DUESSELPORE Webserver manual

1. Install and configure webserver

1.1. System requirement

Minimum requirement

```
* CPU: 2.0 GHz (64bits) 4 cores or higher
* Memory: 8 GB or higher
* Diskdrive: 200 GB free space
* Window 10, Linux (Ubuntu) or Mac
```

Recommend configuration

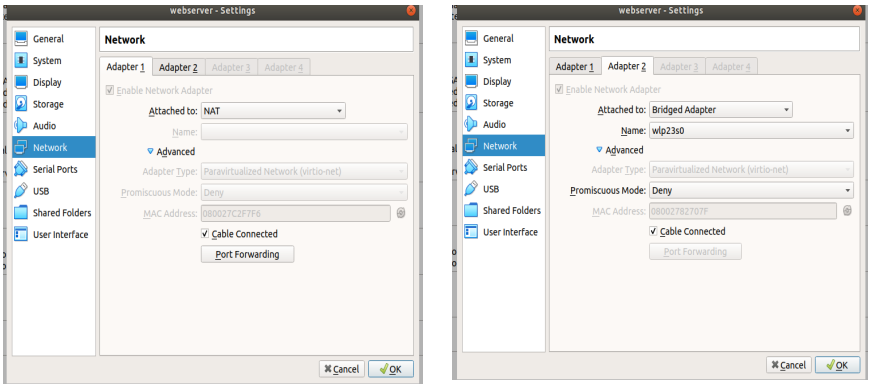
```
* CPU: 3.0 GHz (64bits) 8 cores or higher
* Memory: 16 GB or higher
* Diskdrive: 1000 GB free space
* Window 10, Linux (Ubuntu) or Mac
```

1.2. Installation

1.2.1 Download and install VMWare Note: For inexperienced Linux user our software are tested with current version pipeline. We do not recommend upgrading the version on Linux Virtual machine. The webserver may crash when new software is updated

- Download and install Virtualbox (VB) installation and VirtualBox 6.1.22 Oracle VM VirtualBox Extension Pack from https://www.virtualbox.org/wiki/Downloads. Already tested Virtualbox version 6.1.22 on Ubuntu 18.04 and Window 10.
- Download the webserver.ova image file from this address

After installing VB and its Extension Pack, open VB >File> Import Appliance to select webserver.ova downloaded file then setup configuration based on your machine configuration. By default, our webserver uses 4 cores CPU, 8 GB RAM. We recommend using 8 CPUs, 16 GB RAM, HDD is auto allocated, Therefore when your data increases, the image file size also increases. We recommend deploying VB image in the partition that has at least 200 GB (depend on the number of users and datasize TB volume are highly recommended). Configure the network interface on your host site (your primary OS): Before we start the Virtual machine in Virtual box configuration panel, we configure two network interfaces as in the figure below. The first network interface to internet (NAT) and the second interface to our host machine.



VM Network interface configuration

1.2.2. Login and configure webserver After booting up our guest OS, login your Virtual Machine (VM) with this credential:

```
* user name: ag-rossi (preset)
* password 123456
```

Open the terminal, and we can get our webserver IP address by this command on the guest terminal. When you want to use only the Human genome.

```
$setup_webserver light
$run_server
```

If you want to use RNASeq for other organisms:

```
$setup_webserver full
$run_server
```

The program will download all reference genomes, genome annotation and other required packages. It also sets your IP address into allowed IP list of webserver. IP address is printed out from the printout messages.

NOTE If you cannot access your webserver, please check the IP address. You may have to update your new IP address when your webserver IP address changed. Open /home/ag-rossi/projects/duesselpore/NGS_webserver/settings.py and add your webserver IP address in ALLOW_HOSTS list — ### 2.2.

Using webserver #### 2.2.1. Access webserver Now you can use your webserver within your Local Area Network (LAN) with normal web browser (e.g. Firefox or Google Chrome) http://{Your IP address}:8000/duesselpore.

2.2.2. Data preparation Normal users can upload fastq files as ONE compressed zip file: each subfolder contains several replicas with one experimental condition. NOTE: files and folders' name must contain only alphabetic and numeric characters. Below is an example of data separated in two different conditions 'condition1' and 'condition2'.

```
fastq/(folder)
   condition1 (subfolder)
       condition1_replica1.fastq (single fastq file)
       condition1_replica2.fastq (single fastq file)
   condition2 (subfolder)
        condition2_replica1.fastq (single fastq file)
        condition2_replica2.fastq (single fastq file)
        condition2_replica3.fastq (single fastq file)
```

How to merge multiple fastq files into a single file: On Linux terminal:

```
$ cat /path/to/fastq/files/*.fastq > /your/new/location/output.fastq
```

On Window command prompt (path syntax is different):

```
$ type \path\to\fastq\files\*.fastq> \your\new\location\output.fastq
```

2.2.3. Setup running parameter: First, select one group among your groups as the reference group. Select gene (transcriptome) counting method, then select the differential expression algorithm which you want to analyse. Setting up other parameter of analysis function. There are some optional parameters e.g. ReadCountMinThreshold, Logfold, adjPValueThreshold. Submit and wait for the result. Advance user can customized the RNA.R code to develop a new workflow.

2.2.4. Get the results: After computation completed, all the result are downloaded from browser. We export the interactive html file for some plots.