# **Basestack**

Release 2.0.0

**Brian Merritt** 

Sep 06, 2023

## CONTENTS

1	Getting Started	1	
2	About	3	
3 Contents			
	3.1 Installation	5	
	3.2 Usage		
	3.3 API	103	

CHAPTER

ONE

### **GETTING STARTED**

Installation

Usage

API

#### CHAPTER

### TWO

### ABOUT

#### ••• Basestack 8 Basestack Search for a module. Ξ Basestack 2 What can icie 34 Mit Main Marcha 2.0.0 Basestack do? . CHECK RELEASE ¢ View demo O Ocker is running \$ Version 20.10.8 Å Kernel ŕ٩ Learn to use 5.10.47-linuxkit Basestack like a pro Driver overlav2 Running Containers View Tutorials ₽ Data /var/lib/docker Questions? Contact here Brian Merritt (brian.merritt@jhuapl.edu) MemAvailable (GB) $\sim$ 14.15 Third Party Applications Memory Nextstrain Total Mem (GB) 34.36 Artic Network Using Mem (GB) IGV 13.06 Available Mem (GB) 21.30 Backend Port Change the backend server port to connect to Processor ... 5003 CPU Brand Core™ i9-9880H Cores 16 ¢: Physical Cores ()Manufacturer Intel® Virtualization Support true \$

Get Basestack

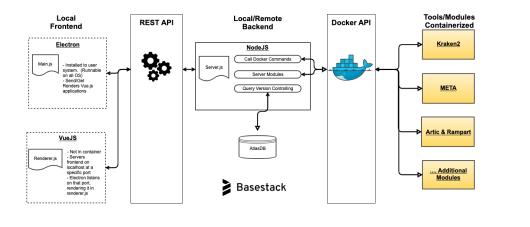
#### Basestack is a platform for rapid and real time analysis of Nanopore generated sequencing data.

Basestack comes bundled with a series of both cutting edge public tools, some modified by JHU built phylogenetic and genomic pipelines, all integrated into a user interface ( that is quickly and seamlessly installable on all standard laptops on any of the major operating systems used today. Ultimately, the suite provides users with a

means of quickly and accurately generating consensus sequences from viral data within minutes of starting. In addition, we provide a system for automated updates once the application is installed for any future changes or additions to the application that is fundamentally built upon the Docker Container ecosystem.

Please follow the sidebar links to review both installation steps for your distribution and additional information on the underlying pipelines.

Basestack's recent releases can be found at: https://github.com/jhuapl-bio/Basestack/releases/





#### CHAPTER

### THREE

### CONTENTS

### 3.1 Installation

Get The Latest Basestack Release

#### 3.1.1 Windows

#### Docker

#### Docker Download + Install

1. Head over to the Docker website to download the necessary package

In order to run docker you must be able to support virtualization from your CPU. This feature must also be enabled within your BIOS and Windows Features.

See more of Virtualization Disabled - Windows for more details

If you're unsure whether docker is supported by your specific cpu, please visit and input your specific model number:

- Intel
- AMD

Type Your Model Number, e.g. T6500 into the product search bar

Intel<sup>®</sup> Core<sup>™</sup>2 Duo Processor T6500 2M Cache, 2.10 GHz, 800 MHz FS8

Specifications	Essentials		Export specifications
Essentials Performance Specifications	Product Collection	Legacy Intel® Core® Processors	
Supplemental Information	Code Name	Products formerly Penryn	
Package Specifications Advanced Technologies	Vertical Segment	Mobile	
Security & Reliability	Processor Number 🕐	T6500	
Product Images	Status	Discontinued	
	Launch Date 👔	Q2'09	
Technical Documentation	Lithography 🕐	45 nm	
	Performance Specifications		
	# of Cores 👔	2	
	# of Threads 🕐	2	
	Processor Base Frequency 🕐	2.10 GHz	
	Cache 🕐	2 MB L2 Cache	
	Bus Speed (?)	800 MHz	
	TDP (?)	35 W	
	Supplemental Information		
	Embedded Options Available (?)	No	
	Datasheet	View now	
	Package Specifications		
	TJUNCTION (2)	105°C	
	Processing Die Size	107 mm <sup>2</sup>	
	# of Processing Die Transistors	410 million	
	Advanced Technologies		
	Intel® Turbo Boost Technology # 🕧	No	
	Intel® Hyper-Threading Technology # 👔	No	
	Intel® Virtualization Technology (VT-x) = 👔	No	
	Intel® 64 = 🕐	Yes	

In this example above, you can see that Vt-x (Virtualization) is not supported. This will be a **Yes** if it is supported. To find the cpu model on Windows:

Add to Compare

The instrument							
All Apps Documents Settings	Photos	More 🗸	Feedback ····		_		Î
Best match				10		_	
					Value		
System Information	$\rightarrow$				Microsoft Windows 10 Enterprise		
Desktop app				Version	10.0.17763 Build 17763		
Apps		· · · · · · · · · · · · · · · · · · ·			Not Available		
					Microsoft Corporation		
System Configuration	>	System Information		· · · · · · · · · · · · · · · · · · ·	MERRIBB1-WL1		
Control Panel	<ul> <li></li> </ul>	Desktop app			Dell Inc.		
Control Panel	/				Latitude 5591		
Intel® Graphics Control Panel	>				x64-based PC		
-		⊏" Open			0819		
NVIDIA Control Panel	>	Run as administrator			Intel(R) Core(TM) i7-8850H CPU @ 2.60GHz, 2592 Mhz, 6 C		
Settings					Dell Inc. 1.2.5, 6/22/2018 3.1		
Settings	/	Run as different user		Embedded Controller V			
Settings (7+)		Open file location	BIOS Mode UEFI				
Folders (12+)		-⊐ Pin to Start		BaseBoard Manufacturer			
Polders (12+)		Ha Pin to Start			07TTKR		
-🗁 Pin		-🛱 Pin to taskbar			X01		
					Mobile		
					On		
				PCR7 Configuration	Elevation Required to View		
				Windows Directory	C:\Windows		
					C:\Windows\system32		
				Boot Device	\Device\HarddiskVolume1		
					United States		
System Information					Version = "10.0.17763.1911"		

### Install Docker Desktop on Windows

Estimated reading time: 6 minutes

Docker Desktop for Windows is the Community version of Docker for Microsoft Windows. You can download Docker Desktop for Windows from Docker Hub.

Download from Docker Hub

By downloading Docker Desktop, you agree to the terms of the Docker Software End User License Agreement and the Docker Data Processing Agreement.

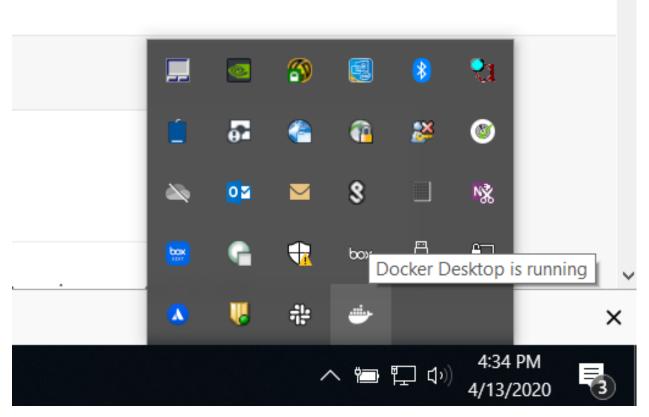
#### 2. Choose Get Docker

Explore	Docker Desktop for Windows		Using 1 of 1 private repositories. <u>Get more</u>
Edition	By Docker	Ktop for Windows way to get started with Docker on Windows	Get Docker Desktop for Windows Docker Desktop for Windows is available for free. Requires Microsoft Windows 10 Professional or Enterprise 64-bit. For previous versions get Docker Toolbox. By downloading this, you agree to the terms of the <u>Docker</u> Software End User License Agreement and the <u>Docker Data</u> <u>Processing Agreement (DPA)</u> .
Descri	ription Reviews Resour	rces	

3. Choose Save File from the prompt

Opening Docker Desktop Installer.exe		×
You have chosen to open:		
Docker Desktop Installer.exe		
which is: EXE file (916 MB)		
from: https://download.docker.com		
Would you like to save this file?		
	Save File	Cancel

4. Once you've installed docker for Windows, you can start it at the **Quick Launch** by search **Docker**. You can also view it on your right-hand-bottom tray by right-clicking



5. Here Docker provides a GUI environment to manage your system. You can allocate or limit resources to your containers as well as set networking settings if you'd like. We use default values for our app

			×
Settings			$\times$
			Â
-	🗄 General	Resources Advanced	
12	Resources	CPUs: 2	
4	ADVANCED     FILE SHARING     PROXIES     NETWORK      Docker Engine	Memory: 2.00 GB	
>. (	Command Line	Swap: 1 GB Disk image size: 64 GB (16.8 GB used)	•
Docker runni	ing	Cancel Apply & Restar	t

6. **OPTIONAL** Choose Local drives to share with containers. Useful if you're storing data on an external drive.

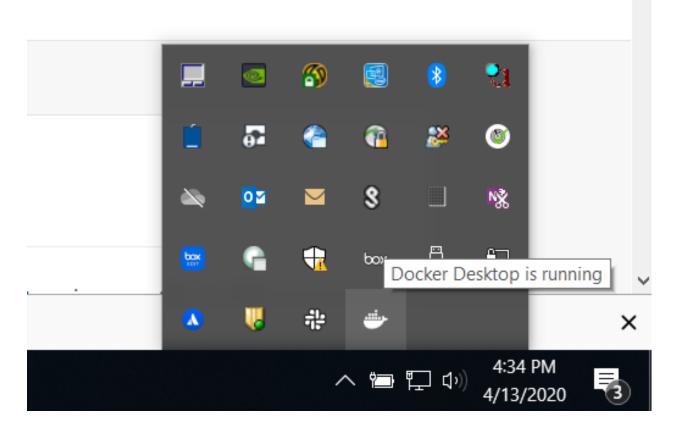
		🐡 docker	🔅 🎄 🔒 bmerritt1762 🦷	
Settings				×
	- contrain	<b>Resources</b> File sharing These drives (and their subfolders) containers. You can check the <u>doc</u>	) can be bind mounted into Docker <u>umentation</u> for more details.	
	ADVANCED • FILE SHARING PROXIES NETWORK	Select the local drives you want to C D	be available to your containers.	
>.				

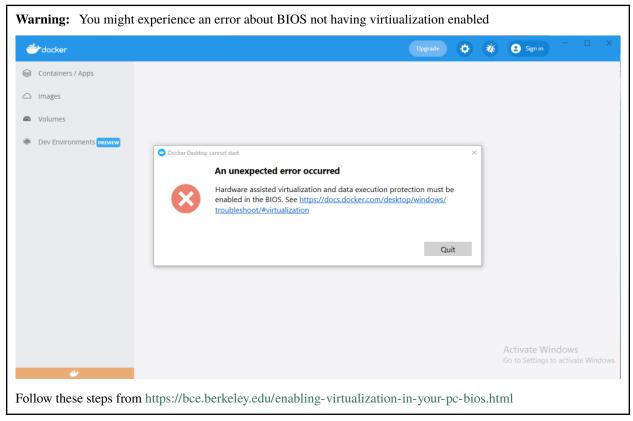
7. Main image that allows you to manage specific containers

	进 docker	¢ 🐐	bmerritt1762	- 🗆 X
Q Search				Sort by 🗸
artic_consensus artic RUNNING				
• Docker running				

### **Confirm Docker is Running**

In your taskbar (lower-right), if you hover over the icon you should see the message displayed below. Right-clicking will give additional options





See here Virtualization Disabled - Windows for more information

#### **Reboot your computer**

- Right when the computer is coming up from the black screen, press Delete, Esc, F1, F2, or F4. Each computer manufacturer uses a different key but it may show a brief message at boot telling you which one to press. If you miss it the first time, reboot and try again. It helps to tap the key about twice a second when the computer is coming up. If you are not able to enter the BIOS via this method, consult your computer's manual.
- In the BIOS settings, find the configuration items related to the CPU. These can be in under the headings Processor, Chipset, or Northbridge.
- Enable virtualization; the setting may be called VT-x, AMD-V, SVM, or Vanderpool. Enable Intel VT-d or AMD IOMMU if the options are available.
- Save your changes and reboot.
- Delete any existing VMs (Machine > Remove \*\* and select \*\* Delete all files) and re-import the .ova file (following step 4 and subsequent steps of the installation instructions).
- Check if your system supports Virtualization
- If you are unable to find the Virtualization settings in your BIOS it may mean that your laptop does not support it. If you want to try to find this out yourself, then you can try:

On Windows, download and run a Microsoft utility. You can also download utilities to check if your CPU is capable of virtualization, if not enabled. Hyper-V must be disabled in order for VirtualBox to run 64-bit guest operating systems. Visit the "turn Windows feature on or off" application and make sure Hyper-V is not checked.

On Linux, open a terminal window and run:

egrep -q 'vmx|svm' /proc/cpuinfo && echo yes || echo no

#### **Basestack**

#### **Install Main**

Download Basestack from Releases

- You will select the item labeled <Basestack-Version>.Setup.exe
- 1. Double-click Basestack-Version.Setup.exe
- 2. Follow the prompts for installing the software. Choose defaults unless otherwise needed.

#### 3.1.2 MacOS

#### Docker

#### Docker for Mac

In order to run docker you must be able to support virtualization from your CPU. This feature must also be enabled within your BIOS and Windows Features.

If you're unsure whether docker is supported by your specific cpu, please visit and input your specific model number:

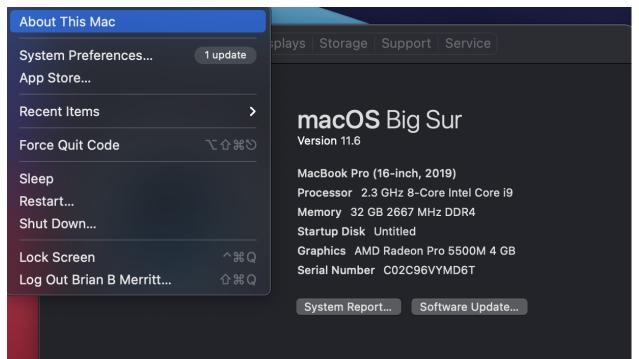
• Intel

#### • AMD

Type Your Model Number, e.g. T6500 into the product search bar

Intel <sup>®</sup> Core <sup>™</sup> 2 Duo Pr 2M Cache, 2.10 GHz, 800 MHz FS8	ocessor T6500	Ad
Specifications	Essentials	Export specific
Essentials	Product Collection	Legacy Intel® Core® Processors
Performance Specifications Supplemental Information	Code Name	Products formerly Penryn
Package Specifications	Vertical Segment	Mobile
Advanced Technologies Security & Reliability	Processor Number 👔	T6500
Deaduct Images	Status	Discontinued
Product Images	Launch Date 👔	Q2'09
Technical Documentation	Lithography 🕐	45 nm
	Performance Specifications	
	# of Cores 🕐	2
	# of Threads 👔	2
	Processor Base Frequency 🕐	2.10 GHz
	Cache 🕐	2 MB L2 Cache
	Bus Speed 👔	800 MHz
	TDP 🕐	35 W
	Supplemental Information	
	Embedded Options Available (?)	No
	Datasheet	View now
	Package Specifications	
	LAUNCTION (2)	105°C
	Processing Die Size	107 mm <sup>2</sup>
	# of Processing Die Transistors	410 million
	Advanced Technologies	
	Intel® Turbo Boost Technology # 🕧	No
	Intel® Hyper-Threading Technology # 👔	No
	Intel® Virtualization Technology (VT-x) = 👔	No
	Intel® 64 = 7	Yes

In this example above, you can see that Vt-x (Virtualization) is not supported. This will be a **Yes** if it is supported. On Mac you can find this value by



#### Instructions

https://docs.docker.com/docker-for-mac/install/

#### **Basestack**

Download Basestack from Releases

- You will select the item labeled <Basestack-Version>.dmg
- 1. Double-click <Basestack-Version>.dmg
- 2. Follow the prompts for installing the software. Choose defaults unless otherwise needed.

#### 3.1.3 Linux

#### Docker

#### Instructions

#### Install Docker

In order to run docker you must be able to support virtualization from your CPU. This feature must also be enabled within your BIOS and Windows Features.

If you're unsure whether docker is supported by your specific cpu, please visit and input your specific model number:

- Intel
- AMD

Type Your Model Number, e.g. T6500 into the product search bar

Intel<sup>®</sup> Core<sup>™</sup>2 Duo Processor T6500 <sup>2M Cache, 2.10 GHz, 800 MHz FSB</sup> Add to Compare

Specifications	Essentials		Export specifications
Essentials	Product Collection	Legacy Intel <sup>®</sup> Core <sup>™</sup> Processors	
Performance Specifications Supplemental Information	Code Name	Products formerly Penryn	
Package Specifications	Vertical Segment	Mobile	
Advanced Technologies Security & Reliability	Processor Number 👔	T6500	
	Status	Discontinued	
Product Images	Launch Date (7)	Q2'09	
Technical Documentation	Lithography 🕐	45 nm	
	Performance Specifications		
	# of Cores ①	2	
	# of Threads 👔	2	
	Processor Base Frequency 🕐	2.10 GHz	
	Cache 🕐	2 MB L2 Cache	
	Bus Speed 👔	800 MHz	
	TDP 🕐	35 W	
	Supplemental Information		
	Embedded Options Available 🕐	No	
	Datasheet	View now	
	Package Specifications		
	TJUNCTION (2)	105°C	
	Processing Die Size	107 mm <sup>2</sup>	
	# of Processing Die Transistors	410 million	
	Advanced Technologies		
	Intel® Turbo Boost Technology # 🕜	No	
	Intel® Hyper-Threading Technology # 🕐	No	
	Intel® Virtualization Technology (VT-x) = 👔	No	
	Intel® 64‡ 🕜	Yes	

**width** 100% In this example above, you can see that Vt-x (Virtualization) is not supported. This will be a **Yes** if it is supported. On Linux you can find the processor by typing on the command line:

🖒 Mouse & Touchpad	Memory	47.0 GiB
Keyboard Shortcuts	Processor	AMD® Ryzen 7 3700x 8 core processor × 16
🛱 Printers	Graphics	NV162
📋 Removable Media	Disk Capacity	5.0 TB
🔏 Color		
Region & Language	OS Name	Ubuntu 20.04.2 LTS
🕯 Universal Access	OS Type	64-bit
العدم ال	GNOME Version	3.36.8
★ Default Applications	Windowing System	X11
🕚 Date & Time	Software Updates	>
✦ About		

Make sure that you select the appropriate distribution for your machine. If you are unsure of your distribution use  $lsb\_release -a$  from the command line to check your distro.

**Note:** You can retrieve a basic install script for a fresh environment that can be found here It is primarily intended for newly flashed Ubuntu systems. This script will work for both ARM64 and AMD64 processors

It will set up Docker, Download Basestack, CUDA, and MinKNOW

> lsb_release -	9
No LSB modules	are available.
Distributor ID:	Ubuntu
Description:	Ubuntu 20.04.2 LTS
Release:	20.04
Codename:	focal

#### **Docker Installation**

Choose ONE option

- A. Rootless RECOMMENDED
  - https://docs.docker.com/engine/security/rootless/
    - \* If you already have *docker* installed, see documentation on [*docker context*](https://docs.docker. com/engine/security/rootless/#client) to switch between rootless and rootful
- B. Rootful (gives root access, use if you already have docker installed or use it regularly)
  - https://docs.docker.com/engine/install/ubuntu/
    - \* Required to map you user permissions appropriately for generated files.
    - \* Recommended for most rootful-specific personal systems running Docker
  - Post-Installation Steps:
    - 1. Create Docker group
      - \* sudo groupadd docker
    - 2. Add your user to the docker group
      - \* sudo usermod -aG docker \$USER
    - 3. Ensure all root-created files map as your user id in docker containers and volumes (Do both of these)
      - \* 1. sudo sed -i "1s/^/\$USER:\$(id -u):1\n/" /etc/subuid
      - \* 2. sudo sed -i "1s/^/\$USER:\$(id -g):1\n/" /etc/subgid
    - 4. Create Docker container namespace CHOOSE ONE
      - \* a. echo \$(jq -arg user "\$USER" '. += {"userns-remap": \$user}'
        /etc/docker/daemon.json) > ~/daemon.json && sudo mv ~/daemon.json
        /etc/docker/daemon.json
        - · If you dont have the file already created (isn't created by default)
      - \* b. Manually add your user by following the instructions here: https://docs.docker.com/engine/security/userns-remap/.

• You can disable the *userns-remap* functionality by deleting the *daemon.json* file described above or removing the line attributed to your user

- 5. Check that the subgid and subuid files are correct. Order of these lines matters in that the *<username>:<uid>:1* must come first in each file
  - \* 1. cat /etc/subuid
    - -<username>:<uid>:1-<username>:100000:65536
  - \* 2. cat /etc/subgid

-<username>:<uid>:1 -<username>:100000:65536

- 6. Restart Docker
  - \* a. sudo service docker restart
  - \* **b**. OR Restart your computer/session
- 7. Ensure that permissions are appropriate
  - \* 1. docker run -v /tmp:/opt/tmp nginx touch /opt/tmp/test.txt

#### \* 2.ls -lht /tmp/test.txt

• ^ ensure that ownership is your uid/gid or username:group

```
Listing 1: Full block of code for Option B (Rootful)
```

Open a terminal and type docker info. You should see information about your docker service

```
(base) brian-home@ubuntu:~$ docker info
Client:
 Debug Mode: false
Server:
 Containers: 14
  Runnina: 0
  Paused: 0
  Stopped: 14
 Images: 5
 Server Version: 19.03.8
 Storage Driver: overlay2
 Backing Filesystem: <unknown>
  Supports d_type: true
  Native Overlay Diff: true
 Logging Driver: json-file
 Cgroup Driver: cgroupfs
 Plugins:
 Volume: local
 Network: bridge host ipvlan macvlan null overlay
 Log: awslogs fluentd gcplogs gelf journald json-file local logentries splunk syslog
 Swarm: inactive
 Runtimes: runc
 Default Runtime: runc
 Init Binary: docker-init
 containerd version: 7ad184331fa3e55e52b890ea95e65ba581ae3429
 runc version: dc9208a3303feef5b3839f4323d9beb36df0a9dd
 init version: fec3683
 Security Options:
  apparmor
  seccomp
  Profile: default
 Kernel Version: 4.15.0-118-generic
 Operating System: Ubuntu 18.04.1 LTS
 OSType: linux
 Architecture: x86 64
 CPUs: 4
 Total Memory: 5.086GiB
 Name: ubuntu
 ID: 3WKY:JXQC:F6AY:XAWE:OHWP:SPYH:JN4Y:VNER:XBIM:YELW:W6AJ:NGCD
 Docker Root Dir: /var/lib/docker
 Debug Mode: false
 Username: bmerritt1762
 Registry: https://index.docker.io/v1/
 Labels:
 Experimental: false
 Insecure Registries:
  127.0.0.0/8
 Live Restore Enabled: false
WARNING: No swap limit support
(base) brian-home@ubuntu:~$
```

**Rootful:** - */var/lib/docker* is the Docker Root Dir. YOU MUST correctly utilize the *userns-remap* configuration described above for this to work

Rootless: - \$HOME/.local/share/docker (or something similar in \$HOME) will be the Docker Root Dir.

Additionally, for Docker Rootless only, you'll need to adjust the socket that Basestack is connecting to directly within the System tab of the application. This value will be wherever your *docker.sock* file is made.

	Log Streams	Docker is installed and running	Advanced 🛱
<b>1</b> =	Contact		navancea 🕌
۰	System	/run/user/1000/docker.sock	Update Socket Current Socket Configuration: /run/user/1000/docker.sock
			Carrent Socket Configuration, July aser, 1000, abereilsber

If you're unsure where that is run: docker context ls and it will be the DOCKER ENDPOINT value sans the unix://

> docker co	ontext ls			
NAME	DESCRIPTION	DOCKER ENDPOINT	KUBERNETES ENDPOINT	ORCHESTRATOR
default *	Current DOCKER_HOST based configuration	unix:///run/user/1000/docker.sock		swarm

#### **Basestack**

#### **Install Main**

Download Basestack from Releases

- You will select the item labeled <Basestack-Version>.<arch>.AppImage
- Double-click <Basestack-Version>.<arch>.AppImage
- 2. Follow the prompts for installing the software. Choose defaults unless otherwise needed.
- 3. chmod ugo+x <Basestack-Version>.<arch>.AppImage

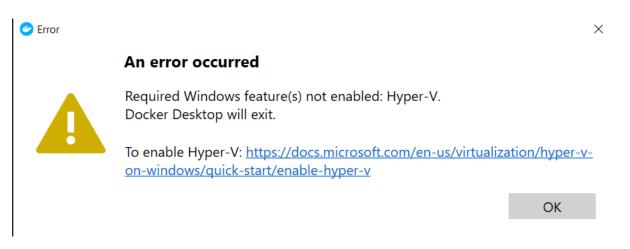
### 3.1.4 Troubleshooting

#### General

- Request or view feature changes at our [issue tracker](https://github.com/jhuapl-bio/Basestack/issues)
- If you run into issues with the online install, you may want to download (or otherwise obtain) the offline install package
  - Using the above download links, download the appropriate docker images you'd like e.g. basestack\_consensus.tar.gz (~5.2GB)
  - Run docker load < basestack\_consensus.tar.gz from the command line
- See below Appendices for more detailed installation instructions.

#### Windows

#### Hyper-V Not Enabled - Windows



If you are on older Windows distributions, you may experience an error when attempting to start docker on how HyperV is not enabled.

