Jun 07, 2023

Clinical metagenomic sequencing - CSF RNA and DNA Illumina MiSeq

DOI
https://dx.doi.org/10.17504/protocols.io.j8nlkk7wdl5r/v1

Gert Marais[1], Ziyaad Valley-Omar[1], Kathleen Bateman[2], Kate McMullen[2], Dieter van der Westhuizen[1], Moepeng Maseko[1], Suzaan Marais[2], Diana Hardie[1], Adrian Brink[1]

[1]University of Cape Town, National Health Laboratory Service; [2]University of Cape Town

Gert Marais: Medical Microbiology;
Ziyaad Valley-Omar: Medical Virology
Kathleen Bateman: Neurology
Kate McMullen: Neurology
Dieter van der Westhuizen: Chemical Pathology
Moepeng Maseko: Medical Virology
Suzaan Marais: Neurology
Diana Hardie: Medical Virology
Adrian Brink: Medical Microbiology

Ruan Marais

OPEN ACCESS

DOI: https://dx.doi.org/10.17504/protocols.io.j8nlkk7wdl5r/v1

Protocol Citation: Gert Marais, Ziyaad Valley-Omar, Kathleen Bateman, Kate McMullen, Dieter van der Westhuizen, Moepeng Maseko, Suzaan Marais, Diana Hardie, Adrian Brink 2023. Clinical metagenomic sequencing - CSF RNA and DNA Illumina MiSeq. **protocols.io** https://dx.doi.org/10.17504/protocols.io.j8nlkk7wdl5r/v1

Protocol status: Working
We use this protocol and it's working

Created: August 15, 2022

Last Modified: June 07, 2023

Protocol Integer ID: 68673

Keywords: Clinical metagenomics, CSF, central nervous system infections, unbiased cerebrospinal fluid metagenomic, csf rna, commercial kit protocol

Funders Acknowledgements:
Oppenheimer Generations Research and Conservation
Grant ID: N/A
Poliomyelitis Research Foundation
Grant ID: 21/68

Disclaimer

DISCLAIMER – FOR INFORMATIONAL PURPOSES ONLY; USE AT YOUR OWN RISK

The protocol content here is for informational purposes only and does not constitute legal, medical, clinical, or safety advice, or otherwise; content added to **protocols.io** is not peer reviewed and may not have undergone a formal approval of any kind. Information presented in this protocol should not substitute for independent professional judgment, advice, diagnosis, or treatment. Any action you take or refrain from taking using or relying upon the information presented here is strictly at your own risk. You agree that neither the Company nor any of the authors, contributors, administrators, or anyone else associated with **protocols.io**, can be held responsible for your use of the information contained in or linked to this protocol or any of our Sites/Apps and Services.

Abstract

This protocol was developed to evaluate unbiased cerebrospinal fluid metagenomics at a referral centre in South Africa. Commercial kit protocols are generally used unmodified and limited use of in-house developed methods or reagents is made. This maintains the simplicity and accessibility of the protocol which can serve as a base for evaluation of process improvements.

Guidelines

Please note: This protocol is intended for research use only.

1) This protocol functions optimally with low cell-count CSF specimens with minimal human DNA. High cell-count specimens should ideally be evaluated using sequencing protocols that include general microbiome enrichment or targeted enrichment and/or human DNA depletion.

2) The number of individuals performing the clinical metagenomic sequencing (CMS) protocol and the number of days over which it is performed determines the order of steps.

3) Sodium hypochlorite is the preferred cleaning solution

4) The preferred run size is 10 samples and 2 controls though the cellularity of used CSF may alter the preferred run size due to increased quantity of human DNA. Master mix volumes are calculated and presented here for this run size.

5) The total hands-on time (including incubation and thermocycler time) for the protocol is 14 hours and 11 minutes. This assumes a single operator and 10 samples with 2 controls. The time estimate excludes sequencing time (generally 48-56 hours with the MiSeq v3 600-cycle kits) and bioinformatics time. It is generally advisable to perform the protocol over 2 days.

6) Prior to use samples should be stored at -70 °C and defrosted immediately prior to use.

Materials

Reagents

- NUCLISENS easyMAG extraction reagents **bioMérieux**
- RNA Clean & Concentrator™-5 **Zymo Research** **Catalog #**R1015
- SuperScript™ IV First-Strand Synthesis System **Thermo Fisher Scientific** **Catalog #**18091050
- Second strand synthesis kit **Invitrogen - Thermo Fisher** **Catalog #**A48571
- iTaq Universal SYBR® Green One-step kit **Bio-Rad Laboratories** **Catalog #**1725150
- Illumina DNA Prep (M) Tagmentation (24 Samples) **Illumina, Inc.** **Catalog #**20018704
- Illumina Nextera DNA CD indexes (24 indexes) **Illumina, Inc.**
- MiSeq v3 Sequencing Reagents (600 cycles) **Illumina, Inc.** **Catalog #**MS-102-3003
- PhiX Control v3 **Illumina, Inc.** **Catalog #**FC-110-3001
- Qubit 1X dsDNA HS Assay Kit **Thermo Fisher Scientific** **Catalog #**Q33230
- Agencourt AMPure XP **Beckman Coulter** **Catalog #**A63880
- Viral Multiplex Reference 11/242 NIBSC **National Institute of Biological Standards and Control (NIBSC)** **Catalog #**11/242

Optional (- High Sensitivity D5000 ScreenTape **Agilent Technologies** **Catalog #**5067-5592)

