimun read length:
75

Output quality encoding:
sanger

- Truncate the end of sequences
- Filter short sequences
- Filter sequence under average quality rate
- convert quality format

Alignment by Mosaik 2.1.33

Quality reads filter (FastQ)

Map paired-end with Mosaik

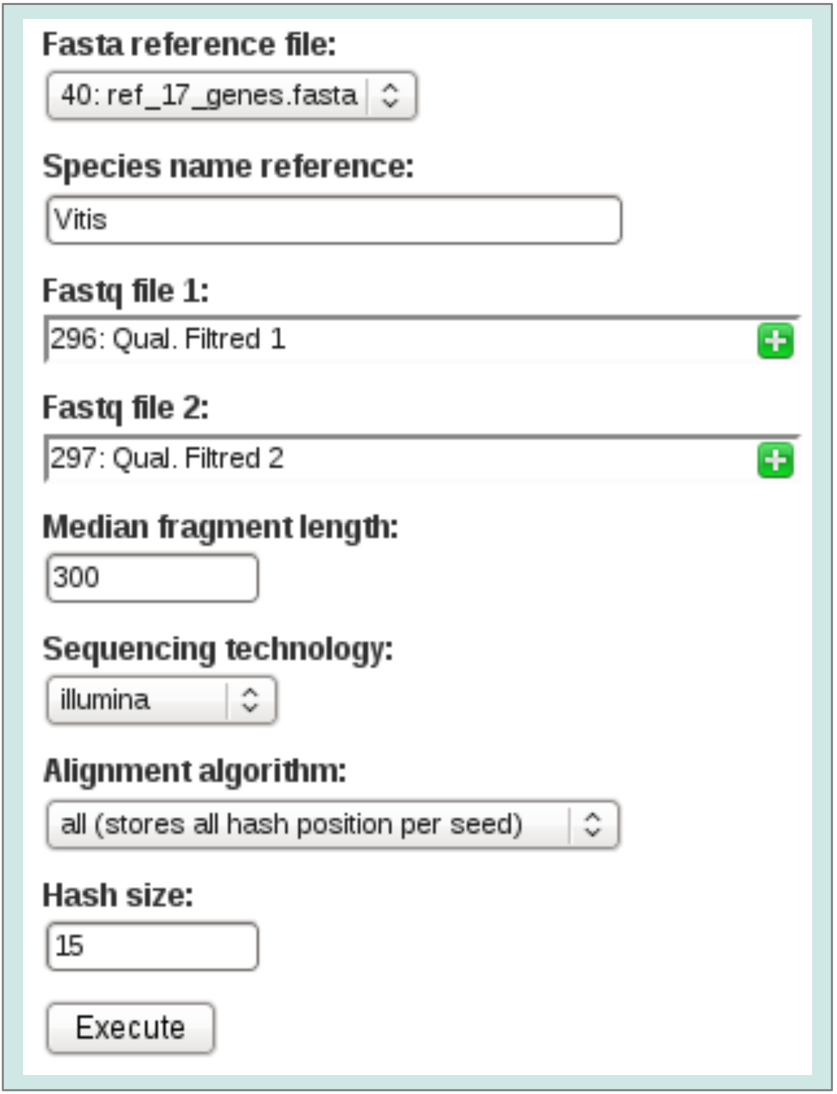
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



Fasta reference file: 40: ref_17_genes.fasta

Species name reference: Vitis

Fastq file 1: 296: Qual. Filtred 1

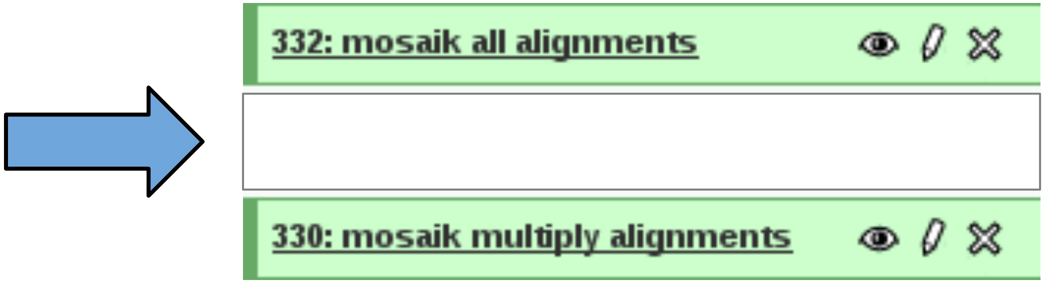
Fastq file 2: 297: Qual. Filtred 2

Median fragment length: 300

Sequencing technology: illumina

Alignment algorithm: all (stores all hash position per seed)

Hash size: 15



332: mosaik all alignments

330: mosaik multiply alignments

Alignment by Mosaik 2.1.33

Quality reads filter (FastQ)

Map paired-end with Mosaik

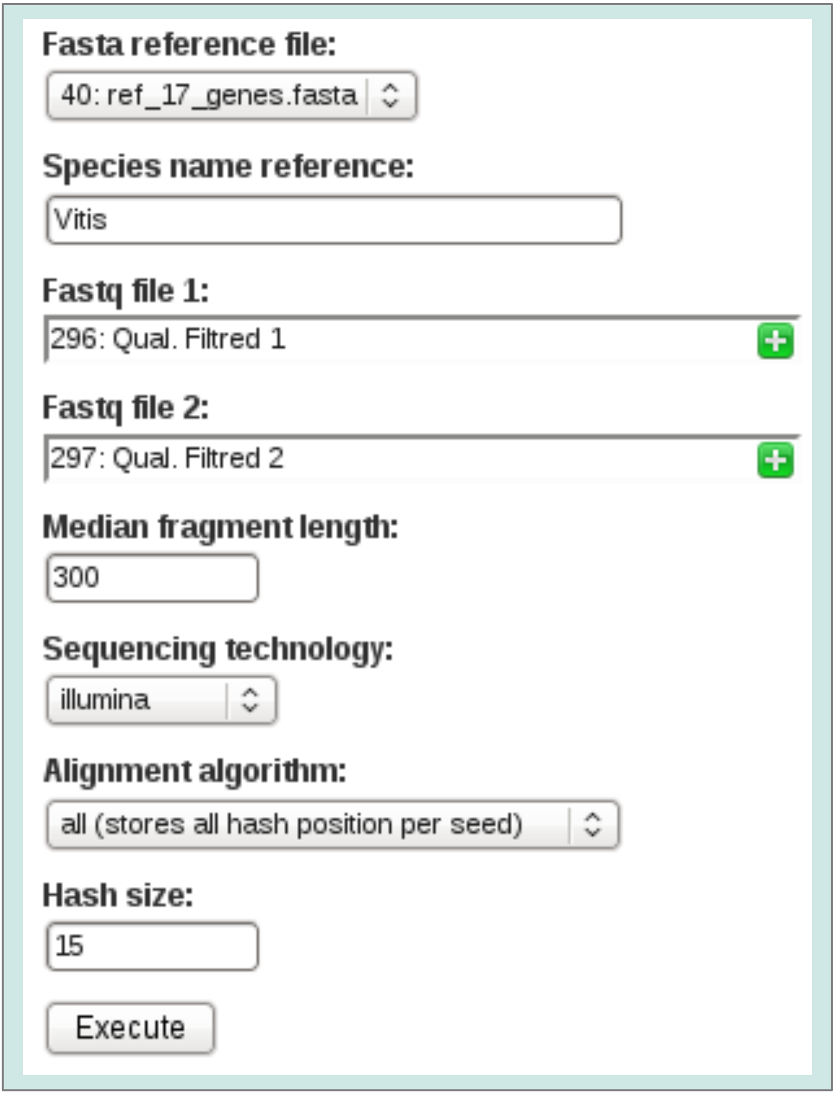
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



Fasta reference file:
40: ref_17_genes.fasta

Species name reference:
Vitis

Fastq file 1:
296: Qual. Filtred 1

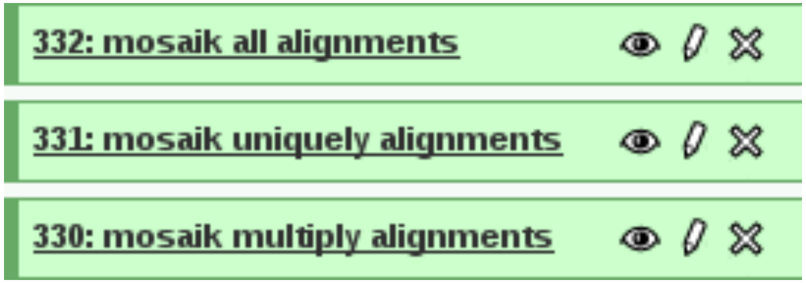
Fastq file 2:
297: Qual. Filtred 2

Median fragment length:
300

Sequencing technology:
illumina

Alignment algorithm:
all (stores all hash position per seed)

