

ЗМІСТ – ГІПЕРПОСИЛАННЯ НА ПРОЦЕДУРУ

- [Запуск робочого процесу аналізу hqSNP \(5.1.1\)](#)
- [Запуск робочого процесу аналізу kSNP3 \(5.1.2\)](#)
- [Перегляд та завантаження матриці hqSNP \(5.2.3\)](#)
- [Завантаження та перегляд дерева hqSNP \(дендрограми\) \(5.2.4\)](#)
- [Перегляд та завантаження матриці kSNP3 core SNP \(5.2.5\)](#)
- [Завантаження та перегляд дерева kSNP3 core SNP \(дендрограми\) \(5.2.6\)](#)

1. **МЕТА:** Описати процедуру виконання філогенетичного аналізу даних повногеномного секвенування (WGS), що будуть використовуватися для епідагляду PulseNet International (PNI) за допомогою хмарної платформи Terra.Bio.

2. **СФЕРА ЗАСТОСУВАННЯ:** Ця процедура застосовується до всього персоналу PulseNet, який використовує платформу Terra.Bio для виконання філогенетичного аналізу даних WGS з метою виконання заходів епідагляду в мережі PulseNet International. Дана СОП охоплює аналіз hqSNP та kSNP3 з використанням даних короткого зчитування Illumina. Завантаження послідовностей та метаданих до Terra.Bio, оцінка якості послідовностей, робочі процеси складання та генотипування, а також завантаження послідовностей до NCBI охоплюються стандартною операційною процедурою SOP PNID01 (Стандартна операційна процедура PulseNet International для аналізу даних короткого зчитування WGS Illumina з використанням платформи Terra.Bio).

3. ВИЗНАЧЕННЯ/ТЕРМІНИ:

- 3.1 **Основний SNP:** одонуклеотидний поліморфізм, присутній у геномній області, яка є спільною для всіх вирівняних геномів набору запитів.
- 3.2 **Покриття:** середня кількість зчитувань, що включає даний нуклеотид у реконструйованій послідовності.
- 3.3 **CSV:** значення, розділені комами.
- 3.4 **Дендрограма:** розгалужена діаграма, що відображає відносини подібності між групами об'єктів.
- 3.5 **DeNovo assembly:** Збірка послідовності, згенерована з коротких необроблених зчитувань без використання еталонного геному.
- 3.6 **FASTA:** текстовий формат для представлення нуклеотидних або пептидних послідовностей, в якому пари основ або амінокислоти представлені однолітерними кодами. Послідовність у форматі FASTA починається з однорядкового опису, за яким ідуть рядки даних послідовності. Рядок опису відрізняється від даних послідовності символом «більше ніж» («>») у першому стовпці.
- 3.7 **FASTQ:** текстовий формат для зберігання як біологічної послідовності, так і відповідних їй показників якості.
- 3.8 **hqSNP:** високоякісний одонуклеотидний поліморфізм. Поліморфізм одного нуклеотиду, який було ідентифіковано шляхом зіставлення необроблених даних з еталонним геномом і перевірено за допомогою таких критеріїв, як: покриття

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 2 з 30

послідовності, якість послідовності, частота популяції та алелі. Хоча ці параметри будуть відрізнятися залежно від контексту та застосування SNP-визначення, типовими мінімальними критеріями є показник покриття 10-20X, >Q20 та частота 95%.

- 3.9 kSNP3:** Робочий процес одонуклеотидного поліморфізму без еталонного геному, в якому нуклеотиди непарної довжини k (к-мери) витягуються з de-novo збірки набору запитів послідовності. Витягнуті k-мери з одного геному потім порівнюються з k-мерами з інших геномів у наборі запитів і перевіряються на наявність змінних середніх основ, які використовуються для побудови філогенезу.
- 3.10 Lyve-SET:** конвеєр hqSNP, розроблений Katz *et al.* (2017), який використовується Відділом лабораторії кишкових захворювань Центру CDC, і вбудований у робочий процес опису мови (WDL) для використання на Terra.Bio.
- 3.11 MEGA:** Молекулярний аналіз еволюційної генетики (Molecular Evolutionary Genetics Analysis). Безкоштовне програмне забезпечення для побудови та перегляду філогенетичних дерев, спочатку розроблене Університетом штату Пенсильванія, доступне за [адресою https://www.megasoftware.net/](https://www.megasoftware.net/).
- 3.12 NWK:** Формат файлів Newick. Текстовий формат для представлення дерев у формі, придатній для зчитування комп'ютером, із використанням вкладених дужок і ком.
- 3.13 PNI:** Організація PulseNet International
- 3.14 Матриця SNP:** таблиця 2x2, що містить парні номери SNP між усіма парами зразків, включених в аналіз.
- 3.15 SOP:** Стандартна операційна процедура.
- 3.16 Terra.Bio:** хмарна платформа для аналізу послідовностей, розроблена Інститутом Броуда Массачусетського технологічного інституту та Гарвардським університетом і яка використовується компанією Theiagen Genomics (Хайлендс-Ранч, штат Колорадо, США) для надання платної загальної платформи для лабораторій громадського здоров'я для розміщення, аналізу та обміну даними (Libuit *et al.*, 2023).
- 3.17 TSV:** Значення, розділені табуляцією.
- 3.18 WGS:** Повногеномне секвенування.

4. ОБОВ'ЯЗКИ:

- 4.1 Персонал PulseNet виконує філогенетичний аналіз та виявлення кластерів, за необхідності, на коротких послідовностях Illumina, згенерованих для міжнародного епідагляду PulseNet з використанням платформи Terra.Bio.

5. ПРОЦЕДУРА:

- 5.1 **Виконайте філогенетичний аналіз.** Для філогенетичного аналізу доступні два варіанти: **(5.1.1) аналіз hqSNP**, залежний від еталонної послідовності, який підходить тоді, коли очікується, що послідовності будуть відносно близько пов'язані і можуть бути проаналізовані за допомогою однієї еталонної послідовності; **(5.1.2) аналіз kSNP3** без еталонної послідовності, який тоді підходить, коли очікується, що послідовності, які включені в аналіз, будуть настільки віддалено пов'язані, що

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 3 з 30

використання єдиної еталонної послідовності неможливе або їх спорідненість невідома.

ПРИМІТКА: послідовності, включені в один і той же аналіз *hqSNP* або *kSNP3*, повинні бути присутніми в **одній і тій же** таблиці даних.

5.1.1 Запустіть робочий процес аналізу *hqSNP*

ПРИМІТКА: *Lyve-SET* генерує дерево лише за умови достатньої різноманітності (*SNP*) у наборі зразків. Набір зразків, що складається з технічних реплікатів одного і того ж штаму, не має достатньої різноманітності для генерації дерева і призведе лише до створення матриці парних *SNP*. Щоб згенерувати дерево, додайте до набору зразків винятковий (непов'язаний) штам.

5.1.1.1 Створіть **набір** зразків для філогенетичного аналізу:

ПРИМІТКА: якщо ваш аналіз *hqSNP* містить лише послідовності, які були проаналізовані одночасно за допомогою робочого процесу *TheiaProk*, ви можете використовувати набір, створений автоматично робочим процесом *TheiaProk*. Номенклатура набору відповідає формату *TheiaProk_Illumina_PE_PHB_рік-місяць-датаГодина-хвилина-секунда*. Якщо ви бажаєте включити послідовності, завантажені та проаналізовані в різний час, вам потрібно виконати наведені нижче інструкції, щоб створити набір вручну.

CDC_ATCC_Sequences_set_id	CDC_ATCC_Sequences	kimp3_core_snp_matrix	kimp3_core_snp_matrix_st
<input type="checkbox"/> TheiaProk_Illumina_PE_2023-03-28T16:25:03	ATCC-17802-C1_VP, ATCC-17802-C2_VP, ATCC-17802-C3...	[9 entities]	
<input type="checkbox"/> TheiaProk_Illumina_PE_2023-03-29T16:27:29	O3-98ah-MI, C6472ah-MI, D5480ah-MI, H8394ah-MI, K34...	[5 entities]	
<input type="checkbox"/> TheiaProk_Illumina_PE_2023-08-04T11:52:52	ATCC-33560-C1-1_Prep_Seq, ATCC-33560-C10_Prep_Se...	[13 entities]	
<input type="checkbox"/> TheiaProk_Illumina_PE_PHB_2024-05-23T11:57:36	ATCC-10708-C1_SE, ATCC-10708-C2_SE, ATCC-10708-C3...	[3 entities]	
<input checked="" type="checkbox"/> TheiaProk_Illumina_PE_PHB_2024-05-24T11:50:18	O3-98ah-MI, C6472ah-MI, D5480ah-MI, H8394ah-MI, K34...	[5 entities]	
<input type="checkbox"/> Vibrio_colony_picks	ATCC-17802-C1_VP, ATCC-17802-C2_VP, ATCC-17802-C3...	[4 entities]	

5.1.1.1.1 У розділі «Terra Workspaces» виберіть вкладку «Data», а потім виберіть таблицю даних, що вас цікавить, наприклад «CDC_ATCC_Sequences».

5.1.1.1.2 На вкладці «data» виберіть послідовності, які потрібно включити до філогенетичного аналізу.

5.1.1.1.3 У спадному меню «Edit» виберіть «Save selection as set».

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 4 з 30

ATCC Sequences Id	amrfinderplus_amr_genes	amrfinderplus_virulence_genes	plasmidfinder_plasmids
33560-C1_C3	No AMR genes detected by NCBI-A...	No VIRULENCE genes detected by N...	No plasmids detected in database
33560-C2_C3	No AMR genes detected by NCBI-A...	No VIRULENCE genes detected by N...	No plasmids detected in database
33560-C3_C3	No AMR genes detected by NCBI-A...	No VIRULENCE genes detected by N...	No plasmids detected in database
ATCC-33560A_C3	No AMR genes detected by NCBI-A...	No VIRULENCE genes detected by N...	No plasmids detected in database
ATCC-33560B_C3	No AMR genes detected by NCBI-A...	No VIRULENCE genes detected by N...	No plasmids detected in database
ATCC-51812-C1_SE	mdsA,mdsB	iroC,iroB,sodC1,sinH	IncFIB(S),IncFII(S)
ATCC-51812-C2_SE	mdsA,mdsB	iroC,iroB,sinH,sodC1	IncFIB(S),IncFII(S)
ATCC-51812-C3_SE	mdsA,mdsB	iroB,iroC,sinH,sodC1	IncFIB(S),IncFII(S)
BAA-679A-C1_LM	lin,fosX	No VIRULENCE genes detected by N...	No plasmids detected in database
BAA-679A-C2_LM	lin,fosX	No VIRULENCE genes detected by N...	No plasmids detected in database
BAA-679A-C3_LM	lin,fosX	No VIRULENCE genes detected by N...	No plasmids detected in database
BAA-679A-1 M			

5.1.1.1.4 У спливаючому вікні вкажіть назву набору, наприклад «Salmonella_colony_picks» або «Salmonella_cert_labid_initials». **ПРИМІТКА:** пробіли та тире не допускаються.

Create a CDC_ATCC Sequences set

Set name (required)

Salmonella_colony_picks

CANCEL SAVE

5.1.1.1.5 Натисніть «Save» (скріншот вище).

5.1.1.1.6 Новостворений набір повинен з'явитися в таблиці даних «Set», наприклад, «CDC_ATCC Sequences_set».

