SouthGreen bioinformatics platform

TOGGLe

Tools for Generic NGS analysis

A framework to quickly build pipelines and to perform large-scale NGS analysis

toggle@ird.fr

Bioversity International · cirad · INRA SCIENCE & IMPACT · IRD Institut de Recherche pour le Développement FRANCE · Montpellier SupAgro

TOGGLE

- A toolbox to perform large-scale NGS analyses

19 modules, 88 functions
40 open-source tools

https://github.com/SouthGreenPlatform/TOGGLE

SouthGreen bioinformatics platform

Data preprocessing

Fastqc, Cutadapt
FastxTrimmer
Stack process_radstats

Structural Variations

MindTheGap,
BreakDancer, Pindel

RNA-seq Assembly

TransAbyss, Trinity
TGI-CL

TOGGLE

ReadCount

Htseq-count

Mapping

Bwa aln, sampe/ samse
Bwa mem
Tophat2

SNP calling/ filtering

SAMtools, GATK,VarScan,SNPEff

SAM/BAM management

picardTools, SAMtools, GATK

https://github.com/SouthGreenPlatform/TOGGLE

How to perform an analysis with TOGGLE ?

A command-line based pipeline framework

A single command line

toggleGenerator.pl -d DIR-c FILE -o DIR

SouthGreen bioinformatics platform

What does TOGGLE need to run ?

- An input directory (with fastq, sam/bam, vcf files)
- The name of output directory used to store the data generated by the analyses
- A unique and simple configuration file to design the pipeline and define software parameters.
- Optional arguments : reference file, annotation…