Hash size:
15



332: mosaik all alignments

331: mosaik uniquely alignments

330: mosaik multiply alignments

Mosaik alignment filter

Quality reads filter (FastQ)

Map paired-end with Mosaik

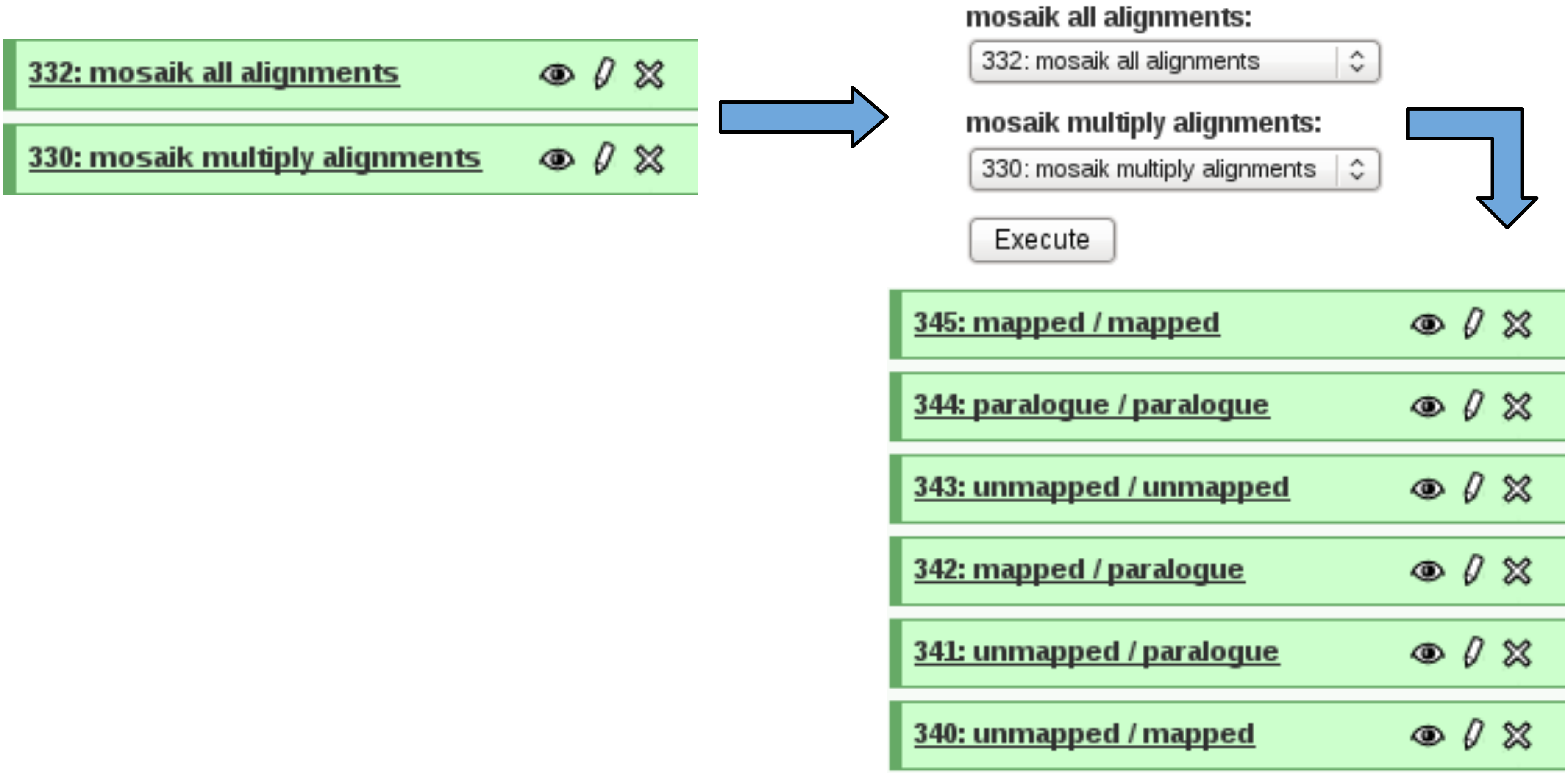
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



332: mosaik all alignments

330: mosaik multiply alignments

mosaik all alignments: 332: mosaik all alignments

mosaik multiply alignments: 330: mosaik multiply alignments

345: mapped / mapped

344: paralogue / paralogue

343: unmapped / unmapped

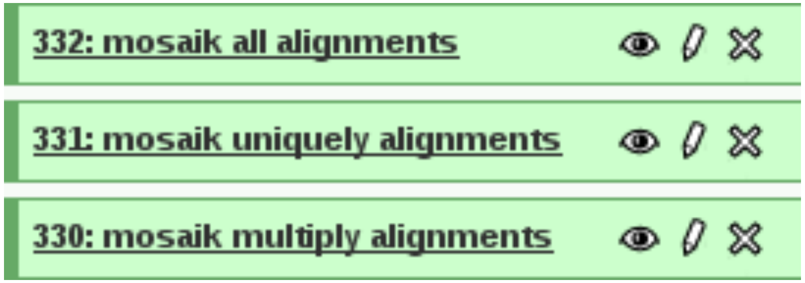
342: mapped / paralogue

341: unmapped / paralogue

340: unmapped / mapped

Mosaik alignment filter

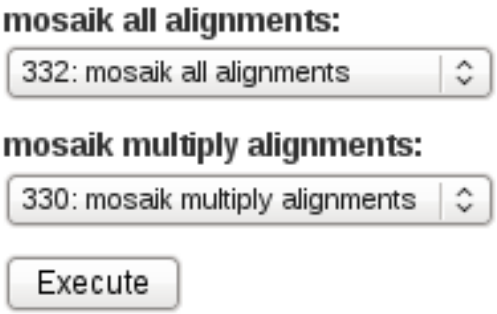
Quality reads filter (FastQ) | Map paired-end with Mosaik | **Mosaik Align. filter** | coverage and depth | Freebayes | IDfixe | Pretty Freebayes



332: mosaik all alignments

331: mosaik uniquely alignments

330: mosaik multiply alignments

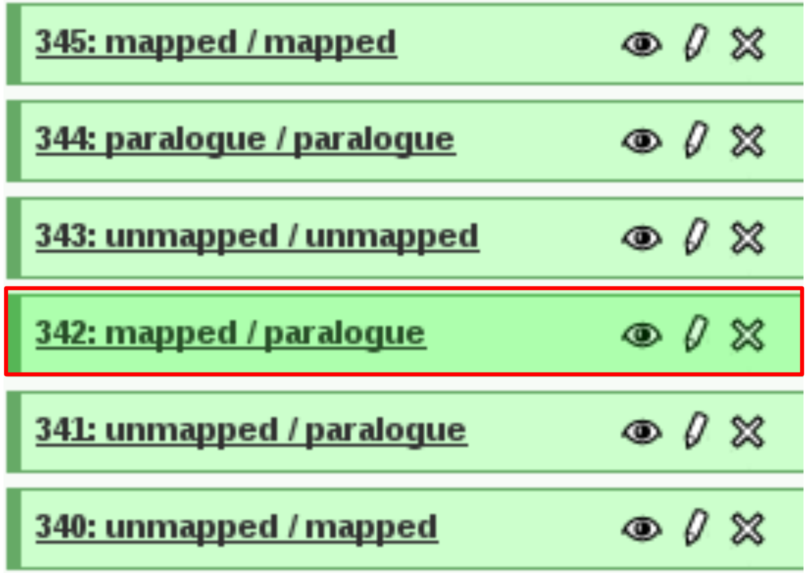


mosaik all alignments:

332: mosaik all alignments

mosaik multiply alignments:

330: mosaik multiply alignments



345: mapped / mapped

344: paralogue / paralogue

343: unmapped / unmapped

342: mapped / paralogue

341: unmapped / paralogue

340: unmapped / mapped

Sequence alignment control

Quality reads filter (FastQ)