CDC_ATCC Sequences set Id	CDC_ATCC Sequences	lyveset_alignment_fasta	lyveset_docker
CjeJuni-dataset-230104	ATCC-33560A_C3, ATC... (2 entities)		
CjeJuni-dataset-230106	ATCC-33560A_C3, ATC... (2 entities)		
Listeria_colony_picks	BAA-679A-C1_LM, BA... (6 entities)	quayio1staphb/	
LM_colony_picks_we_outlier	BAA-679A-C1_LM, BA... (5 entities)	quayio1staphb/	
Lyve_SET_2023-02-10T17-22-53	BAA-679A-C1_LM, BA... (5 entities)	quayio1staphb/	
Salmonella_colony_picks	ATCC-51812-C1_SE, A... (3 entities)	quayio1staphb/	
TheiaProk_illumina_PE_2023-01-03T16-21-59	ATCC-17802_VP, ATC... (6 entities)		
TheiaProk_illumina_PE_2023-01-03T16-42-40	ATCC-17802_VP, ATC... (6 entities)		
TheiaProk_illumina_PE_2023-01-03T17-01-04	ATCC-17802_VP, ATC... (6 entities)		
TheiaProk_illumina_PE_2023-01-04T12-44-34	ATCC-17802_VP, ATC... (6 entities)		
TheiaProk_illumina_PE_2023-01-19T17-35-22	ATCC-17802_VP, ATC... (6 entities)		

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

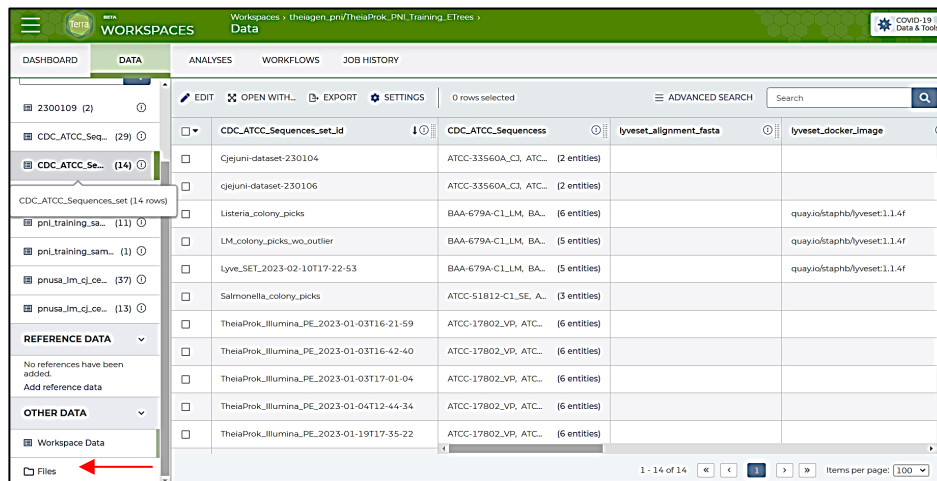
Дата набрання чинності:

Сторінка 5 з 30

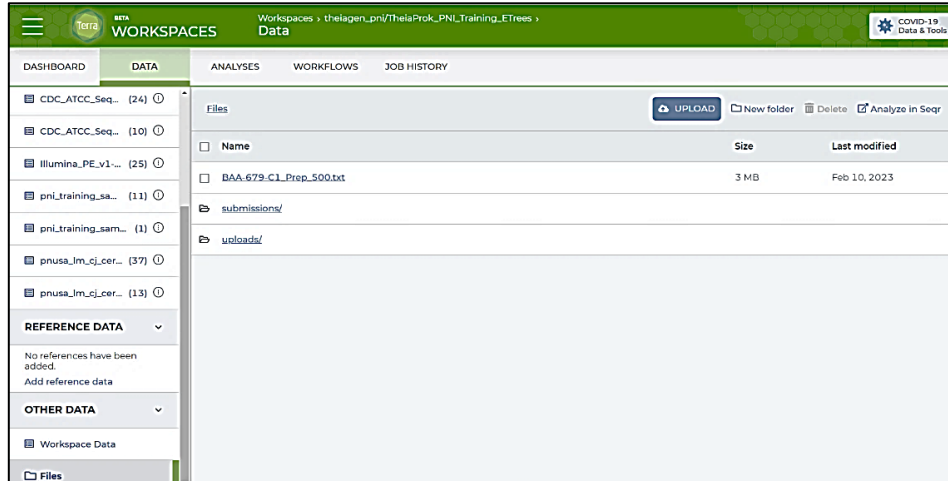
5.1.1.2 Завантажте зовнішню еталонну послідовність, якщо потрібно.

ПРИМІТКА: Цей крок *не* потрібен, якщо посилання вже завантажено до папки «Files» або використовується внутрішня референсна збірка, створена Terra (крок 5.1.1.3).

5.1.1.2.1 У вкладці «Data» прокрутіть до кінця і натисніть «Files» в розділі «Other data».



5.1.1.2.2 На екрані «Files» натисніть «Upload».



5.1.1.2.3 Перейдіть до місця, де збережено файл FASTA.

5.1.1.2.4 Виберіть файл, виділивши його, і натисніть «Open».

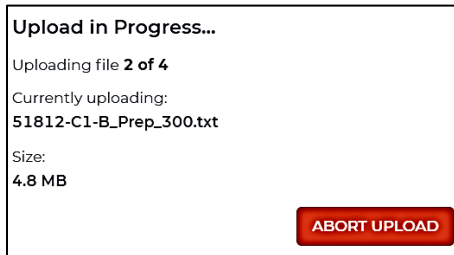
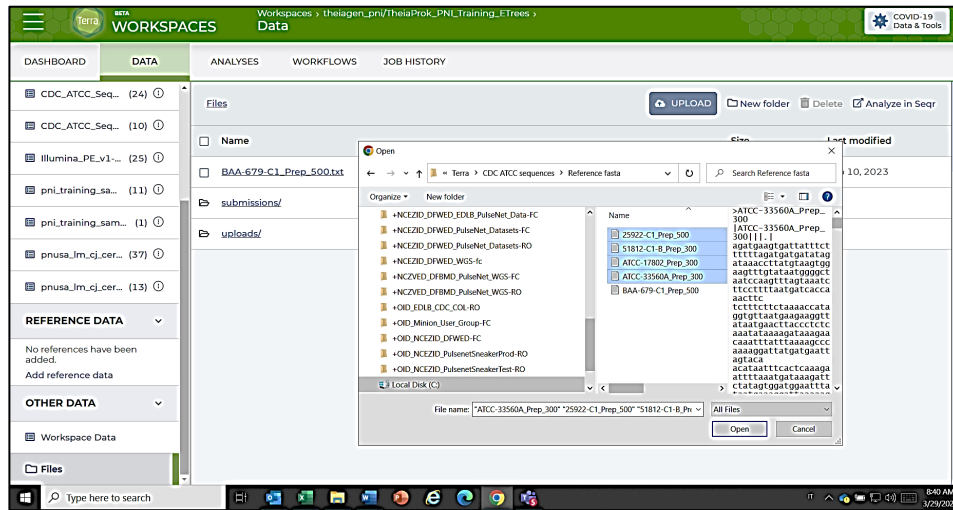
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

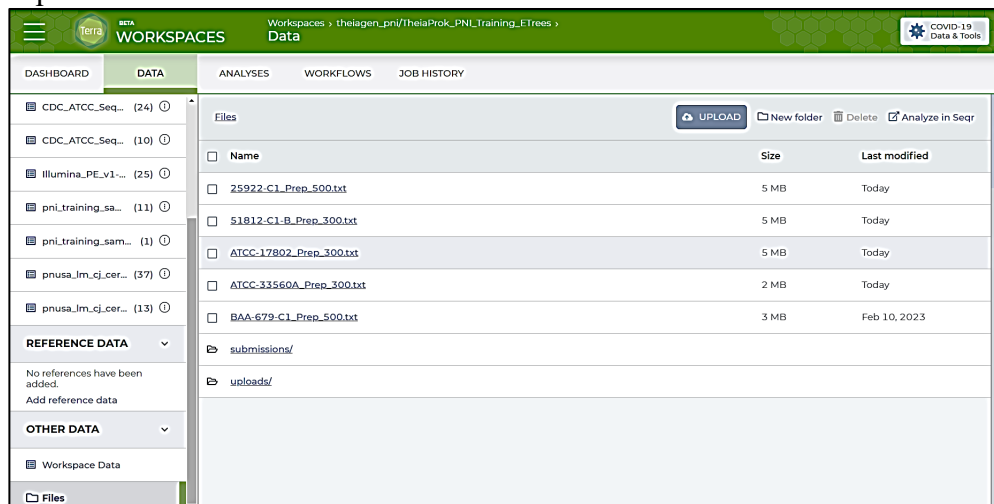
Дата набрання чинності:

Сторінка з 30



5.1.1.2.5 З'явиться спливаюче вікно «Upload in progress».

5.1.1.2.6 Після завершення завантаження файл FASTA повинен бути видимий на екрані «Files».

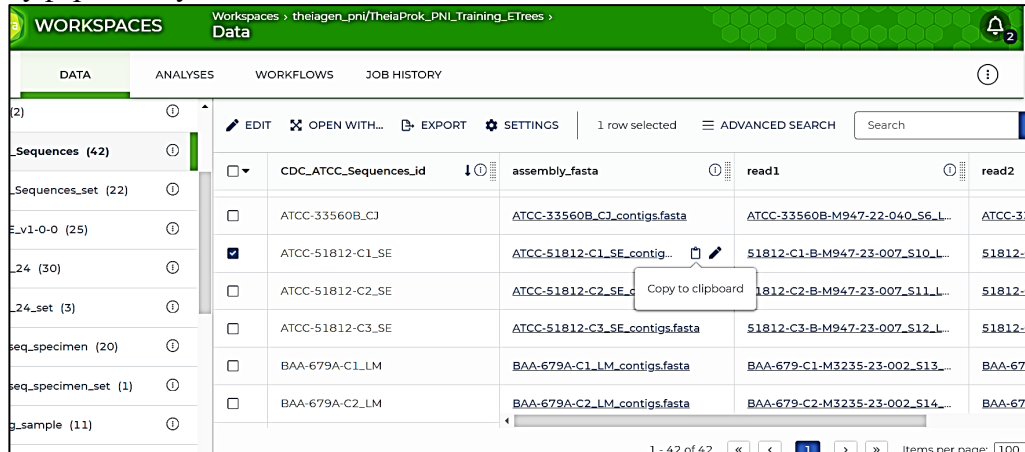


5.1.1.3 Виберіть еталонну збірку, створену Terra, якщо це доречно.

ПРИМІТКА: Цей крок *не* потрібен, якщо використовується еталонний FASTA, завантажений у папку «Files» в кроці 5.1.1.2.