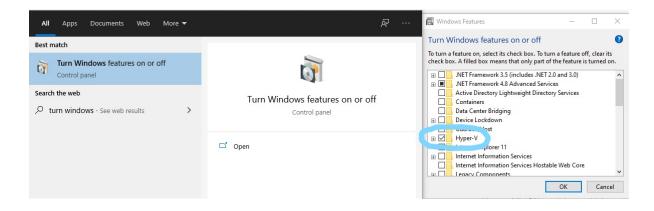
#### **Enable Hyper-V in Basestack**

To enable it within Basestack select: System -> Windows Services -> Hyper-V -> Enable Hyper-V. A window will appear prompting admin rights and then it will automatically being the enable process. See more below.

53 E	Basestack							
Edit	System	View	Window	Check	for Updates	Lo	gs and Info	Help
		esh Serv	/er					Hom
~		cer Site art App				٨/	alcon	ne to
	Wind	dows Se	rvices		Hyper-\		Disable	Hyper-V
•	Close	e	Ctr	l+W	WSL2	<u> </u>	Enable	Hyper-V

#### Enable Hyper-V in Windows System

**Alternatively** you can enable it within the Host system itself by searching for "Turns Windows features on or off" and selecting "Hyper-V". This will require a computer restart



#### WSL2 Not Installed - Windows

The error (seen below) is often shown for newer Windows OS types. If this occurs, you may have different variants. In the included example, I have the option to enable WSL or use Hyper-V.

🗢 Docker Desktop	p - Install WSL 2
	WSL 2 is not installed
	Install WSL using this PowerShell script (in an administrative PowerShell) and restart your computer before using Docker Desktop:
	Enable-WindowsOptionalFeature -Online -FeatureName \$("VirtualMachinePlatform", "Microsoft-Windows-Subsystem-Linux")
	Use Hyper-V Stop Docker

Sometimes, another window will appear regarding installing WSL.

#### Install WSL2 from External Sources

Please follow that link

Make sure to perform AT LEAST step 4. Once WSL2 is installed/enabled, please restart Docker Desktop

#### Install WSL2 in Basestack

Alternatively Basestack allows users to download WSL directly.

To Download then Install it within Basestack do:

1. System -> Windows Services -> WSL2 -> Download WSL2

2. System -> Windows Services -> WSL2 -> Install WSL2

	े в	asestack						
	Edit	System	View	Window	Check	for Updates	Log	s and Info Help
		Refre	esh Serv	/er				Home
ĺ		Dock	ker Site					
	*	Rest	art App			۸/ ا	_	come to Base
I		Wind	dows Se	rvices	►	Hyper-		
	\$	Close	e	Cti	rl+W	WSL2	►	1. Download WSL2
			ICV					2. Install WSL2

You can then attempt to restart Docker Desktop. This also may require a system restart.

If you are still experiencing issues, attempt to enable virtualization from Basestack:

3. System -> Windows Services -> WSL2 -> Turn WSL On

4. System -> Windows Services -> WSL2 -> Enable Virtualization

5. System -> Windows Services -> WSL2 -> Set WSL2

Or from "Turn Windows features on or off". This is also a good way double check that it is now enabled.

📷 Windows Features	—		Х
Turn Windows features on or off			?
To turn a feature on, select its check box. To turn check box. A filled box means that only part of the		•	
TFTP Client			~
🗹 📙 Virtual Machine Platform			
Windows Hypervisor Platform			
Windows Identity Foundation 3.5			
🛞 🖂 🔣 Windows PowerShell 2.0			
🗄 🗌 📙 Windows Process Activation Service			
Windows Projected File System			
Windows Sandbox			
🗹 📊 Windows Subsystem for Linux			
Windows TIFF IFilter			
✓ Work Folders Client			
			<b>Y</b>
	ОК	Cance	el

#### You will need to restart your PC/Laptop after doing this!

#### Switching between HyperV and WSL2 instance

When inside Docker-Desktop, hit the settings (cog) icon at the top-right of the page. Then, select General Tab and tick/un-tick the *Use the WSL 2 based engine*. Be aware that when using HyperV you may need to adjust resources to accommodate your system appropriately.

	Upgrade 🐼 🛷 🔁 Sign in 👘 🗆 🗙
Settings	×
Image: GeneralImage: ResourcesImage: Docker EngineImage: Experimental FeaturesImage: ResourcesImage: Res	General         Image: Automatically check for updates         Paid Team plans allow IT-managed organizations to disable checking for updates.         Image: Start Docker Desktop when you log in         Image: Expose daemon on tcp://localhost:2375 without TLS         Exposing daemon on tCP without TLS helps legacy clients connect to the daemon. It also makes yourself vulnerable to remote code execution attacks. Use with caution.         Image: Use the WSL 2 based engine         WSL 2 provides better performance than the legacy Hyper-V backend. Learn more.         Image: Send usage statistics         Send usage statistics         Send weekly tips         Image: Open Docker Desktop dashboard at startup
	Cancel Apply & Restart
<b>*</b>	

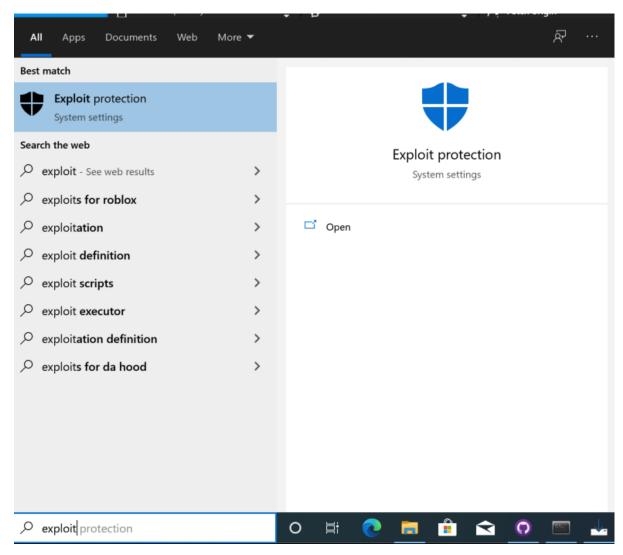
#### WSL2 error on Docker Start

If you're still experiencing issues after install WSL2 (also making sure you've attempted to use both installation methods), there may be an issue with your system's firewall configuration. To alleviate this, you can modify some settings within Windows by following:

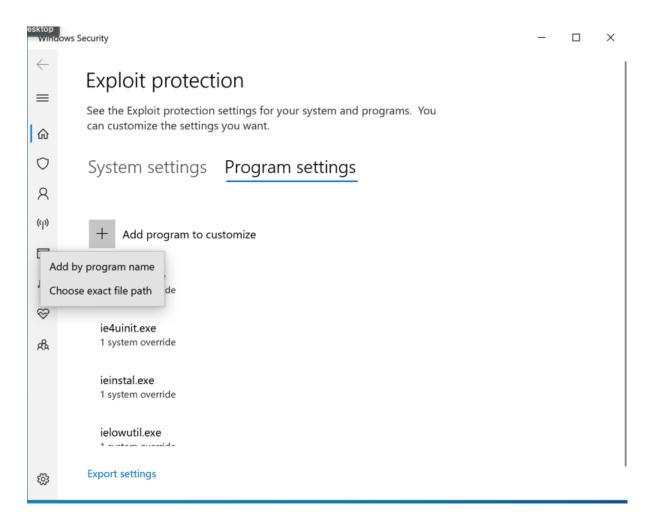
- 1. navigate to "Start" button,
- 2. type in "Exploit protection" and run it as administrator,

3. once in, nawigate to: "Program settings" "Add program to customise", adding the two below separatelly, in each case, disabling by unticking: "Code flow guard (CFG)"

```
C:\WINDOWS\System32\vmwp.exe
C:\WINDOWS\System32\vmcompute.exe
```



Select choose exact file path



Copy + paste these 2 commands one-by-one then apply changes

C:\WINDOWS\System32\vmwp.exe C:\WINDOWS\System32\vmcompute.exe

Open				;	×	-	[	×
$\leftarrow$ $\rightarrow$ $\checkmark$ $\uparrow$ 🗢 This PC >	~	Ū	Search This PC	ر	ρ			1
Organize *								
<ul> <li>Quick access</li> <li>Desktop</li> <li>Downloads</li> <li>Documents</li> <li>Pictures</li> <li>Movies</li> <li>Music</li> <li>OneDrive</li> <li>This PC</li> </ul>	(7) 3D Objects Desktop Documents Downloads				*			
C:\Windows\System 1 system override C:\Windows\System 1 system override		e ∨ 	All files Open	Cancel				
Export settings								

Wind	lows Security	Program settings: vmwp.exe	-	×
=	Explc See the F	Also allow loading of images signed by Microsoft Store		
	can custo Syster	<b>Control flow guard (CFG)</b> Ensures control flow integrity for indirect calls.		
8		Override system settings		
((1))	+ A	Off Use strict CFG		
	0 syste			
旦	Syster	Data Execution Prevention (DEP) Prevents code from being run from data-only memory pages.		
$\otimes$	1 syste	Override system settings		
<i>1</i> 84	C:\Wii 1 syste	On Enable ATL thunk emulation		
	C:\Wi 1 syste	Disable extension points		
	r syste	Changes require you to restart vmwp.exe		
£03	Export se	Apply Cancel		

#### Finally, restart Docker Desktop

\*Credit to this solution

#### **Virtualization Disabled - Windows**

In order for either of the above to work, you need to ensure that **virtualization** is enabled in your firmware. Some processors do so by default, others do not. If you are having issue with starting Docker despite following either of the options above, please see below.

First, check that your CPU can support virtualization by viewing the model on Intel/AMD product page(s)

You can first check if it is enabled by going into the **Task Manager** and seeing if the Virtualization attribute is enabled.

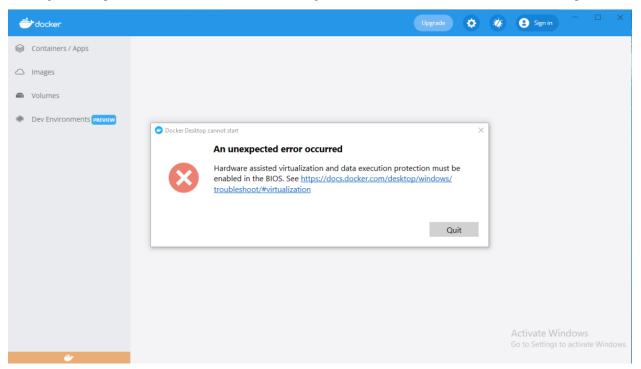
All Apps Documents Web More 🔻	R	P File Options View	×
Best match		Processes Performance Apphistory Startup Users Details Services	
Task Manager App		CPU AMD Ryzen 7 3700X 8-Core Processor 5 Utilization 1000	
Search the web	Task Manager	> Memory 7.0/47.9 GB (15%)	
✓ task manager - See web results	App	O Disk 0 (L)	
Settings (1)		t HDD 0% 60 seconds	0
	Copen     Run as administrator     Open file location	O         Disk 1 (D2) SSD         Utilization Speed         Bars speed         359 GHz           SSD         2%         3.57 GHz         Scelaria         1           0%         Processes         Treads Hundles         Cores         8           0         Disk 2 (H; C)         294         5102         1255090         Treads Hundles         Logical processor: 16	
	-t⊐ Pin to Start -t⊐ Pin to taskbar	SSD 0% Up time L1 cohe 31245 0% 0% 0% 0% 0% 0% 0% 0% 0% 0% 0% 0% 0% 0	Ŷ

If it is not, open up **Command Terminal** and type: systeminfo. Scroll to the bottom of the output and check if the Firmware has it enabled for Hyper-V requirements.

Hyper-V Requirements:	[06]: Bluetooth Device (Personal Area Network) Connection Name: Bluetooth Network Connection 2 Status: Media disconnected VM Monitor Mode Extensions: Yes Virtualization Enabled In Firmware: Yes Second Level Address Translation: Yes Data Execution Prevention Available: Yes	
C:\Users\bmer<>systeminfo		~

If not, you will need to enable Virtualization in your BIOS. This process will look different based on everyone's system. You should try to follow the instructions in this [link](https://www.thewindowsclub.com/disable-hardware-virtualization-in-windows-10). Choose your manufacturer type.

You might also experience an error about BIOS not having virtualization enabled as soon as Docker attempts to start



Follow these steps from https://bce.berkeley.edu/enabling-virtualization-in-your-pc-bios.html:

#### **Reboot your computer**

- Right when the computer is coming up from the black screen, press Delete, Esc, F1, F2, or F4. Each computer manufacturer uses a different key but it may show a brief message at boot telling you which one to press. If you miss it the first time, reboot and try again. It helps to tap the key about twice a second when the computer is coming up. If you are not able to enter the BIOS via this method, consult your computer's manual.
- In the BIOS settings, find the configuration items related to the CPU. These can be in under the headings Processor, Chipset, or Northbridge.
- Enable virtualization; the setting may be called VT-x, AMD-V, SVM, or Vanderpool. Enable Intel VT-d or AMD IOMMU if the options are available.
- Save your changes and reboot.

- Delete any existing VMs (Machine > Remove \*\* and select \*\* Delete all files) and re-import the .ova file (following step 4 and subsequent steps of the installation instructions).
- Check if your system supports Virtualization
- If you are unable to find the Virtualization settings in your BIOS it may mean that your laptop does not support it. If you want to try to find this out yourself, then you can try:

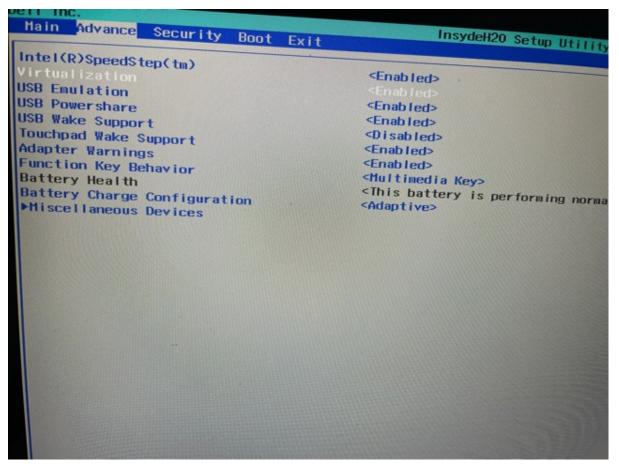
On Windows, download and run a Microsoft utility. You can also download utilities to check if your CPU is capable of virtualization, if not enabled. Hyper-V must be disabled in order for VirtualBox to run 64-bit guest operating systems. Visit the "turn Windows feature on or off" application and make sure Hyper-V is not checked.

On Linux, open a terminal window and run:

egrep -q 'vmx|svm' /proc/cpuinfo && echo yes || echo no

Typically, though, to enter BIOS you must restart the computer and while it is booting hit **DEL** or **F2** or sometimes **F12**. This process is usually very quick so be ready. When it is booting, you may be able to catch the necessary keys flash.

The default BIOS should look like the one below. In there, head to the **Advanced** tab and check if **Virtualization** is present. If so, enable it, save changes, and restart. If not, try to search in other tabs or open up some options that have further submenus within them as there is no guarantee it will be directly on the base **Advanced** tab.



On AMD CPU's if you don't see virtualization it may be labeled as SVM in the Advanced tab

	OS Utility	- Advanced M	lode			
riday 07:23	S⊷   ⊕ ∈	nglish 💷 MyFav	vorite(F3) 20 Qf4	an Control(F6)		F11) [2] Search(F9)
My Favorites	Main	Ai Tweaker	Advanced	Monitor	Boot Tool	Exit
Total L3 Cache p PSS Support NX Mode SVM Mode SMT Mode Core Leveling N	g @ 3595 MH MHZ Level: 87010 r core che: 32 KB/8-0 : 512 KB/8-0 ber Socket: 3	42 1100 mV 013  3-way way 2 MB/~way			nabled nabled isabled isabled nabled utomatic mode	
			2.17.1246. Copyri	Last ght (C) 2019 Ame	Modified EzMode	(F7) -E Hot Keys

If the option is not present in the BIOS that means that your CPU does not support Virtualization and Docker won't be able to properly run on your system.

docker\_support

## **Operation not permitted - Windows**

If you receive an error about operation not being permitted and you're reinstalling or updating Basestack (see image), attempt to uninstall Basestack and reinstall it.

#### **Operation not permitted error**



## **Uninstall Basestack**

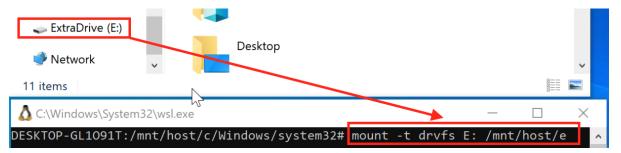
All Apps Documents Web M	ore 🔻		Ap Ap	ps & teatures	
Best match			0	Anaconda3 2021.05 (Python 3.8.8 64-bit)	7/22/2021
Add or remove programs System settings		E C C C C C C C C C C C C C C C C C C C	Î	App Installer Microsoft Corporation	<b>48.0 KB</b> 10/6/2021
Settings	>	Add or remove programs System settings		Basestack 1.1.1	252 MB 10/6/2021
冬 Add, edit, or remove other users ② Link your phone	>	C Open	Ŧ	Calculator Microsoft Corporation	16.0 KB 7/29/2021
Search the web	,		<u></u>	Camera Microsoft Corporation	16.0 KB 7/22/2021
✓ add - See web results	>		0	Cortana	16.0 KB

## **Docker Pipelines Do Not Run With External Drives**

If you need to use an external Drives within the modules AND are using Docker-Desktop for Windows with WSL2, you will need to likely mount your external drive within WSL first.

This is done in 2 steps: 1. Start wsl from the command line by typing wsl into the quicklaunch and starting it - Make sure Docker is running on your system before starting this - A terminal window will appear at the start. Done fear you can follow step 2 easily by copying it in the terminal and just changing the <drive\_letter> (See below for more info)

2. mkdir -p /mnt/host/<drive\_letter> && mount -t drvfs <drive\_letter> /mnt/host/ <drive\_letter> - This process maps your new drive letter to the exact letter in windows. For instance in the example my flash drive is E: and /mnt/host/e is what is it mapped to - <drive\_letter> is whatever the letter is from your system. For example the example shown here is E: but yours may (likely to) differ



Alternatively, you can switch WSL2 instance to HyperV by following this [step](#switching-between-hyperv-and-wsl2-instance)

#### I/O timeout

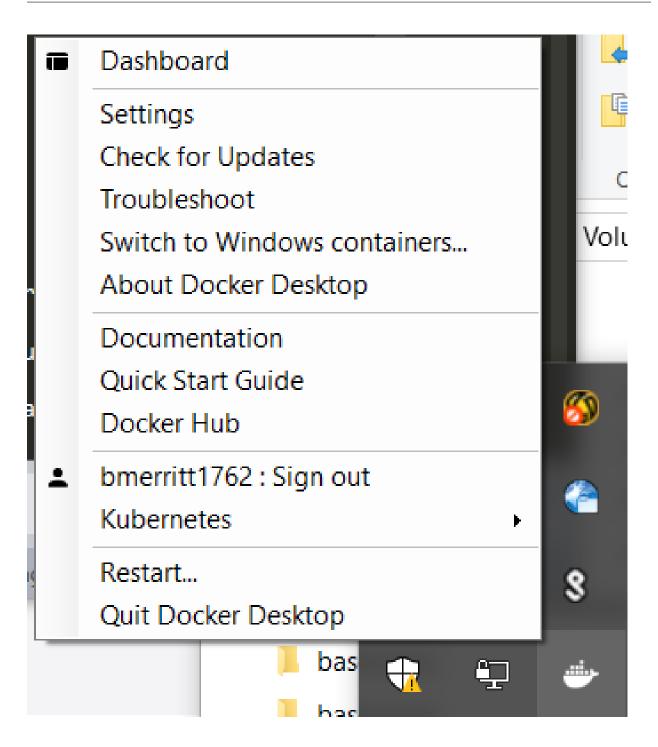
If installing as the offline method, sometimes you may retrieve and I/O exception as an error message. This is typically because you've tried to send docker to build too many images from large files in a short period of time. To fix this, you'll need to rerun docker a specific way

B. (#1)	Recommended for first-time installs or after an up	date nas	request of a dependency	the income the 1 day area
All Apps Documents Settings Photos	More ✓ Feedback …	Action View Help		:
Hyper-V Manager         →           Desktop app         →           Apps		yper-v Manager MERRIBB1-WL1	Virtual Machines Name State DockerDesktopVM Running	Actions Quick Create New New
Hyper-V Quick Create > Folders (2+)	Hyper-V Manager Desktop app		< Checkpoints	Hyper-V Settings Virtual Switch Manager Virtual SAN Manager
Settings (1) Documents (5+)	다 Open 중 Run as administrator , Run as different user		The selected virtua	Stop Service     Remove Server
	☐ Open file location -₽= Pin to Start		DockerDesktopVM Created: 10.	Refresh     View     Help     DockerDesktopVM
3	-Bi Pin to taskbar		Configuration Version: 9.0 Generation: 2 Notes: No	Connect
P hvp			Summary Memory Networking	Save     Pause

• Simply seach for Hyper-V Manager in your toolbar, select the VM (usually your username is in its name), and then *Turn Off.* Docker will then shut down and you can restart it

## **Docker Connection**

If you receive an error that you couldn't connect to docker, please try to restart via the taskbar



## 'You are not allowed to use Docker, you must be in the "docker-users" group' - Windows

In Basestack select System -> Windows Services -> Add User Docker-Users. When completed you should see that either you're already a part of that group **OR** you've been successfully added.

8 S. E	lasestack							
Edit	System	View	Window	Check fo	r U	pdates	Logs and Info	Help
	Refre	esh Sen	/er					
	Dock	er Site						
~	Oper	n Deskt	ор					
_	Resta	art App						
i I	Chec	k Dock	er Installed	I				
	Орег	n Termi	nal					
۲	Wind	dows Se	ervices		►	Нур	per-V	- •
	Close	e		Ctrl+W		WS	L2	•
*	- PREASE					Ade	d User Docker-U	sers
	Log Stre	ams				Ор	en Powershell	

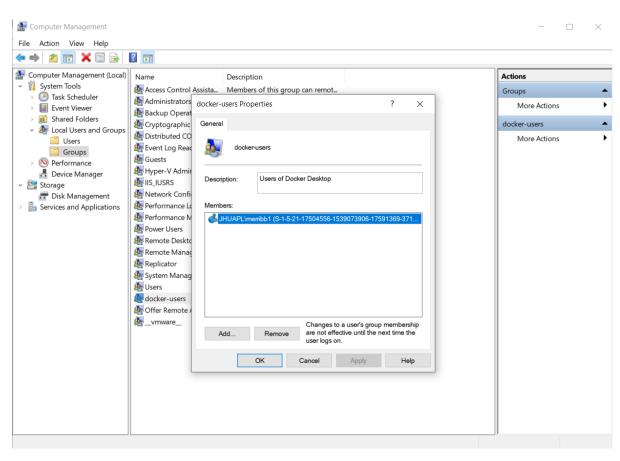
Alternatively if the above does not work try the following:

≡	All Apps Documents Settings	Photos	More V Feedback
ሴ	Best match		
	Computer Management Desktop app	$\rightarrow$	
	Settings		
	Administrative Tools	>	Computer Management Desktop app
			□       Open         □       Run as administrator         □       Run as different user         □       Open file location         -□       Pin to Start         -□       Pin to taskbar

• In the Windows search taskbar (bottom left icon), find Computer Management

🔚 Computer Management			- 0	×
File Action View Help				
🔶 🔿 📶 ដ 📄	?			
Computer Management (Local)  V 🎁 System Tools	Name Maccess Control Assista	Description Members of this group can remot	Actions Groups	
A Scheduler     Task Scheduler     Task Scheduler     Event Viewer     Stared Folders     Schedulers	Administrators Backup Operators Cryptographic Operat	Administrators have complete and Backup Operators can override sec Members are authorized to perfor	More Actions docker-users	) •
<ul> <li>Local Users and Groups</li> <li>Users</li> <li>Groups</li> <li>Performance</li> <li>Device Manager</li> <li>Storage</li> <li>This Management</li> <li>Services and Applications</li> </ul>	Distributed COM Users     Event Log Readers     Guests     Hyper-V Administrators     IIS_IUSRS     Network Configuratio     Performance Log Users     Performance Monitor     Power Users     Remote Desktop Users     Remote Management     Replicator	Members are autonized to perfor- Members are autonized to launch, ac Members of this group can read e Guests have the same access as m Members of this group have comp Built-in group used by Internet Inf Members in this group can have s Members of this group can have s Members of this group can access Power Users are included for back Members in this group are grante Members of this group are grante Members of this group are manag Users are prevented from making Users are prevented from making Users of Docker Desktop Members in this group can offer R VMware User Group	More Actions	,

• Select (left-side) System Tools -> Local Users and Groups -> Groups



• Double click docker-users and see if your name is there, if not:

🛃 Computer Management					—	$\times$
File Action View Help						
🗢 🄿 🙍 📷 🗙 🗐 🔒	?					
Computer Management (Local)	Name Full	DefaultAccount Properties		? ×	Actions	
<ul> <li>W System Tools</li> <li>O Task Scheduler</li> </ul>	🛃 A1pha8rav0 👼 APLr00t	General Member Of Profile			Users	<b>•</b>
> 🛃 Event Viewer	DefaultAcco	Member of:			More Actions	•
<ul> <li>         Istared Folders     </li> <li>         Istared Folders     </li> <li>         Istared Folders     </li> </ul>	WDAGUtility	A docker-users			DefaultAccount	-
Users Groups		System Managed Accounts Group			More Actions	•
> 🔊 Performance 🛃 Device Manager	Select Groups		×			
v 🔄 Storage	Select this object type:					
<ul> <li>Disk Management</li> <li>Services and Applications</li> </ul>	Groups		Object Types			
Services and Applications	From this location:					
	MERRIBB1-WL1		Locations			
	Enter the object names	s to select ( <u>examples</u> ):				
	docker-users		Check Names			
	Advanced	ОК	Cancel	nbership me the		
		usonoga				
		OK Cancel	Apply	Help		

- Select the Users folder right about where you clicked Groups
- Select the name of your user

#### • enter docker-users into the object field and add.

- You will need to log out and back into your account for this to take effect

## Linux

#### Permisson denied (Linux)

Please ensure that you follow the correct Docker Installation here to using userns-remap

Note that this will map all of your processes INSIDE the docker containers to your user id if used properly. You will need sudo to delete any files or folders that are causing issues.