Map paired-end with Mosaik

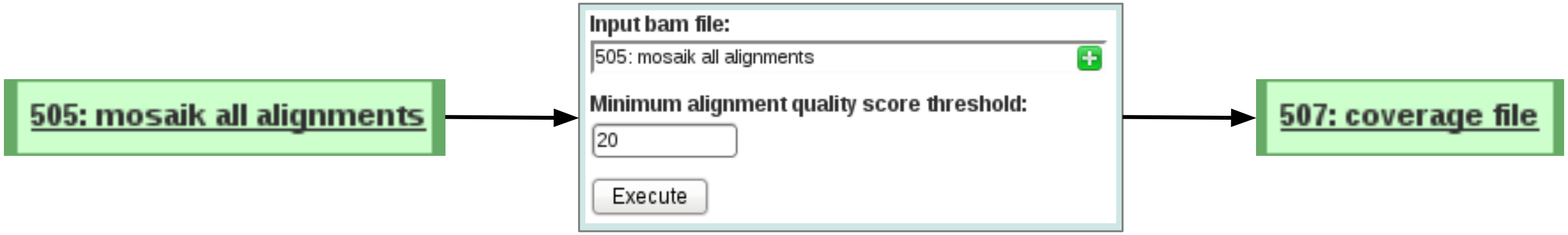
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



505: mosaik all alignments

Input bam file: 505: mosaik all alignments

Minimum alignment quality score threshold: 20

507: coverage file

```
+----------------------------------------------------------------+
|                     magic number (1668445300)               8b |
+----------------------------------------------------------------+
+----------------...-++--------++--------------------------------+
|   individual_name  ||  \0 1b ||              nb_ref         4b |
+----------------...-++--------++--------------------------------+

/  +----...-++--------+
|  |ref_name||  \0 1b |
\  +----...-++--------+
   +-----------------------------------------------------------------+
   |                       ref_offset                             8b |
   +-----------------------------------------------------------------+
   +---------------------------------+  \
   |     size_ref                 4b |  | * nb_ref
   +---------------------------------+  /

+----------------+
|     depth   2b |  * ( ref_offset[ nb_ref-1 ] - ref_offset[ 0 ] )
+----------------+
```

- Memory leak eliminate with Valgrind
- ansi C language
- Internal Framework for test driving

Sequence alignment control

Quality reads filter (FastQ)

Map paired-end with Mosaik

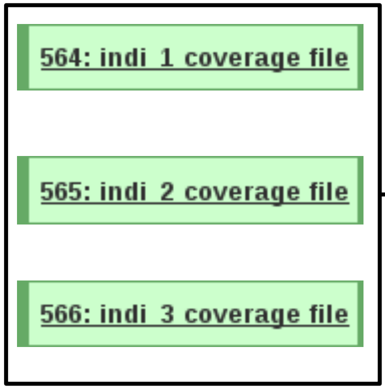
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

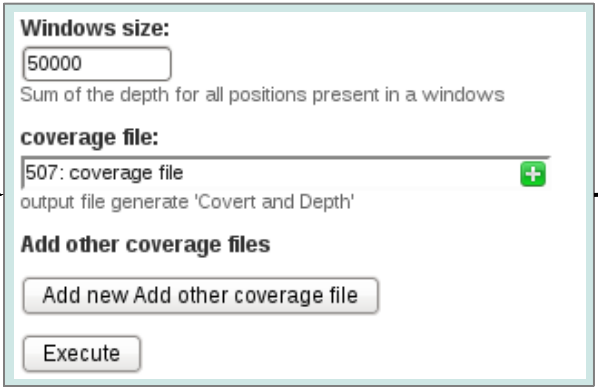
Pretty Freebayes



564: indi_1 coverage file

565: indi_2 coverage file

566: indi_3 coverage file



Windows size:

50000

Sum of the depth for all positions present in a windows

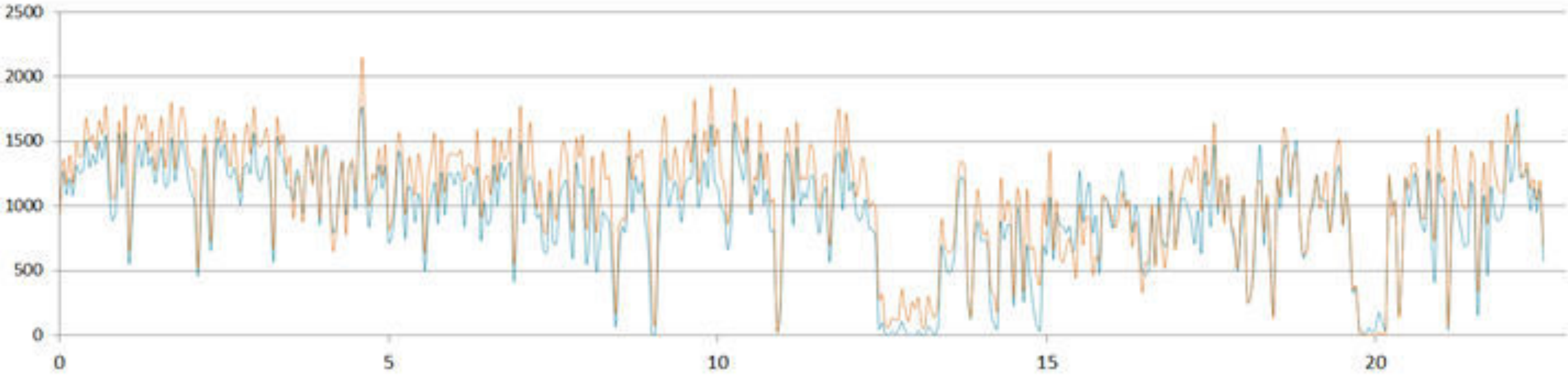
coverage file:

507: coverage file

output file generate 'Covert and Depth'

Add other coverage files

Ref	Pos	Indi_1	Indi_2	Indi_3
chr1	10000	212522	213150	212854
chr1	20000	209954	212431	211192
chr1	30000	105892	215241	206841
chr1	40000	152921	182152	175890
chr2	10000	211139	215226	215759
chr2	20000	225521	231155	224845
chr2	30000	212050	218751	216568
chr3	10000	211925	220015	213952



Sequence alignment control

Quality reads filter (FastQ)

Map paired-end with Mosaik

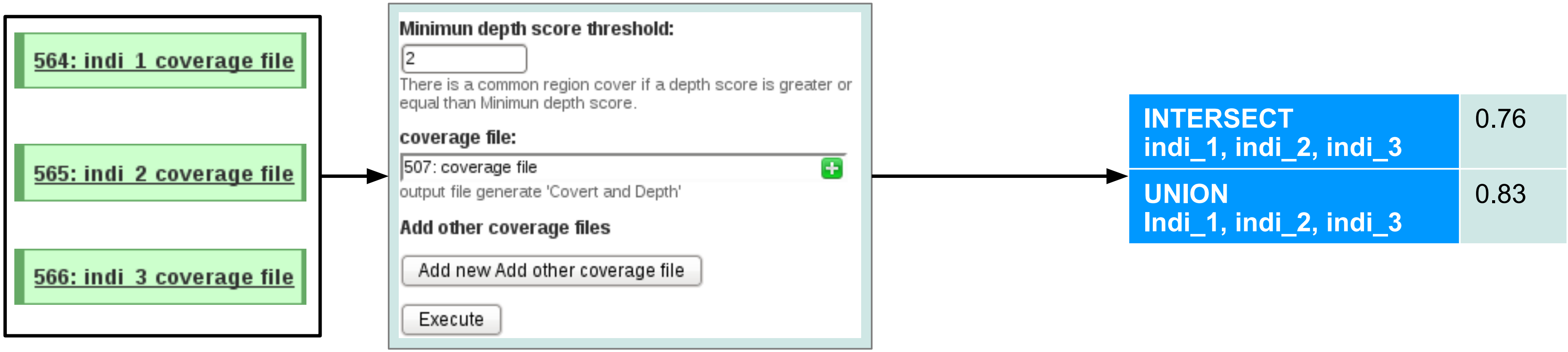
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



564: indi_1 coverage file

565: indi_2 coverage file

566: indi_3 coverage file

Minimun depth score threshold:

2

There is a common region cover if a depth score is greater or equal than Minimun depth score.

coverage file:

507: coverage file

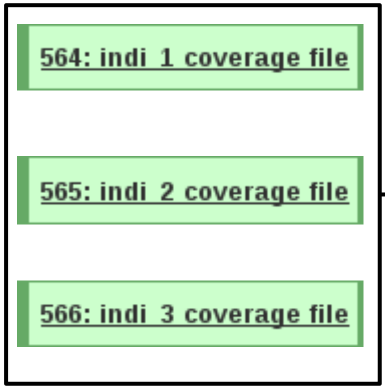
output file generate 'Covert and Depth'

