TOGGLE-3: a tool for on the fly pipelines creation and performing robust large-scale NGS analyses
Sébastien Ravel, Christine Tranchan-Dubreuil, Cécile Monat, Gautier Sarah, Julie Orjuela-Bouniol, Francois Sabot

To cite this version:
Sébastien Ravel, Christine Tranchan-Dubreuil, Cécile Monat, Gautier Sarah, Julie Orjuela-Bouniol, et al.. TOGGLE-3: a tool for on the fly pipelines creation and performing robust large-scale NGS analyses. JOBIM 2016: Journées ouvertes de biologie informatique et mathématiques, Jun 2016, Lyon, France. 437 p., 2016, JOBIM 2016. hal-02738889

HAL Id: hal-02738889
https://hal.inrae.fr/hal-02738889
Submitted on 2 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
TOGGLE-3: a tool for on the fly pipelines creation and performing robust large-scale NGS analyses

Sébastien Ravel ¹ †, Christine Tranchan-Dubreuil ², Cécile Monat², Gautier Sarah ³, Julie Orjuela-Bouniol ⁴, François Sabot § ²

Web site:
https://github.com/SouthGreenPlatform/TOGGLE

Dear biologist, have you ever dreamed of using the whole power of those numerous NGS tools that your bioinformatician colleagues use through this awful list of command lines ?

Dear bioinformatician, have you ever wished for a really quick way of designing a new NGS pipeline without having to retype again dozens of code lines to readapt your scripts or starting from scratch ?

So, be happy ! TOGGLE is for you !

With TOGGLE (TOolbox for Generic nGs anaLysEs), you can create your own pipeline through an easy and user-friendly approach. Indeed, TOGGLE integrate a large set of NGS softwares and utilities to easily design pipelines able to handle hundreds of samples. The pipelines can start from Fastq (plain or compressed), SAM, BAM or VCF (plain or compressed) files, with parallel (by sample) and global analyses (by multi samples).

Moreover, TOGGLE offers an easy way to configure your pipeline with a single configuration file:
— organizing the different steps of workflow,
— setting the parameters for the different softwares,
— managing storage space through compressing/deleting intermediate data,
— determining the way the jobs are managed (serial or parallel jobs through scheduler SGE and SLURM).

TOGGLE can work on your laptop, on a single machine server as well as on a HPC system, as a local instance or in a Docker machine. The only limit will be your available space on the storage system, not the amount of samples to be treated or the number of steps.

* Intervenant
† Corresponding author: sebastien.ravel@cirad.fr
‡ Corresponding author: christine.tranchant@ird.fr
§ Corresponding author: francois.sabot@ird.fr
TOGGLE was used on different organisms, from a single sample to more than one hundred at a time, in RNAseq, DNAreseq/SNP discovery and GBS analyses.

List of bioinformatics tools included:
- BWA : bwaAln, bwaSampe, bwaSamse, bwaIndex, bwaMem
- PicardTools : picardToolsMarkDuplicates, picardToolsCreateSequenceDictionary, picardToolsSortSam, picardToolsAddOrReplaceReadGroup, picardToolsValidateSamFile, picardToolsCleanSam, picardToolsSamFormatConverter
- GATK : gatkBaseRecalibrator, gatkRealignerTargetCreator, gatkIndelRealigner, gatkHaplotypeCaller, gatkSelectVariants, gatkVariantFiltration, gatkReadBackedPhasing, gatkUnifiedGenotyper, gatkBaseRecalibrator, gatkPrintReads
- Fastqc : fastqc
- FastxToolkit : fastxTrimmer
- Tophat : bowtie-build, bowtie2-build, tophat2
- Snpeff : snpeffAnnotation
- Cutadapt : cutadapt
- Graphviz v2.xx (optional)

References


Mots clés : NGS, Toolbox, On the fly pipeline, Flexible, RNASeq, GBS, SNP, Scheduler