## 3.1.5 Uninstall

#### Uninstall

#### **Docker Image(s):**

- In Basestack, go to Module Install
- Select: Remove Docker Images (trash-bin icon)
  - This will only remove Basestack-specific images

## Basestack

#### Windows

Add or remove programs -> Select Basestack -> Uninstall

#### Mac

Drag ~/Library/Application Support/Basestack to the Bin

#### Linux

- 1. Remove the Basestack.AppImage Folder or Executable
- 2. Remove the directory: ~/.config/Basestack

## **User Data**

- Mac OS: ~/Library/Application Support/Basestack (taken from the name property in package.json)
- Windows: C:\Users\<you>\AppData\Local\Basestack Name
- Linux: ~/.config/Basestack

# 3.2 Usage

Note: Test data for most modules can be found here from a Gdrive location

## 3.2.1 Modules

## Pavian

Pavian is a tool designed to analyze report files from a variety of classification modules, namely, for Basestack, Mytax uses Kraken2

Note: This module contains a UI to be displayed from the rendering button in the variables table

#### **Parameters**

#### report file

[File] Report file from Kraken2 run. See here. To be input INSIDE the running visualization. See images below

#### Returns

#### **Pavian Visualization**

[User interface] The tool simply requires you to input your kraken.report (from mytax or otherwise) into the home-page.

Uploaded sample set	≡					5 PC	Pavian metagenomi	ics data explorer	¢
Data Selection     Uploaded sample set     Results Overview	files. Pleas	e note that curren	analysis of metagenomics classificat ly the default Centrifuge report forma tion please go to https://github.com/	it is not supported. To generate a cor					
Sample	Upload fi	es Example	data					Data Source	
<ul> <li>∠ Comparison</li> <li>★ Alignment viewer</li> </ul>			files from the local computer. If selec le that has at least the columns 'Nam		ase try with a differe	ent browser. With each samp	ole set, you may also include m	eta-data with a colon-	
About	Browse.	. sample_meta	igenome.fastq.report Upload complete						
& Bookmark state Generate HTML report	Added sam	ple set <b>Uploaded</b>	sample set with 1 valid reports in tot	al.					
@fbreitw, 2021	Available	sample sets						-	-
	🖲 Uplo	aded sample se	t		View results	🖋 Rename sample set	Remove sample set		
	F	ormatOK Include	Name	ReportFile		ReportFilePa	th		
	1		sample_metagenome.fastq.report	1 - 0 1 1		dc97061a0603d56fcc73b8d	d/sample_metagenome.fastq.	report	
	You can sp	cify which sample	es to include as well as their names. B	e sure to save the table to make the o	hanges persistent.				
	Save tab	e							

Select Sample in the left-hand navigation drawer to view diagrams and information of your report

Sankey visualization	Table Text			
Hover over a node to se	e the abundance of the	e taxon in other samples	Configure Sankey	
<sup>1</sup> Archaea <sup>1</sup> Euryar	chaeota			
	330 Bacteroida	Bacteroides	45 Bacteroides dorei Bacteroides ovatus Bacteroides vulgatus	
2.07k 50 Actino Bacter	bacteria 32 36 Tannerellao 113 Lachnospir 371	216 316	es 76 [Eubacterium] rectale 109 Lactococcus piscium	
594 Firmic	Streptococ	Sheptococcus		
Bacteria <sup>3</sup> Fusob <sup>1</sup> Nitros 927	Comamona	Arcobacter Acidovorax Adaceae	36 Aeromonas media 32 Arcobacter butzleri	
	bacteria 41 346 Moraxellac	Acinetobacter	<ul> <li>Acinetobacter johnsonii</li> <li>Moraxella osloensis</li> </ul>	
<sup>32</sup> Viruses		<sup>30</sup> Phix174microv	<i>v</i> irus	
D P	F	G	S	
Save Network				

#### **NCBI Scrubber**

Note: Only works on interleaved R1/R2 fastq file OR NanoPore reads.

NCBI scubber removes human reads from sequencing data

#### **Parameters**

#### input

[File] Your input fastq file. Can be compressed but you must check "decompressed" if so

#### interleaved

[Boolean] Is the file interleaved? (TRUE/FALSE, Illumina)

#### decompressed: Boolean

Check if your file is compressed. You can decompress it directly in the UI here to the name: decompressed.fastq in the same path as the input file

#### Returns

#### **Filtered File**

[.filtered.fastq] removed of human reads

#### **Mytax**

Mytax is a tool designed to utilize the *Kraken2 <https://ccb.jhu.edu/software/kraken2/>* classification tool in order to identify (and visualize) taxonomic designations of fastq files. It includes a flu-kraken (from JHUAPL) and minikraken (from developers of kraken) set of databases in the installed images.

Note: This module contains a UI to be displayed from the rendering button in the variables table

#### **Parameters**

#### fastq file

[*File*] FastQ file (single) obtained from any sequencing run. You can consolidate multiple fastq files into one if you want to run a long sequencing run.

#### nodes

[*Directory*] Location of the taxonomy files. Default is contained within the `jhuaplbio/basestack\_mytax image and pre-loaded on run

## database

[Directory] Directory to use as the database for Kraken2. Defaults to Flukraken pre-loaded in the Docker image.

#### memory-mapping`

[*Option*] Choose to pre-load the database (fast, requires high RAM) or run on the local filesystem (slow, only use with low RAM available)

## Returns

#### **Kraken-Report File**

[.report]

- Used in Pavian and for additional 3rd-party applications
- Contains a summary of the run

#### **Kraken-JSON**

[.json]

• Formatted file uploaded/used in the Mytax Sunburst Visualization Procedure.

There are 2 primary procedures for this module:

## 1. Create Report Files

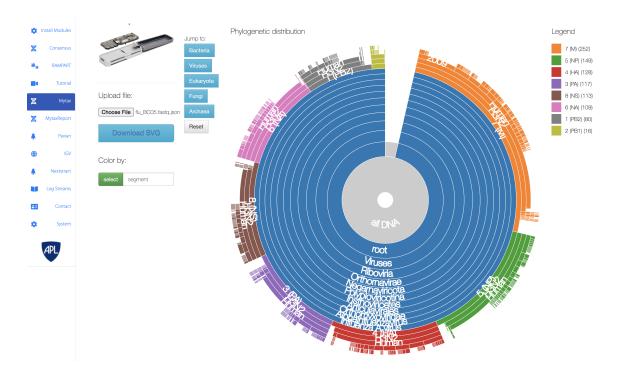
- Run kraken-classify to create kraken.report and kraken.json files from a single fastq file
- 2. View the Visualization
  - View the output JSON file from Step 1

			Basestack				Q Search		
X	Basestack		C Mytax	)			nome		
8			<b>(</b> ) 1	Kraken2 Mytax2			Desktop minikraken2_		
2	Mytax		Version select	Procedure select		•	i demux-fastq_pass →	i covid >	
e	Kraken2 Report				🕑 Creat	e Report JSON	i medaka → metagenome →	sample_metagenome.fastq sample_metagenome.report	
	Inputs						i minimap2 >	sample_metaeport.fullstring	
Ă		Param		Source		Туре		sample_metame.report.out	
ñ	F	ASTA/Q File 🧧	/Users/mernome.fastq	(69 B)	1 files (69 B in to	file			
	Nodes.dmp file, ş	pathered from ncbi taxonomy 🧧	default Select an item			-			
	Cla	ssifier database 🧧	Flukraken2-2021			-			
				Rows per page: 10	1-3 of 3	NO O			
	Output Locations				mDELE	TE OUTPUTS			s T
		label		Status		Remove			
		Report From classification, in kraken-s	tyle 🕤	1/1		<u>ت</u>			
		JSON Visualization File 🥹		1/1		۵.			
				Rows per page:	5 1-2 of	2			
*	MOPEN LOG FOLDER								
	Initiating logging object								
\$	stdout: done.	880. AME							
۲	stdout: stdout: <b>6</b> 7500 sequences (1.51 M	bp) processed in 0.771s (583.8 Kseq/m, 117.92 Hbp/m).	1932 sequences classified (25.76%) 1	568 sequences unclassified (74.24%)					

## **WIP Procedures**

- Download and create/update custom databases
- Utilize other classifiers or aligners
  - 1. Centrifuge
  - 2. Kraken2
  - 3. BLAST

## Display



## Mytax version 2 (Metagenomics)

Warning: This module is under construction and is in alpha-release. Scheduled full release of v1.0 in Oct. 2022

Please see relevant links in the listed modules for more information on the underlying mechanisms and corresponding papers (if existent)

## **Parameters**

• Samplesheet (.csv): *file* 

Contains a mapping of metadata and a single sample per row. Explanations of the possible columns for Basestack are seen below:

Column	Description
	Description
Name	
sample	Custom sample name. This entry will be identical for multiple sequencing libraries/runs from the same
	sample. Spaces in sample names are automatically converted to underscores (_).
format	[directory, run, file] Is it a run directory of files that need to be demux'd, an already full directory of
	files for Oxford, or a single file (or paired files)
path_1	Full path to FastQ file for Illumina short reads 1 OR OXFORD reads. File has to be gzipped and have
	the extension ".fastq.gz" or ".fq.
path_2	Full path to FastQ file for Illumina short reads 2. File has to be gzipped and have the extension ".fastq.gz"
	or ".fq.
kits	What default barcode kit to use for demux. Only applies to those with the "run" format
pattern	Regex matching for the names of the folders that are made on demux. Default is barcode[0-9]+
plat-	Platform used, [illumina, oxford]
form	
database	Kraken2 database path (root level folder)
com-	TRUE/FALSE for gunzipped files
pressed	

## Table 1: Samplesheet Description

sam-	path_1	path_2	for-	plat-	database	com-	pattern	kits
ple			mat	form		pressed		
covid_ru	nfastq_pass		run	OX-			barcode[0-	EXP-
				ford			9]+	NBD103
NB03	./NB11		direc-	OX-	minikraker	TALSE		
			tory	ford				
ERR123	ERR123_R1.fast	qEgRR123_R2.fas	qfglæ	illu-	flukraken2	TRUE		
				mina				

## Table 2: Example Samplesheet

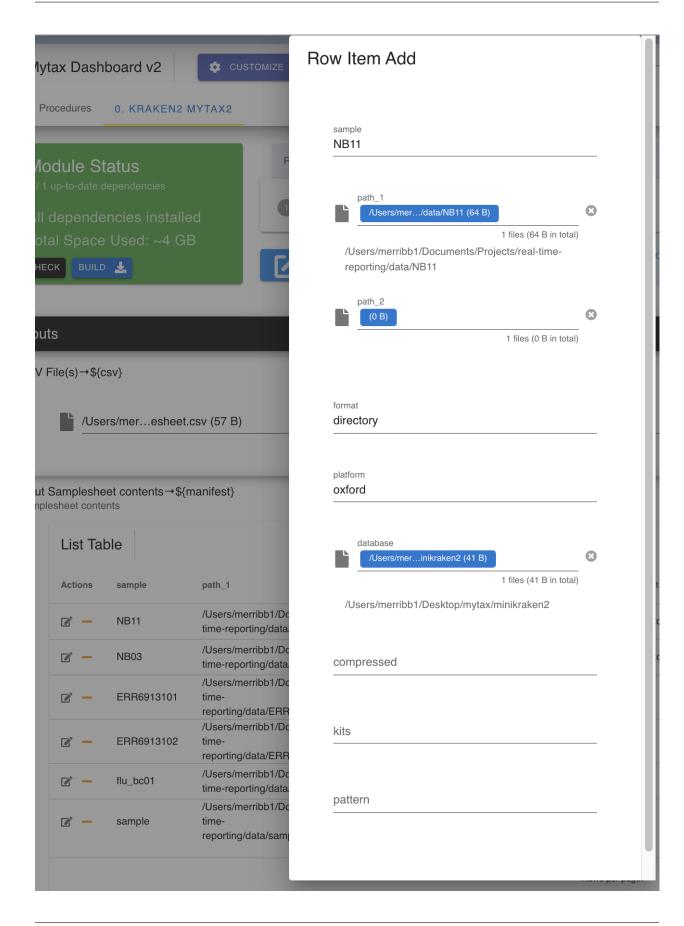
## Returns

- 1. Barcode Folders OPTIONAL. This will only be generated if you input a run in the data sheet and opt to set barcoding (demux) on it from the Data Sheet view
- 2. Kraken2 Classification Reports. Each will be output in the same folder as the sample of interest as full.report. Full.report is the aggregation of all reports for all fastqs attributed to a sample

## Display

🄰 Base	stack		S	earch for an entry	to Basestack's I	brary		Home	. A	<b>≡</b> · %	6 T -	Conditional Fo		·
Mytax Dash	hboard v2	CUSTOMIZE Version	📎 🞸 Latest	v1		*	C O	× Paste ✓	Font	Alignment Nur	mber 🛒	Format as Tab Cell Styles *		Cells f
Procedures	0. KRAKEN2 M	YTAX2					🔅 RESET DEFAU			ome features might $f_X$ sample	be lost if you	u save this wo	orkbook in	
Module S		Kraken2 Report2	oard for kraken2 classifications					A 1 sample 2 covid_run 3 NB11 4 NB03	B path_1 p /Users/merrib /Users/merrib	C D hath_2 format hb1/Docume run hb1/Docume directory hb1/Docume directory	E platform oxford oxford	/Users/merr	G compressed p b ibb1/Desktop/r ibb1/Desktop/r	arcode[0-9] EX nytax/minikrał
	encies installeo e Used: ~4 GB ▣ 🛓		DPEN AT 8098	A This proce	edure utilizes a n	endered UI, select the "	"Click Me"	5 ERR691310	/Users/merr / /Users/merr /	Users/merr file Users/merr file b1/Docume file b1/Docume file	illumina illumina oxford oxford illumina	/Users/merr /Users/merr /Users/merr /Users/merr	ibb1/Desktop/r ibb1/Desktop/r ibb1/Desktop/r ibb1/Desktop/r ibb1/Desktop/r	nytax/flukrake nytax/flukrake nytax/flukrake nytax/minikrak
nputs							-	11 12 13						
/Us	ers/meresheet.c	sv (57 B)					0	17 18						
-	ieet contents→\${m tents					1 files (57 B in s	iotal)	18 19 20 21 22 23 24 25 26						
nput Samplesh samplesheet cont	ieet contents→\${m tents		path_2	form	at platform		iotal)	18 19 20 21 22 23 24 25 26 27 28 29						
nput Samplesh amplesheet cont List Ta	ieet contents→\$(m lents able	anifest}	path.2	form		NEW IT	TEM	18 19 20 21 22 23 24 25 26 27 28 29 30 31						
nput Samplesh samplesheet cont List Ta Actions	eet contents →\${m tents able sample	anifest}	μπ.2		ctory oxford	NEW IT	rem tep/mytax	18 20 21 22 23 25 25 26 27 28 29 30 31 31 32 33 34						
nput Samplesh amplesheet cont List Ta Actions	eet contents \${m able sample NB11	anlest) puth. 1 Alsers/membel/Documents/Projects/real- tione-reporting/data/R011 Auges/membel/Documents/Projects/real- Auges/membel/2015/11/staga 2015/11/staga 2015/	/Users/merribb1/Documents/Pro time- reporting/data/ERR6913101_2.f	direc direc ojects/real- file fastq.gz	ctory oxford	NEW IT database /Users/merribb1/Deskt	tetal) TEM top/mytaxi	18 20 22 22 23 24 25 25 25 25 26 27 27 28 29 30 31 32 33 34 33 34 35 35 37						
nput Samplesh samplesheet cont List Ta Actions & -	eet contents \${m tents able sample NB11 NB03	path. 1. Alsen/him/the1/Documents/Project/him- ine-reporting/data/NB11 Alsen/him/the1/Documents/Project/him- time-reporting/data/NB03 Alsen/him/the1/Documents/Project/him- time-	/Users/merribb1/Documents/Pro time-	direc direc ojects/real- fiastq.gz ojects/real- file	ctory oxford	database Alsens/merribb1/Deskt	tenal) TEM kop/mytax/ kop/mytax/	18 20 21 22 23 24 25 26 27 26 27 28 29 30 31 32 33 34 35 55 57 38 39						
nput Samplesh amplesheet cont List Ta Actions C - C -	teet contents→\$(m tents sable NB11 NB03 ERR6913101	anifest) amin.1 User/inviribil Documento/Projects/read- time-reporting/data/R611 User-inviribil Cocumento/Projects/read- time-reporting/data/R613 User-inviribil Documento/Projects/read- reporting/data/EFR6913101_11.1mst gz Users/inviribil Documento/Projects/read- time-reporting/data/EFR6913102_11.fastig gz Users/inviribil Documento/Projects/read- time-reporting/data/EFR6913102_11.fastig gz Users/inviribil Documento/Projects/read- time-reporting/data/EFR6913102_11.fastig gz	AUsers/merribb1/Documents/Pro time- reporting/data/ERR0913101_2.1 /Users/merribb1/Documents/Pro time-	direc direc ojects/real- fiastq.gz ojects/real- file	ztory oxford ztory oxford illumina	NEW IT database Alsers/menibb1/Deskt Alsers/menibb1/Deskt	rend nophnytax nophnytax nophnytax	18 19 20 21 22 22 23 24 25 26 27 28 30 31 32 33 34 34 35 35 36 35 39 40 41 42						
nput Samplesheet contraining amplesheet contr	able sample NB11 NB03 ERR6913102	anifest) Jash, 1 Alarsrikentbi 1.Documenta/Projectureation ne-reporting Idata/B11 Alarsrikentbi 1.Documenta/Projectureation- encerosting Idata/B03 Alarsrikentbi 1.Documenta/Projectureation- tencerosting Idata/B13 (1), 1 Istalig 21 Alarsrikentbi 1.Documenta/Projectureation- reporting Idata/BERBB/31 (2), 1 Istalig 21 Alarsrikentbi 0.Documenta/Projectureation- reporting Idata/BERBB/31 (2), 1 Istalig 21 Alarsrikentbi 0.Documenta/Projectureation- tencerosting Idata/BERBB/31 (2), 1 Istalig 21 Alarsrikentbi 0.Documenta/Projectureation- reporting Idata/BERBB/31 (2), 1 Istalig 21 Alarsrikentbi 0.Documenta/Projectureation- Alarsrikentbi 0.Documentbi 0.D	AUsers/merribb1/Documents/Pro time- reporting/data/ERR0913101_2.1 /Users/merribb1/Documents/Pro time-	direc direc ojects/real- fiastq.gz file fastq.gz	ctory oxford ctory oxford illumina illumina	NEW IT database Users/meribb1/Deskt Asers/meribb1/Deskt Asers/meribb1/Deskt	tendi) TEM kopimytaw kopimytaw kopimytaw kopimytaw	18 20 21 22 23 24 25 26 27 26 27 26 27 28 29 31 31 31 32 33 34 33 34 35 37 37 38 39 30 36 37 37 38 39 30 20 20 20 20 20 20 20 20 20 20 20 20 20						

If you need to edit a specific field directly within the UI, you can do so by selecting the pencil icon or adding a New Item



Within the dashboard, you can view all reports being generated in real time as the sunburst, sankey, table, or tops tab. These will be automatically started as soon as you access the website url and basestack has loaded the model as running.

Real Time Nanopore Report Analysis					DATA SHEET 🕞
Samples-	Sample Reports	RESET 🛟	Update Legend rank S1	Legend Placement Position <b>- side</b>	÷
30 Max Depth of Tax Tree 0 Min Depth of Tax Tree	TOP SUNBURST	SANKEY TABLE	Lagend Sample: flu_bc01 8 (NS) (42,2700 %), taxid: 10000 7 (M) (30,700 %), taxid: 10000C 5 (NP) (340 %), taxid: 10000C		Show all at           Legend           Sample: test           1 (PB2) (12.500 %), taxid: 1000           2 (PB1) (12.500 %), taxid: 1000           3 (PA) (12.500 %), taxid: 1000
0.08 Min Percent in Sample Tas Rank Codes K, R, R1, U, P, G, D, D1, O, C, S, F, S1, ~			3 (PA) (5 300 %), taxid: 100000 4 (HA) (3 900 %), taxid: 100000 6 (NA) (3 800 %), taxid: 100000		4 (HA) (12.500 %), taxid: 10000 5 (NP) (12.500 %), taxid: 10000 6 (NA) (12.500 %), taxid: 10000
52, 53, 54					

Warning: If the app doesn't start automatically classifying, try to refresh the page

At the top of the page on the right, you can adjust your samplesheet.

**Warning:** If run with Basestack, this will be limited in scope since you are running on data mounted with Docker. Any changes must be made with Basestack regarding adjusting data that is to be seen not from the Data Sheet directly.

OWNLOAD C	sv			PAUS	E UPDATES FLUSH RESTART REPORT RUN			NEW ITEM	ADVANCE
				reporting/data/INBU3					
0	/	• •	ERR6913101	/Users/merribb1/Documents/Projects/real-time- reporting/data/ERR6913101_1.fastq.gz	/Users/merribb1/Documents/Projects/real-time- reporting/data/ERR6913101_2.fastq.gz	file	illumina	/Users/merribb1/Desktop/mytax/minikraken2	
0	/	• •	ERR6913102	/Users/merribb1/Documents/Projects/real-time- reporting/data/ERR6913102_1.fastq.gz	/Users/merribb1/Documents/Projects/real-time- reporting/data/ERR6913102_2.fastq.gz	file	illumina	/Users/merribb1/Desktop/mytax/minikraken2	
0	/	• •	flu_bc01	/Users/merribb1/Documents/Projects/real-time- reporting/data/flu_BC01.fastq		file	oxford	/Users/merribb1/Desktop/mytax/flukraken2	
0	/	• •	sample	/Users/merribb1/Documents/Projects/real-time- reporting/data/sample_metagenome.fastq		file	oxford	/Users/merribb1/Desktop/mytax/minikraken2	
0	/	• •	test	/Users/merribb1/Documents/Projects/real-time- reporting/data/test2.fastq		file	illumina	/Users/merribb1/Desktop/mytax/flukraken2	
Barcodes	Barcod		covid_run	/Users/merribb1/Documents/Projects/real-time- reporting/data/example-run		run	oxford	/Users/merribb1/Desktop/mytax/minikraken2	
0	/	• •	barcode01	/Users/merribb1/Documents/Projects/real-time- reporting/data/example-run/demultiplexed/barcode01		directory	oxford	/Users/merribb1/Desktop/mytax/flukraken2	true
0	/	• •	barcode03	/Users/merribb1/Documents/Projects/real-time- reporting/data/example-run/demultiplexed/barcode03		directory	oxford	/Users/merribb1/Desktop/mytax/flukraken2	true
				File /Users/merribb1/Documents/Projects/real	l-time-reporting/data/example-run/demultiplexed/barco	de11/fastq_runi	i_77d58da2fa	1dfd2379616fbe43b369b2c89c93f5_0.fastq.gz has	
/data/Sample	sheet.csv		0 files	(0 B in total)	already been seen for sample:	barcode11			SHOW FU
								Rows per page: 10 ▼ 1-10 of 11	<

## **Basestack Consensus**

#### **Parameters**

#### Long Read Run Directory

[*Dir*] Run directory from a MinION, GridION, Mk1C, etc. Must contain several files or folders in it. See manifest, sequencing summary and fastq\_pass below

#### primers

[*Text/Dir*] Choice of Custom (directory) or pre-loadable options for primer set used (Artic only supported pre-loaded primer set currently)

#### barcoding

[*Text/File*] Select barcoding configuration used during demultiplexing. If demultiplexing didn't take place, any are allowed

#### basecalling

[*Text/File*] Select one of many supported basecalling configurations during the Basecalling step (creating fastq files from fast5)

#### fastq\_pass

[*Dir*] Directory of fastq files (can be demultiplexed or not). All fastq files to be analyzed MUST be decompressed (no .gz or .zip format allowed)

#### manifest

[List] Contains your ID to barcode mapping in a .tsv (tab-separated) format.

#### **Sequencing Summary**

[*File*] Not inputable. Ensure that it is in the top-level directory (root, same level as the Run Directory). It is required to run some portions. It is output at the end of every Basecalling step from Guppy

**Note:** Within manifest, one entry must contain the NTC (case-sensitive, no template control). If you don't have a NTC, select NB00 as the barcode and NTC as the id.

## Returns

#### Consensuses

[./artic-pipeline/4-draft-consensus/]

- Complete FASTA files will be output as ... complete.fasta for each barcode
- This folder and other sibling folders will contain other file formats such as .vcf and .bam for other down-stream analysis pipelines.