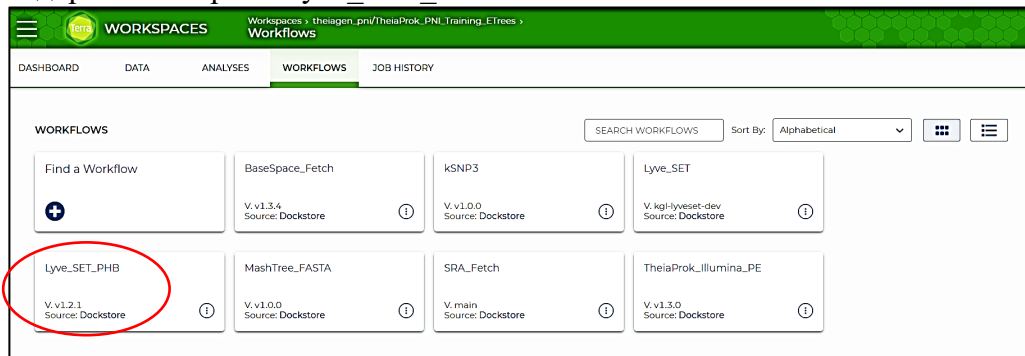
5.1.1.3.1 На вкладці «Data» перейдіть до таблиці даних, що містить потрібну еталонну послідовність.

5.1.1.3.2 У стовпці «assembly_fasta» натисніть на піктограму «Copy to clipboard» для відповідної збірки, щоб скопіювати шлях до місця розташування збірки в буфер обміну.

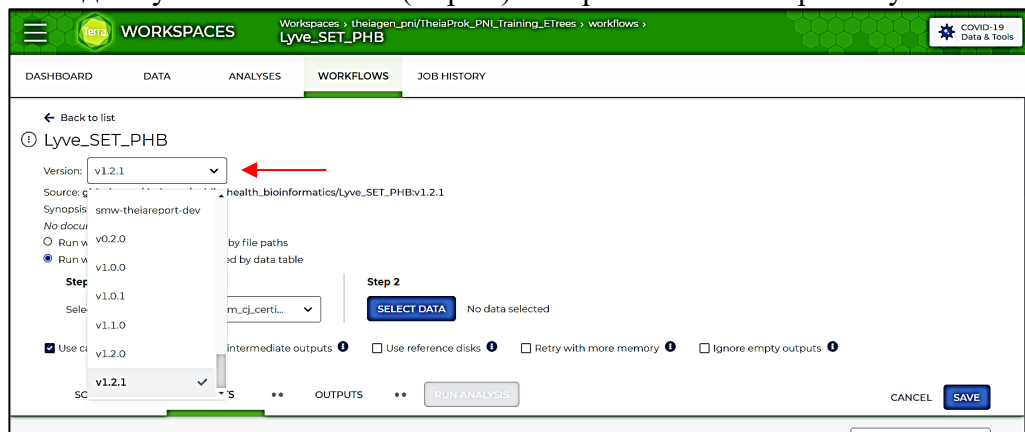


5.1.1.4 Налаштуйте параметри для робочого процесу аналізу Lyve-SET hgSNP.

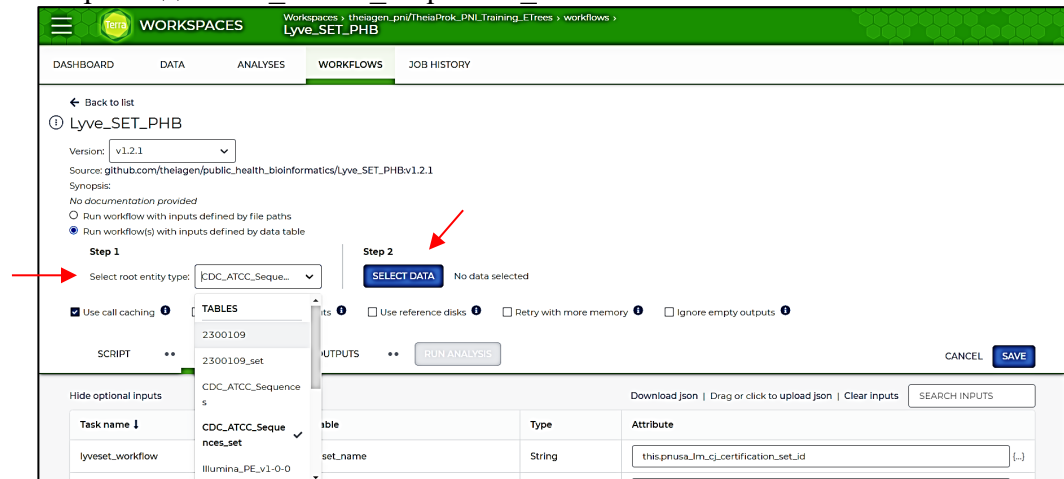
5.1.1.4.1 На вкладці «Workflows» виберіть робочий процес «Lyve_SET_PHB». Відкриється екран «Lyve_SET_PHB».



5.1.1.4.2 У спадному меню «Version» (Версія) виберіть останню версію Lyve-SET.

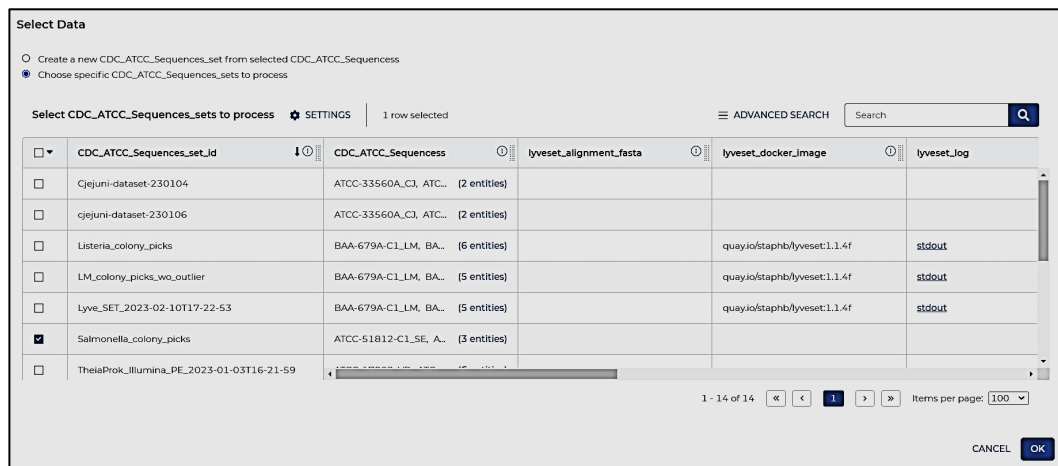


5.1.1.4.3 У кроці 1 з випадного меню «Select root entity type» виберіть таблицю **набору** даних, де знаходиться набір зразків, створений у кроці [5.1.1.1](#), наприклад «CDC ATCC Sequences set».

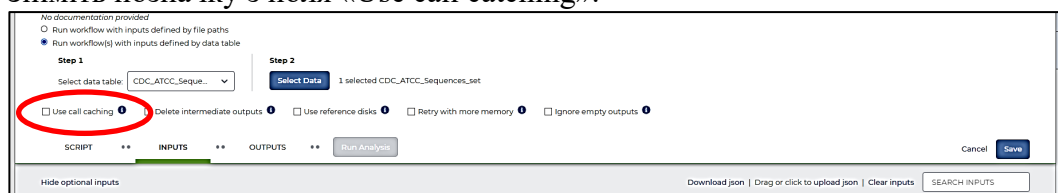


5.1.1.4.4 У кроці 2 натисніть «Select data» (скріншот вище). Це направить вас до таблиці даних, вибраної на попередньому кроці.

5.1.1.4.5 Виберіть бажаний набір зразків, наприклад «Salmonella_colony_picks», і натисніть «ОК».



5.1.1.4.6 Зніміть позначку з поля «Use call caching».



5.1.1.4.7 На вкладці «Inputs» необхідно заповнити наступні «Variables» в полях «Attribute»:

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 9 з 30

ПРИМІТКА: Коли ви заповнюєте стовпець «Attribute», натискання всередині комірки відкриє спадне меню атрибутів, які ви можете вибрати, щоб уникнути помилок.

Task name	Variable	Type	Attribute
lyveset_workflow	dataset_name	String	this.CDC_ATCC-Sequencess.set_id
lyveset_workflow	read1	Array[File]	this.CDC_ATCC-Sequencess.read1
lyveset_workflow	read2	Array[File]	this.CDC_ATCC-Sequencess.read2
lyveset_workflow	reference_genome	File	"gs://fc-b23cdded-78b1-4bfa-944d-95e427aea43e/51812-C1-B_Prep_300.txt"
lyveset_workflow	samplename	Array[String]	this.CDC_ATCC-Sequencess.CDC_ATCC-Sequence_id

5.1.1.4.7.1 Dataset_name: «This.data table set name_id», наприклад, «This.CDC_ATCC-Sequencess.set_id»

5.1.1.4.7.2 Read1: «This.data table set names.read1», наприклад, «This.CDC_ATCC-Sequencess.read1»

5.1.1.4.7.3 Read2: «This.data table set names.read2», наприклад, «This.CDC_ATCC-Sequencess.read2»

ПРИМІТКА: додаткова літера «s» після назви набору таблиць даних для читання 1 і 2 є **ОБОВ'ЯЗКОВОЮ**.

5.1.1.4.7.4 Reference_genome: Можна використовувати зовнішню еталонну збірку, завантажену в папку «Files», або внутрішню еталонну збірку, створену Terra. **ПРИМІТКА:** використовуйте лапки навколо шляху до збірки.

5.1.1.4.7.4.1 **Референсна збірка з папки «Files»:** натисніть на піктограму «Browse bucket files» (Переглянути файли кошика) у полі атрибуту reference_genome (скріншот вище в кроці 5.1.1.4.7.3). З'явиться спливаюче вікно зі списком завантажених референсних геномів. Натисніть на відповідну референсну збірку. Вікно закриється автоматично.

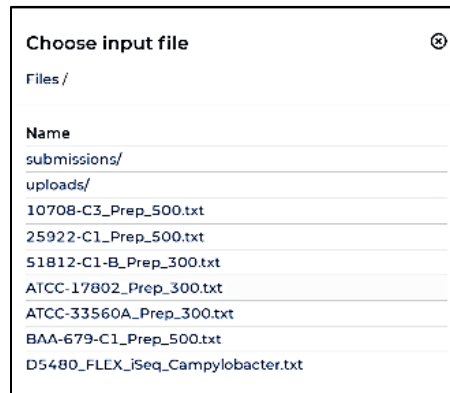
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 10 з 30



5.1.1.4.7.4.2 **Референсна збірка, створена Terra:** Вставте шлях, скопійований у буфер обміну в кроці 5.1.1.3.2, у поле атрибуту «reference_genome».



5.1.1.4.7.5 **samplename:** введіть назву вашої таблиці даних у форматі: **this.data_table_names.data_table_name_id**, «this.CDC_ATCC_Sequences.CDC_ATCC_Sequences_id» (скріншот вище в кроці 5.1.1.4.7).

ПРИМІТКА: подвійний формат імен є **ОБОВ'ЯЗКОВИМ**.

