



## SBG API scripts for running RNA-Seq Alignment - STAR workflow on CGC environment

### Scripts

- star.py
- config.yaml
- requirements.txt
- requirements3.txt

### Instructions for using scripts

Before the scripts can be used, there are python packages which need to be installed. These are listed in [requirement.txt](#) (and [requirements3.txt](#) in case Python 3 is used). To install the requirements following command should be invoked:

```
$ pip install -r requirements.txt  
($pip install -r requirements3.txt for Python 3)
```

Basic information about the targeted project, workflow, inputs and similar, can be adjusted in [config.yaml](#) file. The scripts are written so that the logs from the standard output are preserved and the path to the file with this information can be also changed through [config.yaml](#) (by default it's saved in *star.log* file in the current directory). Inputs are configured assuming usage of RNA-Seq Alignment - STAR workflow and can be changed in case of using other pipelines. **However, users should be aware that switching to another workflow may require understanding and modifying of the *star.py* script along with configuration file.**

# SevenBridges

101 MAIN ST, 1st FLOOR  
CAMBRIDGE, MA 02142



There are three major steps (script modes) that can be used throughout the running process:

1. `$ python star.py create`

This command should be invoked just once at the beginning and it will create task drafts for all input files.

2. `$ python star.py run`

This command starts the execution of at maximum 200 tasks. In case there are 200 task already running, the command won't do anything. Otherwise, if there are less than 200 running tasks, the command will start that many tasks, so the total number of running ones is 200.

3. `$ python star.py status`

Invoking status command gives basic information about the task status for each task in the project. It may be useful to check how many tasks completed the execution, so the new runs can be triggered.