#### **Report of Run**

[./artic-pipeline/report.pdf]

• Contains important information about your samples (each barcode), lineage information, mutations, etc.

## **Running Consensus Generation and Reporting**

Consensus Generation is the main feature of this application and is used to generate a report of a run directory that was generated from a MinION run. It has multiple steps but is designed to be very automated once a job is submit for analysis.

If you haven't already done so, download and unzip the Test data folder

## Starting a Run

•••							
Z	Basestack	AUHL 🕱	PL - Artic Consensus			< > test-o	
e i					Favorites	test-data2	
Z	JHUAPL - Artic Consensus	⊠ 1	Artic Consensus - JHUAPL	0	<ul> <li>Recents</li> </ul>		
		Version select	Procedure select		🙏 Applications		
	Seport				Desktop	20200514_2000_ X3_FAN97e74b4	demux- fastq_pass
	- Inner de			-	merribb1  Documents		_
	Inputs		Source	Туре			
Å			Source	Type	Locations Box	minimap2	reference
Ä	Run Name 🧧	Test		string	⊟ Basestack 1.1.0 ≜		
		DRAG FOR	LDER HERE		🖂 Basestack 1.1.4 🛛 ≜		
	Input Long-Reads Full Run Directory			dir	⊟ Basestack 1.1.5		
		/Users/merribb1/Desktop/test-data2/20200514_2000_X3_	_FAN44250_697674b4		Basestack 1.1.1 ≜		
	Input Long-Reads FASTQ Directory	fastq_pass		dir	GIMP 2.10 Install ≜ Basestack 2.0.0 ≜		
		Select an item		Gill			
	Primer Configuration	nCoV-2019/V3			Tags • p12		
		Select an item		_	Red		
	Basecalling Configuration	dna_r9.4.1_450bps_hac.cfg 			Orange		
					Yellow		
	Barcode Configuration 🗧	barcode_arrs_nb96.cfg Select an item			Blue		
		List Table	NEW ITEM		Purple		
					<ul> <li>Gray</li> <li>All Tags</li> </ul>		
		barcode id	Actions				
*	Input Manifest	NB01 NTC	2 -	list			
		NB03 Sample3	B, —				
\$		NB11 Sample11	2 <b>–</b>				
۲			Rows per page: 5 1-3 of 3				
	Run Config File 🧧			file			
I							

- 1. Select the appropriate test folder first. This folder is either included in the *test-data* folder in the source of this application OR you can retrieve it within the install location of the app. For example, in *C:Program FilesBasestackclientdatatest-data*.
  - You can either drag + drop it into the *Run Folder* field or select it by left clicking and browser to the directory location on your computer
- 2. Once a folder has been input, you should see the text fields populate and turn green. If any field is marked as read, it is invalid and should be looked at further for proper formatting. These files should be valid for the test dataset. If you want to use your own data please follow the formatting in *run\_config.txt*, *run\_info.txt*, and *manifest.txt*. These formats are like:

Note: You may skip this portion if you'd like

- 1. run\_config.txt
- This should be 3 rows that dictate the primer (first), basecalling workflow (second), and barcoding cfg (third/last). Separate by tab.

Example:

Target (Unlisted)	Config (Unlisted)
primers	nCov-2019/V3
basecalling	dna_r9.4.1_450bps_hac.cfg
barcoding	barcode_arrs_nb96.cfg

#### 2. manifest.txt

Example:

Barcode	Sample
NB01	NTC (always required)
NB02	MDHP-00058
NB03	MDHP-00059
	•
•	•

Note: If you don't have an NTC (NOT RECOMMENDED EXPERIMENTALLY), set NTC as NB00

- This should be any number of rows that contain barcode on the left ALWAYS and the sample code on the right. A no-template-control (NTC) must always be specified for a report to be completed. Separate by tab.
- You are allowed to input your own custom values for each of the 3 files where the app will overwrite that corresponding file on a job submission. That means you can populate these fields by either directly modifying the files OR by inputting them into the input fields
  - These values are currently not validating to their greatest extent so take care to correctly input values and delimit them with tabs if doing this manually.

Lastly, there are three files that are made following a successful sequencing (and basecalling) run. These three are

#### 1. Sequencing Summary REQUIRED

- This file is made following **basecalling**. It contains the mapping and summary stats of all fast5 to fastq generations and must be present in the run directory for report generation
- If using CLI or stand-alone **basecalling** you will likely need to move this file from the fastq output savepath **to** the base run directory.
- 2. Throughput....csv **OPTIONAL**
- 3. Drift Correction OPTIONAL

Note: Future updates of Basestack will prevent the job from commencing if the sequencing summary is not present

## Starting the process

JHUAPL - Artic Consensus	I     Image: Artic Consensus - JHUAPL       Version select     Procedure select	- 0
1 Report		
Inputs		
Param	Source	Туре
Run Name 🧧	Test	string
Input Long-Reads Full Run Directory	/Users/merribb1/Desktop/test-data2/20200514_2000_X3_FAN44250_e97e74b4	dir
Input Long-Reads FASTQ Directory	fastq_pass Select an item	dir
Primer Configuration	nCoV-2019/V3 Select an item	
Basecalling Configuration	dna_r9.4.1_450bps_hac.cfg Select an item	
Barcode Configuration	barcode_arrs_nb96.cfg	

1. Once everything is staged, you should see all items update accordingly based on information in the directory.

Basestack	Z JHUAPL - Artic Consensus	
JHUAPL - Artic Consensus	Image: Market state     Image: Market st	0

2. Hit Start in the upper right-hand corner to start consensus generation.

**Note:** Depending on your method of installing Docker on Windows, you may receive a notification for docker to share a folder. Hit okay to allow the pipeline to continue. If you run Basestack as an admin, this error will be avoided. You can also opt to share the Basestack folder and sub-folders in the Docker Desktop on Windows as well (see how to do this in the next 2 images)

			<u> </u>	docker			<b>\$</b>	bmerrit	t1762	
ettings										
	🚟 General			5 File sharing						
	I Resources			ies (and their su u can check the		-		nted into Docker ls.		
	ADVANC • FILE SHA PROXIES	RING		erribb1\Docume ata\histories\tes		<sup>ıb\</sup> ⊝				
	NETWOR			erribb1\Docume ata\histories\Te		<sup>ıb\</sup> Θ				
	<ul> <li>Docker En</li> <li>Command</li> </ul>	-		erribb1\Docume ata\histories\trs		<sup>ıb\</sup> ⊝				
	Kubernete		C:\Users\me	erribb1\Docume	nts\GitHu					
			unpacked ta	ata\pipelines\sc	iserver\co	OVI				
-		<del></del>			bmorritt17	DVI	X			
ettings		dock			hmorritt17/	62	× vuild > win-unpact	ked v 🗸	Search win-ur	npacked
-	표 General	C:\Users\merribb1\Dr	er		New folder	62	x puild > win-unpact	Date modified	Туре	§== •
-	표 General I® Resources		er pcuments\GitHub\	chi     chi       Felect Folder       ←     →       Organize ▼       O       →     Quick acce       O     >	konserveitet 17     k	mp > client > t	vuild > win-unpact		Type File folder File folder	§== •
-	Resources     ADVANCED     FILE SHARING	C:\Users\merribb1\Do	er Documents\GitHub\	Felect Folder       ←     →       Organize ▼       ○        ↓        Box	New folder	mp > client > t Name 1 data 1 locales	vuild > win-unpac	Date modified 10/2/2020 12:37 A 10/2/2020 12:32 A	Type 	§== •
-	Resources     ADVANCED     FILE SHARING     PROXIES     NETWORK	C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr	er ocuments\GitHub\ ocuments\GitHub\ ocuments\GitHub\		New folder	mp > client > b lame l data l locales resources	build > win-unpad	Date modified 10/2/2020 12:37 A 10/2/2020 12:32 A 10/2/2020 12:33 A	Type 	§== •
	Resources     ADVANCED     FILE SHARING     PROXIES	C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr	er pocuments\GitHub\ pocuments\GitHub\ pocuments\GitHub\ pocuments\GitHub\	Image: Second Secon	New folder	mp > client > b iame data locales resources switshader	xuild > win-unpac	Date modified 10/2/2020 12:37 A 10/2/2020 12:32 A 10/2/2020 12:33 A	Type 	
	Resources     ADVANCED     FILE SHARING     PROXIES     NETWORK     Docker Engine	C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr C:\Users\merribb1\Dr	er couments/GitHub/ couments/GitHub/ couments/GitHub/ couments/GitHub/ couments/GitHub/ couments/GitHub/	elect Folder       elect Folder       ← → ↓ ↑       Organize ↓       Organize ↓       → ↓ Quick acce       → ↓ ⊕ Docume       → ↓ ⊕ Docume       → ↓ ⊕ Docume       → ↓ ⊕ Docume       → ↓ ⊕ Music	New folder	mp > client > b kame data locales resources switshader	xuild > win-unpad	Date modified 10/2/2020 12:37 A 10/2/2020 12:32 A 10/2/2020 12:33 A	Type 	iii • Size

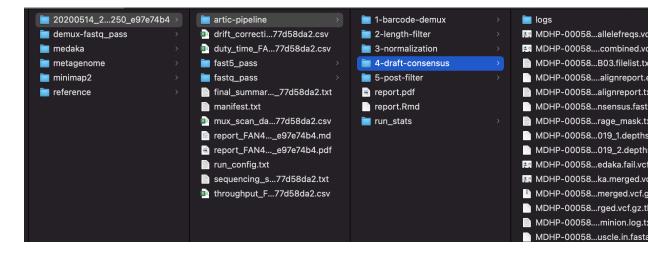
**Note:** Simply select the folder that contains the *Basestack.exe* file by selecting the plus-mark and navigation and selecting it within the browser. In this example it is: ... *buildwin-unpacked* 

## **Checking Logs and Status**

BOPEN LOG FOLDER
stdout: [2022-02-15 20:21:39] RUN data: sequencing run folder: [1;36m/opt/data[Um
stdout: [2022-02-15 20:21:39] RUN data: recording software version numbers
stdout: 🕏 [2022-02-15 20:21:39] RUN data: guppy_barcoder, part of Guppy basecalling suite, (C) Oxford Nanopore Technologies, Limited. Version 4.2.2+effbaf8
stdout: [2022-02-15 20:21:39] RUN data: run configuration file: /opt/data/run_config.txt
stdout: [2022-02-15 20:21:39] RUN data: run manifest file: /opt/data/manifest.txt
stdout: [2022-02-15 20:21:39] RUN data: inputs: fastq_directory: /opt/data/fastq_pass, arrangements files: barcode_arrs_nb96.cfg
stdout: [2022-02-15 20:21:39] RUN data: output demultiplex directory: /opt/data/artic-pipeline/1-barcode-demux
stdout: [2022-02-15 20:21:39] RUN data: processing pipeline output
stdout: >[2022-02-15 20:21:39] RUN data: Starting guppy demux module 1
stdout:
stdout: ONT Guppy barcoding software version 4.2.2+effbaf8 input path: /opt/data/fastq_pass save path: /opt/data/artic-pipeline/l-barcode-demux arrangement files: barcode_arrs_nb96.cfg lamp arr. files: barcode_arrs_lamp8.cfg barcode_arrs_lamp96.cfg min. score front: 60 min. score rear: 60
stdout:

**Note:** You can see the output of the run in the *Log Window* container on the bottom of the page. You can also see the *Output(s)* table begin to change as modules are completed for your run. The final module is the report generation module and should always be 1/1 when complete

## **Final Report**

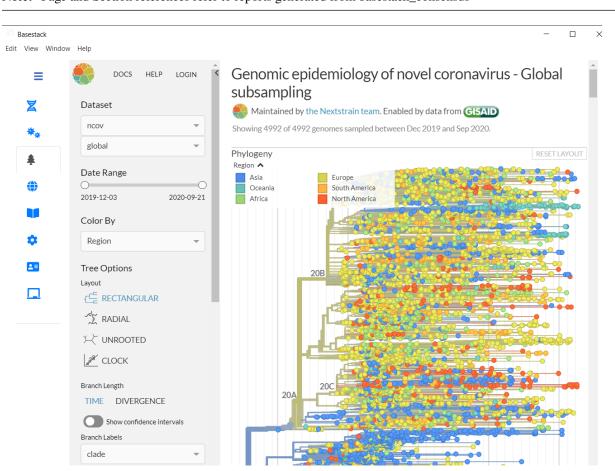


**Note:** Once complete, you can view the pdf report by clicking the pdf file icon link underneath the final row's status of 1/1. You can also traverse to any of the module directories by hitting the link text on the first column for each module. In this example, I've chosen *Report Generation* as my link which is a top-level view of all modules, as well as the *report.pdf* location. Open this pdf to see your report either from the folder or the *pdf* link on the left-most column to see your results!

## Nextstrain

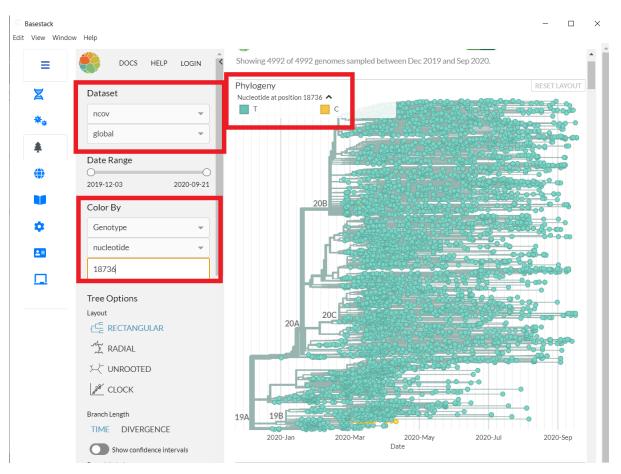
After the app has generated consensus genomes from the previous 2 commands, you can view mutations or SNPs in your run by selecting the Nextstrain tab at the left-hand side. These mutations can be viewed around Page 11 or section 5.3

Let's go into our report we just generated and traverse to section 5.3 or page 10-11. The tables provided are all reported mutations against the reference.



**Note:** Page and Section references refer to reports generated from basestack\_consensus

- Position 18736 (for example) is a reported SNP for T to C mutation. The annotatin is a missense\_variant (see all annotation types in the description of this figure for the report). Though nextstrain we can view this mutation across all samples available that have been sequences and input into nextstrain's website.
- Now that we've selected our mutation to view let's move forward through the interactive site



To View a specific mutation (and this one for example) reported in the table, select underneath Color By

Genotype -> nucleotide -> [Your number here]

Also, be sure that the dataset is nov and global. You can change the layout of the visualization(s) with the Tree Options parameters

Note: Requires Internet. Also available at https://nextstrain.org/ncov/global

#### IGV

Note: Requires Internet. Also available at https://igv.org/app/

IGV is an interactive environment that allows you to view a genome and see any annotations at a specific position.

These plots are also available in the Consensus Generation Pipeline Reports

← → C	v.org/app/ ···· 🗵 🏠	
Genome 🔹 Tracks 🔹 Session 🔹	Share Bookmark Save SVG Help <del>*</del>	
Human (GRCh38/hg38) Human (CRCh37/hg19)		Cursor Guide Track Labels
Human (hg18) Mouse (GRCm38/mm10) Mouse (NCBI37/mm9)	5         6         7         8         9         10         11         12         13         14         15         16         17         18         19         20	0 21 22 X Y
Rat (RGCS 6.0/rn6) Gorilla (gorGor4.1/gorGor4)		*
Chimp (SAC 2.1.4/panTro4)		
Bonobo (MPI-EVA panpan1.1/panPan2) Dog (Broad CanFam3.1/canFam3)		
Cow (UMD_3.1.1/bosTau8) Cow (ARS-UCD1.2/bosTau9)		
Pig (SGSC Sscrofa11.1/susScr11)		
Zebrafish (GRCZ11/danRer11) Zebrafish (GRCZ10/danRer10)		
C. elegans (ce11) D. melanogaster (dm6)		
S. cerevisiae (sacCer3)		
S. pombe (ASM294v2) Sars-CoV-2 (ASM985889v3)		

1. First, we are going to select our genome of interest, Sars-CoV-2.

$\overleftarrow{\leftarrow} \rightarrow$	ଟା	Ъ		0		https	://igv	.org	/app	/															••••	⊌	☆			١I١	•	۹	≡
Geno	me 🔻	Tra	acks		Sessi	ion `		Sh	are		Boo	kmar	k	Sa	ve S	SVG	н	elp '	•														
IGV	ASM98	5889v3	NC	_0455	512.2	~	NC	_045	512.	2:18,	721-	18,759	Q	39 b	p				С	ursor G	Guide	Cen	ter Lir	e	Track	Labels		) -				• (	
	$\square$																															)	
	bp		18,725		_		8,730		_		-	735 bp			18	,740 bp				745 bp			18	750				8,755	-		1		
	C A	т	C A	т	тс	т	•	т	т	G	G	• т	. 1		G	^		^	c	G 1	r o	Т	^	т	^	<b>^</b> 1	r c	c	G	т	тт	1	¥
	Anno	tations																														14	2
	)	) )	)	)	)	)	>		-										>	>	•	•	•	•	>	•	>	>	>	)	)	>	
								na	me	ORF	-1ab							~															
								typ	be C	DS																							
								ID	cds	-YP	0097	24389	).1																				
								Pa	rent	ger	ne-Gl	J280_	gp01																				
								Db	xref	Ge	nbanl	k:YP_	0097	24389	9.1,0	GenelD	:4374																
								No	te p	p1al	5%3E	3 trans	lated	d by -1	l ribo	osoma	l fra																
								ex	cept	ion	ribos	omal	slipp	age																			
								gb	key	CDS	3																						
								ge	ne (	ORF	1ab																						
								loo	cus_t	ag	GU28	30_gp(	01																				
								pro	oduc	t O	RF1a	b poly	prote	ein																			
								pro	otein	_id	YP_	00972	4389	.1																			
								ро	sitio	n N	C_04	5512.2	2:266	5-21,5	55																		
								na	me:	OR	F1ab							-															
																		Ŷ															
igv ig	v.org																									U	JCS	ban	Die	go	<b>ĕ</b> I	BRO	AD

2. For our example at position 18736, we can see that it belongs to ORF1ab from the reference and it has a T. As we saw in the report that T was change to a C when compared to the reference. We can also see what protein this

position is a part of as well as supplemental information like the *protein\_id* or *Dbxref* (references for external database resource)

## GAMMA

GAMMA is a tool designed to analyze gene allele mutations for microbes, primarily bacterial genomes.

Note: This module outputs only a text file for analysis

## **Parameters**

#### **FASTA file**

[File] FASTA or genome file to check for AMR alleles

#### output

[String] Text string to have as a prefix for the output file(s)

#### Gene Database file

[File] . fsa file that contains the mapping of genes linked to AMR

#### Returns

GAMMA file : Tab-separated file

Note: Below output gathered from here

The default output of GAMMA is a tab-delimited file with a ".gamma" extension with 15 columns:

- 1. Gene The name of the closest matching gene (target) from the database. If there are ambiguous gene matches (i.e., multiple target matches with the same number of non-degenerate codon changes, basepair changes, and transversions), the gene match will be appended with a "‡".
- 2. Contig The name of the contig on which the match was found.
- 3. Start The start position of the sequence matching the gene on the contig.
- 4. Stop The end position of the sequence matching the gene on the contig.
- 5. Match\_Type The type of the gene match based on the translation of the sequence (i.e., the protein sequence). Can be native (for identical amino acid sequences to the target), mutant (for nonsynonymous mutations), truncation (for nonsense mutations), indels (for insertions/deletions), nonstop (for a missing stop codon), contig edge (for matches that are truncated at the start or stop of a contig), or a combination of multiple types (i.e., indel truncation).
- 6. Description A short description of the match type.
- 7. Codon\_Changes The count of the non-degenerate codon changes in the sequence versus the closest match from the datbase.
- 8. BP\_Changes The count of the basepair changes in the sequence versus the closest match from the datbase.
- 9. Transversions The count of basepair changes that are transversions (i.e., purine to pyrimidine or vice versa, such as an A -> C or a T -> G)
- 10. Codon\_Percent The percent (expressed as a decimal value) of the degenerate codon similarity between the query and match sequence. Gene matches with large insertions may show a negative value.

- 11. BP\_Percent The percent (expressed as a decimal value) of the basepair similarity between the query and match sequence. Gene matches with large insertions may show a negative value.
- 12. Percent\_Length The percent (expressed as a decimal value) of the length of the target covered by the matching sequence, maximum of 1.
- 13. Match\_Length The length (in basepairs) of the matching sequence.
- 14. Target\_Length The length (in basepairs) of the target sequence.
- 15. Strand The sense of the strand (+ or -) on which the match is found.

#### Rampart

Rampart is an annotation tool provided by the Artic Network that gives quick, but less accurate reports for each of your demuxed samples. It is ideal for a quick look into what the annotations for each of your barcodes are but should not be fully relied on, instead opt for the pipeline that was just discussed.

Note: This module contains a UI to be displayed from the rendering button in the variables table

#### **Parameters**

#### **Fastq Dir**

[Dir] Run annotation of SARS-nCoV-2 sequences within the RAMPART suite

#### **Annotated Clear**

[*Option*] Remove your annotated folder if you want to do a fresh run. Annotations will be removed from the same directory as the Fastq Dir

#### Returns

#### Annotations

[./annotations]

- Select Click Me! button to render RAMPART in the UI
- 1. Ensure you've loaded a run that has been bookmarked from the previous tab (described above).
- 2. Select one of the included primer-schemes from the drop-down list. For this example, the data is ncov-related so we will choose Default Genome fasta for SARS-nCoV-2.
- 3. Select the Click Me! button to render Rampart in a new window