Add other coverage files

INTERSECT **indi_1, indi_2, indi_3**	0.76
UNION **Indi_1, indi_2, indi_3**	0.83

Sequence alignment control

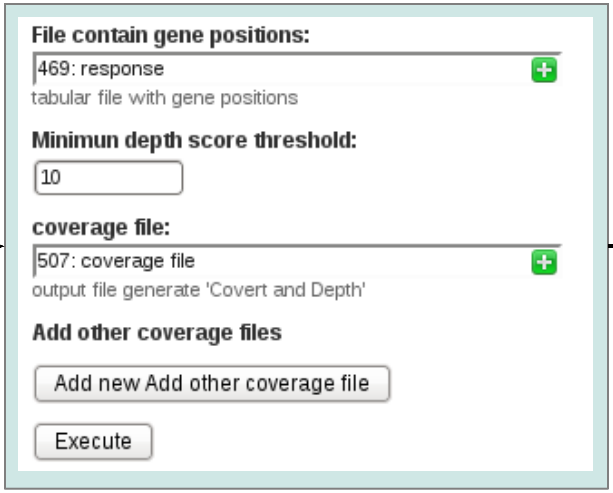
- Quality reads filter (FastQ)
- Map paired-end with Mosaik
- Mosaik Align. filter
- coverage and depth
- Freebayes
- IDfixe
- Pretty Freebayes



564: indi_1 coverage file

565: indi_2 coverage file

566: indi_3 coverage file



File contain gene positions:

469: response

tabular file with gene positions

Minimun depth score threshold:

10

coverage file:

507: coverage file

output file generate 'Covert and Depth'

Add other coverage files

Gene	Indi_1	Indi_2	Indi_3
bar	0.539	0.898	0.782
foo	1.0	1.0	1.0
foo_bar	0.225	0.345	0.651

SNP (Freebayes 0.9.4)

Quality reads filter (FastQ)

Map paired-end with Mosaik

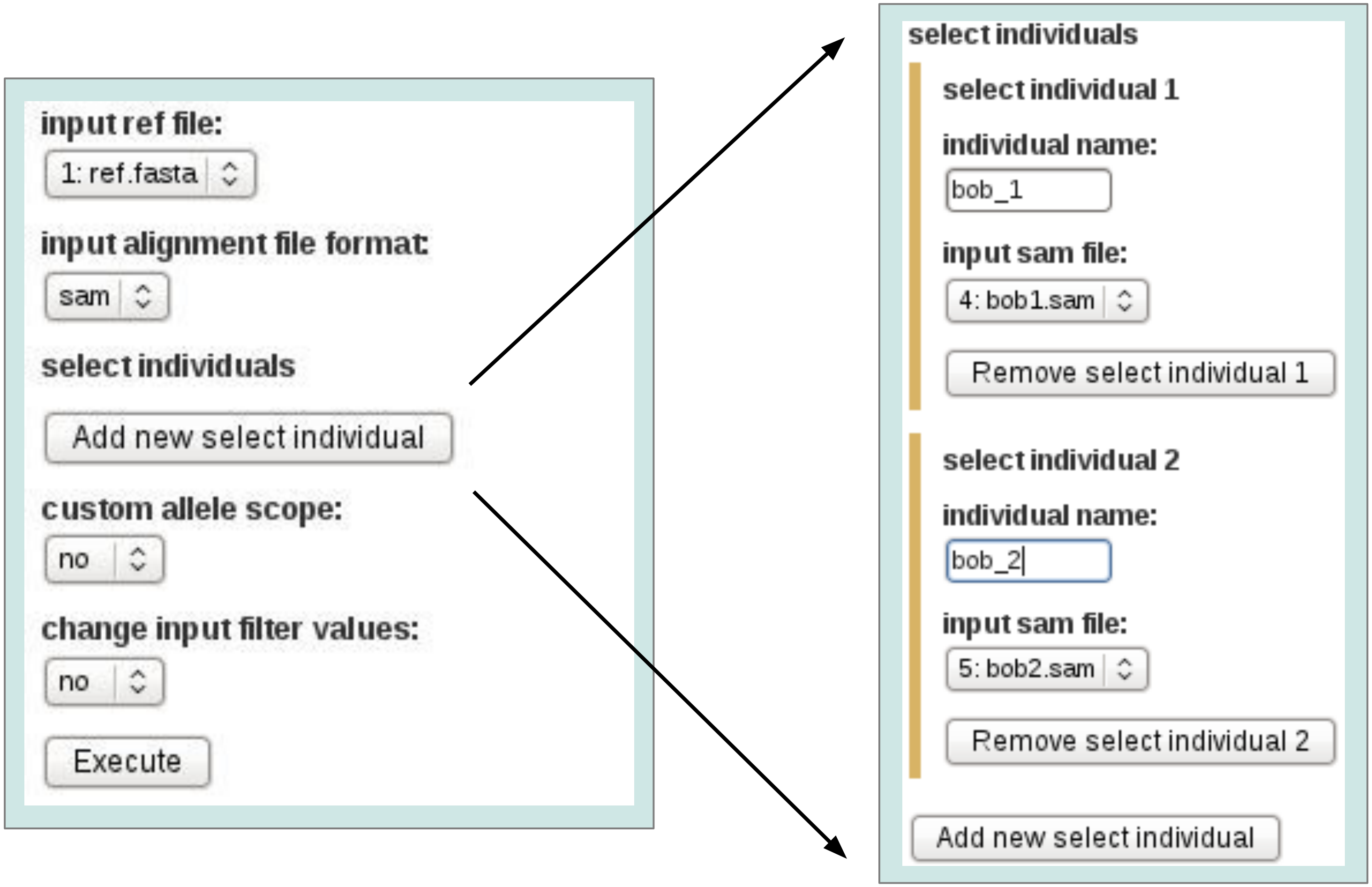
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



input ref file: 1: ref.fasta

input alignment file format: sam

select individuals

Add new select individual

custom allele scope: no

change input filter values: no

select individuals

select individual 1

individual name: bob_1

input sam file: 4: bob1.sam

Remove select individual 1

select individual 2

individual name: bob_2

input sam file: 5: bob2.sam

Remove select individual 2
Add new select individual

SNP (Pretty Freebayes)

uali filte

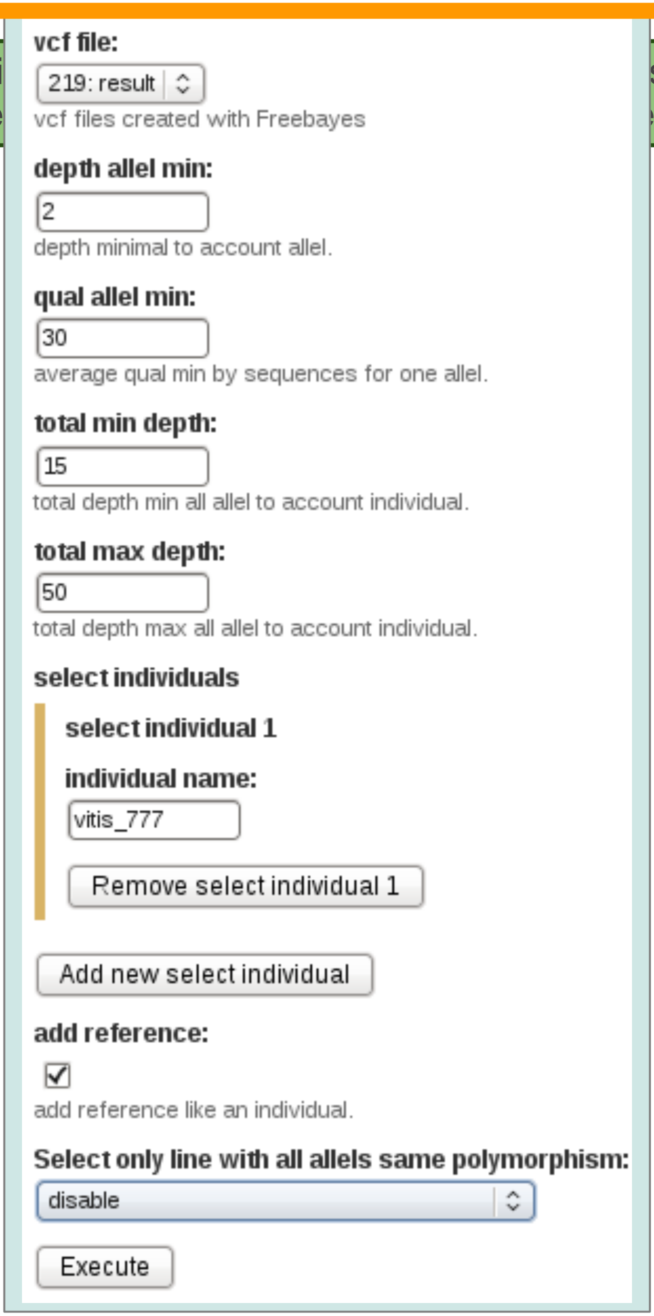
saik Align. er

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



vcf file:

219: result

vcf files created with Freebayes

depth allel min:

2

depth minimal to account allel.

qual allel min:

30

average qual min by sequences for one allel.

total min depth:

15

total depth min all allel to account individual.

total max depth:

50

total depth max all allel to account individual.

select individuals

select individual 1

individual name:

vitis_777

add reference:

☑

add reference like an individual.