5.1.1.4.7.6 **allowedFlanking:**

5.1.1.4.7.6.1 *Escherichia, Listeria, Salmonella*: 5

5.1.1.4.7.6.2 *Vibrio parahaemolyticus, V. vulnificus*: 75

5.1.1.4.7.6.3 *Campylobacter*: 100

5.1.1.4.7.6.4 *V. cholerae*: залишити порожнім

5.1.1.4.7.7 **min_alt_frac:**

5.1.1.4.7.7.1 *Listeria, V. cholerae*: 0,75

5.1.1.4.7.7.2 *Campylobacter, Escherichia, Salmonella, Vibrio*: 0,95

5.1.1.4.7.8 **min_coverage:**

5.1.1.4.7.8.1 *Listeria, V. cholerae*: 10

5.1.1.4.7.8.2 *Campylobacter, Escherichia, Salmonella, V. parahaemolyticus, V. vulnificus*: 20

5.1.1.4.7.9 **mask_cliffs:**

5.1.1.4.7.9.1 *Escherichia, Salmonella*: true

5.1.1.4.7.9.2 *Campylobacter, Listeria, Vibrio*: залишити порожнім

5.1.1.4.7.10 **mask_phages:**

5.1.1.4.7.10.1 *Campylobacter, Escherichia, Salmonella, V. Parahaemolyticus, V. vulnificus*: true

5.1.1.4.7.10.2 *Listeria, V. cholerae*: залишити порожнім

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 11 з 30

lyveset	allowedFlanking	Int	5
lyveset	cpu	Int	Optional
lyveset	disk_size	Int	Optional
lyveset	docker_image	String	Optional
lyveset	downsample	Boolean	Optional
lyveset	fast	Boolean	Optional
lyveset	mapper	String	Optional
lyveset	mask_cliffs	Boolean	true
lyveset	mask_phages	Boolean	true
lyveset	memory	Int	Optional
lyveset	min_alt_frac	Float	0.95
lyveset	min_coverage	Int	20
lyveset	nomatrix	Boolean	Optional
lyveset	nomsa	Boolean	Optional
lyveset	notrees	Boolean	Optional

ПРИМІТКА: замість створення вручну і заповнення пунктів 5.1.1.4.7.6-5.1.1.4.7.10, ви можете використовувати змінну «presets», ввівши попередньо встановлений організм у лапках, наприклад, «salmonella_enterica». Це забезпечить використання всіх правильних параметрів аналізу для даного організму. Попередньо встановлені параметри доступні для:

- *listeria_monocytogenes*
- *salmonella_enterica*
- *escherichia_coli*

lyveset	nomatrix	Boolean	Optional
lyveset	nomsa	Boolean	Optional
lyveset	notrees	Boolean	Optional
lyveset	presets	String	"salmonella_enterica"
lyveset	read_cleaner	String	Optional
lyveset	sample_sites	Boolean	Optional
lyveset	snpcaller	String	Optional

5.1.1.4.8 На вкладці «Outputs» натисніть «Use defaults» для «Атрибутів», а потім натисніть «Save».

ПРИМІТКА: кнопка «Save» відображається тільки в тому випадку, якщо ви змінили вхідні дані порівняно з попереднім запитом.

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 12 з 30

SCRIPT ** INPUTS ** **OUTPUTS** ** RUN ANALYSIS

Output files will be saved to
Files / submission unique ID / lyeset_workflow / workflow unique ID

References to outputs will be written to
Tables / CDC_ATCC_Sequences_set
Fill in the attributes below to add or update columns in your data table

CANCEL **SAVE**

Download json | Drag or click to upload json | Clear outputs | SEARCH OUTPUTS

Task name ↓	Variable	Type	Attribute Use defaults
lyeset_workflow	lyeset_alignment_fasta	File	this.lyeset_alignment_fasta [-]
lyeset_workflow	lyeset_docker_image	String	this.lyeset_docker_image [-]
lyeset_workflow	lyeset_log	File	this.lyeset_log [-]
lyeset_workflow	lyeset_pairwise_matrix	File	this.lyeset_pairwise_matrix [-]
lyeset_workflow	lyeset_pooled_snps_vcf	File	this.lyeset_pooled_snps_vcf [-]
lyeset_workflow	lyeset_raxml_tree	File	this.lyeset_raxml_tree [-]
lyeset_workflow	lyeset_wf_analysis_date	String	this.lyeset_wf_analysis_date [-]
lyeset_workflow	lyeset_wf_version	String	this.lyeset_wf_version [-]

5.1.1.4.9 Натисніть «Run analysis». З'явиться спливаюче вікно «Confirm launch», в якому ви можете ввести необов'язковий опис. Натисніть «Launch».

No documentation provided

○ Run workflow with inputs defined by file paths

● Run workflow(s) with inputs defined by data table

Step 1 Select data table: CDC_ATCC_Sequences_set

Step 2 Select Data 1 selected CDC_ATCC_Sequences_set

Use call caching Delete intermediate outputs Use reference disks Retry with more memory Ignore empty outputs

SCRIPT ** INPUTS ** **OUTPUTS** ** **Run Analysis**

Hide optional inputs

Download json | Drag or click to upload json | Clear inputs | SEARCH INPUTS

Task name ↓	Variable	Type	Input value
-------------	----------	------	-------------

Confirm launch

Output files will be saved as workspace data in:
us US (multi-region) ⓘ

Running workflows will generate cloud charges. ⓘ
How much does my workflow cost? ⓘ
Set up budget alert ⓘ

Describe your submission (optional):
Salmonella colony picks

This will launch 1 analysis.

CANCEL LAUNCH

5.1.1.4.10 З'явиться екран «Workflow statuses», де ваші завдання спочатку будуть вказані як «Queued» або «Launching».

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 13 з 30

The screenshot shows the 'Job History' page in Terra Workspaces. It displays details for a specific submission: Submission 7bf21a76-0880-4317-b1e3-2924fd885e01. Key information includes: Workflow Statuses (Submitted: 1), Workflow Configuration (theiagen.pn/lyve.SET.PHB), Submitted by (eja.trees@theiagen.cloud), Total Run Cost (N/A), Data Entity (Salmonella_colony_picks), Submission ID (7bf21a76-0880-4317-b1e3-2924fd885e01), Call Caching (Enabled), Comment (Salmonella colony picks), Delete Intermediate Outputs (Disabled), Use Reference Disks (Disabled), and Retry with More Memory (Disabled). Below this, there is a 'WORKFLOWS' section with a search bar and a table listing workflow runs.

Data Entity	Last Changed	Status	Run Cost	Messages	Workflow ID	Links
Salmonella_colony_picks (CDC_ATCC_Sequences_set)	Mar 30, 2023, 3:07 PM	Queued	N/A			

5.1.1.4.11 Перейдіть на вкладку «Job history», щоб перевірити статус вашого завдання. Успішно виконане завдання позначається зеленою галочкою.

The screenshot shows the 'Job History' page with a table listing workflow runs. The table has columns for Submission (click for details), Data entity, No. of Workflows, Status, Submitted, Submission ID, Comment, and Actions.

Submission (click for details)	Data entity	No. of Workflows	Status	Submitted	Submission ID	Comment	Actions
lyve.SET.PHB Submitted by eja.trees@theiagen.cloud	Salmonella_colony_picks (...)	1	Done	Mar 30, 2023 3:33 PM	8d525d45-80d1-4810-8a8b-18ba755a16bd	Salmonella colony picks w...	ⓘ
lyve.SET.PHB_gg Submitted by eja.trees@theiagen.cloud This job succeeded	Salmonella_colony_picks (...)	1	Done	Mar 30, 2023 3:21 PM	23ad3e6a-fd66-43e6-9713-a2b5d35ae061	Salmonella colony picks w...	ⓘ
lyve.SET.PHB_J4 Submitted by eja.trees@theiagen.cloud	Vibrio_colony_picks (CDC_...)	1	Done	Mar 30, 2023 3:15 PM	b64700c-c193-4809-958b-e21a45218765	Vibrio colony picks	ⓘ
lyve.SET.PHB_Cle7PY-ro7M Submitted by eja.trees@theiagen.cloud	Salmonella_colony_picks (...)	1	Done	Mar 30, 2023 3:07 PM	7bf21a76-0880-4317-b1e3-2924fd885e01	Salmonella colony picks	ⓘ
lyve.SET.PHB_Q07ZMm3K3VM Submitted by eja.trees@theiagen.cloud	LM_colony_picks_wo_outli...	1	Done	Mar 30, 2023 12:18 PM	648085c4-f6a5-4b71-b4c2-2a45629984e8	LM colony picks without o...	ⓘ
lyve.SET.PHB_Q07ZMm3K3VM				Mar 29, 2023	09c0f112-82e7-		ⓘ

5.1.2 Запустіть робочий процес аналізу kSNP3

5.1.2.1 Створіть набір зразків для філогенетичного аналізу, дотримуючись інструкцій у кроці 5.1.1.1.

ПРИМІТКА: На відміну від аналізу *hqSNP*, для успішного аналізу *kSNP3* необхідна генетична різноманітність у наборі тестових послідовностей. Наприклад, включення в аналіз лише колоній одного штаму призведе до невдачі аналізу, якщо не буде включено генетично відмінну зовнішню групу.

Create a CDC_ATCC_Sequences set

Set name (required)

CANCEL **SAVE**

5.1.2.2 Налаштуйте параметри робочого процесу аналізу kSNP3:

5.1.2.2.1 На вкладці «Workflows» виберіть процес «kSNP3_PHB». Відкриється екран «kSNP3_PHB».

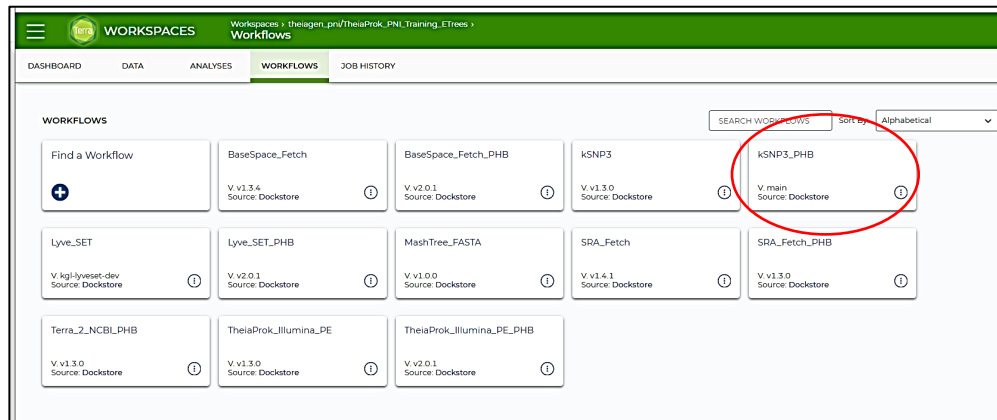
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

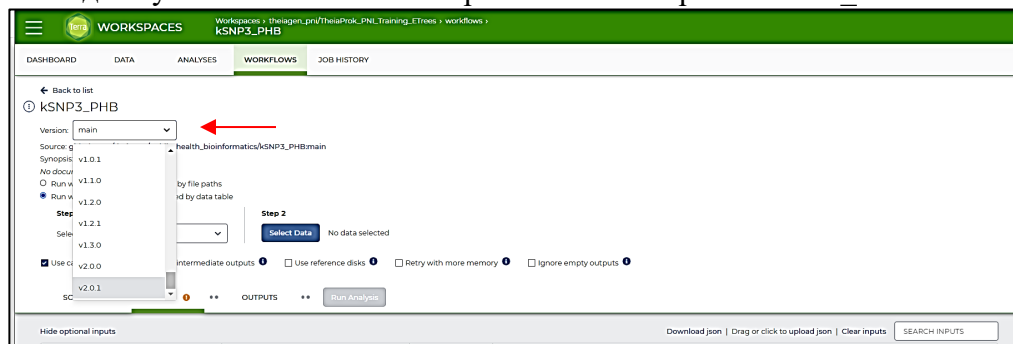
Версія № 01

Дата набрання чинності:

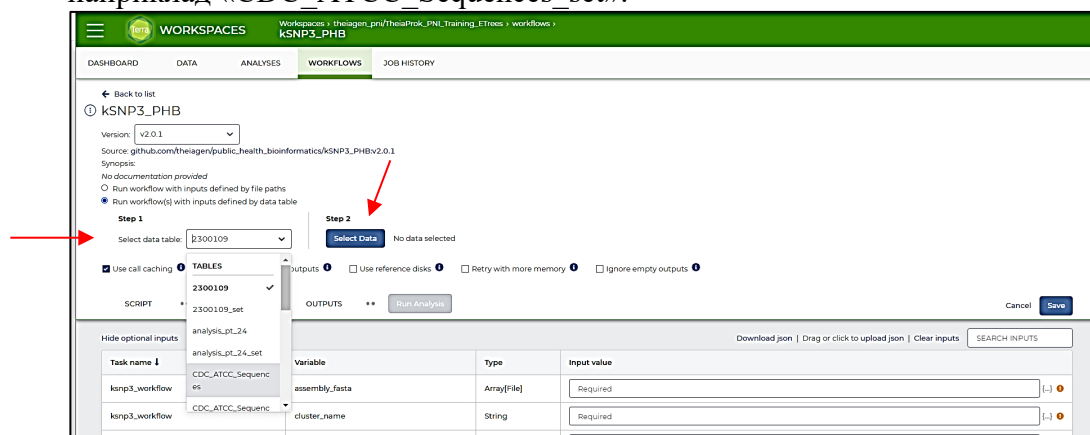
Сторінка 14 з 30



5.1.2.2.2 У спадному меню «Version» виберіть останню версію kSNP3_PHB.



5.1.2.2.3 У кроці 1 з випадного меню «Select root entity type» виберіть таблицю набору даних, де знаходиться набір зразків, створений у кроці 5.1.2.1, наприклад «CDC_ATCC_Sequences_set».



5.1.2.2.4 У кроці 2 натисніть «Select data» (скріншот вище). Це приведе вас до таблиці даних, вибраної на попередньому кроці.

5.1.2.2.5 Виберіть бажаний набір зразків, наприклад «Salmonella_colony_picks_with_outgroup», і натисніть «OK».

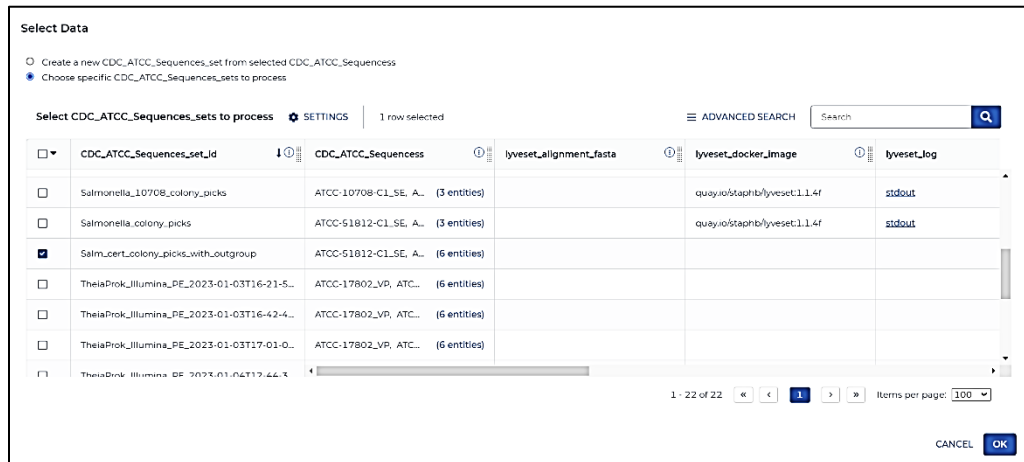
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

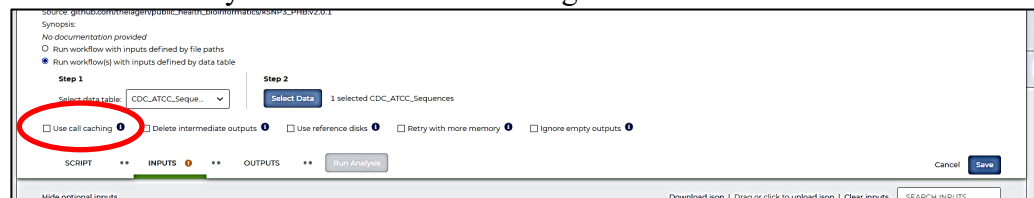
Версія № 01

Дата набрання чинності:

Сторінка 15 з 30

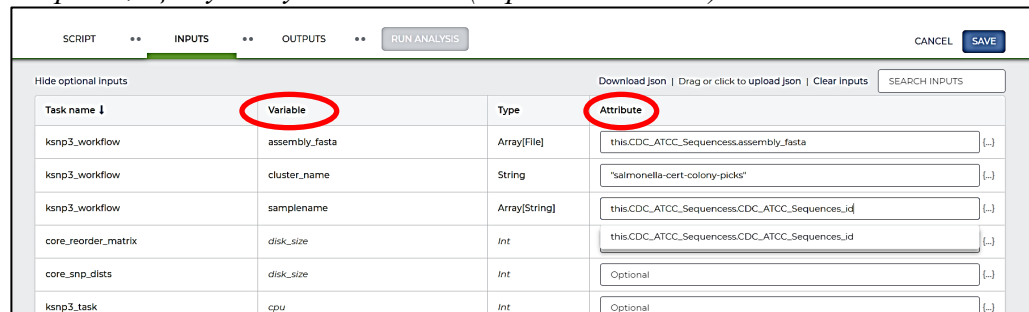


5.1.2.2.6 Зніміть позначку з поля «Use call catching».



5.1.2.2.7 На вкладці «Input data» необхідно заповнити наступні «Variables» в полях «Attribute»:

ПРИМІТКА: Коли ви заповнюєте стовпець «Attribute», натискання всередині комірки відкриє спадне меню атрибутів, які ви можете вибрати, щоб уникнути помилок (скріншот нижче).



5.1.2.2.7.1 assembly_fasta: «**This.data table names.assembly_fasta**», наприклад, «This.CDC_ATCC_Sequences.assembly_fasta».

ПРИМІТКА: додаткова літера «s» після імені таблиці даних є **ОБОВ'ЯЗКОВОЮ**.

5.1.2.2.7.2 cluster_name: будь-яке значення, взяте в лапки, наприклад, «salmonella-cert-colony-picks-outgroup».

5.1.2.2.7.3 samplename: «**This.data table names.data table name_id**», наприклад, «This.CDC_ATCC_Sequences.CDC_ATCC_Sequences_id».

ПРИМІТКА: подвійний формат імен є **ОБОВ'ЯЗКОВИМ**.

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 16 з 30

Variable	Type	Attribute
assembly_fasta	Array[File]	this.CDC_ATCC_Sequencess.assembly_fasta
cluster_name	String	"salmonella-cert-colony-picks"
samplename	Array[String]	this.CDC_ATCC_Sequencess.CDC_ATCC_Sequences_id
disk_size	Int	this.CDC_ATCC_Sequencess.CDC_ATCC_Sequences_id
disk_size	Int	Optional

5.1.2.2.8 На вкладці «Outputs» натисніть «Use defaults» для «Attributes», а потім натисніть «Save».

ПРИМІТКА: кнопка «Save» відображається лише в тому випадку, якщо ви змінили вхідні дані порівняно з попереднім поданням.

SCRIPT ** INPUTS ** **OUTPUTS** ** RUN ANALYSIS

Output files will be saved to
Files / submission unique ID / ksnp3_workflow / workflow unique ID

References to outputs will be written to
Tables / CDC_ATCC_Sequencess_set

Fill in the attributes below to add or update columns in your data table

Download json | Drag or click to upload json | Clear outputs | SEARCH OUTPUTS

Task name ↓	Variable	Type	Attribute Use defaults
ksnp3_workflow	ksnp3_core_snp_matrix	File	this.ksnp3_core_snp_matrix (-)
ksnp3_workflow	ksnp3_core_tree	File	this.ksnp3_core_tree (-)
ksnp3_workflow	ksnp3_core_vcf	File	this.ksnp3_core_vcf (-)
ksnp3_workflow	ksnp3_docker	String	this.ksnp3_docker (-)

CANCEL **SAVE**

5.1.2.2.9 Натисніть «Run analysis». З'явиться спливаюче вікно «Confirm launch», в якому ви можете ввести необов'язковий опис. Натисніть «Launch».

SCRIPT ** INPUTS ** **OUTPUTS** ** **RUN ANALYSIS**

Output files will be saved to
Files / submission unique ID / ksnp3_workflow / workflow unique ID

References to outputs will be written to
Tables / CDC_ATCC_Sequencess_set

Fill in the attributes below to add or update columns in your data table

Download json | Drag or click to upload json | Clear outputs | SEA

Task name ↓	Variable	Type	Attribute Use defaults
ksnp3_workflow	ksnp3_core_snp_matrix	File	this.ksnp3_core_snp_matrix

Confirm launch

Output files will be saved as workspace data in:
us-central1 (lowa) ⓘ

Running workflows will generate cloud charges. ⓘ
How much does my workflow cost? ⓘ
Set up budget alert ⓘ

Describe your submission (optional):

ksnp analysis of Salmonella cert strain colony picks

This will launch 1 analysis.

CANCEL **LAUNCH**

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

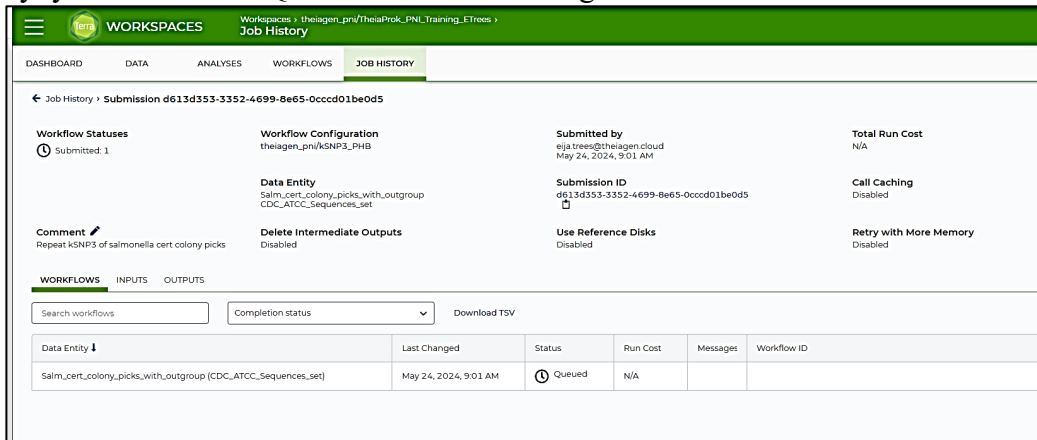
Док. № PNID02

Версія № 01

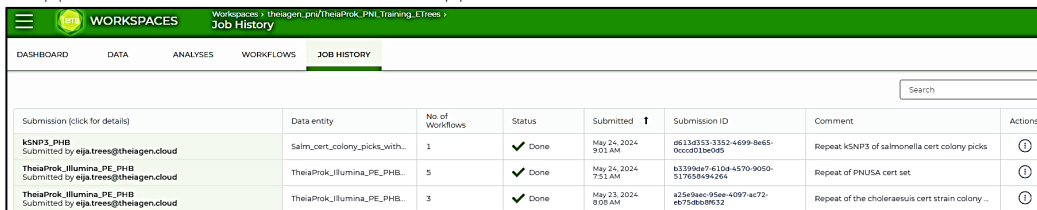
Дата набрання чинності:

Сторінка 17 з 30

5.1.2.2.10 З'явиться екран «Workflow statuses», де ваші подані завдання спочатку будуть вказані як «Queued» або «Launching».