				Favorites	
Rampart		AMPART	<b>O</b> ×	<ul><li>AirDrop</li><li>Recents</li></ul>	
1 RAMPART				Applications	тхт
puts				merribb1	run_config.txt
Param	Source		Type	Documents Locations	
Served Port			render	🖮 Box 🖂 Basestack 1.1.0	d fast5_pass 2
Protocol JSON     T	Artic Default Protocol Select an item		dir	<ul> <li>Basestack 1.1.4</li> <li>Basestack 1.1.5</li> </ul>	A
Genome Reference	Default Genome fastas for SARS-nCoV-2 Select an item			<ul> <li>Basestack 1.1.1</li> <li>⊖ GIMP 2.10 Install</li> </ul>	drift_correction_F M AN4425da2.csv
🚱 Clear Annotated 🏾 🗧	Clear Annotated Select an item			Basestack 2.0.0	
	DRAG FOLDER HERE	]		Tags • p12	
🕜 Fastq Directory ा 📷	/Users/merribb1/Desktop/test-data2/20200514_2000_X3_FAN44250_e97e74b4/	astq_pass	dir	<ul><li>Red</li><li>Orange</li></ul>	
	L	Rows per page: 10 1-5 of 5	O O H	<ul><li>Yellow</li><li>Blue</li></ul>	
				<ul><li>Purple</li><li>Gray</li></ul>	
utput Locations	label Sta		TE OUTPUTS	<ul> <li>Gray</li> <li>All Tags</li> </ul>	
	Annotations      O				
		Rows per page: 5 1-1 of	1		
Initiating logging object	reetary' emabled	RANFART dateson running	•		
ZOPEN LOG FOLDER Initiating logging object stdout: @Flag: "Clearing annotation d Basestack	Research	RAMPART Read Assignmen	RA t, Mapping, a	IMPART nd Phylogenetic Analysi	s in Real Time
Taitlating Logging object asdout: @riag: "Clearing association d Basestack	Basestack		RA t, Mapping, a 5.558Z sulating rate   [	nd Phylogenetic Analysi Data last received 08s ago	s in Real Time
Initiating logging object stdout: @Flag: 'Clearing annotation d.	Basestack	RAMPART Read Assignmen Experiment: Started @ 2022-02-15T20:54-54 18753 reads mapped   19526 processed   cal	RA t, Mapping, a 5.558Z sulating rate   [	nd Phylogenetic Analysi Data last received 08s ago mapped, 19526 processed) ds/sec	
Initiating logging object andwork @riag: "Clearing anostation d Basestack mpart I RAMPART US	Essestack	RAMPART Read Assignmen Experiment: Started @ 2022-02-15T20:54-54 18753 reads mapped   19526 processed   cal	RA t, Mapping, a .5582 culating rate   [ -113864s, 18753 r ed reads • Rea	nd Phylogenetic Analysi Data last received 08s ago mapped, 19526 processed) ds/sec 23 Mapped Reads / 1 14000 12000	
Initiating logging object andourt #Fing: "Clearing anostation d Basestack mpart RAMPART	Essestack	RAMPART Read Assignmen Experiment: Started @ 2022-02-15120-154-55 18753 reads mapped 19526 processed   cal	R# t, Mapping, a 55582 119884s, 18753 n rd reads rd reads 	nd Phylogenetic Analysi Data last received 08s ago mapped, 19526 processed) ds/sec 23 Mapped Reads / 3 13000 12000	Sample 20 Reference i r.Car2019W
Initiating Logging object products #Filey "Clearing anostation of Basestack mpart PRAMPART JUS Proven Served Port @ Probool JSON = A	Essestack	RAMPART Read Assignmen Experiment: Started @ 2002-02-15720-54.56 18753 matis mapped [19526 processed [ and @ server messages 15:56:15 new data (tr Read Depth 21 Marphone) Read Depth 21 Marphone)	R# t, Mapping, a 55582 119884s, 18753 n rd reads rd reads 	nd Phylogenetic Analysi Data last received OBs ago mapped, 19526 processed) ds/sec \$ Mapped Reads / 1 10000- 10000- 400-	Sample 🕄 Reference i ricovooryw
Initiating logsing object anderst #Fings "Clearing anostation of Basestack mpart RAMPART US Bernet Port Served Port	Basestack	RAMPART Read Assignmen Experiment: Started @ 2002-02-15720-54-56 18753 reads mapped   19526 processed   call © Server messages 155.615 nev data te Read Dupti © Mupped 1000 Mupped	R# t, Mapping, a, .5582 culating rate   [ 13844, 13753 xd reads • Rea xd xd xd	nd Phylogenetic Analysi Data last received OBs ago mapped, 19526 processed) ds/sec 55 Mapped Reads / 12000 10000 acco 4000 2000	Sample 20 Reference i r.Car2019W
Initiating loging object rideut: #Tiep: "Clearing anotation di Basestack mpart RAMPART Param Served Port Poteon JSON = A Clearone Reference = C	It Default Protocol ist an time afault Genome fastas for SARS-mCoV-2	RAMPART Read Assignmen Expression and States (2002-02-15120-54-56 10753 reads mapped) 19526 processed (adl © Server messages 15-561) nev date (1 Read Dupt) (1 Mapped) (1	RA t, Mapping, a. 25582 culating rate   C 13844, 13753 ed cads 2555 reads m	nd Phylogenetic Analysi Data last received OBs ago mapped, 19526 processed) ds/sec \$ Mapped Reads / 1 10000- 10000- 400-	Sample C Reference : 
Initiating loging object rideut: #Tiep: "Clearing anotation di Basestack mpart RAMPART Param Served Port Poteon JSON = A Clearone Reference = C	Researce	RAMPART Read Assignment Dis753 reads mapped 19526 processed of all of server reasons i 155.01 ° nev data te read Droft ° ° Server reasons i 156.01 ° nev data te read Droft ° ° Server reasons i 156.01 ° nev data te reasons i 156.01 ° 160.01 ° 160.00 ° 160.	R# t, Mapping, a. 5582 113844, 18753 ad reads 2555 reads m 2196 reads m 14002 reads m	nd Phylogenetic Analysi Data last received OBs ago mapped, 19526 processed of sec on on-two on on-two on-	Sample CC Reference : 
Excitating logging object addet: #Fiag: "Clearing antotation d Basestack npart RAMPART Served Port @ Protocol_SON	Resentack	RAMPART Read Assignmen Experiment: Started & 2002-02-15720.5454 19535 reads mapped   19526 processed   cal © Server messages 1:55:15 new data be Read Dupt	R/ t, Mapping, a, 5582 culating rate   C 13844, 1973 e 0 0 0 19845, 1973 e 2555 reads m 2196 reads m 14002 reads m 0 reads m	nd Phylogenetic Analysi Data last received 08s ago mapped, 1922 processed do/sec	Sample 10 Reference ncovzotryw. d <sup>20</sup> d <sup>30</sup> 0 reads/sec 0 reads/sec 0 reads/sec 0 reads/sec
Initiating logsing object seducts #Flags "Clearing anotaction of Basestack mpart RAMPART US Served Post © Protocol JSON = © Genome Reference = © Genome Reference = © Genome Reference =		RAMPART Read Assignment Dis753 reads mapped 19526 processed of all of server reasons i 155.01 ° nev data te read Droft ° ° Server reasons i 156.01 ° nev data te read Droft ° ° Server reasons i 156.01 ° nev data te reasons i 156.01 ° 160.01 ° 160.00 ° 160.	R/ t, Mapping, a, 5582 culating rate   C 13844, 1973 e 0 0 0 19845, 1973 e 2555 reads m 2196 reads m 14002 reads m 0 reads m	nd Phylogenetic Analysi Data last received 08s ago mapped, 1922 processed do/sec	Sample 10 Reference ncovzotryw. d <sup>20</sup> d <sup>30</sup> 0 reads/sec 0 reads/sec 0 reads/sec 0 reads/sec
Initiating logsing object padents #Flags "Clearing anotaction d Basestack mpart RAMPART Braven Post © Protocol JSON = P © Genome Releience = 0 © Genome Releience		RAMPART Read Assignmen Experiment: Started & 2002-02-15720.5454 19535 reads mapped   19526 processed   cal © Server messages 1:55:15 new data be Read Dupt	R/ t, Mapping, a, 5582 culating rate   C 13844, 1973 e 0 0 0 19845, 1973 e 2555 reads m 2196 reads m 14002 reads m 0 reads m	nd Phylogenetic Analysi Data last received 08s ago mapped, 1922 processed do/sec	Sample 10 Reference ncovzotryw. d <sup>20</sup> d <sup>30</sup> 0 reads/sec 0 reads/sec 0 reads/sec 0 reads/sec

## FastQC

FastQC is a tool designed to view the quality and do control on raw sequencing data, primarily for illumina reads

Note: This module outputs an html file that is a report of your raw sequencing data

## **Parameters**

#### **FASTQ-containing Directory**

[Directory] Directory that contains one or more fastq files

#### Returns

Report file : *HTML file* 

# *Report* **Report**

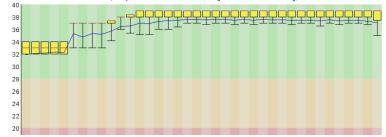
#### Summary

Basic Statistics
Per base sequence quality
Per sequence quality scores
OPer base sequence content
Per sequence GC content
Per base N content
Sequence Length Distribution
Sequence Duplication Levels
Overrepresented sequences
Adapter Content

Øв	Basic Statistics								
	Measure	Value							
	Filename	full.fastq							
	File type	Conventional base calls							
	Encoding	Sanger / Illumina 1.9							
	Total Sequences	293832							
	Sequences flagged as poor quality	0							
	Sequence length	35-151							
	%GC	38							

## Per base sequence quality

Quality scores across all bases (Sanger / Illumina 1.9 encoding)



FastQC Report				
Summary	Overrepresented sequences			
•	Sequence	Count	Percentage	Possible Source
Basic Statistics	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	3832	1.3041465871654552	No Hit
Per base sequence quality	ทพทพทพทพทพทพทพทพทพทพทพทพทพทพทพทพทพทพ	NNN 879	0.2991505349995916	No Hit
Per sequence quality scores	GTATATGAGATCTCTCAAAGTGCCAGCTACAGTTTCTGTTTCTTCAC	CTG 451	0.15348906858340822	No Hit
	GTGTAAGACGGGCTGCACTTACACCGCAAACCCGTTTAAAAACGATT	GTG 349	0.11877535462441124	No Hit
Per base sequence content	CTCCAAGGGTGTTCACTTTGTTTGCAACTTGCTGTTGTTGTTGTAA	CAG 347	0.11809469356639168	No Hit
Per sequence GC content	ACTTTAGACTGACTCTTGGTGTTTATGATTACTTAGTTTCTACACAG	GAG 308	0.10482180293501049	No Hit
Per base N content	GATTAAAGATTGCTATGTGAGATTAAAGTTAACTACATCTACTTGTG	CTA 303	0.1031201502899616	No Hit
Sequence Length Distribution				
Sequence Duplication Levels				
Overrepresented sequences	Adapter Content			
Adapter Content		% Adapter		
	100		Illumi	na Universal Adapter
	90			ha Small RNA 3' Adapter
	90			na Small RNA 5' Adapter Ira Transposase Sequeni
	80			Small RNA Adapter

## NanoPlot

NanoPlot is a tool designed to view the quality and do control on raw sequencing data from long-read runs

Note: This module outputs an html file that is a report of your raw sequencing data

#### **Parameters**

#### • input file: *file*

Can be one of these list:

fastq file [file ...]

Data is in one or more default fastq file(s).

2. fasta file [file ...]

Data is in one or more default fasta file(s).

3. fastq\_rich file [file ...]

Data is in one or more fastq file(s) generated by albacore or MinKNOW with additional information concerning channel and time.

4. fastq\_minimal file [file ...]

Data is in one or more fastq file(s) generated by albacore or MinKNOW with additional information concerning channel and time. Minimal data is extracted swiftly without elaborate checks.

5. summary file [file ...]

Data is in one or more summary file(s) generated by albacore or guppy.

6. bam file [file ...]

Data is in one or more sorted bam file(s).

Note: Input information taken from the --help output of NanoPlot

• barcoding file: optional - file Must only be given when using the summary file option

## **Returns**

Report file : HTML file

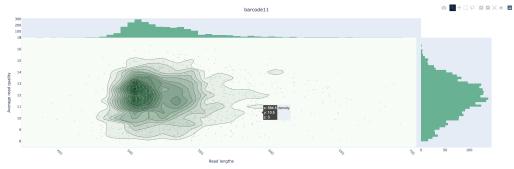
Summary Statistics Plots

Report issue on Github

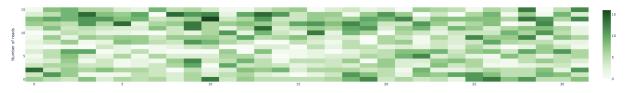
# **NanoPlot reports**

Summary statistics

		•					
General summary	barcode11	unclassified	barcode03	barcode01	None	None	None
Active channels	479.0	490.0	500.0	3.0	None	None	None
Mean read length	526.4	495.5	513.3	422.7	None	None	None
Mean read quality	11.7	12.2	12.4	10.8	None	None	None
Median read length	519.0	492.0	508.0	423.0	None	None	None
Median read quality	11.8	12.3	12.4	10.8	None	None	None
Number of reads	2,683.0	4,192.0	12,648.0	3.0	None	None	None
Read length N50	521.0	493.0	508.0	423.0	None	None	None
STDEV read length	29.9	23.1	33.6	7.8	None	None	None
Total bases	1,412,359.0	2,077,197.0	6,492,511.0	1,268.0	None	None	None
Number, percentage and megabases of reads above quality cutoffs					None	None	None
>Q5	2683 (100.0%) 1.4Mb	4192 (100.0%) 2.1Mb	12648 (100.0%) 6.5Mb	3 (100.0%) 0.0Mb	None	None	None
>Q7	2683 (100.0%) 1.4Mb	4192 (100.0%) 2.1Mb	12648 (100.0%) 6.5Mb	3 (100.0%) 0.0Mb	None	None	None
>Q10	2278 (84.9%) 1.2Mb	3888 (92.7%) 1.9Mb	12073 (95.5%) 6.2Mb	2 (66.7%) 0.0Mb	None	None	None
>Q12	1199 (44.7%) 0.6Mb	2426 (57.9%) 1.2Mb	7933 (62.7%) 4.1Mb	0 (0.0%) 0.0Mb	None	None	None
>Q15	33 (1.2%) 0.0Mb	93 (2.2%) 0.0Mb	281 (2.2%) 0.1Mb	0 (0.0%) 0.0Mb	None	None	None
Top 5 highest mean basecall quality scores and their read lengths					None	None	None



barcode11



## **Viral Recon**

nf-core/viralrecon is a bioinformatics analysis pipeline used to perform assembly and intra-host/low-frequency variant calling for viral samples. The pipeline supports both Illumina and Nanopore sequencing data. For Illumina short-reads the pipeline is able to analyse metagenomics data typically obtained from shotgun sequencing (e.g. directly from clinical samples) and enrichment-based library preparation methods (e.g. amplicon-based: ARTIC SARS-CoV-2 enrichment protocol; or probe-capture-based). For Nanopore data the pipeline only supports amplicon-based analysis obtained from primer sets created and maintained by the ARTIC Network.\*

\*Pulled from [https://nf-co.re/viralrecon{]}(https://nf-co.re/viralrecon)

Note: The modules runs nextflow on the backend and thus utilizes Docker within Docker.

#### **Parameters**

#### Fastq Dir

[Dir] Basecalled Fastq files

#### Fast5 Dir

[Dir] Fast5 files directory from which you received the basecalled fastq directory of files from

## Returns

#### Consensus

[./viralrecon/medaka|nanopolish]

• Consensus FASTA files are made for both assembly processes

#### MultiQC Report: ./viralrecon/multiqc/multiqc\_report.html

• HTML files that has information of your run

1. Ensure you've loaded a run with a fastq and fast5 directory specified

		Contraction > Desktop	
Select an item  Select FOLDER OR DRAG HERE  Artic Minion Caller  Specific nf-core revision of viralrecon  Specific nf-core revision of viralrecon	Favorites Favorites AirDrop Recents Applications	Desktop Name V 10 test-data2 V 10 20200514_2000N44250_e97e74b4	Date Nov Oct
2.4.1 Fastq pass directory generated from Guppy GPU/CPU basecalling   SELECT FOLDER OR DRAG HERE	<ul> <li>Desktop</li> <li>merribb1</li> <li>Documents</li> <li>Locations</li> <li>Box</li> </ul>	<pre>&gt; fastq_pass &gt; NanoPlot &gt; demultiplexed</pre>	Dec Nov Oct
Fast5 pass directory generated from sequencing, pre-basecall	<ul> <li>Network</li> <li>Tags</li> <li>p12</li> <li>Red</li> </ul>	<ul> <li>artic-pipeline</li> <li>consensus</li> <li>manifest.txt</li> <li>fast5_pass</li> </ul>	Aug Aug Aug Jul
Select your sequencing summary file output from basecalling from duppy. Ger	erated alo Orange	O files (0 B in total)	

- 2. Select one of the included primer-schemes from the drop-down list. For this example, the data is ncov-related so we will choose Default Genome fasta for SARS-nCoV-2.
- 3. Select one of the basecaller options *medaka* or *nanopolish*
- 4. Select the Play button button to start the pipeline

#### **Bamstats**

Bamstats is a tool designed compute mapping stats from a BAM file, often output from alignment methods

Note: This module outputs a zipped folder that contains your necessary information

#### **Parameters**

• input BAM: file

This is often output from alignment methods like bwa, bowtie2, or minimap2. Additionally, you can receive BAM files from modules within Basestack such as:

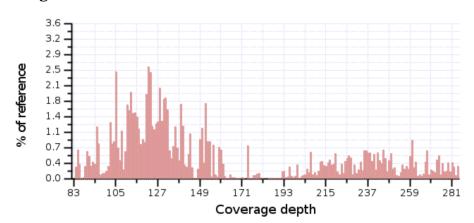
- 1. Basestack Consensus
- 2. Medaka
- 3. Nanopolish

## Returns

Bamstats Report zipped : bamstats\_report.zip, HTML-containing directory

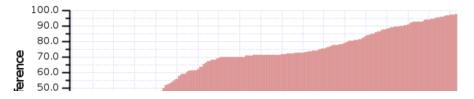
# MN908947.3: Coverage distribution

Id	N	Mean	Median	sd	q1	q3	2.5% percentile	97.5% percentile	Min	Max
MN908947.3	29,903	157.91	132.00	60.61	115.00	214.00	84.00	284.00	0.00	320.00



# Histogram

# **Cumulative histogram**



## **Artic Field Bioinformatics**

Artic is a tool designed for analyzing and creating consensuses from viral nanopore sequencing reads. These also are tied to the use of amplicon schemes (tiled)

**Note:** This module will output multiple consensuses per barcode of interest. It takes in a raw fastq\_pass directory and outputs *.vcfs*, *.fasta*, and *.bam* files for further use

## Medaka

## **Parameters**

• input run folder: *directory* 

Must contain a fastq\_pass directory or a custom inputted one (See below)

• Primer: option or Directory

Artic Primer set or a custom one (directory) which contain the genome.fasta and necessary BED file for the primer set

- Normalize Coverage: Number
- Barcode Configuration: option

Which barcode kit you used for demux. Select any for non-barcoded sample

• FASTQ Dir: Directory, optional

Select your own custom *fastq\_pass* directory to analyze and demux

• Medaka Model: option

FAST or HAC used during basecalling

## Returns

- 1. Medaka Consensus files: FASTA files containing your consensuses for each barcode
- 2. Medaka VCF files: Variant files containing your variant calls for each barcode
- 3. Medaka BAM files: BAM files containing your alignment information for each barcode

Artic		1		Medaka			- 0
Artic		Version select		Procedure select			- 0
Demultiplex Barcodes			/Plex			- 😢 Med Enabl	aka ed 🗸 🎽
nputs		O	utput Locations		â	DELETE O	JTPUTS
nput Long-Reads Full Run Directory Directory that contains your fastq_pass, fast5_pass, and sequencing	Required dir	•	Label	Access	Completed		Remove
ummary.txt file. Output directory from a Oxford Nanopore run			BAM file		3		Ē
Users/merribb1/Desktop/test-			Variants		2		Ē
ata2/20200514_2000_X3_FAN44250_e97e74b4		-	Consensus		2		Ī
Primer Configuration Default Primer Scheme to use for nanopolish	Deviced			Rows per page:	5	1-3 of 3	
CoV-2019/3	Required	-	PEN LOG FOLDER				
elect an item	_						
arcode Configuration /hich barcoding configuration you used. If no demux took place, selec ny	ct Required	•					
arcode_arrs_nb96.cfg	_	_					
elect an item							
elect an item		8 ≔ (	□	₫ 🖉 ⊕•	Q Search		
		₩ 📰 (	□□	∧ Date Modifie		Size	Kind
> consensus_medaka Name barcode01_consensus_medaka.alignreport.er		8 ≔ (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM	61 bytes	Document
Name barcode01_consensus_medaka.alignreport.er barcode01_consensus_medaka.alignreport.txt		88 ∷≣ (		A Date Modifie Mar 2, 202 Mar 2, 202			Document Plain Text
consensus_medaka           Name           barcode01_consensus_medaka.alignreport.er           barcode01_consensus_medaka.alignreport.txt           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.merged.vcf.gz		88 ≔ (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes	Document Plain Text gzip coa Document
consensus_medaka           Name           barcode01_consensus_medaka.alignreport.er           barcode01_consensus_medaka.alignreport.txt           barcode01_consensus_medaka.areged.vcf.gz           barcode01_consensus_medaka.areged.vcf.gz           barcode01_consensus_medaka.areged.vcf.gz           barcode01_consensus_medaka.areged.vcf.gz		88 ☷ (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB	Document Plain Text gzip coa Document Plain Text
consensus_medaka           Name           barcode01_consensus_medaka.alignreport.er           barcode01_consensus_medaka.alignreport.txt           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.minol.gt.xt           barcode01_consensus_medaka.nCoV-2018_1.hdf		88 := (		Date Modifie Mar 2, 202 Mar 2, 202 Mar 2, 202 Mar 2, 202 Mar 2, 202 Mar 2, 202 Mar 2, 202	d 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB	Document Plain Text gzip coa Document Plain Text TextEdit
consensus_medaka Name barcode01_consensus_medaka.alignreport.er barcode01_consensus_medaka.alignreport.txt barcode01_consensus_medaka.merged.vcf.gzt barcode01_consensus_medaka.merged.vcf.gztbi barcode01_consensus_medaka.minion.log.txt barcode01_consensus_medaka.nCoV-2019_1.hdf barcode01_consensus_medaka.nCoV-2019_1.vcf barcode01_consensus_medaka.nCoV-2019_2.hdf barc		88 := (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 6 KB	Document Plain Text gzip coa Document Plain Text TextEdit vCard TextEdit
consensus_medaka Name barcode01_consensus_medaka.alignreport.er barcode01_consensus_medaka.alignreport.txt barcode01_consensus_medaka.merged.vcf.gz barcode01_consensus_medaka.micol.gt.tt barcode01_consensus_medaka.nCoV-2019_1.hdf barcode01_consensus_medaka.nCoV-2019_2.hdf barco		# III (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 6 KB 249 bytes	Document Plain Text gzip coa Document Plain Text TextEdit vCard TextEdit vCard
consensus_medaka           Name           barcode01_consensus_medaka.alignreport.er           barcode01_consensus_medaka.alignreport.txt           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.morged.vcf.gz           barcode01_consensus_medaka.morged.vcf.gz           barcode01_consensus_medaka.nCoV-2019_1.hdf           barcode01_consensus_medaka.nCoV-2019_2.xdf           barcode01_consensus_medaka.nCoV-2019_2.xcf           barcode01_consensus_medaka.nCoV-2019_2.xcf           barcode01_consensus_medaka.nCoV-2019_2.xcf		* = (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 6 KB 249 bytes 315 bytes	Document Plain Text gzip coa Document Plain Text TextEdit vCard TextEdit vCard vCard vCard vCard
consensus_medaka Name barcode01_consensus_medaka.alignreport.er barcode01_consensus_medaka.alignreport.txt barcode01_consensus_medaka.merged.vcf.gz.tbi barcode01_consensus_medaka.merged.vcf.gz.tbi barcode01_consensus_medaka.ncOv-2019_1.hdf barcode01_consensus_medaka.nCOv-2019_1.hdf barcode01_consensus_medaka.nCOv-2019_2.vcf barcode01_consensus_medaka.nCOv-2019_2.vcf barcode01_consensus_medaka.nCOv-2019_2.vcf barcode01_consensus_medaka.nCOv-2019_2.vcf barcode01_consensus_medaka.primersivef barcode01_consensus_medaka.primersivef barcode01_consensus_medaka.primersivef barcode01_consensus_medaka.primersivef.txt barcode01_consensus_medaka.txt barcode01_consensus_medaka.txt barcode01_consensus_medaka.txt barcode01_consensus_medaka.txt barcode01_co			□	<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 6 KB 249 bytes 315 bytes Zero bytes 389 bytes	Document Plain Text gzip coa Document Plain Text VCard VCard VCard VCard VCard Plain Text Document
consensus_medaka           Name           barcode01_consensus_medaka.alignreport.er           barcode01_consensus_medaka.alignreport.txt           barcode01_consensus_medaka.merged.vcf.gz           barcode01_consensus_medaka.merged.vcf.gz.tbi           barcode01_consensus_medaka.molov-2019_1.tdf           barcode01_consensus_medaka.nCoV-2019_1.tdf           barcode01_consensus_medaka.nCoV-2019_2.tdf           barcode01_consensus_medaka.nCoV-2019_2.tdf           barcode01_consensus_medaka.nCoV-2019_2.tdf           barcode01_consensus_medaka.primersitereport.txt           barcode01_consensus_medaka.primersitereport.txt           barcode01_consensus_medaka.primertrimmed.rg.sorted.bam		88 = (		<ul> <li>Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 6 KB 249 bytes 315 bytes Zero bytes 389 bytes 24 bytes	Document Plain Text gzip coa Document Plain Text TextEdit vCard TextEdit vCard VCard Plain Text Document Document
		88 := (		<ul> <li>&gt; Date Modifie</li> <li>Mar 2, 202</li> </ul>	d 2 at 2:48 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:49 PM 2 at 2:48 PM 2 at 2:48 PM 2 at 2:49 PM	61 bytes 145 bytes 254 bytes 72 bytes 3 KB 6 KB 249 bytes 315 bytes Zero bytes 389 bytes 24 bytes 337 bytes	Document Plain Text gzip coa Document Plain Text TextEdit vCard TextEdit vCard VCard Plain Text Document Document

# Nanopolish

Artic	A 1 Version	n select	Nanopolish Procedure select		
Complex Barcodes Enabled		GuppyPlex			Nanopolish
Inputs		Output Locat	ions	DEL	ETE OUTPUTS
Input Long-Reads Full Run Directory	Required dir 📀	Label	Access	Completed	Remove
Directory that contains your fastq_pass, fast5_pass, and sequencing summary.txt file. Output directory from a Oxford Nanopore run		BAM file		3	Ē
1 Drag Folder here		Variants		3	Ē
/Users/merribb1/Desktop/test- data2/20200514_2000_X3_FAN44250_e97e74b4		Consensus	-	3	<u> </u>
Sequencing Summary File	Optional		-	0	ш
Generated at the end of basecalling in the directory			Rows per page:	5 1-3 c	of 3
sequencing_summary file default		COPEN LOG FOLD	DER		•
Select an item		1	_		
Must be present in root level of directory					
Primer Configuration					
Default Primer Scheme to use for nanopolish	Required 📿				
nCoV-2019/3	Hequireu 💟				
Select an item					
Normalise coverage					
1000000					
He Normalise down to moderate coverage to save	equired number 🤣				
runtime					
Fast5 directory	Optional dir 🔪				
Select the fast5 pass directory from your run					
▲ Drag Folder here ②					
null/fast5_pass					
🗸 🛅 consensus_nanopolish				Mar 2, 202	2 at 2:51 PM
barcode11_consensus_nanopolish.mi					2 at 2:51 PM
barcode11_consensus_nanopolish.mu		3			2 at 2:51 PM
barcode11_consensus_nanopolish.mu					2 at 2:51 PM
barcode11_consensus_nanopolish.co					2 at 2:51 PM
barcode11_consensus_nanopolish.pro					2 at 2:51 PM
barcode11_consensus_nanopolish.co					2 at 2:51 PM
barcode11_consensus_nanopolish.co					2 at 2:51 PM
barcode11_consensus_nanopolish.co		.txt.nCov-2019_2.d	eptns		2 at 2:51 PM
barcode11_consensus_nanopolish.pa	-				2 at 2:51 PM
barcode11_consensus_nanopolish.pa barcode11_consensus_nanopolish.fai					2 at 2:51 PM
Barcode I I_consensus_nanopolisn.rai					2 at 2:51 PM 2 at 2:51 PM
Barcode l'I_consensus_nanopolish.me					2 at 2:51 PM 2 at 2:51 PM
barcode11_consensus_nanopolish.pri		Ltxt			2 at 2:51 PM
en barcode 11_consensus_nanopolish.nC					2 at 2:51 PM
Barcode11_consensus_nanopolish.nC					2 at 2:51 PM
barcode11_consensus_nanopolish.pri					2 at 2:51 PM

# **Parameters**

• input run folder: *directory* 

Must contain a fastq\_pass directory or a custom inputted one (See below)

• Primer: option or Directory

Artic Primer set or a custom one (directory) which contain the genome.fasta and necessary BED file for the primer set

- Normalize Coverage: Number
- Barcode Configuration: option

Which barcode kit you used for demux. Select any for non-barcoded sample

• FASTQ Dir: Directory, optional

Select your own custom *fastq\_pass* directory to analyze and demux

• Sequencing Summary File: File, exists

Selects the Sequencing summary file in the root of your run directory

**Note:** Oftentimes, this file can be found in the fastq\_pass directory. Move it one level up to the root run directory

# Returns

- 1. Nanopolish Consensus files: FASTA files containing your consensuses for each barcode
- 2. Nanopolish VCF files: Variant files containing your variant calls for each barcode
- 3. Nanopolish BAM files: BAM files containing your alignment information for each barcode

# IVAR

# IVAR

iVar is a computational package that contains functions broadly useful for viral amplicon-based sequencing. Additional tools for metagenomic sequencing are actively being incorporated into iVar. While each of these functions can be accomplished using existing tools, iVar contains an intersection of functionality from multiple tools that are required to call iSNVs and consensus sequences from viral sequencing data across multiple replicates. We implemented the following functions in iVar: (1) trimming of primers and low-quality bases, (2) consensus calling, (3) variant calling - both iSNVs and insertions/deletions, and (4) identifying mismatches to primer sequences and excluding the corresponding reads from alignment files.

