Instructions

*** This instruction applies to SEQC v0.2.5*

Prerequisite

Docker

Install Docker version 2 (Engine version 18+). You need at least macOS Sierra 10.12 or newer macOS such as Mojave.

Python

Have Python 3 on your computer.

AWS Credentials

Configure AWS credentials:

```
$ aws configure
```

Ensure the `.aws` directory (which contains your AWS credentials and configuration) is located at your home directory (e.g. `/home/john/.aws`)

Make sure your EC2 key pair file (*.pem) is NOT accessible by others. You can do this by running this command:

```
$ chmod 400 /path/my-key.pem
```

How to Install

Note that the steps described here are only tested on Mac.

Run the following commands from your Bash terminal:

```
mkdir -p seqc-0.2.5
tar xzf seqc-0.2.5.tar.gz -C seqc-0.2.5
cd seqc-0.2.5
```

If you run `tree`, you should see something like this:

```
$ tree
.
├── config
│   └── jobs.template.yml
├── instructions.md
├── seqc-progress.sh
├── seqc-submit.sh
└── seqc_submit_mjobs.py

1 directory, 5 files
```

How to Submit Multiple Jobs (Multiple Samples)

Input Configuration

Jump start by duplicating the template:

```
$ cp config/jobs.template.yml config/jobs.yml
```

Edit `jobs.yml`:

```
$ nano config/jobs.yml
```

```
jobs:
  - job: 1
    ami-id: ${PLACE_AMI_ID_HERE}
    platform: ten_x_v2
    user-tags:
      Job: 1
      Project: 10178
      Sample: my-pbmc1
    index: s3://.../genomes/hg38_long_polya/
    barcode-files: s3://.../barcodes/ten_x_v2/flat/
    genomic-fastq: s3://.../pbmc1/genomic/
    barcode-fastq: s3://.../pbmc1/barcode/
    upload-prefix: s3://.../pbmc1/seqc-results/
    output-prefix: pbmc1
    email: chunj@mskcc.org
    star-args: "runRNGseed=0"
  - job: 2
    ami-id: ${PLACE_AMI_ID_HERE}
    platform: ten_x_v2
```

```
    user-tags:
      Job: 2
      Project: 10178
      Sample: my-pbmc2
    index: s3://.../genomes/hg38_long_polya/
    barcode-files: s3://.../barcodes/ten_x_v2/flat/
    genomic-fastq: s3://.../pbmc2/genomic/
    barcode-fastq: s3://.../pbmc2/barcode/
    upload-prefix: s3://.../pbmc2/seqc-results/
    output-prefix: pbmc2
    email: chunj@mskcc.org
    star-args: "runRNGseed=0"
```

If you want to specify any of the SEQC parameters, you can add a new line to the job description using the same format. For example, to specify `--min-poly-t=0` and `--no-filter-low-coverage`, add the following two lines:

```
min-poly-t: "0"
no-filter-low-coverage: ""
```

Job Submission

```
$ python seqc_submit_mjobs.py --help
usage: seqc_submit_mjobs.py [-h] --config PATH_YAML_INPUT --pem
                            PATH_EC2_KEYPAIR [--key-name EC2_KEYPAIR_NAME]
                            [--dry-run]

optional arguments:
  -h, --help            show this help message and exit
  --config PATH_YAML_INPUT, -c PATH_YAML_INPUT
                        path to jobs.yaml
  --pem PATH_EC2_KEYPAIR, -k PATH_EC2_KEYPAIR
                        path to AWS EC key pair file (*.pem)
  --key-name EC2_KEYPAIR_NAME, -n EC2_KEYPAIR_NAME
                        the name of your AWS EC2 key pair
  --dry-run             Dry run (i.e. don't actually submit the job)
```

```
$ python seqc_submit_mjobs.py \
    --pem ~/dpeerlab-chunj.pem \
    --config config/jobs.yml
```

```
./logs/jobs.001.log

SEQC: 2019-04-24 18:03:26: Created new security group: sg-0bea8fd60b2706360 (name=SEQC-72
```

```
SEQC: 2019-04-24 18:03:27: Enabled ssh access via port 22 for security group sg-0bea8fd60
SEQC: 2019-04-24 18:03:28: instance i-0684839987b018f94 created, waiting until running
SEQC: 2019-04-24 18:03:44: instance i-0684839987b018f94 in running state
SEQC: 2019-04-24 18:03:44: connecting to instance i-0684839987b018f94 via ssh
SEQC: 2019-04-24 18:04:37: Formatting and mounting /dev/xvdf to /home/ec2-user
SEQC: 2019-04-24 18:04:40: Successfully mounted new volume onto /home/ec2-user.
SEQC: 2019-04-24 18:04:40: setting aws credentials.
SEQC: 2019-04-24 18:06:12: SEQC setup complete.
SEQC: 2019-04-24 18:06:12: instance login: ssh -i <path to your key file> ec2-user@18.232
SEQC: 2019-04-24 18:06:12: connecting to instance i-0684839987b018f94 via ssh
--
./logs/jobs.002.log

SEQC: 2019-04-24 18:06:24: Created new security group: sg-0dc16ec0bfe3c83cf (name=SEQC-28
SEQC: 2019-04-24 18:06:25: Enabled ssh access via port 22 for security group sg-0dc16ec0b
SEQC: 2019-04-24 18:06:26: instance i-0fb1b1a8ca8f9451e created, waiting until running
SEQC: 2019-04-24 18:06:42: instance i-0fb1b1a8ca8f9451e in running state
SEQC: 2019-04-24 18:06:42: connecting to instance i-0fb1b1a8ca8f9451e via ssh
SEQC: 2019-04-24 18:07:39: Formatting and mounting /dev/xvdf to /home/ec2-user
SEQC: 2019-04-24 18:07:42: Successfully mounted new volume onto /home/ec2-user.
SEQC: 2019-04-24 18:07:42: setting aws credentials.
SEQC: 2019-04-24 18:09:11: SEQC setup complete.
SEQC: 2019-04-24 18:09:11: instance login: ssh -i <path to your key file> ec2-user@3.81.4
SEQC: 2019-04-24 18:09:11: connecting to instance i-0fb1b1a8ca8f9451e via ssh
```