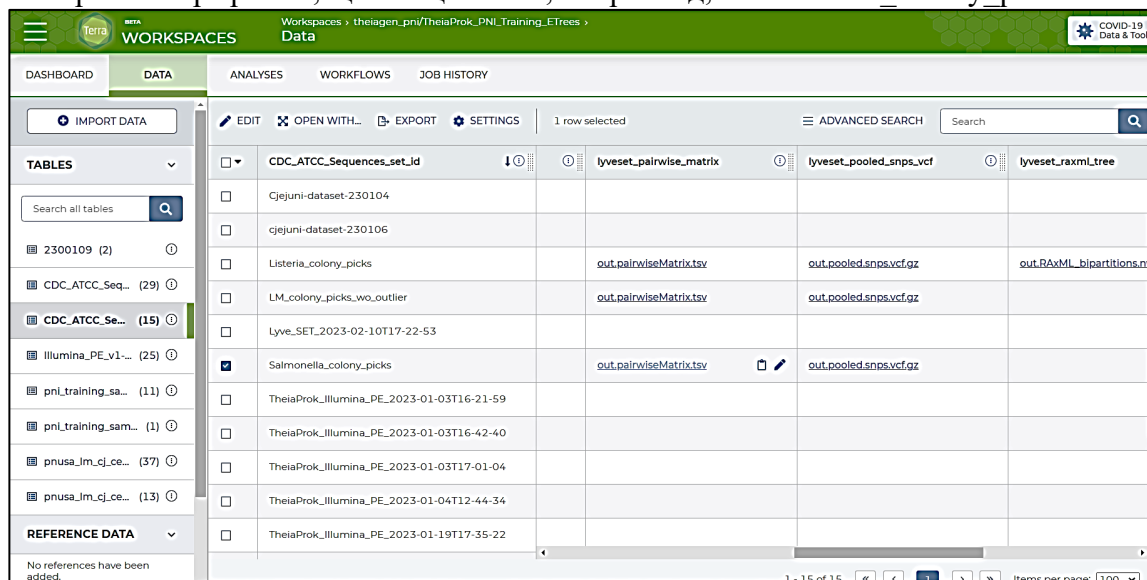
5.1.2.2.11 Перейдіть на вкладку «Job history», щоб перевірити статус вашого завдання. Успішно виконане завдання позначається зеленою галочкою.



5.2 Перегляньте результати філогенетичного аналізу. Ви можете переглянути та експортувати матриці SNP і дерева/дендрограми. Кроки для перегляду результатів **hqSNP** (5.2.3-5.2.4) і **kSNP3** (5.2.5-5.2.6) дещо відрізняються.

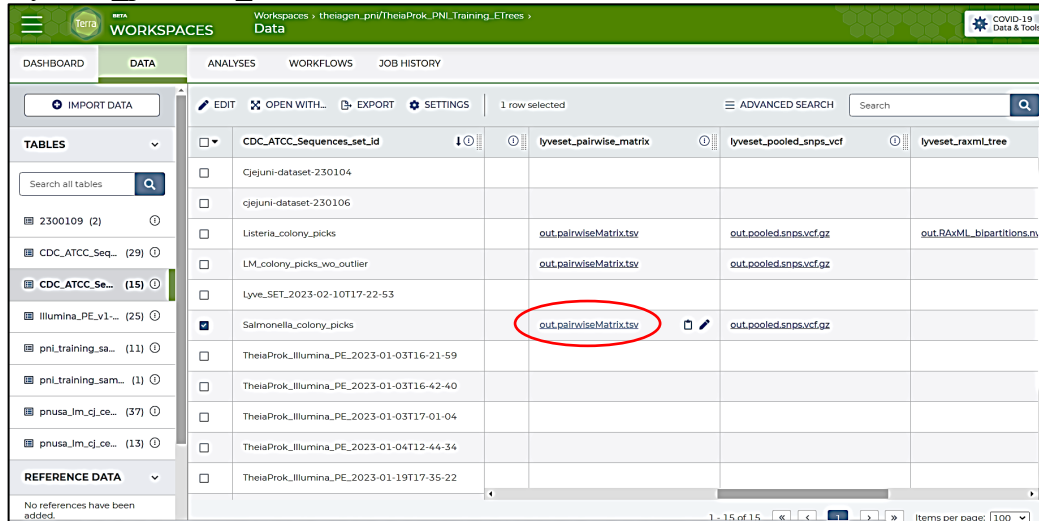
5.2.1 Виберіть таблицю даних, що вас цікавить, «CDC_ATCC_Sequences_set».

5.2.2 Виберіть набір зразків, що вас цікавить, наприклад, «Salmonella colony picks».

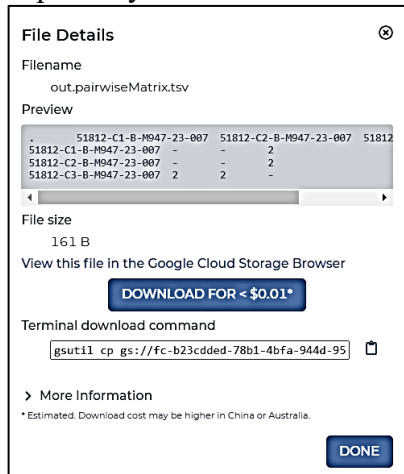


5.2.3 Перегляньте та завантажте матрицю **hqSNP**

5.2.3.1 Клацніть на посилання tsv-файлу для матриці в колонці «lyveset_pairwise_matrix».



5.2.3.2 З'явиться спливаюче вікно «File details». У цьому вікні матрицю можна переглянути без завантаження.



5.2.3.3 Щоб завантажити матрицю, натисніть «Download for <\$0,01*» і потім натисніть «Done» (скріншот вище).

5.2.3.4 Завантажений файл tsv з'явиться у верхньому правому куті екрана.

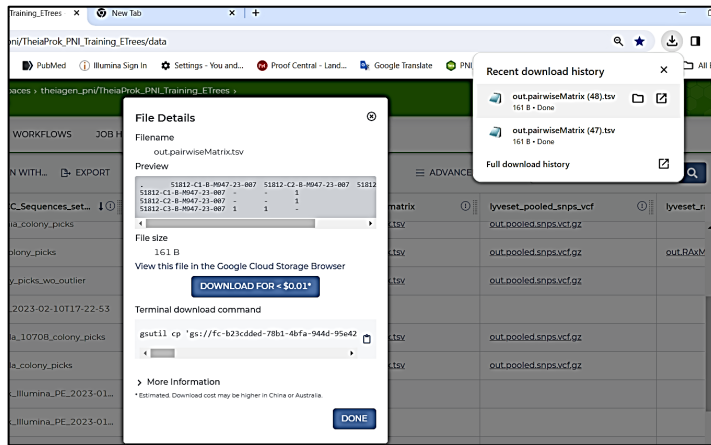
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

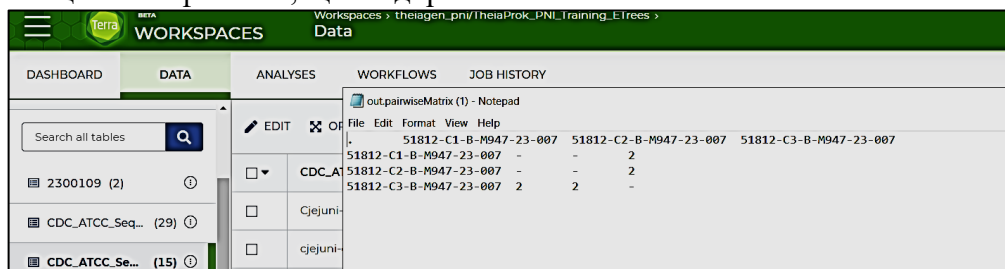
Версія № 01

Дата набрання чинності:

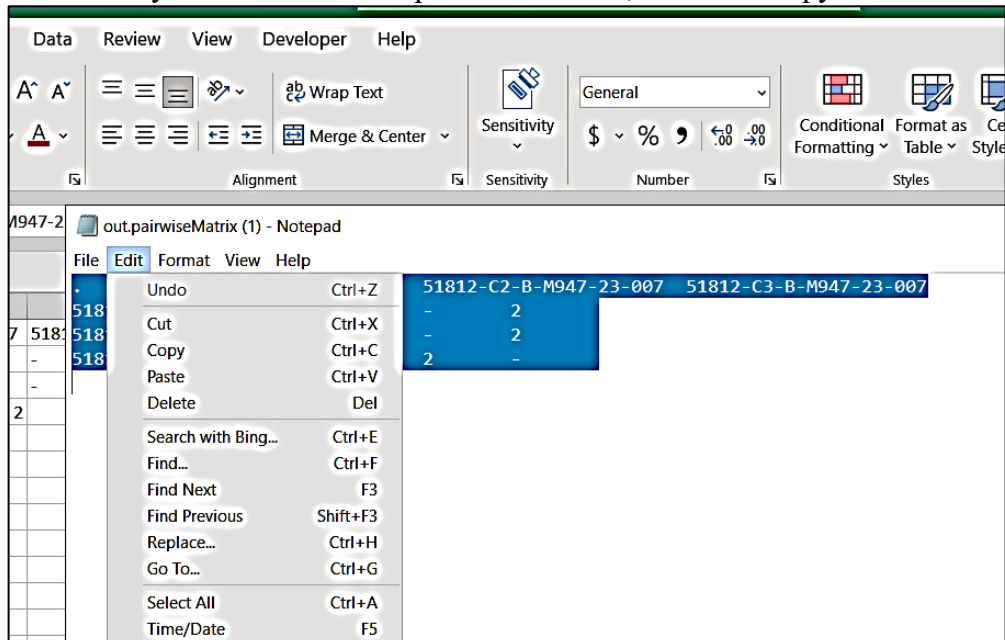
Сторінка 19 з 30



5.2.3.5 Клацніть на файл tsv, щоб відкрити його.



5.2.3.6 У спадному меню «Edit» виберіть «Select all», а потім «Copy».



5.2.3.7 Вставте вміст файлу tsv в аркуш Excel.

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 20 з 30

	A	B	C	D	E
1	.	51812-C1-B-M947-23-007	51812-C2-B-M947-23-007	51812-C3-B-M947-23-007	
2	51812-C1-B-M947-23-007	-	-		2
3	51812-C2-B-M947-23-007	-	-		2
4	51812-C3-B-M947-23-007	2	2	-	
5					
6					

5.2.4 Завантажте та перегляньте дерево **hqSNP** (дендрограму)

ПРИМІТКА: *Lyve-SET* генерує дерево лише за умови достатньої різноманітності (SNP) у наборі зразків. Набір зразків, що складається з технічних реплікатів одного і того ж штаму, не має достатньої різноманітності для генерації дерева і дає лише матрицю SNP. Щоб згенерувати дерево, додайте до набору зразків (непов'язаний) штаму, що має випадające значення.