**Note:** Definition of IVAR gathered directory from the above-mentioned github link

# **Parameters**

- BAM file: file
  - Alignment BAM file to search for coding regions from. Example pipelines/commands that generate a *bam* file are
    - \* Minimap2
    - \* Medaka
    - \* Nanopolish
    - \* Basestack Consensus
- Reference GFF3: option or File
  - Coding region list that you want to use the alignment searching against
  - Optional, only used for the Variant service
- Reference FASTA: File
  - Your original reference FASTA file during alignment. Must also be linked to the GFF3 described above

# Returns

- 1. IVAR Consensus FASTA: file
- 2. IVAR Output: Tab-separated file

**Note:** Contains list of GFFs that map to variants that were called in the BAM file

• • •									parcodeC	3_consensus_me	daka.sor	ted.ivar.tsv							
REGION POS	REF	ALT	REF DP	REF RV	REF OU	AL ALT_DP	ALT RV	ALT OU	AL ALT FR	EO TOTAL DP PVAL	PASS	GFF FEATURE	REF CO	DON	REF AA	ALT C	ODON	ALT_AA	
1N908947.3	23	G	+AT	2 -	1	24	1	0 -	20 -	0.1666676	0.375	FALSE NA	NA	NA	NA -	NA -			
N908947.3	34	A	т	94	54	27	4	1	27	0.0408163	98	0.104496 FALSE	NA	NA	NA	NA	NA		
N908947.3	40	с	+T	83	56	26	8	0	20	0.0551724	145	0.00252449	TRUE	NA	NA	NA	NA	NA	
908947.3	50	с	+T	121	53	43	9	0	20	0.0588235	153	0.00484424	TRUE	NA	NA	NA	NA	NA	
908947.3	75	Ċ	-T	93	64	27	8	0	20	0.0519481	154	0.00314154	TRUE	NA	NA	NA	NA	NA	
908947.3	75	ć	+T	93	64	27		0	20	0.0454545	154	0.00707821	TRUE	NA	NA	NA	NA	NA	
V908947.3	76	Ť	Α	77	38	26	3	1	29	0.0375 80	0.13291	8 FALSE NA	NA	NA	NA	NA			
V908947.3	78	т	G	93	55	28	3	1	29	0.03125 96	0.21163	1 FALSE NA	NA	NA	NA	NA			
N908947.3	78	т	+A	93	55	28	6	0	20	0.0465116	129	0.0282057	FALSE	NA	NA	NA	NA	NA	
908947.3	84	C	-TG	80	42	25	5	0	20	0.0396825	126	0.0409386	FALSE	NA	NA	NA	NA	NA	
908947.3	96	ċ	-A	75	30	26	5	ø	20	0.0384615	130	0.0315931	FALSE	NA	NA	NA	NA	NA	
N908947.3	98	č	-T	19	6	29	17	0	20	0.130769 130	2.27973	e-12 TRUE	NA	NA	NA	NA	NA		

Field	Description
REGION	Region from BAM file
POS	Position on reference sequence
REF	Reference base
ALT	Alternate Base
REF_DP	Ungapped depth of reference base
REF_RV	Ungapped depth of reference base on reverse reads
REF_QUAL	Mean quality of reference base
ALT_DP	Ungapped depth of alternate base.
ALT_RV	Ungapped deapth of alternate base on reverse reads
ALT_QUAL	Mean quality of alternate base
ALT_FREQ	Frequency of alternate base
TOTAL_DP	Total depth at position
PVAL	p-value of fisher's exact test
PASS	Result of p-value <= 0.05
GFF_FEATURE	ID of the GFF feature used for the translation
REF_CODON	Codong using the reference base
REF_AA	Amino acid translated from reference codon
ALT_CODON	Codon using the alternate base
ALT_AA	Amino acid translated from the alternate codon

**Note:** Description of each field was gathered from the Manual

## Minimap2

# Minimap2

Minimap2 performs alignment for genomic and spliced nucleotide files

## **Parameters**

• FASTQ/A file: file

Raw Reads file to run alignment against a reference on.

• **Reference FASTA:** *File* Your original reference FASTA file during alignment. Must also be linked to the GFF3 described above

# Returns

- 1. BAM: file
- 2. SAM: file

Minimap2	Version select		Minimap2 Procedure select		0
Minimap2 Alignment	Samtools Enabled	S BAM from SAM		S S	amtools Index
Inputs		Output Locations		DELET	E OUTPUTS
FASTA/Q File	Required file 🥏	Label	Access	Completed	Remove
Letter the second secon		BAM File 😮		1/1	Ē
Reference to align against Must be in FASTA format. This is the genome you want to align your sequences	Required file 🥏	Alignment File 👔		1/1	Ē
again to the in FASTA format. This is the genome you want to align your sequence		BAI File 🥹		1/1	Ē
1 files (29.9 kB in total)			Rows per page:	5 1-3 of 3	
		<pre>"sam-/opt/data/calls_to_tr "Metworkingconfig': { "End "staphb/samtools", "Cmd': \$(bam)" }, "WorkingDir": " starting the container sam Initiating logging object. % { "name": "samtools_index_ "AutoRemove": false, "Priv"</pre>	tools_view_sam_to_bam_minimap2	alls_to_draft.bam.bam. lse, "OpenStdin": true -5 -F4 -b \${sam}   sa 80/tcp": {} }, "HostCo sers/merribbl/Desktop/	, "Image": mtools sort > nfig": { test-

🔴 🕒 📄 test_covid.fastq.sam
@SQ SN:MN908947.3 LN:29903
<pre>@PG ID:minimap2 PN:minimap2 VN:2.24-r1122 CL:minimap2 -a -o /opt/data/</pre>
test_covid.fastg.sam /opt/ref/nCoV-2019.reference.fasta /opt/data/test_covid.fastg
0033cf58-a90e-43cf-aa8b-8238cd262d89 0 MN908947.3 9784 60
67S22M4D250M1I69M4D2M1I39M37S * 0 0
ATTGTACTTCGTTCAGTTACGTATTGCTAAGGTTAACCTGGTAACTGGGACACAAGACTCCAGCACCTACTTTTGAAGAAGCTGCGCTGCCTTTTTGTTAA
ATAAAGAAATGTATCTAAAGTTGCGTAGTGATGTGCTATTACCTCTTACGCAATATAATAGATACTTAGCTCTTTATAATAAGTACAAGTATTTAGTGGA
GCAATGGATACAACTAGCTACAGAGAAGCTGCTTGTTGTCATCTCATAAAGGCTCTCAATGACTTCAGTAACTCAGGTTCTGATGTTCTTTACCAACCA
ACAAACCTCTATCACCTCAGCTGTTTTGCAGAGTGGTTTTTAGAAAAATGGCATTCCCATCTGGTAAAGTTGAGGGTTGTATGGTACAAGTAACTTGTGGT
ACAACCTGTAACGGTCTTTGGCTTGATGACGTAGTTTACTGTCCAAGGTGCTGGAGTCTTGTCCCAGTTACTGCTACTAACCTT \$%
%'')24=>76981320'''-133454=8=B@]B?9789/./'&%%,38]]><<76;;12?>?=::<@@B?A00;6/;0),7:8;7.*%)1;>?>>?G?
F>45470634367=94338877<6*4::;876350>=87798//88=:8579@J]]G= .2-3<]: 69=]=?A=@>=>16?
67CDDC<434;=45= <c?<@?dbaa767888554,]]] 113:9:="">&gt;=&gt;?D@&lt;<?:8((()8>961?]]0&lt;&gt;&gt;B&lt;=?;3*38&gt;&gt;=&lt;89A<?</td></c?<@?dbaa767888554,]]]>
9:ABA>>>F <b=53*-4623330400?00;f?;>B9:86:-/.,)23===42,15;=@&gt;?7BC;-=C]EDD32266=;:968=97**53</b=53*-4623330400?00;f?;>
=>B@A>?:9 =5411-8 78,&%'/))*./*(%%\$\$\$*,,05: <ccf;==;7433?6 :3779<?="">**;&lt;359&gt;@@5489AB776A&lt;421(%</ccf;==;7433?6>
%&%&&&&*)''\$&% NM:i:12 ms:i:723 AS:i:716 nn:i:0 tp:A:P cm:i:58 s1:i:350 s2:i:0 de:f:0.0155
rl:i:0
00942f38-a6d3-4abc-bfd2-fba1785243b2 16 MN908947.3 10999 60
49S20M1D12M1D42M2I13M1I16M1I23M2I43M2D8M1D60M1D55M2D40M1I2M3I11M3D41M66S * 0 0
ACATGTATTAAGAAGGTTAACCTGGTAACTGGGACAAGACTCCAGCACCTACACCACTGGTTGTTACCACAATTTTGACTCACTTTTAGTTTAGTCCA
GAGTACTCAATGGTCTTTGTTCTTTTTTTTGTATGAAAAATGCCTTTTTACCTTTTTGCTATGGGTATTATTGCTAAGTGTCTGCTTTTGCAATGATGT
TTGTCAAACATAAGCATGCATTTCTGTTTGTTTGTTACCTTCTTGCCACTGTAGCTTATTTTAATATGGTCTATATGCCTGCTAGTTGGTGATGCGTA
TTATGACATGGTTGGATAAGGCTGATACTAGTTTGTCTGGTTTTGCCAAAAGACTGTGTTATGTATG
TACAAC &(((/11,*(''.+)24=C<@57333==8916414:861,.79652)+=]F<884409:;:43%'(',,623352+%\$5:BAA?
457:;>3643720'('+1),+(*515325.:2>?AEDDE<;4222@?DC@<434LEEK<=9?C <aa=?:98:;>:99@99866501++-A/</aa=?:98:;>
3//00)))*./.7> ABCFEGD:9.,,22125+1486+].+-5,((3 >4336465841]CAH><<;:;=:=<43? C<<8437]C>96779]:;6558554;97((())*121005021446;=94((**+)0-
*%&'''':;<>>@DB@88586=K]''*%&@AG::9:]]DA54442124;=<545B=<8]]4.&%&/,+.***\$\$\$&'\$

# Pangolin

### Pangolin

Software package for assigning SARS-CoV-2 genome sequences to global lineages.

# **Parameters**

• FASTA file: file

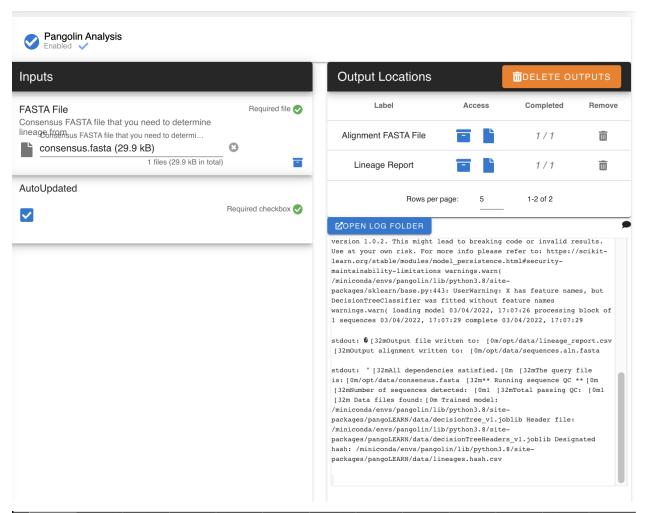
Genome file, can be a multi-sequence FASTA file

• AutoUpdate: *option* 

Note: Pangolin updates so frequently, opt to check for updates before running. Requires Internet access

# Returns

- 1. Pangolin Lineage Report: comma-delimited file
- 2. alignment FASTA: sequences.aln.fasta, contains your alignment from pangolin



	A	В	С	D	E	F	G	н	I.	J	к	L	м
1	taxon	lineage	conflict	ambiguity_s	scorpio_call	scorpio_supp	scorpio_conf	version	pangolin_ver	pangoLEARN	pango_versio	status	note
2	MN908947.3	A.1	0	1				PLEARN-v1.2	3.1.20	2/2/22	v1.2.124	passed_qc	
3													
4													

### MASH

MASH

MASH is for uncovering metagenomic distances using MinHash

### **Parameters**

• FASTA genome file: *file* 

Whatever genome file you want to uncover genomic distances from. Can be multipe entries. Can also be a consensus file

- Reference Genomes: *File* Can be Sketch format .msh or FASTA file.
- Sketch File 1: option

Speeds up the process for the input file 1

• Sketch File 2: option

Speeds up the process for the input file 2

• Winnder Take All: option

Only take the highest (best) distance annotations against the reference for your input file.

# Returns

Inputs		Output Locations		<b>DELETE OU</b>	TPUTS
Genome 1 Selects Mart Just Hangamen (FASSTA), fille	Required file 📀	Label	Access	Completed	Remov
nCoV-2019ence.fasta (30.4 kB) 1 files (30.4 kB in total)	8	Output distance metrics		1/1	
	Required 🛇	Output screen metrics		1/1	Ī
Genome 2 Select your second genome (FASTA) file	Hequirea 💟	Rows per	page: 5	1-2 of 2	
Refseq 1000 genomes		COPEN LOG FOLDER			
Select an item		BOPEN LOG FOLDEN			
Sketch first FASTA file beforehand					
	Optional checkbox 🔨				
Sketch second FASTA file beforehand This is the reference that is compared against from genome file 1					
This is the reference that is compared against from genome file 1	Optional checkbox 🔨				
Employ Winner-Take-All to remove redund					
✓	Optional checkbox 📉				

1. Distance: tab-separated file

# Note:

# Columns:

- 1. Reference-ID
- 2. Query-ID
- 3. Mash-distance

- 4. P-value
- 5. Matching-hashes

See more information on calculaions here

		📄 mash	.dist.tsv	
<pre>GCF_000001215.4_Release_6_plus_IS01_MT_genomic.fna.gz /opt/data/nCoV-2019.reference.fa</pre>	cto	1	1	0/1000
	1	0/1000	1	0/1000
<pre>GCF_000001405.36_GRCh38.p10_genomic.fna.gz/opt/data/nCoV-2019.reference.fasta 1</pre>	1			
GCF_000001515.7 Pan_tro_3.0_genomic.fna.gz/opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
GCF_000001545.4_P_pygmaeus_2.0.2_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta	1	1	0/1000	
GCF_000001635.25_GRCm38.p5_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
GCF 000001735.3 TAIR10 genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
GCF_000001765.3 Dpse 3.0 genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
GCF_000001895.5_Rnor_6.0_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
<pre>GCF_000001905.1_Loxafr3.0_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1</pre>	1	0/1000		
GCF_000001985.1_JCVI-PMFA1-2.0_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta	1	1	0/1000	
GCF_000002035.5_GRCz10_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
<pre>GCF_000002075.1_AplCal3.0_genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1</pre>	1	0/1000		
GCF 000002195.4 Amel 4.5 genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		
GCF_000002235.4 Spur 4.2 genomic.fna.gz /opt/data/nCoV-2019.reference.fasta 1	1	0/1000		

2. Screen: tab-separated file

### Note:

### **Columns:**

- 1. identity
- 2. shared-hashes
- 3. median-multiplicity
- 4. p-value
- 5. query-ID
- 6. query-comment

See more information here

• • •		mash.screen.tsv
0.867054 50/1000 1	1.97626e-323	GCF_000864885.1_ViralProj15500_genomic.fna.gz NC_004718.3 SARS coronavirus, complete genome
0.825882 18/1000 1	2.24222e-109	GCF_000887595.1_ViralProj51751_genomic.fna.gz NC_014470.1 Bat coronavirus BM48-31/BGR/2008, complete genome
0.719686 1/1000 1	7.01082e-06	GCF_000899495.1_ViralProj177902_genomic.fna.gz NC_018871.1 Rousettus bat coronavirus HKU10, complete genome
0.719686 1/1000 1	7.01082e-06	GCF_000894435.1_ViralProj109275_genomic.fna.gz NC_016993.1 Magpie-robin coronavirus HKU18, complete genome
0.719686 1/1000 1	7.01082e-06	GCF_000889395.1_ViralProj59783_genomic.fna.gz NC_014637.1 Cafeteria roenbergensis virus BV-PW1, complete genome
0.719686 1/1000 1	7.01082e-06	GCF_000868045.1_ViralProj18867_genomic.fna.gz NC_009021.1 Bat coronavirus HKU9-1, complete genome
0.719686 1/1000 1	7.01082e-06	GCF_000312225.1_ASM31222V1_genomic.fna.gz [267 seqs] NZ_HE973313.1 Microcystis aeruginosa PCC 9807, whole genome
shotgun sequence []		

## VADR

Viral Annotation DefineR: classification and annotation of viral sequences based on RefSeq annotation

## VADR

VADR is a suite of tools for classifying and analyzing sequences homologous to a set of reference models of viral genomes or gene families. It has been mainly tested for analysis of Norovirus, Dengue, and SARS-CoV-2 virus sequences in preparation for submission to the GenBank database.

# **Parameters**

• FASTA consensus file: file

Sequencing file you want to get viral annotations from in the VADR pipeline

• Reference Genome: *option or File* Can be Sketch format .msh or FASTA file.

## **Returns**

Vadr	Version select	2	Vadr SARS-CoV-2 Procedure select		C
Vadr Enabled V					
Inputs		Output Locations		DELETE	OUTPUTS
FASTA File	Required file 🤣	Label	Access	Completed	Remove
Dne or more FASTA sequences (single file) that you want to annotate with he SARSGRV/27A®nላቄ፤በባይନወደ፤(ያቤያይ ነልታ የተቀ/ that you want to annotate consensus.fasta (29.9 kB)		seqstat		1/1	Î
1 files (29.9 kB in total)		passed table		1/1	Ē
/adr Models	Required 🥏	fasta seqs passed		1/1	Î
Default 1.3.2 for SARS-CoV-2		annoations passed		1/1	Î
Select an item	8	models summary		1/1	Ē
		COPEN LOG FOLDER	Rows per page:	5 1-5 of 7	
		<pre>stdout: \# List and descrip stdout: \# esl-seqstat -a vadr_output.vadr.seqstat # in: vadr_output.vadr.seqstat # in: vadr_output.vadr.sesl. vadr_output.vadr.sesl.set vadr_output.vadr.sell.list vadr_output.vadr.seq # vadr_output.vadr.seq # per-model-segment tabular tabular summary file saved saved in: vadr_output.vadr vadr_output.vadr.set # set for the sec # replaced stretches of Ms output files created in div stdout: ?# Elapsed time: 00</pre>	output for input fasts fi 5 column feature table o bl # 5 column feature ta fail.tb # list of passi # list of failing sequen # fasts file with passing fasts file with failing per-sequence tabular ann- sequence tabular classif feature tabular summary f le saved in: in: vadr_output.vadr.alt alc # per-model tabular manet doctoring tabular summary file (-r) saved metery ./vadr_output/ #	le saved in: itput for passing seque ble output for failing ng sequences saved in: res saved in: feature tables saved in: sequences saved in: totation summary file sa ication summary file saved in: vad_output.vadr.sgm # # alert count tabular summary file saved in: (-s) saved in: vadr_out in: vadr_output.vadr.rp	nces saved sequences : ved in: ved in: put.vadr.ftr per-alert summary file put.vadr.sda

# 1. Table format: Tab-separated annotation

5 column feature table output for passing sequences

			l va
>Feature	e MN90894	/.3	
266	21555	gene	
		gene ORF1ab	
266	13468	CDS	
13468	21555		
		product ORF1ab polyprotein	
		exception ribosomal slippage	
		protein_id MN908947.3_1	
266	13483	CDS	
		product ORF1a polyprotein	
		protein_id MN908947.3_2	
266	805	mat_peptide	
		product leader protein	
		protein_id MN908947.3_1	
266	805	mat_peptide	
		product leader protein	
		protein_id MN908947.3_2	
806	2719	mat_peptide	
		product nsp2	
000		protein_id MN908947.3_1	
806	2719	mat_peptide	
		product nsp2	
		protein_id MN908947.3_2	

2. sgm format: Tab-separated annotation

per-model-segment tabular summary file

••	•									📄 vadr	_outpu	t.vadr.s	gm							
⊭ #idx	seg name	seq len	p/f	model	ftr type	ftr name	ftr idx	num sgm	sgm idx	seg from	seq to	mdl from	mdl to	sgm len	str	trc	5' pp	3' 99	5' gap	3' gap
#	MN908947.3					ORF1ab	1	1	1	266	21555	266	21555	21290		no	?	?	no	no
1.2.1 1.2.2	MN908947.3 MN908947.3	29903	PASS	NC_045512	CDS	ORF1ab_polyprotein ORF1ab_polyprotein	2 2	2 2	1 2	266 13468	13468 21555		13468 21555	13203 8088		no no	?	? ?	no no	no no
1.3.1 1.4.1	MN908947.3 MN908947.3	29903	PASS	NC_045512	gene	ORF1a_polyprotein S	3 4	1 1	1 1	266 21563	13483 25384	266 21563	13483 25384	13218 3822		no no	? ?	? ?	no no	no no
1.5.1 1.6.1	MN908947.3 MN908947.3					surface_glycoprotein ORF3a	5 6	1 1	1 1	21563 25393	25384 26220	21563 25393	25384 26220	3822 828		no no	? ?	? ?	no no	no no
	MN908947.3 MN908947.3					ORF3a_protein E	7 8	1 1	1 1	25393 26245		25393 26245		828 228		no no	? ?	? ?	no no	no no
	MN908947.3 MN908947.3					envelope_protein M	9 10	1 1	1 1	26245 26523	26472 27191	26245 26523	26472 27191	228 669		no no	? ?	? ?	no no	no no
1.11.1	MN908947.3 MN908947.3	29903	PASS	NC_045512	ČDS	membrane_glycoprotein ORF6	11 12	1 1	1 1	26523 27202	27191 27387	26523 27202	27191 27387	669 186		no no	? ?	?	no no	no no

3. sqa format: Tab-separated file

per-sequence tabular annotation summary file

••							🗋 va	dr_o	utput	.vadr	.sqa	
#seg #idx # 1	seg name  MN908947.3	seg len  29903	p/f  PASS	 best model NC_045512	grp Sarbecovirus	sub grp SARS-CoV-2	<u>nfa</u>  54	<u>nfn</u> 0	<u>nf5</u> 0	<u>nf3</u> 0	<u>nfalt</u> 0	

All output files created in directory your directory containing your input FASTA file under vadr\_output

**Note:** See more information `here <https://github.com/ncbi/vadr/wiki/Coronavirus-annotation#many-alertserrors-in-orf3a-orf6-orf7a-orf7b-orf8-and-orf10-do-not-cause-a-sequence-to-fail-1`\_

1. Output printed to screen saved in: my4.vadr.log

- 2. List of executed commands saved in: my4.vadr.cmd
- 3. List and description of all output files saved in: my4.vadr.filelist
- 4. esl-seqstat -a output for input fasta file saved in: my4.vadr.seqstat
- 5. 5 column feature table output for passing sequences saved in: my4.vadr.pass.tbl
- 6. 5 column feature table output for failing sequences saved in: my4.vadr.fail.tbl
- 7. list of passing sequences saved in: my4.vadr.pass.list
- 8. list of failing sequences saved in: my4.vadr.fail.list
- 9. list of alerts in the feature tables saved in: my4.vadr.alt.list
- 10. fasta file with passing sequences saved in: my4.vadr.pass.fa
- 11. fasta file with failing sequences saved in: my4.vadr.fail.fa
- 12. per-sequence tabular annotation summary file saved in: my4.vadr.sqa
- 13. per-sequence tabular classification summary file saved in: my4.vadr.sqc
- 14. per-feature tabular summary file saved in: my4.vadr.ftr
- 15. per-model-segment tabular summary file saved in: my4.vadr.sgm
- 16. per-alert tabular summary file saved in: my4.vadr.alt
- 17. alert count tabular summary file saved in: my4.vadr.alc
- 18. per-model tabular summary file saved in: my4.vadr.mdl
- 19. alignment doctoring tabular summary file saved in: my4.vadr.dcr
- 20. ungapped seed alignment summary file (-s) saved in: my4.vadr.sda
- 21. replaced stretches of Ns summary file (-r) saved in: my4.vadr.rpn