Select only line with all allels same polymorphism:

disable

rname	pos	indi_1	indi_2	indi_3
At1g_	586	C:	C:T	C:
GSVIVG0100127	748	A:	A:	A:T:
GSVIVG0100127	1235	TC:C	C:	C:
GSVIVG0100127	2751	G:	G:	C:

IDfixe - Under the hood

Quality reads filter (FastQ)

Map paired-end with Mosaik

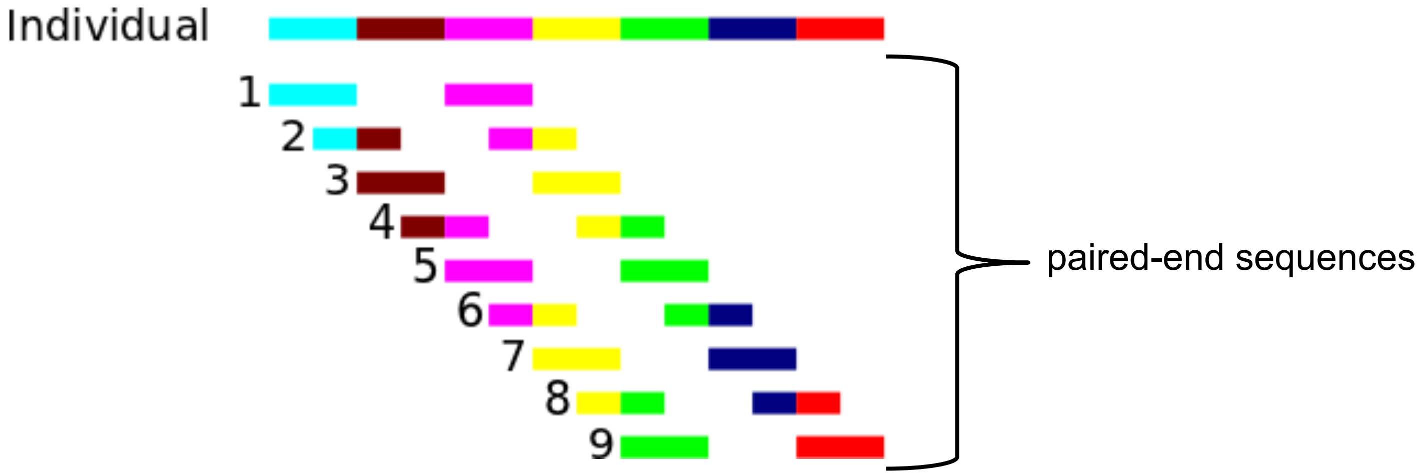
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



IDfixe - Under the hood

Quality reads filter (FastQ)

Map paired-end with Mosaik

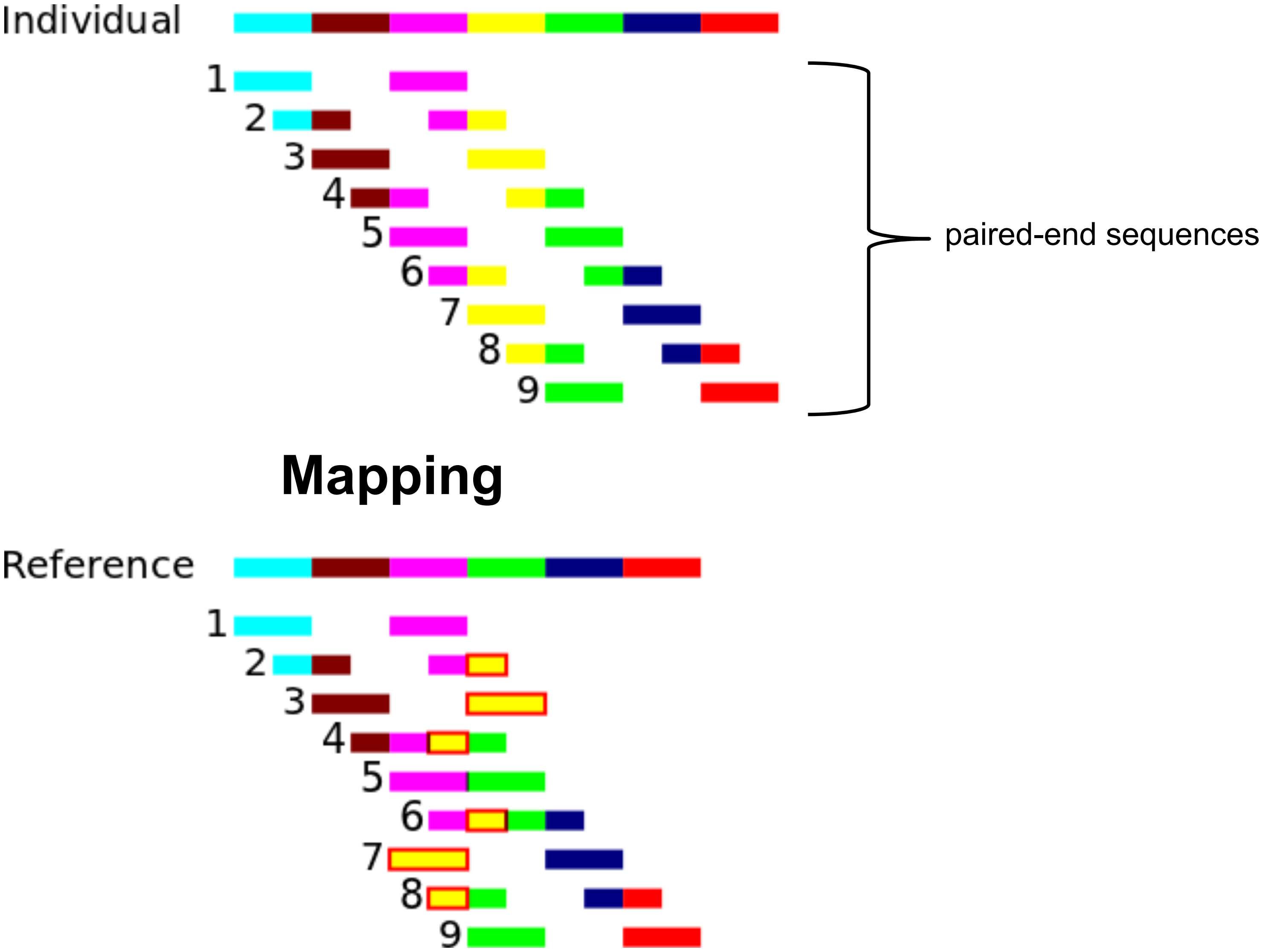
Mosaik Align. filter

coverage and depth

Freebayes

IDfixe

Pretty Freebayes



IDfixe - Under the hood

Quality reads filter (FastQ)