5.2.4.1 Клацніть на посилання файлу **nwk** для дерева в стовпці «**lyveset_raxml_tree**».

Table	File	File	File
CDC_ATCC_Sequences_setLid	lyveset_pairwise_matrix	lyveset_pooled_snps.vcf	lyveset_raxml_tree
Cjjeuni-dataset-230104			
Cjjeuni-dataset-230106			
Listeria_colony_picks	out.pairwiseMatrix.tsv	out.pooled.snps.vcf.gz	out.RAxML_bipartitions.nwk
LM_colony_picks_wo_outlier	out.pairwiseMatrix.tsv	out.pooled.snps.vcf.gz	
Lyve_SET_2023-02-10T17-22-53			
Salmonella_colony_picks	out.pairwiseMatrix.tsv	out.pooled.snps.vcf.gz	
TheiaProk_Illumina_PE_2023-01-03T16-21-59			
TheiaProk_Illumina_PE_2023-01-03T16-42-40			
TheiaProk_Illumina_PE_2023-01-03T17-01-04			
TheiaProk_Illumina_PE_2023-01-04T12-44-34			
TheiaProk_Illumina_PE_2023-01-19T17-35-22			

5.2.4.2 З'явиться спливаюче вікно «File details».

5.2.4.3 Натисніть «Download for <\$0,01*>» і потім натисніть «Done».

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

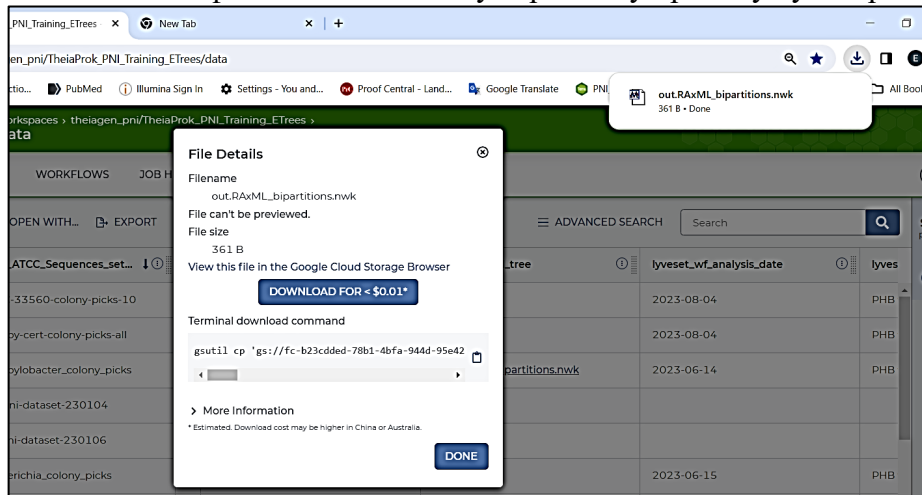
Версія № 01

Дата набрання чинності:

Сторінка 21 з 30

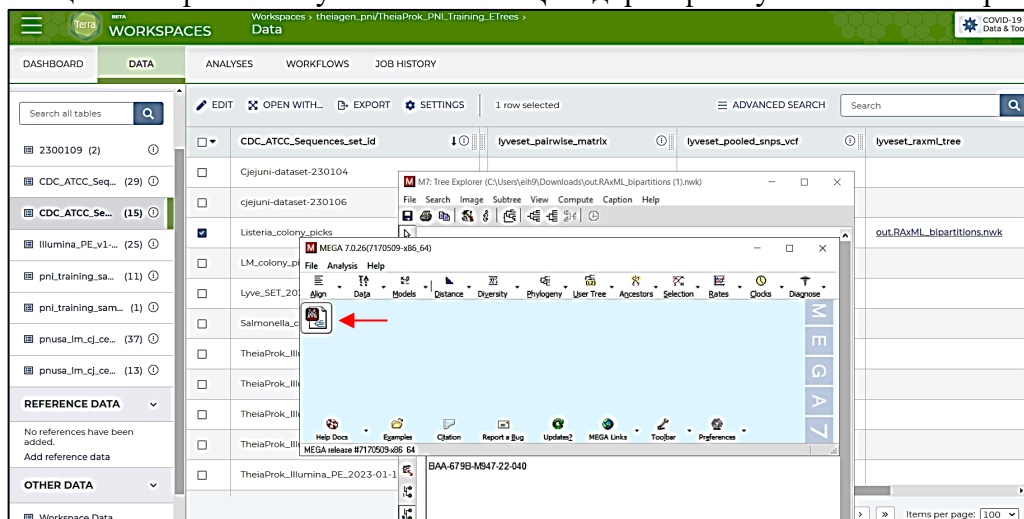


5.2.4.4 Завантажений файл нwk з'явиться у верхньому правому куті екрана.



5.2.4.5 Клацніть на файл нwk. Відкриється програма MEGA (за умови, що MEGA встановлена на вашому комп'ютері).

5.2.4.6 Клацніть на файл нwk у вікні MEGA. Це відкриє файл у MEGA Tree Explorer.



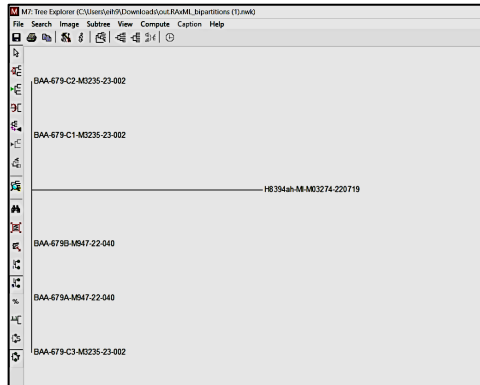
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

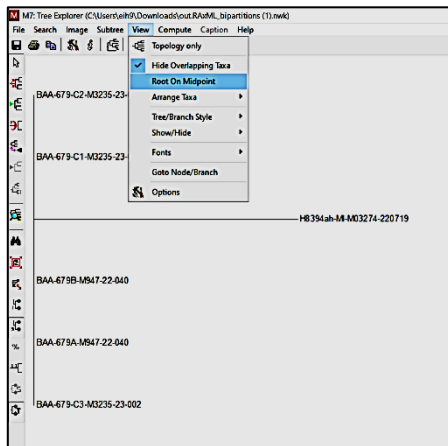
Версія № 01

Дата набрання чинності:

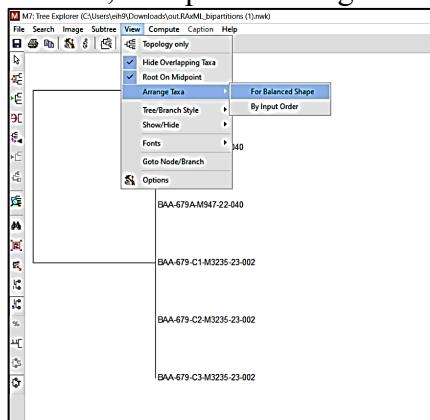
Сторінка 22 з 30



5.2.4.7 Щоб створити корінь дерева, виберіть «Root on midpoint» у спадному меню «View».



5.2.4.8 Для отримання більш збалансованого дерева поверніться до випадного меню «View», виберіть «Arrange taxa» і «For balanced shape».



5.2.4.9 Щоб скопіювати дерево в іншу програму, перейдіть до меню «Image» і виберіть «Copy to clipboard».

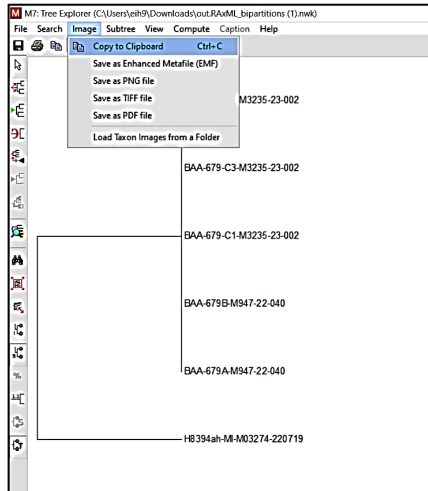
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

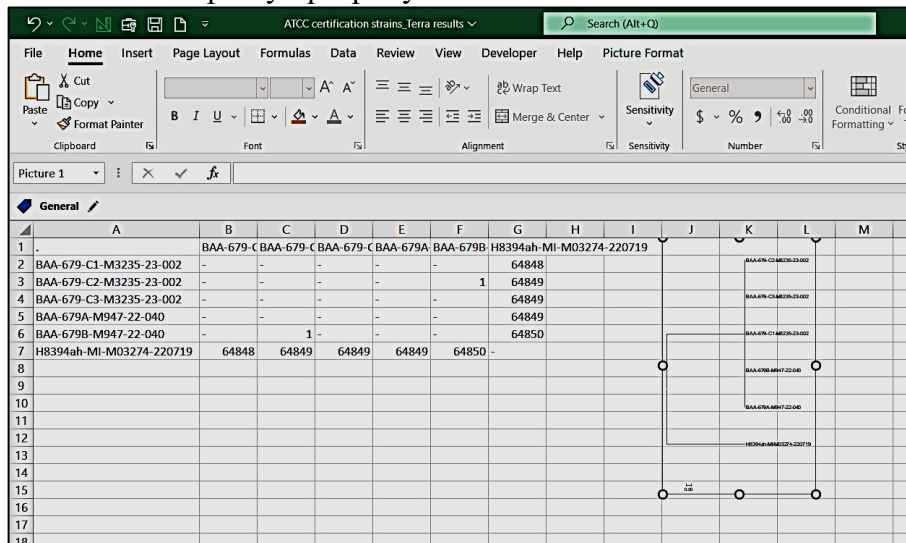
Версія № 01

Дата набрання чинності:

Сторінка 23 з 30



5.2.4.10 Вставте в потрібну програму.



5.2.5 Перегляньте та завантажте матрицю kSNP3 core SNP

5.2.5.1 Клацніть на посилання csv-файлу для матриці в стовпці «kSNP3_core_snp_matrix».

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 24 з 30

TABLES	NAME	ENTITY COUNT	FILE NAME
	CDC_ATCC_Sequences_set...		ksnp3_core_snp_matrix
	Escherichia_colony_picks	(4 entities)	
	Listeria_colony_picks	(6 entities)	
	LM_colony_picks_wo_outlier	(5 entities)	
	Lyve_SET_2023-02-10T17-22-53	(5 entities)	
	Salmonella_10708_colony_picks	(3 entities)	
	Salmonella_colony_picks	(3 entities)	
	Salm_cert_colony_picks_with_out...	(6 entities)	salmonella-cert-colony-picks-outgroup_core_snp_matrix.csv
	TheiaProk_Illumina_PE_2023-01...	(6 entities)	
	TheiaProk_Illumina_PE_2023-01...	(6 entities)	
	TheiaProk_Illumina_PE_2023-01...	(6 entities)	

5.2.5.2 З'явиться спливаюче вікно «File details». У цьому вікні матрицю можна переглянути без завантаження.