### Samtools (Under Construction)

Warning: Under Construction

### Unicycler

A hybrid assembly pipeline for bacterial genomes

### Unicycler

**Note:** This pipeline uses SPAdes

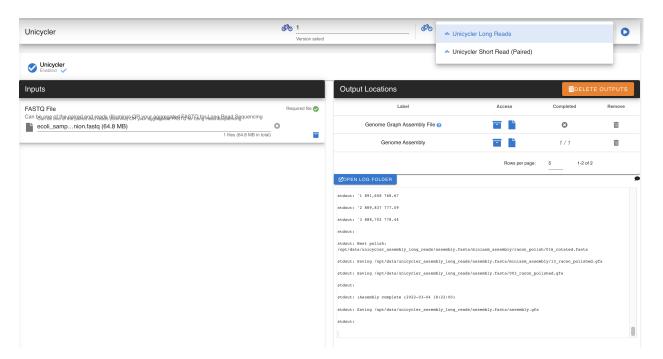
# **Unicycler Long Reads**

## **Parameters**

• FASTQ file: *file* 

Sequencing raw read file. Can be short or long read

## **Returns**



1. Genome Assembly: Tab-separated annotation

Assembly FASTA File

	assembly.fas	sta
>ecoli		
GACATTCCTCCAATTGTTGTTG	GCACAGCGATCCAGGGGGCACTTGCCTGAAAGTTACG	AGGTTCTTCGAA
CGCTATTCCACTGCTCCTTTAT	AGGTACAACAGTACCTTGATGATGGATAAGTCTGAA	ATTTGACGAGAT
AGAATTTAGTGAGGCGCAAGAG	GCTAAATTATCACTGAAGATGATTAATTTAATTACT	AAACCATCAGAT
CATGTGCTTTTGTGGGTCATCG	CCACCGTTGTTGGCCATTCCTGCGTAAATTTGTCTC	GCCGTGTCGCGT
TGCGAGTAGACTTCGTGCTTGT	CAAACTGGAGATTTAACTACTGAGCGAAACGGCATC	CTGGCAGCCGAG
CGCATCCATTCCTAACTTATTA	AAACGCGCGGCGATTATGAGCACCGAGTCGCCTTCG	TTCTTTGCCAGG
TTAGTTGGATTGCTGGAGGTGG	AGACGCCGTATGAGCGAGCGAGACGGTAACCATCGA	TTCATTTATTGT
GCGTTTGAGACACGTTTCGTTG	GCCCCGGGCATTCGCAGGGATGAATCTCTGATAAGT	CGACTGAGCGGA
ATACCATTGAAACGCCTGCTGG	GTTGCCGGTTGGGCCGGTGGTTCCATGTGCAGGGCA	GTTCGCGTAATG
AAAGCAAGTGGGCATGTTATCA	AGCGCTGAGTTCACTATGCTGAGGGGATGGTACGGC	ΑCTACTTAATAT
CCGGTTGATGAACGAGACTGCT	TACTCTTACAACAGGTGCTTATCCCTGCAGCAGAAA	GCCTTCTTATCA
ACAAGCTTTCTGCATTATCTGG	AAATTGACCCACTCTTACCAAATACAAAAGGCAATG	GTGCGGGAAGAT
GGCAGCGAAACTGGATTTGAGC	AATGCTGATGAGACAGGAAACAAAAGAAGACCGCGA	CACGTTGCTACA
ATTGCTGTTGAGCTTTGGAGTC	CGTCAAATATTGGTAAGAAAACCGAAAACCTTTGTA	CCACTTTCTCAG
CCAGCGCATACTGGCGCAAATC	AGTACCCGAGACTGCATCGGGTCGCTCGGCACCGAT	TTGAGAGGTTTA
TTATAAAGCTATTGAGCTGGCG	AATGGTTTCCATGAGAATTGACGCTGCTGTGTAAGG	CGCAACAACGGT
TGAGAACAAGGATAATAATGGG	CAAGCGCGCGGTTTGCCGCAGCACCCTCATTGACCG	ATCTGATTGACG
CCTTACAAAGTCGTATGCTGAC	TGTTCCGGCGTGGCATTAGTGTTGATCGTTCTGATG	TTGGCGGGCGCG

2. GFA File: File

Can be used in 3rd party tools like Bandage

# **Unicycler Short Reads**

### **Parameters**

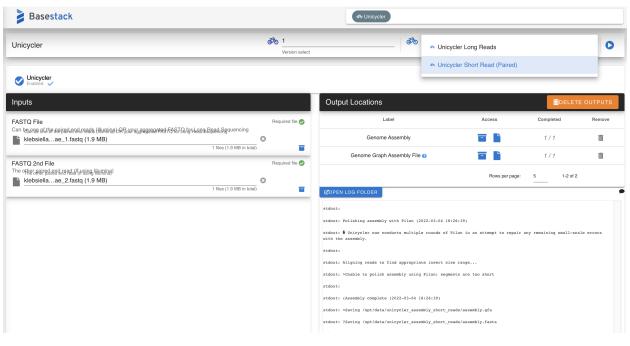
• FASTQ first file: *file* 

First Read file (paired or single-end)

• FASTQ second file: *file* 

Second read file (paired only)

### Returns



1. Genome Assembly: Tab-separated annotation

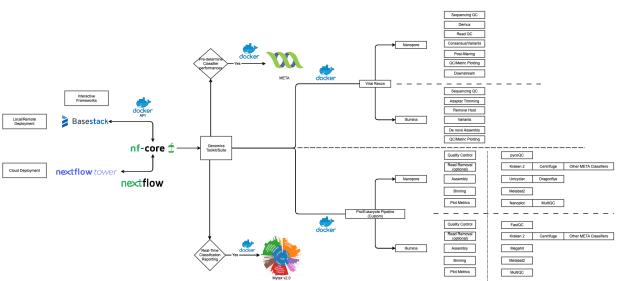
Assembly FASTA File

2. GFA File: File

Can be used in 3rd party tools like Bandage

## TaxTriage (Metagenomics) (Under Construction)

Warning: This module is under construction and is in alpha-release. Scheduled full release of v1.0 in Oct. 2022



Standard diagram for deployment and pipeline development

The pipeline consists of a variety of alignment/classification steps as well as QC and pre-filtering processes. It is designed to be serve as the initial triage step for identifying unknown organisms present in one or more sample types and supports both Illumina or Oxford Nanopore-generated NGS data.

The pipeline is packaged to include basic quality control to making a (potential) de-novo assembly for each organism that is detected in the sample from a filtering a hierarchical perspective. That is, the most prevalent taxonomic IDs at various ranks in the hierarchical chain are reported, binned, and run through a variety of alignment and assembly steps (for lower levels like species). Finally, a set of flags are generated for each taxonomic map that is the most prevalent per sample.

A list of tools used are listed below for each step

- 1. Demultiplex and Gather OPTIONAL, Oxford Nanopore Only
- Artic Guppyplex Aggregate Nanopore reads for downstream analysis
- 2. Quality Control OPTIONAL
- PycoQC Computes metrics and generates interactive QC plots for Oxford Nanopore technologies sequencing data
- 3. Trimming
- Illumina: Trimgalore
- Oxford: Porechop
- 4. Filtering
- Kraken2
- 5. QC Plotting
- Illumina: FastQC

- Oxford: Nanoplot
- 6. Classification (K-mer approach)
- Kraken2
- 7. Alignent Stats
- Illumina: BWAMEM2
- Oxford: Minimap2
- 8. Report Generation
- MultiQC

Please see relevant links in the listed modules for more information on the underlying mechanisms and corresponding papers (if existent)

### **Parameters**

• Samplesheet (.csv): *file* 

Contains a mapping of metadata and a single sample per row. Explanations of the possible columns for Basestack are seen below:

Column	Description
Name	
sample	Custom sample name. This entry will be identical for multiple sequencing libraries/runs from the
	same sample. Spaces in sample names are automatically converted to underscores (_).
single_end	Is the data single or paired end
fastq_1	Full path to FastQ file for Illumina short reads 1 OR OXFORD reads. File has to be gzipped and
	have the extension ".fastq.gz" or ".fq.
fastq_2	Full path to FastQ file for Illumina short reads 2. File has to be gzipped and have the extension
	".fastq.gz" or ".fq.
barcode	TRUE/FALSE, is the row attributed to a demultiplexed barcode folder of 1 or more fastq files or is
	it a single file that is .
from	Directory path of the barcode, only used with the column being set as TRUE in the barcode column
trim	TRUE/FALSE, do you want to run trimming on the sample?
platform	Platform used, [ILLUMINA, OXFORD]
sequenc-	If detected, output plots based on the the sequencing summary file for that sample
ing_summary	

### Table 3: Samplesheet Description

sam-	fastq_1	fastq_2	plat-	from	trim	sequenc-	sin-	bar-
ple			form			ing_summar	y gle_en	dcode
Sam-	AEG588A1_S1_L00	1_ARH_CO388.Aak <u>tof gz</u> L00	1_ <b>IR</b> 2_00	1.Naktq.lgz	FALS	ENULL (or	FALSE	FALSE
ple_1			LU-	(or leave		leave blank)		
			MINA	blank)				
Sam-	ecoli_reads.fastq	NULL	OX-	NULL	FALS	Esequenc-	TRUE	FALSE
ple_2			FORD			ing_summary.	txt	
Sam-	NULL	NULL	OX-	barcode01	TRU	E FALSE	TRUE	TRUE
ple_3			FORD					

### Table 4: Example Samplesheet

For the samples shown above:

- 1. A paired-end run of Illumina data where we DON'T trim anything (no Trimgalore)
- 2. A single-end Oxford Nanopore run where all reads are concatenated to a single fastq file. No barcode. There is a sequencing summary file we want to plot for run statistics/plots
- 3. A single-end Oxford Nanopore run where reads have NOT been demultiplexed and/or aggregated to a single fastq file (like row 2). This will run *artic guppyplex* as well to concatenate all to one fastq file

## Returns

- 1. MultiQC report HTML file
- 2. Variety of intermediate and output results files for the MultiQC report
- Examples: SAM/BAM alignment Filtered FASTQ Files (for downstream use) Assembly (de novo) WIP and is not ready just yet Kraken2 Report(s)

# 3.2.2 Supplemental

### **Guppy Minknow**

### **MinKNOW**

In order to run the MinION sequencer, you first need to download/install the necessary software from Oxford Nanopore's mirror(s).

```
wget -O- https://mirror.oxfordnanoportal.com/apt/ont-repo.pub | sudo apt-key add -
echo "deb http://mirror.oxfordnanoportal.com/apt $(lsb_release -c | awk '{print $2}')-
ostable non-free" | sudo tee /etc/apt/sources.list.d/nanoporetech.sources.list
```

sudo apt-get -y update

```
sudo apt-get install -y minion-nc
```

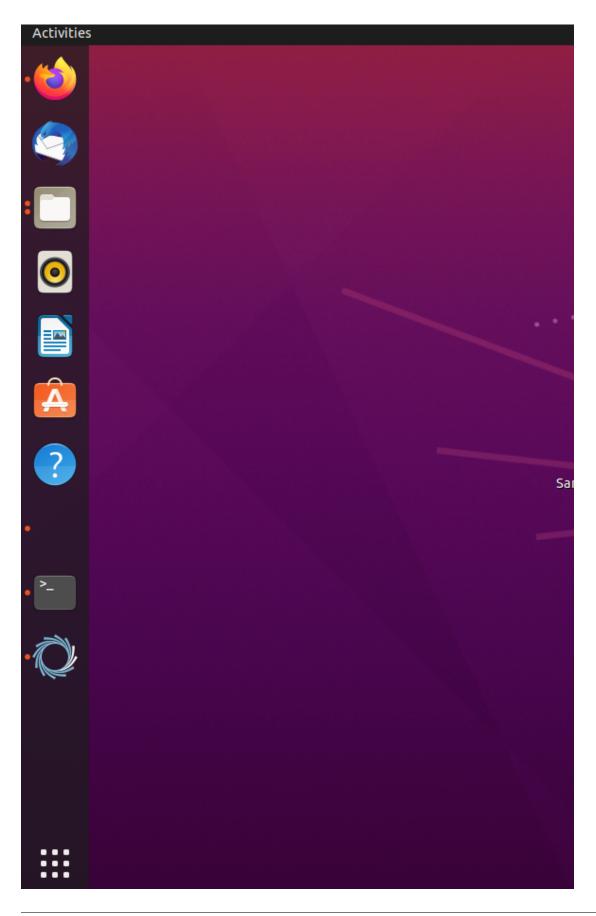
Note: See here

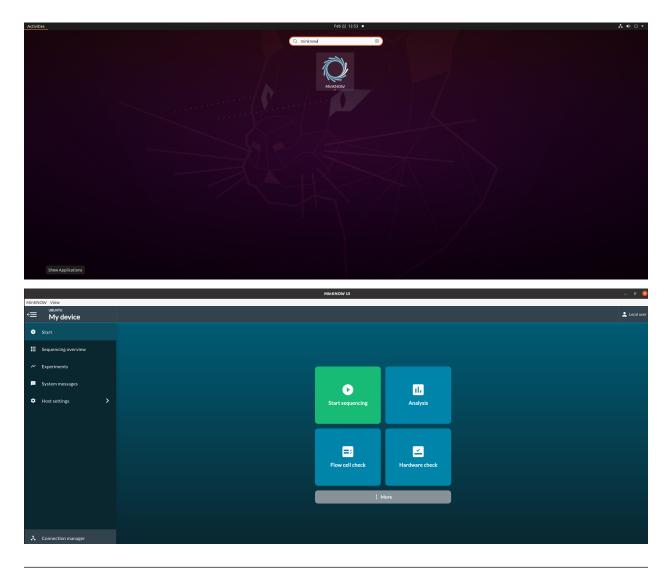
Next, we need to install guppy on your system. Skip this step if you are not using a GPU in your system.

PLEASE NOTE: this option is available only for Linux-based distributions. You have to use CPU-mode for Windows (Fast config basecalling mode)

# Running the app

Select Minknow from the Quick-launch location in Ubuntu. Usually, this is in the bottom-left section of the screen. You should see the MinKNOW icon





**Note:** If you dont set your analysis to run offline, you will see a window that asks you to login. Please follow instructions below to disable online-mode

# Testing Guppy Basecaller from MinKNOW

First, select the Start at the top-left. Then, select Basecalling

Output	
Output folder ⑦	
/var/lib/minknow/data/basecalling	
Compress .fastq files ? Output .fast5 files ?	

Note: By default, all sequencing runs will output to /var/lib/minknow/data on Linux machines.

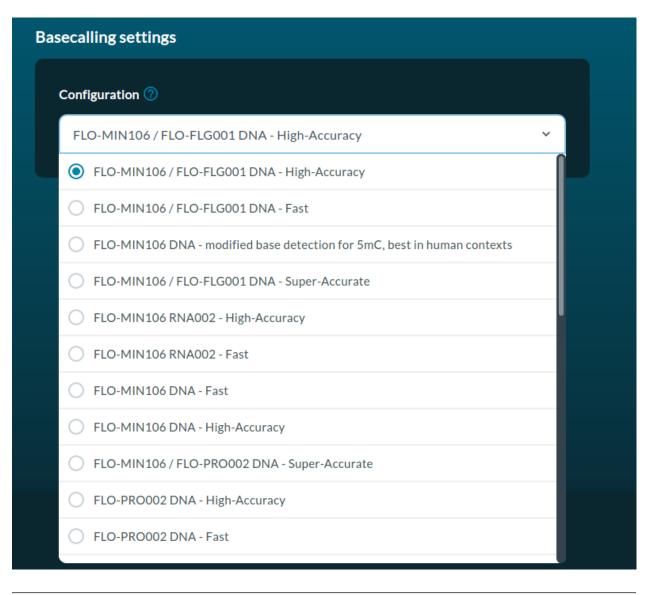
**Warning:** To run basecalling from the UI, you MUST ensure that permissions allow reading + writing for the minknow group on Linux machines

Here, you must select the basecalling configuration you'd like to use

Activities	MinKNOW	Ŧ			12:07 •		▲ ● ① ▼
•🝅	MinKNOW View			MinKNG	IU WC		- * 8
	⇒≡						🔔 Local user
:	•		2.Outout	3.Basecalling	4. Barcoding	5. Allenment	6. Review
0			Ba	secalling settings			
	-			Configuration 🕖			
Â				FLO-MIN106 / FLO-FLG001 DNA - Hig			
	*			FLO-MIN106 / FLO-FLG001 DNA - H	· · ·		
?				FLO-MIN106 / FLO-FLG001 DNA - F	ast		
. >				FLO-MIN106 DNA - modified base de	etection for 5mC, best in human contexts		
				C FLO-MIN106 / FLO-FLG001 DNA - S	uper-Accurate		
 Q				O FLO-MIN106 RNA002 - High-Accura	cy		
				O FLO-MIN106 RNA002 - Fast			
				O FLO-MIN106 DNA - Fast			
				C FLO-MIN106 DNA - High-Accuracy			
				FLO-MIN106 / FLO-PRO002 DNA - S	iuper-Accurate		
				FLO-PRO002 DNA - High-Accuracy			
	*	Back to output		C FLO-PRO002 DNA - Fast			Continue to barcoding >

Most, when running the MinION or MK1C, will use the FLO-MIN106 / FLO-FLG001 DNA options, it is the default option.

Next, pick whether you want the High-Accuracy or Fast configurations (in the names)



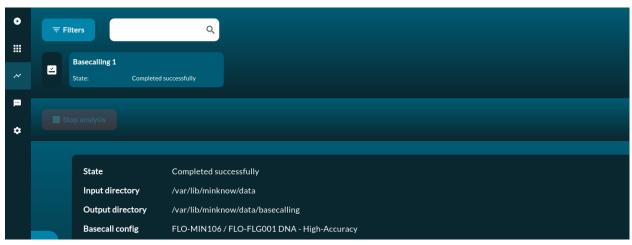
**Note:** High-Accuracy is RECOMMENDED if using a GPU-capable laptop with a CUDA-compatible device. Make sure you set up GPU basecalling as described here *Guppy GPU Basecaller* Fast accuracy is optional but is RECOM-MENDED for non GPU-capable laptops/devices. It will run everything using a CPU.

**Warning:** If sequencing, it is generally not recommended to run basecalling at the same time if using Fast, CPU-only. This is because the system can get overloaded and crash the application

Barcoding settings	
Barcoding kits ⑦ Select any barcoding kits that you used for the run.	
Barcoding kits 🔹	
O EXP-DUAL00	Î
O EXP-NBD103	
O EXP-NBD104	
O EXP-NBD114	
C EXP-NBD196	
O EXP-PBC001	
C EXP-PBC096	
OND-SQK-LP0096M	
OND-SQK-LP0096MA	
OND-SQK-LP0096S	

Note: If you don't see any barcoding kits, go here Barcoding Kits Missing to fix

Move through the rest of the configurations until you finally can hit Start, the green button at the bottom-right of the page



Note: The progress should begin to update as the process goes forward.

### ..warning::

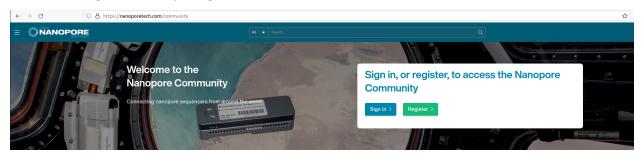
If it errors out, check logs by selecting the right arrow or by looking at all files (recently made) at /var/log/ minknow. Oftentimes it is an issue with permissions

## **Oxford Community Forums**

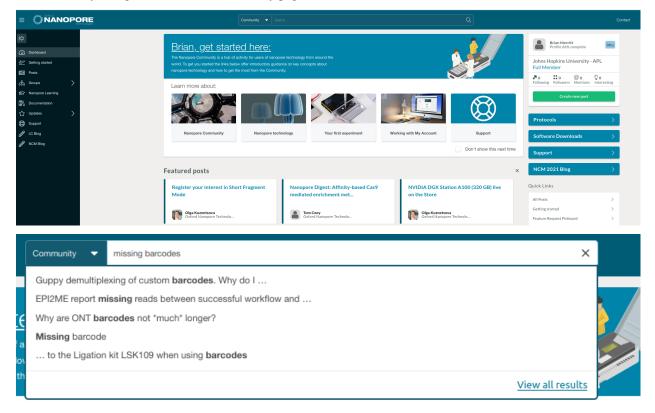
Oxford provides a very active forum for users of ONT software and hardware to interact and post questions or solutions to issues.

Additionally, the Oxford team will oftentimes provide helpful information on their devices or software

First, head to https://community.nanoporetech.com/



You will need to make an Oxford account to do so. Please register for an account as all Oxford hardware users should have the ability to login and view the community pages available on the site



Now, lets run through an example. We discussed earlier (as a warning) that sometimes you may not see the barcoding kits when basecalling your runs This is a common bug that was found with one of the patches for the MinKNOW software, caught by some community members

Simply look at the top of the page and enter your query. In this example, lets type "missing barcodes kits" where you will get some suggested items. Either hit Enter to view all items or select one of the quick-launch items

≡		POR	E	Community   barcoding kits minknow  X	]	
\$			'barcodina kits minknow' - 91 Results	Minknow no barcoding kits available after upgrade to no barcoding kits detected in the minKNOW basecalling	evant 🗸	
14			baccoung kits minknow - at results	Errors after Minknow 21.11.7 update	evant	
(G)	Dashboard			MINKNOW 21.11.7 "stopped with error" when		
<u>&amp;</u>	Getting started			Sequencing samples barcoded with different barcoding kits in		
	Posts		Q Minknow no barcoding kits available after upgrade to 21.11.1	View all results		
品	Groups	>			5	
Ŕ	Nanopore Learning		Greetings, in order to participate to the Q20+ programme, we upgraded M beta release of guppy_duplex. Minknow was upgraded via: sudo ap	linknow from version 21.11.6 to version 21.11.7 and downloaded the	st Updated: 21/1/2022	
£۵	Documentation					
☆	Updates	>	In the minKNOW basecalling and	barcode analysis (UI).		
Ø	Support		Hi, I installed the minKNOW and guppy following the installation instructio	n. A couple weeks ago, basecalling test run was ok. But today, opening	t Updated: 31/12/2020	
B	LC Blog		minKNOW, it can not find any barcoding kits(no drop-do	Las	t Opdated: 31/12/2020	
di di seconda di secon	NCM Blog		Errors after Minknow 21.11.7 update			
-						
			Dear community and nanopore monitors: My team and I have been experi version: 21.11.7. We have already reinstalled the software, cha	encing some issues after updating the Minknow software to its newest	st Updated: 14/2/2022	
			MINKNOW 21.11.7 "stopped with error" when trying to use k	kit SQK-RBK110-96		
			We update to version 21.11.7 in Linux and when trying to set a Flongle set unhelpful error message: Digging around in the logs, we foun	quencing run with barcoding kit SQK-RBK110-96 we got the following	st Updated: 11/2/2022	
			© Sequencing samples barcoded with different barcoding kits i	in the same run		
			Hi all, I am designing my first ONT experiment and I wonder whether is po	ssible to mix libraries prepared with different barcoding kits in the same	ast Updated: 5/8/2021	
ණ	Notifications	>	run I would like to sequence amplicons barcoded with		aas opdaared, 0/0/2021	

This site is an invaludable tool if you're experiencing issues with your software or hardware. Oftentimes, responses to updates or issues are quick, so it is highly recommended to follow these forums from time-to-time

### Viewing downloads

Oxford also provides a set of software links you can download applications from. These are available at https://community.nanoporetech.com/downloads

$\leftarrow \rightarrow G$	C A https://community.nanoporetech	h.com/downloads		
K>	Ĩ.	MinION Software MinION Release 21.11.8 Nanopore sequencing software		Release notes
Setting started IP Posts 규급 Groups >		Windows	Installation guide	Download A
<ul> <li>Nanopore Learning</li> <li>Documentation</li> </ul>			noteination genera	
숫 Updates >		<b>é</b> Mac	Installation guide	Download 🕹
<ul> <li>Support</li> <li>LC Blog</li> <li>NCM Blog</li> </ul>		∆ Linux	Installation guide	Download 💩

## Setting Up Guppy for GPU

See here for more advanced details

If you have a CUDA-capable GPU in your laptop, follow CUDA and the above link to set it up for MinKNOW

$\leftarrow \  \  \rightarrow \  \   G$	O 🛆 https:/	//community.nanoporetech.com/protocols/experiment-companion-minknow/v/mke_1013_v1_revbz_11apr2016/installing-gpu-version-of-guppy-with-minknow-for-	ninion
To play video	, you may need to install the red	quired video codecs.	
≡ QNA	NOPORE	Communey	Co
Ic>     Dashboard       Ic>     Getting started       Image: Started     Image: Started	>	Constant of a sequence / sequencing software / Minknow Installing a GPU version of Guppy with MinKNOW for MinION on Linux	
Documentation Documentation Updates Updates LC Blog NCM Blog	>	MPONTANT The GPU version of Guppy must be installed on the same system to which the MinION is connected.	\$्रेट Eavourite MinKNOW Protocol
		GPU basecalling is supported on NVIDIA GPUs only, and only on Linux and Windows. The installation of the GPU is done at the user's own risk. Misconfiguration of the GPU may result in slow basecalling and/or a large number of skipped reads if the basecall server crashes due to misparameterisation.	Computer requirements for MinKNOW Downloading and installing MinKNOW
		A GPU with at least 8 GB of memory is recommended. GPUs with less than 4 GB of memory may not work.	Installing MinRNOW on Windows Installing MinRNOW on Mac OS X
Notifications     Profile	>	The following commands need to be entered into a terminal. Note that some of them will require superuser	Installing MinKNOW on Linux

**Note:** Running GPU basecalling is not required if you prefer to run it from the command line, detailed here *Guppy GPU Basecaller* 

## **CUDA**

Ensure that your GPU is CUDA-capable first by typing

```
lspci | grep VGA
```

If you see your GPU model, for example: *NVIDIA Corporation TU102 [GeForce RTX 2080 Ti] (rev A1)* then you have a GPU available on your machine. IF you don't see that AND you know there is a GPU in the machine try to install the drivers first.