Map paired-end with Mosaik

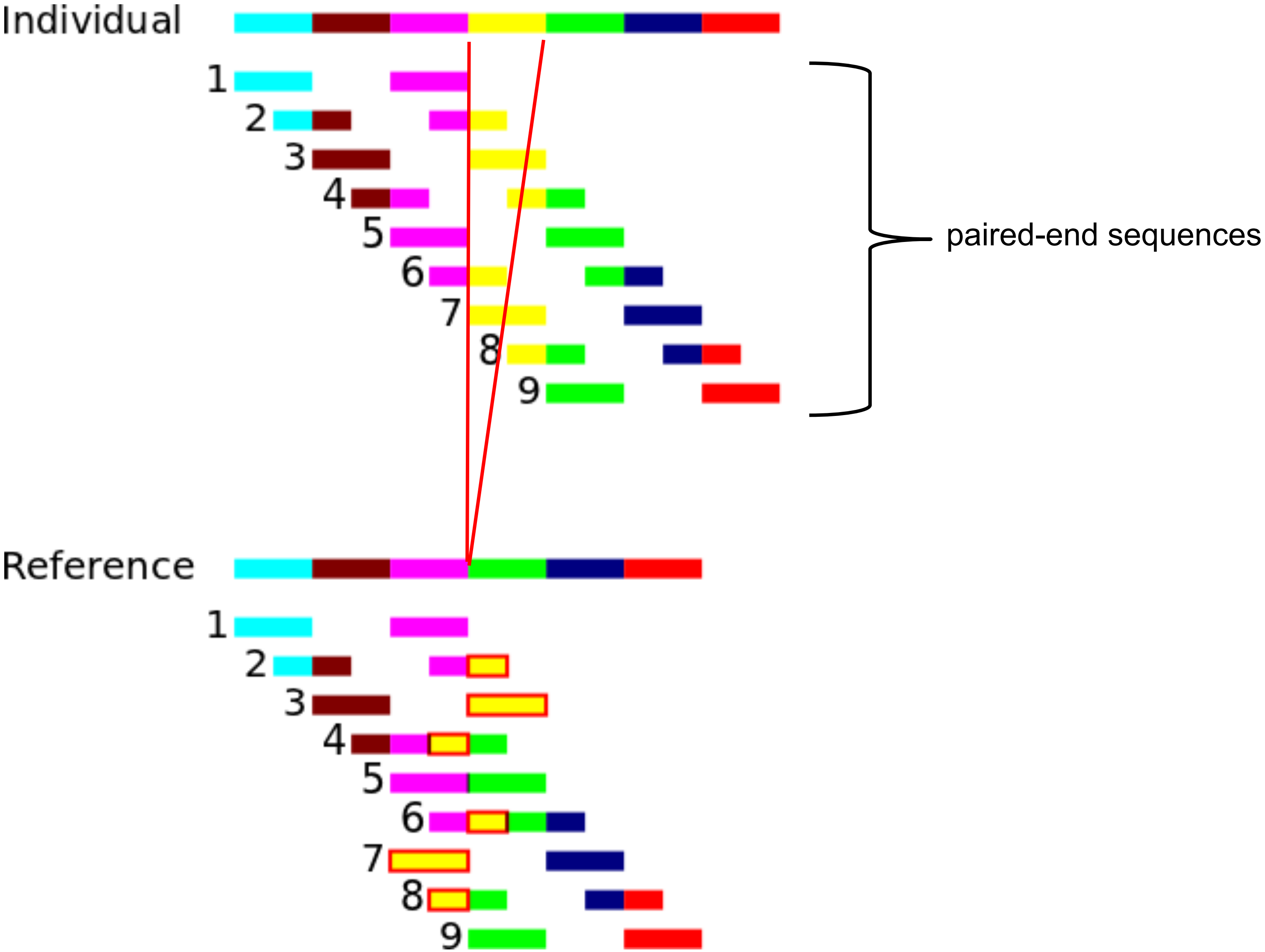
Mosaik Align. filter

coverage and depth

Freebayes

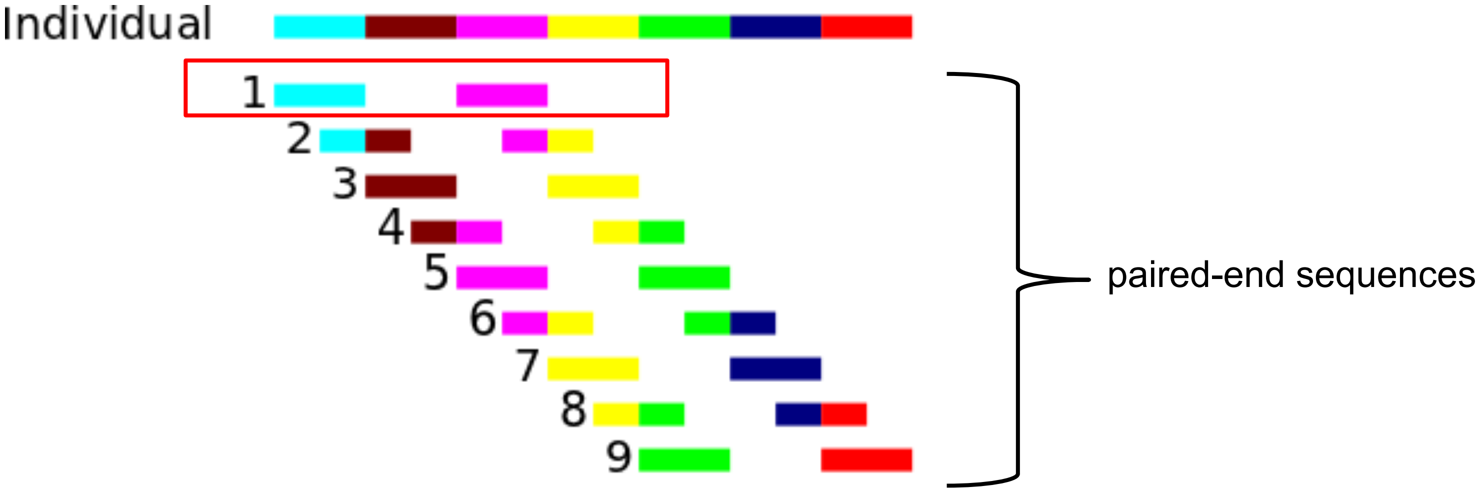
IDfixe

Pretty Freebayes

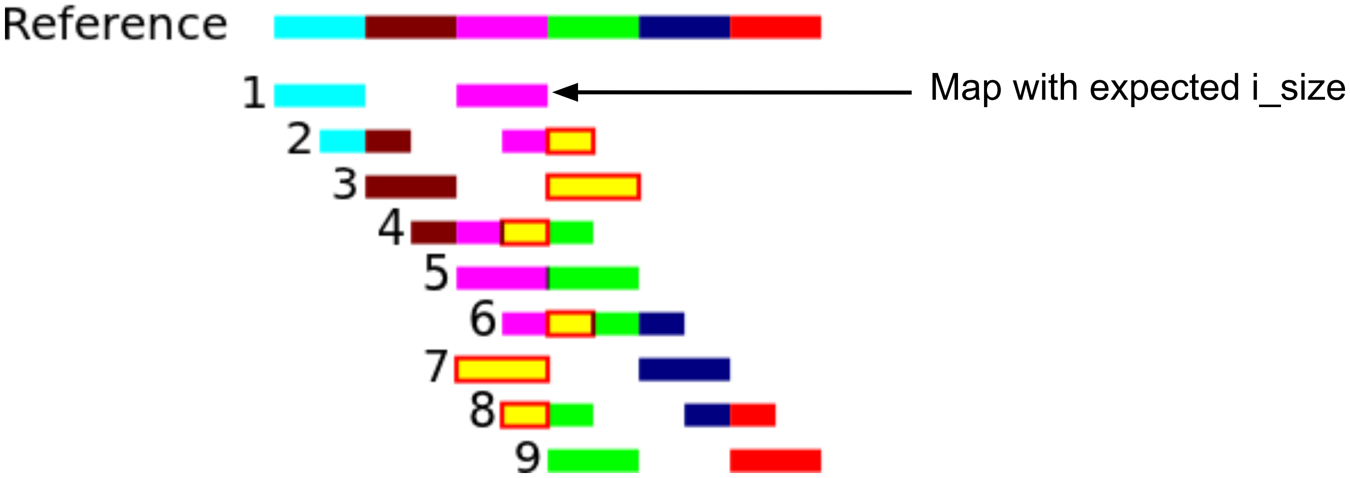


IDfixe - Under the hood

Quality reads filter (FastQ) | Map paired-end with Mosaik | Mosaik Align. filter | coverage and depth | Freebayes | IDfixe | Pretty Freebayes



Mapping



IDfixe - Under the hood

Quality reads filter (FastQ)

