Cogent™ NGS Analysis Pipeline v1.5 Quick Start Guide

The following information is provided as a high-level introduction to the software, also referred to as CogentAP. For more detailed information, please see the Cogent NGS Analysis Pipeline v1.5 User Manual.

Before you begin

A. Supported operating systems
- Linux: CentOS 6.9 or higher, RedHat 7.5 or higher, Ubuntu 17 or higher

B. Hardware requirements
- CPU: 24 cores
- Memory: 64 GB RAM
- Free disk space: at least 500 GB

C. Additional dependencies
- Internet connectivity on the server
- Miniconda3 4.10.2 or higher
- Bash UNIX shell

D. Required input files
- FASTQ files generated by an Illumina® sequencing platform on results from:
  o ICELL8® cx Single-Cell System applications
  o ICELL8 Single-Cell System applications
  o SMARTer® Stranded Total RNA-Seq Kit v3 - Pico Input Mammalian
- An ICELL8 well-list text file, Illumina sample sheet, or similar TDT/CSV format file

Confirm Miniconda3 version

NOTE: For more information on any command in this section, refer to Section IV.A of the User Manual.

1. Verify Miniconda3 is installed and meets or exceeds the required version by typing the following into a terminal window:

   `conda -V`

   If conda is successfully installed, it should return text with the version number.
   Example:

   `conda 4.10.3`

2. Verify the base conda environment can be activated by typing:

   `conda activate`

   If conda is successfully installed, the terminal will look similar to this:

   ```
   $ conda activate
   (base) $
   ```

Type the following command to return to the default Linux prompt.

`conda deactivate`

```
(base) $ conda dectivate
$
```

3. Verify the install location of Miniconda3 is configured in the file `.bash_profile`
   a. For an individual user account, type:

   `more ~/.bash_profile`

   b. Confirm something similar to the following is showing in the file (all on one line):

   ```
   export
   PATH="/home/<USERNAME>/miniconda3/
   bin:$PATH"
   ```

   where `<USERNAME>` is replaced by the username of the account that installed conda,

   If no `.bash_profile` file exists or the line isn't displaying, it will need to be manually created and populated.

Installation

NOTE: For more information on any step in this section, refer to Section IV.B of the User Manual.

1. Sign up to download the installation script from our website. You will receive an email automatically after signing up that will contain a password you will need in Step 4.

   - If a previous version of CogentAP or mappa™ software was previously installed in the same account (or computer) being used for this installation, proceed to Step 2
   - If this is a new install, skip to Step 3

2. (Uninstall the previous version)
   a. If you want to save any files stored in the directory (`CogentAP/` or `mappa/`) of the previous installation, make sure to move them to another folder location before continuing.
   b. Delete the previous installation directory and all subfolders.

(continued on page 2)

3. (Install the new version): Move or copy the installation script downloaded from Step 1 onto the Linux server into the directory location where you want to install.

4. From the same directory location in Step 3, run the command:

```
bash CogentAP_v1.5_Linux64_installer.sh
<AUTHCODE>
```

`<AUTHCODE>` will be replaced by the authorization code included in the sign-up confirmation email.

5. Move into the installed directory, then run the command below to install human genome.

```
bash CogentAP_setup.sh genome_install
hg38
```

NOTE: If other genomes need to be loaded, please refer to Section V.E of the User Manual.

Generate raw FASTQ files

NOTE: Section V.A of the User Manual.

1. Log in to a server that stores the run folder from Illumina sequencing and has the bcl2fastq program installed.

2. Change to a working folder where you want the raw FASTQ files to be located after being generated.

3. Run bcl2fastq with the following syntax template:

```
bcl2fastq -R <RUN_FOLDER> -o <RUN_ID>
--no-lane-splitting --sample-sheet
$COGENT_AP_HOME/config/SampleSheet_dum
my.csv > <RUN_ID>.stdout 2 >
<RUN_ID>.stderr
```

where:

- `<RUN_FOLDER>` is the path to the sequencing run folder

- `<RUN_ID>` is the ID number automatically generated by the Illumina sequencer

The file `SampleSheet_dummy.csv` is stored in the CogentAP `config` folder

NOTE: See Section IV.C of the User Manual for how to set up the `$COGENT_AP_HOME` variable.

4. Move the raw FASTQ files to your preferred storage location. They are typically generated in the `<RUN_ID>` folder and named similar to:

```
Undetermined_S0_R1_001.fastq.gz
Undetermined_S0_R2_001.fastq.gz
```

To run the Cogent NGS Analysis Pipeline

UI:

NOTE: Section V.B of the User Manual.

1. Double-click on the executable (`CogentAP_launcher`) to run it.

2. Fill out the fields in the interface.

3. Click [Start].

Command line:

NOTE: Section V.C of the User Manual.

- To demultiplex (demux)

```
%COGENT_AP_HOME%/cogent demux \
    -i <FASTQ_R1> \
    -p <FASTQ_R2> \
    -b <WELL-LIST> \
    -t <EXP_TYPE> \
    -o <OUTPUT>
```

- To analyze

```
%COGENT_AP_HOME%/cogent analyze \
    -i <DEMUX_RESULT_DIR> \
    -g <GENOME> \
    -t <EXP_TYPE> \
    -o <OUTPUT>
```

where:

- `%COGENT_AP_HOME%` is the path to the directory where CogentAP is installed

- `<FASTQ_R1>` and `<FASTQ_R2>` are the full paths to the FASTQ files generated by an Illumina sequencing platform.

- `<WELL-LIST>` is the full path to the ICELL8 system WellList, Illumina's sample sheet or TDT/CSV format file

- `<EXP_TYPE>` is the experiment type used (e.g., `ICELL8_FLA`, refer to the CogentAP User Manual for more options)

- `<OUTPUT>` is a string; it will be the name of the output folder created by the analysis AND the prefix of all the results files

- `<DEMUX_RESULT_DIR>` is the full path of the demultiplex result directory you specified in `demux` command

- `<GENOME>` is a name of genome build (e.g., `hg38`)

Contact Us	
Customer Service/Ordering	**Technical Support**
tel: 800.662.2566 (toll-free)	tel: 800.662.2566 (toll-free)
fax: 800.424.1350 (toll-free)	fax: 800.424.1350 (toll-free)
web: takarabio.com/service	web: takarabio.com/support
e-mail: ordersUS@takarabio.com	e-mail: technical_support@takarabio.com