



SBG-CGC: web interface

Tiffany Delhomme

IARC course - analysing TCGA data in the cloud

28 feb 2018

Summary

1 Theory

- proposed protocol

2 Application

- getting data
- ex 1 - BAM header
- ex 2 - deconstructSig
- ex 3 - most mutated genes from MAF

web interface to run analyses

proposed protocol

1. create a project



web interface to run analyses

proposed protocol

1. create a project



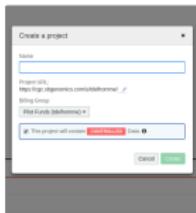
2. query data and add it



web interface to run analyses

proposed protocol

1. create a project



2. query data and add it



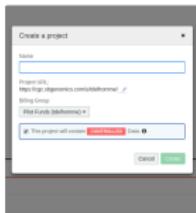
3. create a tool/workflow



web interface to run analyses

proposed protocol

1. create a project



2. query data and add it



3. create a tool/workflow



4. run the analysis



getting data

DEMO

Example 1

BAM headers

Example

extract the header of BAM files

- Software: samtools
- DockerFile: biocontainers/samtools
- Data: Uterine Carcinosarcoma tumor BAM samples
- Command Line: samtools view -H input.bam > fixed-result-name.header

Example 2

mutation signatures

Example

run the existing deconstructSig pipeline on one VCF file.

- Software: deconstructSig 1.8.0
- DockerFile: biocontainers/samtools
- Data: one VCF file
- use default parameters

Example 3

mutated genes from MAF

Example

output the **n** most mutated genes

- DockerFile: ubuntu:latest
- Data: choose your favorite TCGA case
- Script:

```
zcat $1 | grep -v "^#" | awk '{print $1}' |
    sort | uniq -c | sort -k1nr | head -n$2
```

- Inputs: MAF file, number of outputed lines

Example 3bis

mutated genes from MAF

Example

output the **n** most mutated genes

- DockerFile: ubuntu:latest
- Data: choose your favorite TCGA cohort
- **Modify the tool to run it on batch of files**
 - add some MAF files if needed