A simple configuration file

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

$cutadapt
-q 30
-m 35

$bwa mem
-n 5
…

$sge
-q bioinfo.q
-b Y

A simple configuration file

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

Create your own workflow

- The workflow order
- The list of softwares to run

One line = the step followed by the software's name

A simple configuration file

Create your own workflow

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

Step number < 1000

Step number >= 1000

A simple configuration file

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

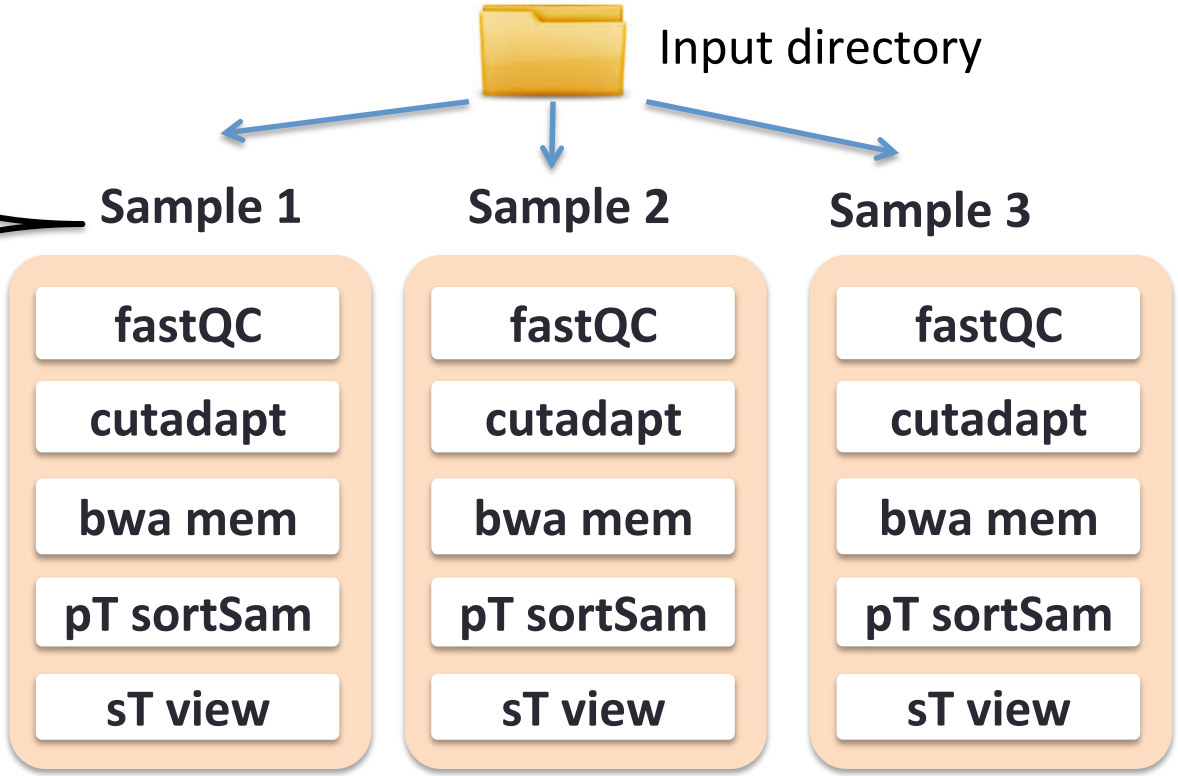
Create your own workflow

Step number < 1000

Parallel analysis by sample

Create your workflow

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration



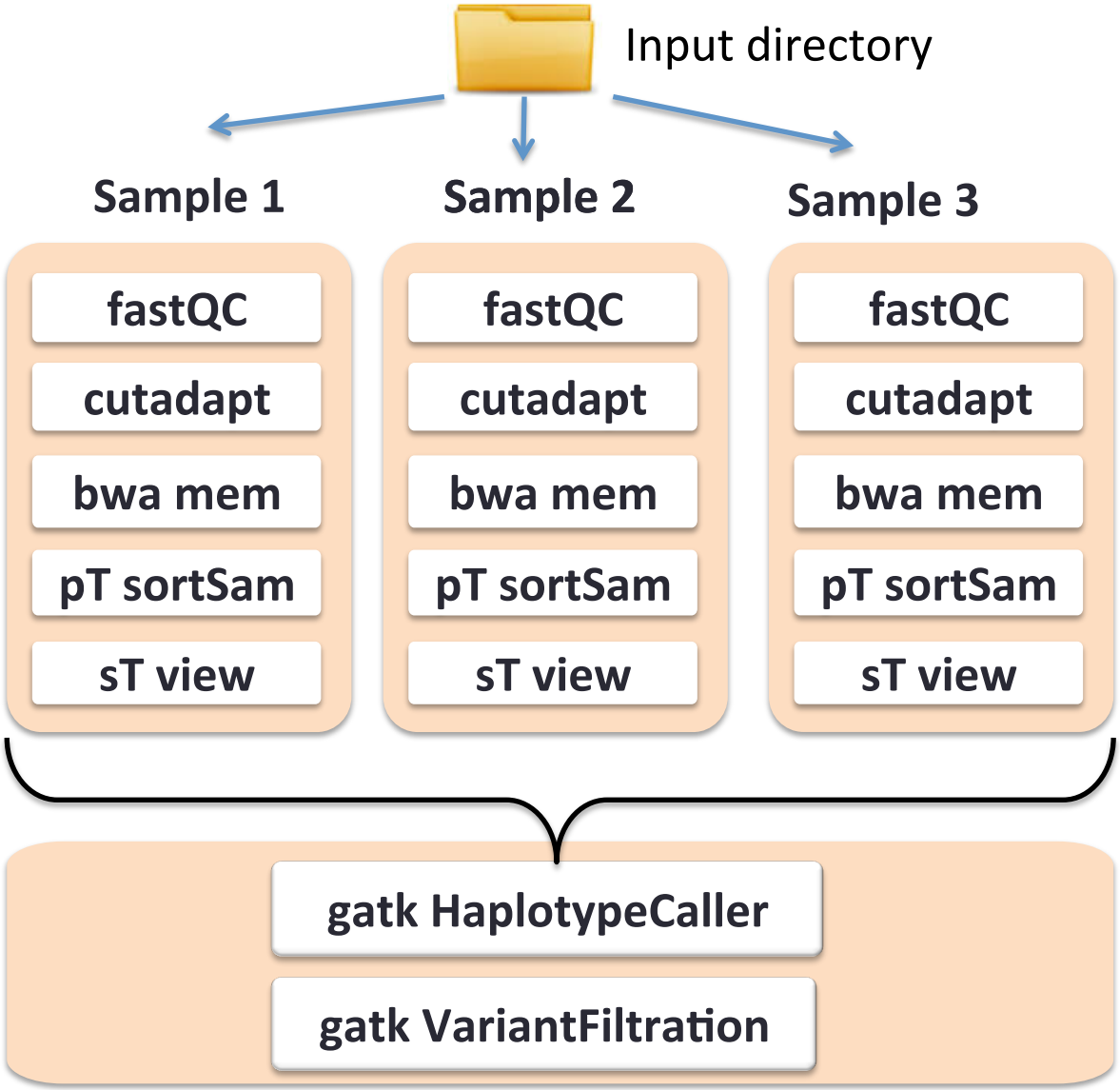
TOGGLE

Create your workflow

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

Step number >= 1000

Global analysis (all samples)



A simple configuration file

$order
1=fastqc
2=cutadapt
3=bwa mem
4=picardToolsSortSam
5=samToolsView
1000=gatkHaplotypeCaller
1001=gatkVariantFiltration

$cutadapt
-q 30
-m 35

$bwa mem
-n 5
…

$sge
-q bioinfo.q
-b Y

Software parameters

One tag per software ($softwareName) followed by the list of options

TOGGLe 's team

IRD Institut de Recherche pour le Développement FRANCE

UMR DIADE

Christine Tranchant-Dubreuil

François Sabot

Cécile Monat

Mawussé Agbessi

Souhila Amazougarene

Abdoulaye Diallo

Laura Helou

Ayité Kougbeadjo

cirad

UMR BGPI

Sébastien Ravel

ADN Agence d'information génétique

Julie Orjuela-Bouniol

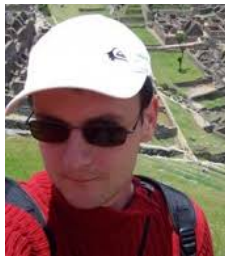
cirad

INRA SCIENCE & IMPACT

Montpellier SupAgro

UMR AGAP

Gautier Sarah



Cédric Farcy

Enrique Ortega-Abboud

Maryline Summo

SouthGreen bioinformatics platform

Thank you
for using TOGGLE !

RESEARCH

TOGGLE: Toolbox for generic NGS analyses

Cecile Monat[1*†], Christine Tranchant-Dubreuil[1†^], Ayité Kougbeadjo[2], Cédric Farcy[2], Mawussé Agbessi[1], Maryline Summo[2] and Francois Sabot[1^]

https://github.com/SouthGreenPlatform/TOGGLE