Once the drivers are installed go to: https://developer.nvidia.com/cuda-downloads.

Select the appropriate distribution values and copy+paste the commands that populate into your terminal, one-by-one.

On my Ubuntu 20.04 (Focal) machine I head to [here]( https://developer.nvidia.com/cuda-downloads?target\_os=Linux&target\_arch=x86\_64&=Ubuntu&target\_version=20.04&target\_type=deb\_local)

then copy + paste

(continues on next page)

(continued from previous page)

```
sudo apt-key add /var/cuda-repo-ubuntu2004-11-6-local/7fa2af80.pub
sudo apt-get update
sudo apt-get -y install cuda
```

You should then reboot your machine for everything to take full effect

Once installed you can confirm that it is working by writing:

nvidia-smi **and** 

nvcc --version

If both commands return a healthy output, you are all set on CUDA.

### **Barcoding Kits Missing**

Newer installs of Minknow will not render barcoding kits appropriately on starting an analysis or sequencing run. To remedy, you must fix 2-3 files

- 1. sudo nano /opt/ont/minknow/conf/app\_conf and replace use\_tcp to true
- 2. sudo systemctl edit guppyd.service and replace --port with 5555 and add --use\_tcp
- 3. sudo mv /etc/systemd/system/guppyd.service.d/override.conf /etc/systemd/system/ guppyd.service.d/override.conf.old (if it exists)
- 4. Finally, run sudo systemctl daemon-reload to make changes.

**Note:** Latest Installs of MinKNOW break GPU-basecalling. There is no fix (we've) discovered that allows it to perform within MinKNOW directly

**Note:** /etc/systemd/system/guppyd.service.d/override.conf may not exist on your system and won't be needed to be changed

### **Guppy GPU Basecaller**

Finally, you need to configure MinKNOW to use a GPU-capable version of guppy and that the guppy basecaller plays nice with the installed MinKNOW you've pulled.

/opt/ont/minknow/guppy/bin/guppy\_basecaller --version

You should see a version, for example for 5.0.13. In MinKNOW 21.11 it should be around 5.1.15 You MUST download the same version by running:

wget https://mirror.oxfordnanoportal.com/software/analysis/ont-guppy\_<version>\_linux64. tar.gz

Make sure to replace the installed version with the values after ont-guppy\_ e.g. wget https://mirror.oxfordnanoportal.com/software/analysis/ont-guppy\_5.1.15\_linux64.tar.gz

Then, we need to replace the guppy version. Let's first save the cpu-only one before replacing as well.

sudo mv /opt/ont/guppy/bin /opt/ont/guppy/bin.sav && sudo mv /opt/ont/guppy/data / ...opt/ont/guppy/data.sav # Save the old guppy just in case tar -xvzf ont-guppy\_5.1.15\_linux64.tar.gz #Decompress guppy. Replace the version number\_ ...with your own sudo cp -r ont-guppy/bin /opt/ont/guppy/bin && sudo cp -r ont-guppy/data /opt/ont/guppy/ ...data # Move the newly downloaded guppy #Disable online need for minknow to ping external servers sudo /opt/ont/minknow/bin/config\_editor --filename /opt/ont/minknow/conf/sys\_conf --conf\_ ...system --set on\_acquisition\_ping\_failure=ignore sudo service minknow restart # Restart minknow

Then, add these two lines to your *\$HOME/.bashrc* 

Note: Add this to your bashrc for the user if you want to run guppy gpu from the command line

**Warning:** If you perform the above steps and you still can't access/see the GPU with nvidia-settings or nvidia-smi, you might need to disable secure boot Here or here are examples of how to do that. You need to enter BIOS to perform this operation

- 1. Boot and press [F2] to enter BIOS.
- 2. Go to [Security] tab > [Default Secure boot on] and set as [Disabled].
- 3. Go to [Save & Exit] tab > [Save Changes] and select [Yes].
- 4. Go to [Security] tab and enter [Delete All Secure Boot Variables] and select [Yes] to proceed.
- 5. Then, select [OK] to restart.

### OR

1. Open the PC BIOS menu:

You can often access this menu by pressing a key while your PC is booting, such as F1, F2, F12, or Esc.

Or

From Windows, hold the Shift key while selecting Restart. Go to Troubleshoot > Advanced Options: UEFI Firmware Settings.

- 2. Find the Secure Boot setting in your BIOS menu. If possible, set it to Disabled. This option is usually in either the Security tab, the Boot tab, or the Authentication tab.
- 3. Save changes and exit. The PC reboots.

**Note:** As of 21.06, MinKNOW requires an additional step to add CUDA capability (GPU processing) to basecalling and can be found here

See here

In short, the fix quoted at the link states that it requires about 10 steps:

1. Use systemctl to edit the existing guppyd service (this will open a text editor with a copy of the existing service file):

sudo systemctl edit guppyd.service --full

Ensure that, if it exists, the override conf doesn't override our changes

sudo mv /etc/systemd/system/guppyd.service.d/override.conf /etc/systemd/system/guppyd. service.d/override.conf.old

2. Edit that new service file to point to your GPU version of guppy, and add the appropriate device flag. You can change any other server arguments at the same time.

For example, change this line in the service file:

ExecStart=/opt/ont/guppy/bin/guppy\_basecall\_server <things>

... to this (make sure you retain the --port argument exactly as it used to be - this is how MinKNOW communicates with the basecall server):

ExecStart=/home/myuser/ont-guppy/bin/guppy\_basecall\_server <things> -x cuda:all

- 3. Save the file and exit the text editor (the filename may look odd, but don't worry systemctl should change it to the correct name later).
- Do the same for /etc/system/guppyd.service.d/override.conf (edit with the addition of -x cuda:all)
- 5. Stop the MinKNOW service, as described in the documentation.
- 6. Stop the guppyd service.
- sudo service guppyd stop
  - 7. Check that guppy is no longer running, as described in the documentation, killing any existing basecall servers as required.
  - 8. Start the guppyd service.

9. Check that the correct version of guppy is running, as described in the documentation. If the guppy basecall server isn't launching correctly, check its log output using journalctl ("-n 100" shows the last 100 entries in the journal) to see what's going wrong:

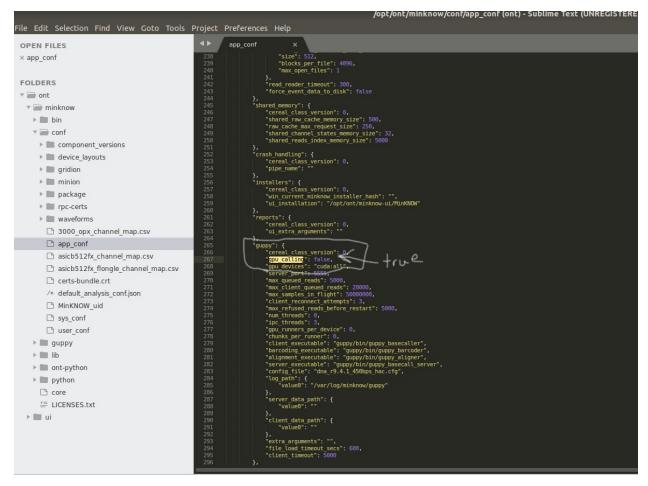
sudo journalctl -u guppyd.service -n 100

10. Start the MinKNOW service.

sudo service minknow restart

You will also need to adjust the configuration file for guppy by modifying /opt/ont/minknow/conf/app\_conf. Adjust the gpu\_calling field to true in the JSON, being careful not to modify/delete any commas or quotations.

sudo service guppyd start



From there you are all set to run basecalling directly within the MinKNOW application.

If you ever experience issues where the UI does not show experiments once started (Basecalling or Sequencing), try:

- 1. Close MinKNOW (UI)
- 2. sudo service minknow restart
- 3. Make sure that a MinION or other Oxford Nanopore devices is plugged in and running
- 4. Restart MinKNOW (UI)
- 5. Re-attempt experiment such as basecalling. Often times experiments will then show up

### **Reduce Runners if GPU basecalling fails**

Note: See here

Occasionally, if you've set up GPU basecalling correctly, but still get errors, this may be due to too many runners being called. You can adjust this easily in the app\_conf file. Adjust the "chunks\_per\_runner" parameter

- If using HAC, set it to "chunks\_per\_runner": 160
- If using SUP, set it to "chunks\_per\_runner": 10

### **Running Command-line GPU Basecalling**

Note: Ensure you've installed a GPU compatible variant by following Guppy GPU Basecaller.

```
guppy_basecaller -x cuda:all -i <fast5_folder> -r -s ./fastq_pass -c dna_r9.4.1_450bps_

→hac.cfg
```

**Note:** If you get an error (or it doesn't exist) about guppy\_basecaller not supporting GPU calling, please place the binaries in your path.

If you set up MinKNOW to use guppy GPU, ensure that it is properly working and setup in /opt/ont/guppy. Follow *Guppy GPU Basecaller* to set this up Ultimately, you need to run the ont-guppy/\* contents into /opt/ont/guppy/

To check your status of your GPU (ensure CUDA is installed by following CUDA) by running nvidia-smi

#### **Reinstalling MinKNOW**

If you experience issues with MinKNOW, one potential solution is to do a purge of MinKNOW and Reinstalling

Notes (below) are gathered from the Oxford Community Forums here

1. First purge MinKNOW and remove its dependencies with these two commands:

sudo apt purge minion-nc

sudo apt autoremove

2. Check that there are no residual config files left over:

dpkg --list | grep -e minknow -e minion -e guppy

3. If Step 2 returns any results, please manually purge those packages like so:

sudo apt purge package1 package2 package3 etc

4. Delete the minknow installation directory:

sudo rm -rf /opt/ont/minknow

5. Ensure there are no Guppy files or folders present in the following directory:

ls -l /etc/systemd/system/

- 6. If any Guppy files or folders are seen in Step 5, please delete them before continuing.
- 7. After rebooting your computer, update the package listing and install MinKNOW:

sudo apt update

sudo apt install minion-nc

8. Once MinKNOW is installed, please open MinKNOW and start a test sequencing run with CPU basecalling to ensure it is working as expected before configuring your GPU. You can start an experiment with a CTC or used flow cell for the purposes of this test.

9. To configure your GPU, please follow Steps 1-15 on the this page of the MinKNOW user guide. Based on user feedback, we've updated it to make a few steps a bit clearer and I've confirmed on my own laptop that these work for this patch.

# **Supplemental**

**Note:** These tools are optional and not needed to run any Basestack feature They are not bundled within Basestack (supplemental to bioinformatic/genomics analysis) and must be downloaded separately. They are (mainly) discussed in the bioinformatics modules of the workshop provided with Basestack

## • BEAST v1.10.4:

- https://beast.community/
- Available in all 3 OS types
- MEGA X
  - https://www.megasoftware.net/
  - Available in all 3 OS types
- AliView v1.26:
  - https://ormbunkar.se/aliview/downloads/linux/linux-version-1.26/
  - Available in all 3 OS types
- MAFFT v7
- https://mafft.cbrc.jp/alignment/software/
- Available in all 3 OS types
- BEAGLE v3.1.2

### build from source (instructions)

- https://github.com/beagle-dev/beagle-lib/
- Available in all 3 OS Types
- TempEst v1.5.3
  - http://tree.bio.ed.ac.uk/software/tempest/
  - Available in all 3 OS Types
- Tracer v1.7.1
  - https://github.com/beast-dev/tracer/releases/tag/v1.7.1
  - Available in all 3 OS Types
- FigTree v1.4.4
  - http://tree.bio.ed.ac.uk/software/figtree/
  - Available in all 3 OS Types

Requires Internet. Also available at https://igv.org/app/

- IQTree v2.1.1
  - https://github.com/iqtree/iqtree2/releases

- You will find the executable in the bin sub-folder. Copy all files in bin folder to your system search path such that you can run IQ-TREE by entering iqtree from the Terminal.
- Available in all 3 OS Types
- FastTree v2.1.11 (double precision variant)
  - http://www.microbesonline.org/fasttree/
  - Available in all 3 OS Types (Less user-friendly)
- Python3
- https://www.python.org/downloads/
- Available in all 3 OS Types
- biopython
  - https://biopython.org/wiki/Download
  - Install after installing/updating Python

### **NVIDIA Jetson Setup**

If you need to set up Basestack, you must install the Nightly build variant of the distribution at: https://github.com/ jhuapl-bio/Basestack/releases/latest

**Note:** Be sure to set up the Nano or Xavier (NX) properly before doing this. See more information here: https://developer.nvidia.com/embedded/downloads

**Note:** Depending on your distribution, you may already have docker installed it seems as NVIDIA is more fully supporting the Docker build toolkit.

To check, run *docker –version*. If so, skip that section of the supplemental software install process (see below) You will definitely need to be sure to run through the post-installation steps, though. See here: *Docker Installation* 

#### Simply follow these 2 steps:

- 1. Download the arm64 AppImage from the above-mentioned releases page
- 2. OPTIONAL: Run the install script. This can be found here

**Note:** The install script described above will ask several questions for setting up your environment based on your needs. Make sure to select the **r** option for any question asking for arm64 or am664

Also, try to follow steps in this link for minknow Or, try this link if you have troubles with CUDA and Guppy

**Warning:** The later versions of MinKNOW can lead to a failure to load the barcoding kits in the UI. If you experience this, follow this procedure *Barcoding Kits Missing* to remedy

**Note:** JetPack (Nviida Jetson only) does not ship with nvidia-smi as a way to monitor gpu usages. You can instead run sudo -H pip install -U jetson-stats to get the jtop command to monitor resources on your jetson device

## **Special Mentions**

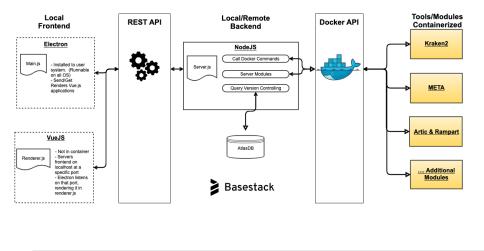
### Mentions ###

- Andrew Rambaut, James Hadfield and Team @ artic-network
- Trevor Bedford and Team @ Nextstrain
- UC San Diego and Broad Inst. @ IGV

### Authors of Basestack ##

• Brian Merritt (brian.merritt@jhuapl.edu)

# 3.3 API





# 3.3.1 Development Setup of Server and App

```
conda activate basestack
npm run dev:server
npm run dev:app (if running the app in parallel)
```

# 3.3.2 Components

### **Modules**

Basestack uses these calls for anything involving modules specifically in the app ecosystem:

## ping POST

Note: To get the status of the server at the specified port, you can use the curl or Postman to test your calls.

curl localhost:5003/server/ping

## getAllCatalog GET

curl localhost:5003/catalog/all/get

Gets all available catalogs, both remote and/or installed

## **Parameters**

### Returns

Array - List of all catalog entries available either remotely situated or locally available

## Example

```
{
 "icon": "dna",
 "title": "Minimap2",
 "tags": [
   "minimap2",
   "alignment",
   "genomics"
 ],
 "status": {
   "installed": true,
   "latest": null,
   "building": true,
   "version": null,
   "running": false,
   "error": null
 },
 "name": "minimap2",
 "remotes": [],
 "modules": [
    {
      "status": {
        "fully_installed": true,
        "latest": null,
        "building": true,
        "version": null,
```

(continues on next page)

(continued from previous page)

```
"partial_install": true,
  "running": false,
  "error": null
},
"name": "minimap2",
```

### getInstalledCatalog GET

curl localhost:5003/catalog/installed/get

Gets all installed catalogs, both remote and/or locally found in the app

### **Parameters**

### Returns

Array - List of installed catalog entries available either remotely situated or locally available

### Example

```
{
 "icon": "dna",
 "title": "Minimap2",
 "tags": [
   "minimap2",
   "alignment",
   "genomics"
 ],
 "status": {
   "installed": true,
   "latest": null,
   "building": true,
   "version": null,
   "running": false,
   "error": null
 },
 "name": "minimap2",
 "remotes": [],
 "modules": [
    {
      "status": {
       "fully_installed": true,
        "latest": null,
        "building": true,
        "version": null,
        "partial_install": true,
```

(continues on next page)

(continued from previous page)

```
"running": false,
"error": null
},
"name": "minimap2",
```

## moduleBuildDependency POST

localhost:5003/module/build/dependency

Installs a single dependency belonging to a module's procedure (version)

### **Parameters**

- 1. dependency Int Dependency index written for a procedure in the configuration of it
- 2. catalog String Catalog name
- 3. module Int Index of the module version for a catalog entry
- 4. procedure Int Index of the procedure in the module

### Returns

Status - Returns successful kickoff of installation or Error

# **Example Body**

```
{
    "procedure": 1,
    "catalog": "mytax",
    "module": 0,
    "dependency": 0
}
```

### procedureBuild POST

```
localhost:5003/procedure/build
```

Installs all dependencies belonging to a procedure

# **Parameters**

- 1. catalog String Catalog name
- 2. module Int Index of the module version for a catalog entry
- 3. procedure Int Index of the procedure to completely install

### Returns

Status - Returns successful kickoff of installation or Error

# **Example Body**

```
{
    "procedure": 0,
    "catalog": "mytax",
    "module": 0
}
```

### moduleBuild POST

localhost:5003/module/build

Installs all procedures and their corresponding dependencies for a given catalog's module

# **Parameters**

- 1. catalog String Catalog name
- 2. module Int Index of the module version for a catalog entry

### Returns

Status - Returns successful kickoff of installation or Error

### **Example Body**

```
{
    "catalog": "mytax",
    "module": 0
}
```

## **Procedures**

Note: To get the status of the server at the specified port, you can use the curl or Postman to test your calls.

### getProcedures GET

curl localhost:5003/procedures/get/:catalog/:module/:token

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

### Returns

Array - List of all procedures that belong to a specific version of a module in the catalog

### Example

```
{
 "status": {
   "error": null,
   "stream": null,
   "running": false,
   "fully_installed": true,
   "partial_install": true
 },
 "dependencies": [
   {
     "target": "jhuaplbio/basestack_mytax",
     "type": "docker",
      "version": "latest",
     "format": "docker",
     "status": {
        "downloading": false,
        "decompressing": false,
        "exists": {
         "version":
→ "sha256:4762c42837c0fcffcb36a08398d7aafff57de72584804e797a63bc28ddd9c9b1"
       },
        "error": null,
       "stream": null,
        "fully_installed": false,
        "partial_install": false,
        "version":
→ "sha256:4762c42837c0fcffcb36a08398d7aafff57de72584804e797a63bc28ddd9c9b1"
     }
   },
```

### runProcedure POST

localhost:5003/procedure/run

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

### **Parameters**

- 1. module Int Index of the version of the module
- 2. catalog String Name of the specific Catalog
- 3. variables Object Object of all unique variable names to assign to 1 or more services
- 4. token String Autogenerated token on startup of Basestack
- 5. procedure Int Index of the procedure for the module. Comprises 1 or more services

## Returns

Status - Success or Error on procedure (1 or more service ) completion

### cancelProcedure POST

localhost:5003/procedure/cancel

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

### **Parameters**

- 1. module Int Index of the version of the module
- 2. catalog String Name of the specific Catalog
- 3. token String Autogenerated token on startup of Basestack
- 4. procedure Int Index of the procedure for the module. Comprises 1 or more services

### Returns

Status - Success or Error on cancelation of the procedure (1 or more service )

### Jobs

Note: To get the status of the server at the specified port, you can use the curl or Postman to test your calls.

## getJob GET

curl localhost:5003/service/get/:catalog/:module/:procedure/:service/:token

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

# Returns

Array - List of all services that belong to a specific version of a module's procedure in the catalog

### Example

```
curl localhost:5003/service/get/mytax/0/1/0/development
{
"status": 200,
"message": "Completed job setting",
"data": {
    "exists": false,
    "fully_installed": true,
    "procedure": {
        "error": null,
        "stream": null,
        "running": false,
        "building": false,
        "fully_installed": true,
        "partial_install": true
    },
    "dependencies": [
        {
            "downloading": false,
            "decompressing": false,
            "dependComplete": true,
            "exists": {
                "version":
```

(continues on next page)

(continued from previous page)

```
→ "sha256:4762c42837c0fcffcb36a08398d7aafff57de72584804e797a63bc28ddd9c9b1"
},
    "error": null,
    "stream": null,
    "fully_installed": fals
```

## startJob POST

localhost:5003/job/start

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

Starts a job, sets necessary variables, designates services and beings the Docker API communication processes. Multiservice procedures will run all services sequentially

### **Parameters**

- 1. module Int Index of the version of the module
- 2. catalog String Name of the specific Catalog
- 3. variables Object Object of all unique variable names to assign to 1 or more services
- 4. token String Autogenerated token on startup of Basestack
- 5. procedure Int Index of the procedure for the module. Comprises 1 or more services
- 6. service List of Int Index of the services for the selected procedure.

### Returns

Status - Success or Error on procedure (1 or more service ) completion

Note: Test data download from Gdrive location

### **Example Body**

```
{
    "service": 0,
    "catalog": "mytax",
    "module":0,
    "procedure": 0,
    "token": "development",
```

(continues on next page)

(continued from previous page)

# cancelJob POST

localhost:5003/job/cancel

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

Stops the running job. Keeps it present in the backend but it no longer runs its services (and cancels any currently running ones)

## **Parameters**

- 1. module Int Index of the version of the module
- 2. catalog String Name of the specific Catalog
- 3. token String Autogenerated token on startup of Basestack
- 4. procedure Int Index of the procedure for the module. Comprises 1 or more services

### Returns

Status - Success or Error on cancelation of the service

### setJob POST

localhost:5003/job/set

Note: Where token is a uniquely generated token on accessing the app. Default development is always generated

Sets the variables and creates a job for a given procedure that is being viewed from the application piece of the app

# **Parameters**

- 1. module Int Index of the version of the module
- 2. catalog String Name of the specific Catalog
- 3. variables Object Object of all unique variable names to assign to 1 or more services
- 4. token String Autogenerated token on startup of Basestack
- 5. procedure Int Index of the procedure for the module. Comprises 1 or more services
- 6. services List of Int Index of the service(s) for the procedure.

## Returns

Status - Success or Error on setting job (1 or more services ) completion

## **Example Body**

```
setJob ``POST``
{
    "token": "development",
    "procedure": 0,
    "catalog": "mytax",
    "module": 0,
    "services": [ 0 ],
    "variables": {
        "file": {
            "source": "/Users/merribb1/Desktop/test-data2/metagenome/sample_metagenome.
→fastq"
        },
        "db": {
            "option": 1
        },
        "nodes": {
            "option": 0
```

# System

Note: To get the status of the server at the specified port, you can use the curl or Postman to test your calls.

## **Pinging the Server GET**

Note: You can use Postman as your API management toolkit for testing the server

curl localhost:5003/server/ping

#### returns

{"status":200,"message":"Server is running at port: 5003"}

### getServerLogs GET

curl localhost:5003/log/system

### Returns

Array - List of all services that belong to a specific version of a module's procedure in the catalog

### Example

curl localhost: 5003/log/system { "status": 200, "message": "Got system log", "data": [ "2022-02-14T23:17:12.235Z [info]: stdout: Indel at position 29094: [0, 75, 0, 0, 0, 0,  $\rightarrow 13, 0] 75'',$ "Indel at position 29130: [0, 0, 1, 77, 0, 4, 15] 78", ···, "2022-02-14T23:17:12.236Z [info]: stdout: Indel at position 29323: [230, 0, 0, 0, 0, \_\_ →35, 18] 230", "Indel at position 29376: [0, 0, 0, 97, 0, 2, 41] 97", ···· . "2022-02-14T23:17:12.237Z [info]: stdout: Indel at position 29385: [0, 125, 0, 0, 0, \_ →21, 41] 125", ···., "2022-02-14T23:17:12.237Z [info]: stdout: Indel at position 29426: [2, 0, 73, 0, 0, **→13, 0] 75",** ···· , "2022-02-14T23:17:12.237Z [info]: stdout: Indel at position 29753: [0, 110, 0, 0, 0, ... →3, 18] 110", ··· . "2022-02-14T23:17:12.238Z [info]: stdout: Indel at position 29775: [92, 0, 0, 0, 0, 0,  $\leftrightarrow 7$ , 16] 92", ···., "2022-02-14T23:17:12.239Z [info]: stdout: Indel at position 29799: [1, 0, 119, 0, 0, (continues on next page)

(continued from previous page)

```
→3, 35] 120",
    "",
    "2022-02-14T23:17:12.277Z [info]: stdout: -[2022-02-14 23:17:12] Starting Module 4.
    →Merging and Allele Frequencies on /opt/data/artic-pipeline/4-draft-consensus/
    →Sample3_NB03.nanopolish.merged.vcf, /opt/data/artic-pipeline/4-draft-consensus/Sample3_
    →NB03.medaka.merged.vcf, /opt/data/artic-pipeline/4-draft-consensus/Sample3_NB03.
    →samtools.vcf",
    "",
    "2022-02-14T23:17:13.124Z [info]: stdout: ^[2022-02-14 23:17:13] SAMPLE Sample3_
    →NB03: Module 4 Samtools and Merging: processing complete",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
    "",
```

Note: Above logs were delivered during a run of a Consensus pipeline procedure