File Details

Filename
salmonella-cert-colony-picks-outgroup_core_snp_matrix.csv

Preview

```
snp-dists 0.8.2, ATCC-51812-C2_SE:c1, ATCC-51812-C3_SE:c1, ATCC-51812-C2_SE:0.2, 1, 29184, 29184, 29184  
ATCC-51812-C3_SE:2, 0, 3, 29184, 29184, 29184  
ATCC-51812-C1_SE:1, 3, 0, 29183, 29183, 29183  
ATCC-10708-C3_SE:29184, 29184, 29183, 0, 0, 0  
ATCC-10708-C2_SE:29184, 29184, 29183, 0, 0, 0  
ATCC-10708-C1_SE:29184, 29184, 29183, 0, 0, 0
```

File size
382 B

View this file in the Google Cloud Storage Browser

DOWNLOAD FOR <\$0,01*

Terminal download command

```
gsutil cp 'gs://fc-b23cdded-78b1-4bfa-944d-95e42'
```

> More Information
* Estimated. Download cost may be higher in China or Australia.

DONE

5.2.5.3 Щоб завантажити матрицю, натисніть «Download for <\$0,01*» і потім натисніть «Done» (скріншот вище).

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

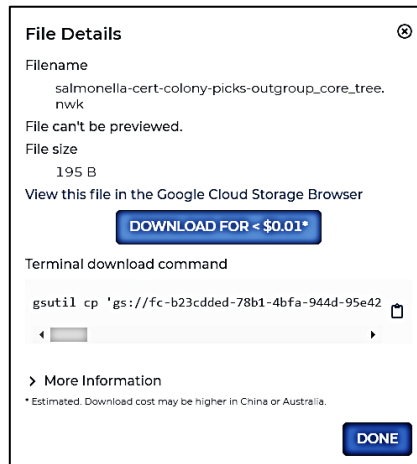
Док. № PNID02

Версія № 01

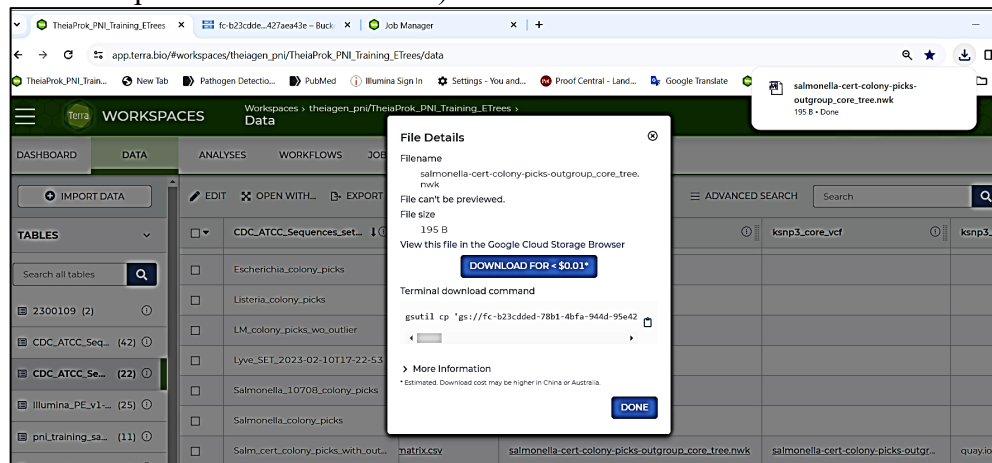
Дата набрання чинності:

Сторінка 26 з 30

5.2.6.3 Натисніть «Download for <\$0,01*» і потім натисніть «Done».



5.2.6.4 Завантажений файл nwk з'явиться у верхньому правому куті екрана. Клацніть на файл, щоб відкрити його в MEGA Tree Explorer (за умови, що на вашому комп'ютері встановлено MEGA).



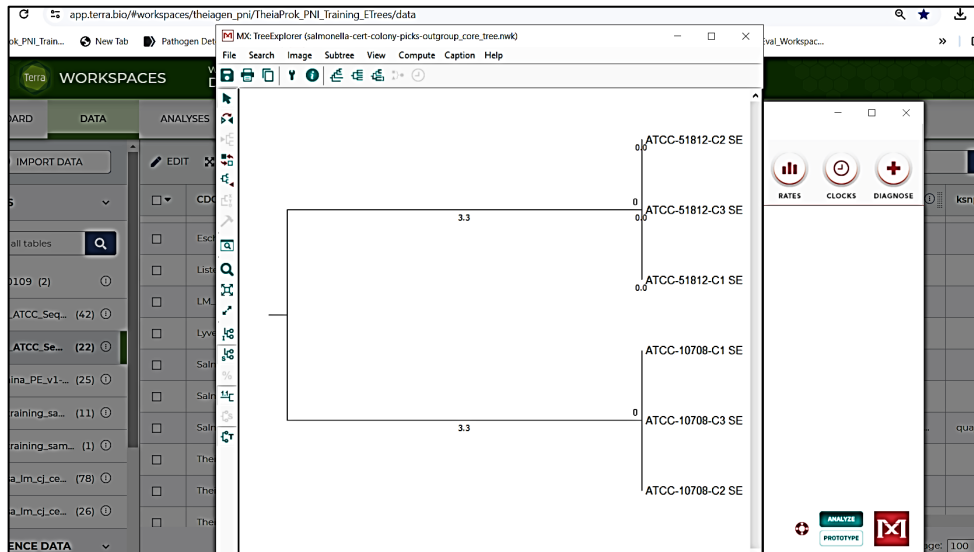
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

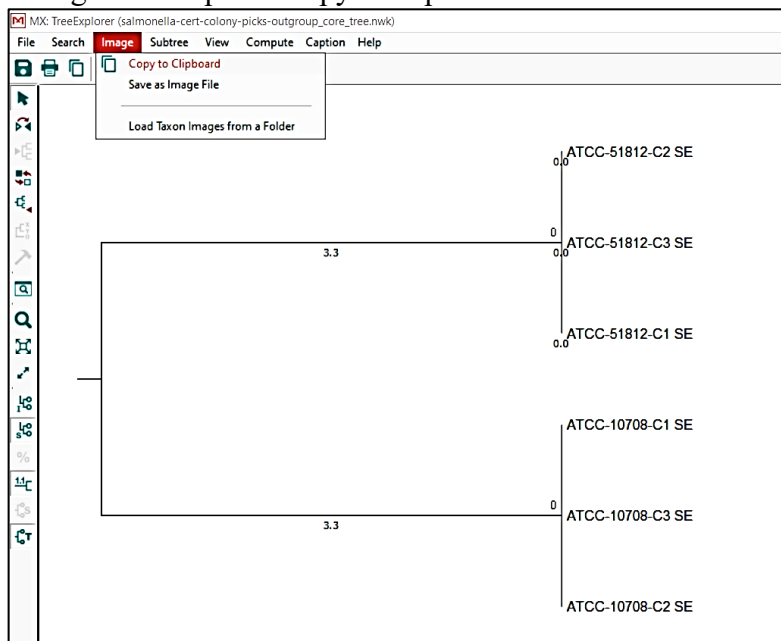
Версія № 01

Дата набрання чинності:

Сторінка 27 з 30



5.2.6.5 Щоб скопіювати дерево в іншу програму, перейдіть до випадного меню «Image» і виберіть «Copy to clipboard».



5.2.6.6 Вставте в потрібну програму.

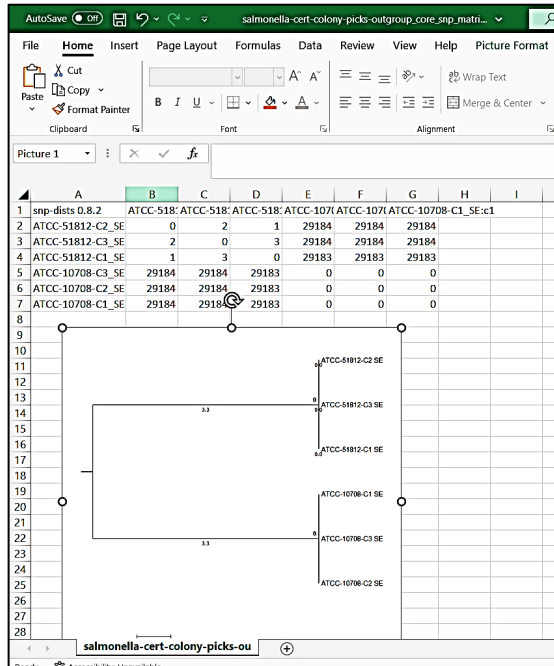
МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 28 з 30



6. БЛОК-СХЕМА:



7. ПОВ'ЯЗАНІ ДОКУМЕНТИ:

- 7.1 **PNID01:** Міжнародна стандартна операційна процедура PulseNet для аналізу даних Illumina Short Read WGS за допомогою платформи Terra.Bio.

8. ВИКОРИСТАНІ ДЖЕРЕЛА:

- 8.1 Katz L.S., Griswold T., Williams-Newkirk A.J., Wager, D., Petkau, A., Sieffert, C., Van Domselaar, G., Deng, X., Carleton, H.A. (2017). A Comparative Analysis of the Lyve-SET Phylogenomics Pipeline for Genomic Epidemiology of Foodborne Pathogens. *Frontiers Microbiol.* 8:375.
- 8.2 Libuit K.G., Doughty E.L., Otieno J.R., Ambrosio F., Kapsak C.J., Smith E.A., Wright S.M., Scribner M.R., Petit III R.A., Mendes C.I., Huergo M., Legacki G., Loreth C., Park D.J., Sevinsky J.R. (2023) Accelerating bioinformatics implementation in public health. *Microbial Genomics* 9:001051.

9. КОНТАКТИ:

- 9.1 Лабораторія Центру CDC USA PulseNet NGS Laboratory: pulsenetngslab@cdc.gov
- 9.2 Координатор з питань забезпечення якості PulseNet International Ейя Тріс: ehyytia-trees@cdc.gov
- 9.3 Theiagen:
- 9.3.1 Загальна електронна адреса для підтримки: support@theiagen.com
- 9.3.2 Мішель Скрібнер: michelle.scribner@theiagen.com
- 9.3.3 Френк Амброзіо: frank.ambrosio@theiagen.com

10. ЗМІНИ: Немає

МІЖНАРОДНА СТАНДАРТНА ОПЕРАЦІЙНА ПРОЦЕДУРА PULSENET ДЛЯ ФІЛОГЕНЕТИЧНОГО АНАЛІЗУ ДАНИХ ПОВНОГЕНОМНОГО СЕКВЕНУВАННЯ ІЗ ВИКОРИСТАННЯМ ПЛАТФОРМИ TERRA.BIO

Док. № PNID02

Версія № 01

Дата набрання чинності:

Сторінка 30 з 30

11. ПІДПИСИ ДЛЯ ЗАТВЕРДЖЕННЯ:

Затверджено: _____ Дата: _____
Персонал PulseNet з контролю якості та забезпечення якості

Затверджено: _____ Дата: _____
Технічний керівник PulseNet WGS

Затверджено: _____ Дата: _____
Міжнародний координатор PulseNet

Затверджено: _____ Дата: _____
Керівник групи реагування та управління спалахами PulseNet

Затверджено: _____ Дата: _____
Керівник лабораторії кишкових захворювань