Map paired-end with Mosaik

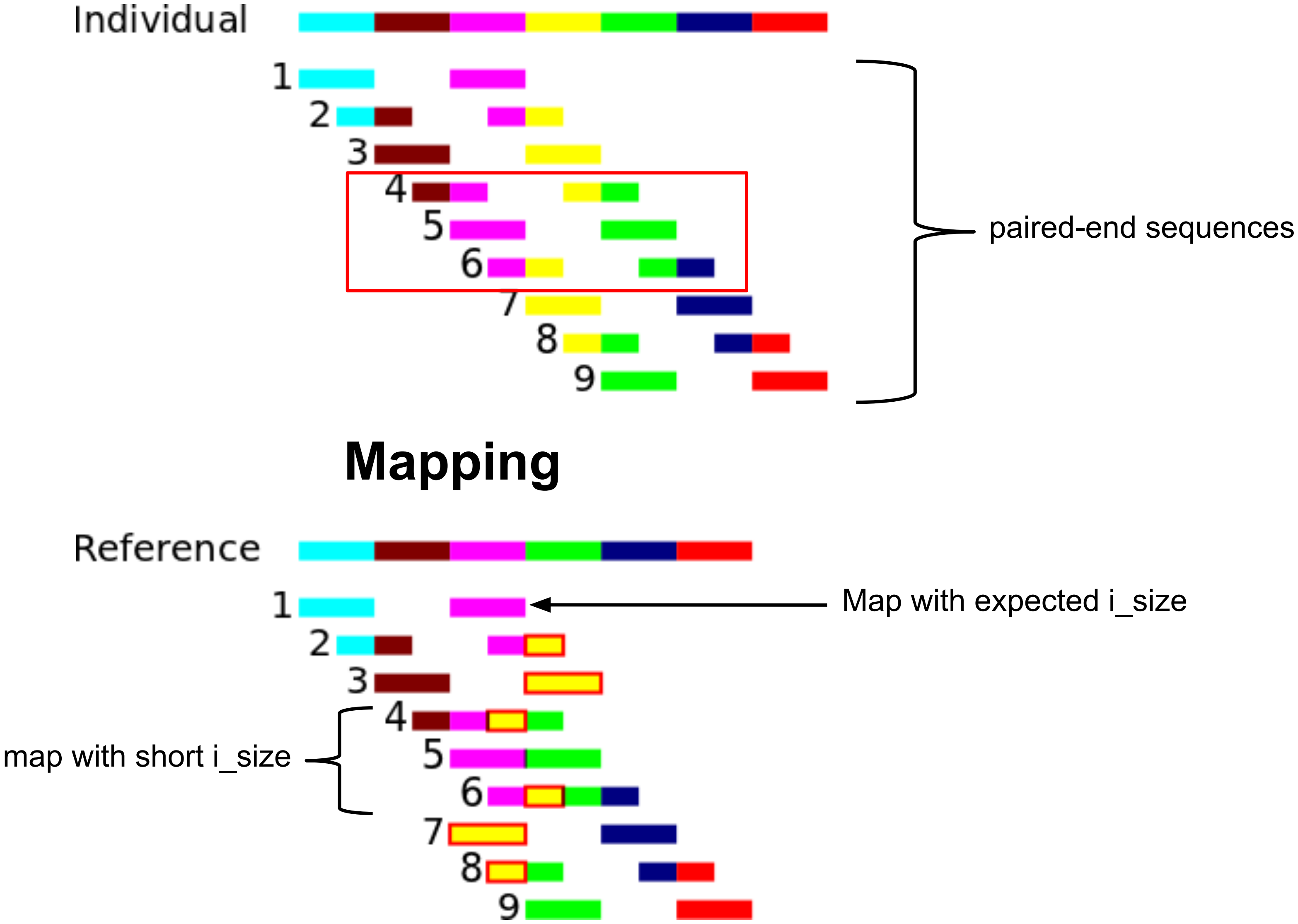
Mosaik Align. filter

coverage and depth

Freebayes

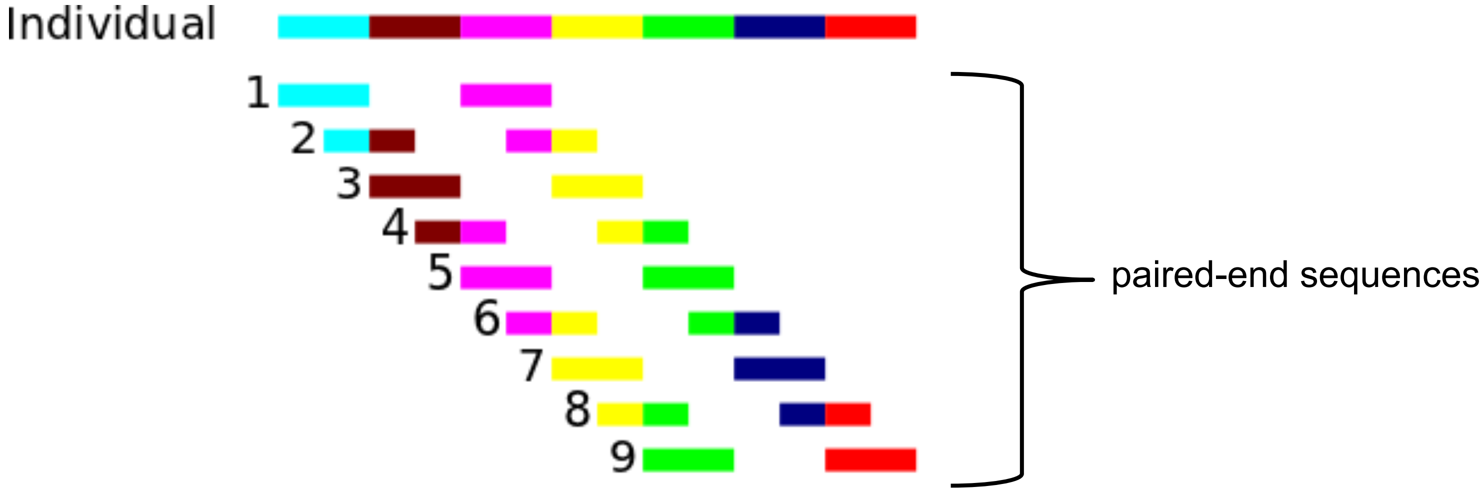
IDfixe

Pretty Freebayes

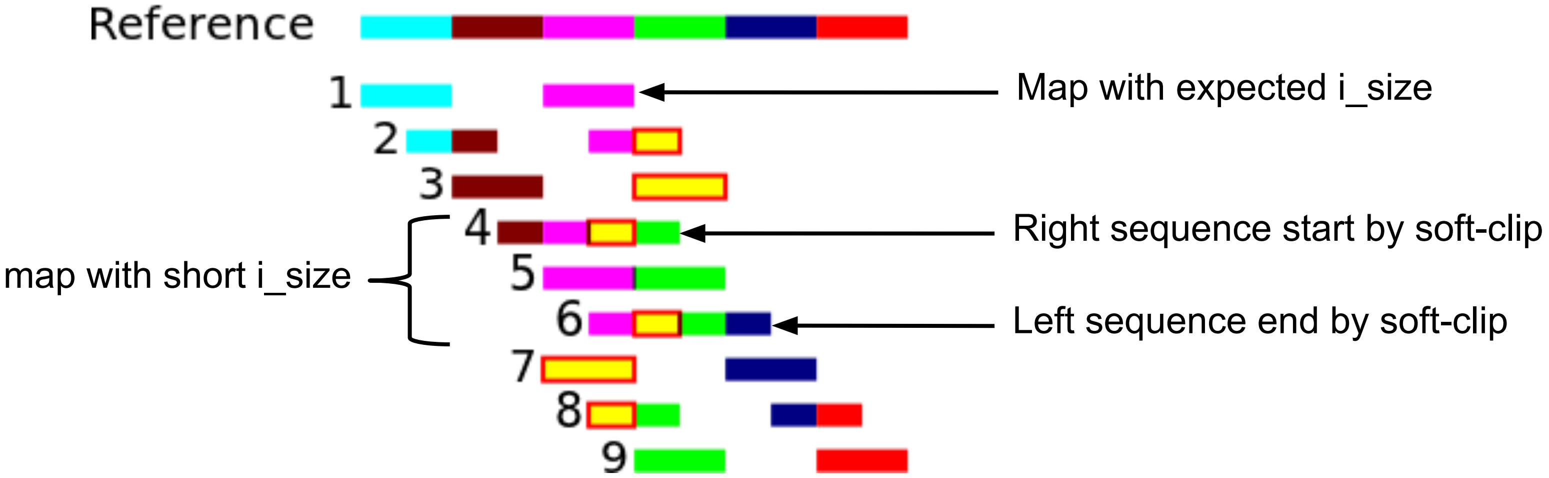


IDfixe - Under the hood

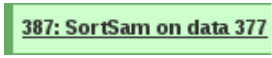
Quality reads filter (FastQ) | Map paired-end with Mosaik | Mosaik Align. filter | coverage and depth | Freebayes | IDfixe | Pretty Freebayes



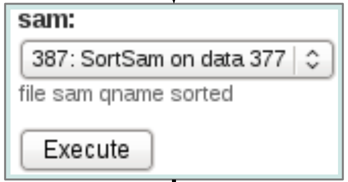
Mapping


Reference
1
2
3
4
5
6
7
8
9
Map with expected i_size
Right sequence start by soft-clip
map with short i_size
Left sequence end by soft-clip

Structural Variation (IDfixe)



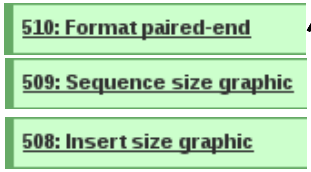
387: SortSam on data 377



sam:

387: SortSam on data 377

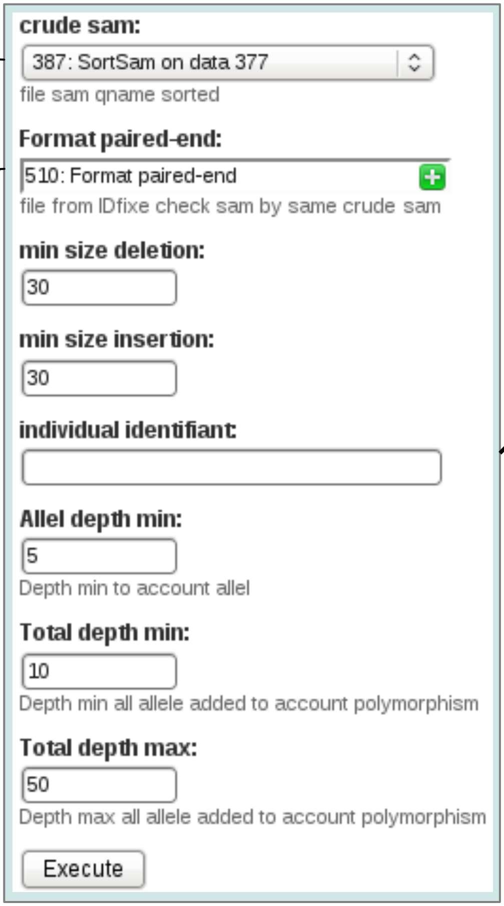
file sam qname sorted



510: Format paired-end

509: Sequence size graphic

508: Insert size graphic



crude sam:

387: SortSam on data 377

file sam qname sorted

Format paired-end:

510: Format paired-end

file from IDfixe check sam by same crude sam

min size deletion:

30

min size insertion:

30

individual identifiant:

Allel depth min:

5

Depth min to account allel

Total depth min:

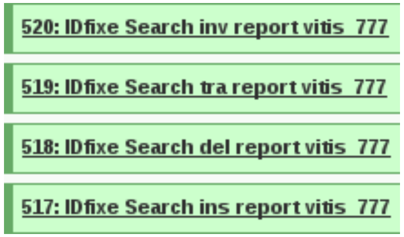
10

Depth min all allele added to account polymorphism

Total depth max:

50

Depth max all allele added to account polymorphism

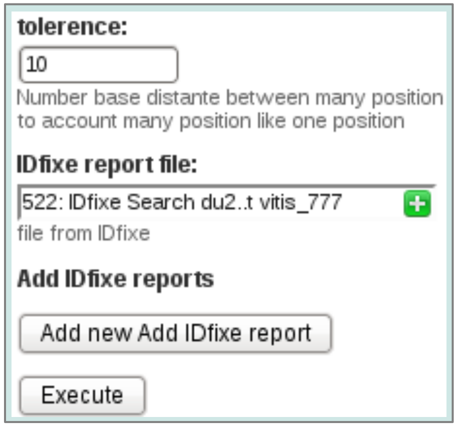


520: IDfixe Search inv report vitis_777

519: IDfixe Search tra report vitis_777

518: IDfixe Search del report vitis_777

517: IDfixe Search ins report vitis_777



tolerence:

10

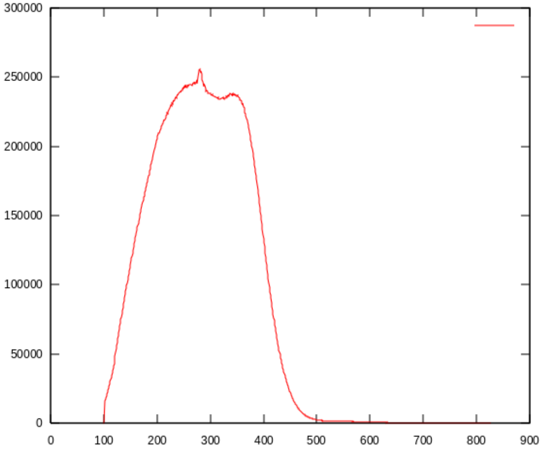
Number base distante between many position to account many position like one position

IDfixe report file:

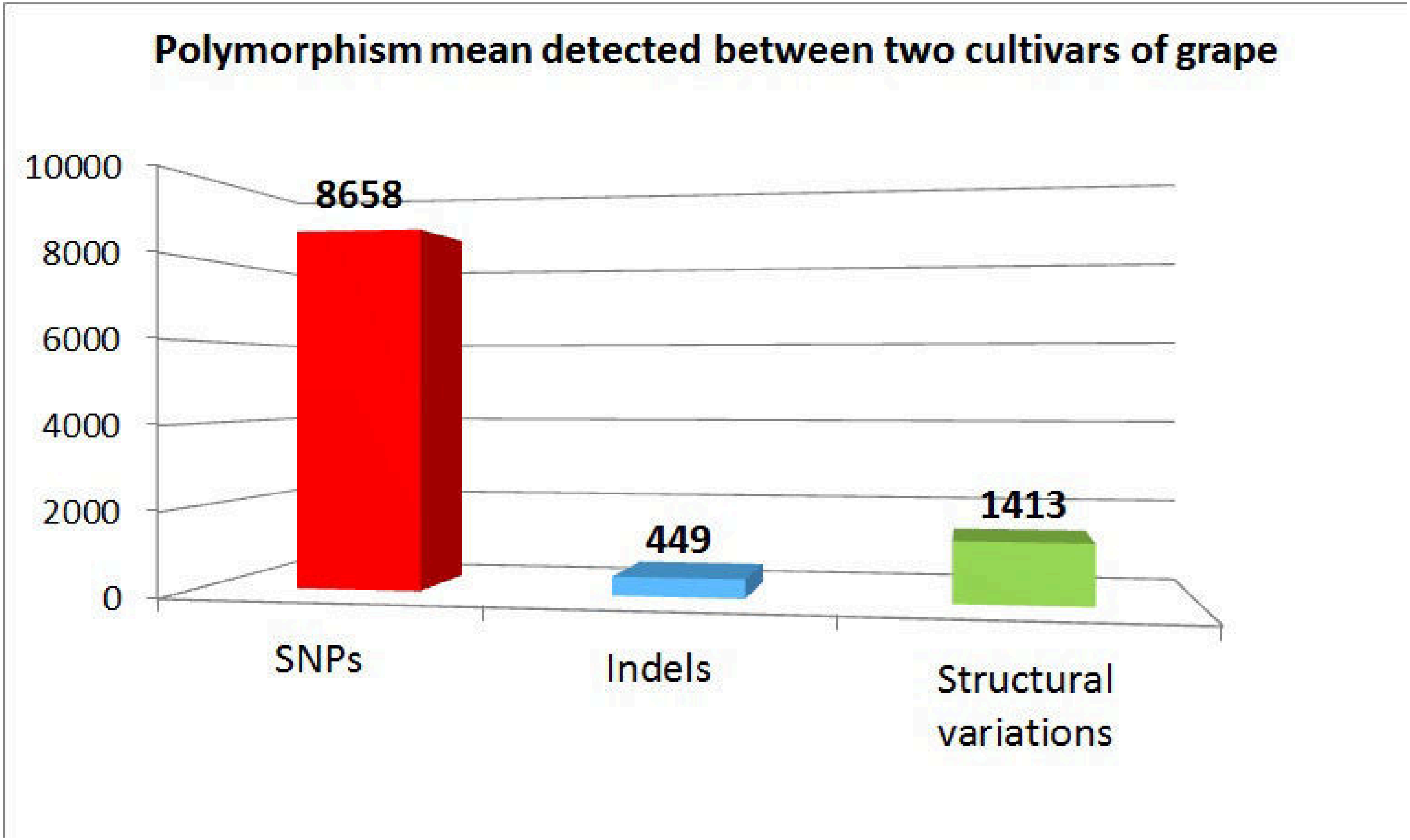
522: IDfixe Search du2..t vitis_777

file from IDfixe

Add IDfixe reports



conclusion



Carrier *et all.*, in prep

Conclusions

Highlights

Statistics:

125 cluster users, and 120 Galaxy users (30% IRD, 40% CIRAD, 20% INRA, 10% others)
Cluster: 1 100 000 jobs / month
Client Galaxy: 2 200 jobs / month, 80 added tools

Training courses:

6 training courses organised since 2 years, +150 researchers and students ;
Sequence analysis (NGS, comparative genomic, annotation...), and Galaxy usage ;
public and private trained people. France, brazil, colombia...

ISO 9001 certification in progress:

mock audit in september 2012 ;
certification audit in december 2012.

South Green©
bioinformatics platform

Thanks for your attention…
Questions ?