Optional (- High Sensitivity D5000 Reagents **Agilent Technologies** **Catalog #**5067-5593)

Custom reagents

Consumables

- Ethanol (100%, Molecular Biology Grade) **Fisher Scientific** **Catalog #**BP2818500
- Sodium Hydroxide 1M solution
- Falcon tube (50 mL)
- Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf** **Catalog #**0030108051
- 96 well LoBind PCR plates Semi-skirted **Eppendorf** **Catalog #**0030129504
- Nuclease-free Water
- Assorted Pipette tips (filter) 10ul 100ul 200ul 1000ul
- Sodium Hypochlorite Solution

Equipment

Equipment

NUCLISENS easyMag — NAME

Nucleic acid extraction — TYPE

bioMerieux — BRAND

N/A — SKU

https://www.biomerieux-usa.com/clinical/nuclisens-easymag — LINK

Equipment

MiSeq — NAME

Sequencer — TYPE

illumina — BRAND

SY-410-1003 — SKU

https://www.illumina.com/systems/sequencing-platforms/miseq/order-miseq.html — LINK

Equipment

Set of micropipettes with rack: 100-1000 µl, 20-200 µl, 2-20 µl, and 0.5-10 µl NAME

Pipettor set TYPE

Pipetman BRAND

QP-1001-07 SKU

https://www.minipcr.com/product/set-of-four-adjustable-volume-micropipettes-rack/ LINK

Can use equivalent Pipettors SPECIFICATIONS

Equipment

CFX96 Touch Real-Time PCR NAME

qPCR TYPE

Bio-Rad BRAND

#1855195 SKU

https://www.bio-rad.com/en-no/sku/1855195-cfx96-touch-real-time-pcr-detection-system?ID=1855195 LINK

Equipment

Magnetic Stand-96

NAME

ThermoFisher Scientific

BRAND

AM10027

SKU

https://www.thermofisher.com/order/catalog/product/AM10027

LINK

Equipment

Qubit Fluorometer

NAME

Fluorometer

TYPE

Invitrogen

BRAND

Q33238

SKU

https://www.thermofisher.com/order/catalog/product/Q33238#/Q33238

LINK

Equipment

Centrifuge	NAME
Benchtop Centrifuge	TYPE
Eppendorf	BRAND
5405000441	SKU
https://online-shop.eppendorf.us/US-en/Centrifugation-44533/Centrifuges-44534/Centrifuge-5425-PF-243560.html	LINK
Any benchtop centrifuge will suffice	SPECIFICATIONS

Equipment

SimpliAmp Thermal Cycler	NAME
PCR	TYPE
Applied Biosystems	BRAND
A24811	SKU
https://www.thermofisher.com/order/catalog/product/A24811	LINK
Any standard PCR thermocycler will suffice	SPECIFICATIONS

Equipment

Plate centrifuge	NAME
Centrifuge	TYPE
Any	BRAND
N/A	SKU

Protocol materials

- Sodium Hydroxide 1M solution
- PhiX Control v3 **Illumina, Inc.** **Catalog #**FC-110-3001
- MiSeq v3 Sequencing Reagents (600 cycles) **Illumina, Inc.** **Catalog #**MS-102-3003
- NUCLISENS easyMAG extraction reagents **bioMérieux**
- Sodium Hypochlorite Solution
- High Sensitivity D5000 ScreenTape **Agilent Technologies** **Catalog #**5067-5592
- Second strand synthesis kit **Invitrogen - Thermo Fisher** **Catalog #**A48571
- Illumina DNA Prep (M) Tagmentation (24 Samples) **Illumina, Inc.** **Catalog #**20018704
- Falcon tube (50 mL)
- Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf** **Catalog #**0030108051
- Viral Multiplex Reference 11/242 NIBSC **National Institute of Biological Standards and Control (NIBSC)** **Catalog #**11/242
- 96 well LoBind PCR plates Semi-skirted **Eppendorf** **Catalog #**0030129504
- SuperScript™ IV First-Strand Synthesis System **Thermo Fisher Scientific** **Catalog #**18091050
- Qubit 1X dsDNA HS Assay Kit **Thermo Fisher Scientific** **Catalog #**Q33230
- High Sensitivity D5000 Reagents **Agilent Technologies** **Catalog #**5067-5593
- Ethanol (100%, Molecular Biology Grade) **Fisher Scientific** **Catalog #**BP2818500
- Nuclease-free Water
- iTaq Universal SYBR® Green One-step kit **Bio-Rad Laboratories** **Catalog #**1725150
- Assorted Pipette tips (filter) 10ul 100ul 200ul 1000ul
- Illumina Nextera DNA CD indexes (24 indexes) **Illumina, Inc.**
- RNA Clean & Concentrator™-5 **Zymo Research** **Catalog #**R1015
- Agencourt AMPure XP **Beckman Coulter** **Catalog #**A63880
- Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf** **Catalog #**0030108051

Nucleic Acid Purification

1h 35m

1

1h 35m

Safety information

Clinical samples should be handled with appropriate biosafety and infection prevention and control precautions.

Perform total nucleic acid extraction using the NUCLISENS easyMAG platform or another similar total nucleic acid extraction platform.

- The preferred run size is 10 samples and 2 controls
- The preferred positive control is the NIBSC Viral Multiplex Reference 11/242
- The preferred negative control is the extraction platform lysis buffer
- The internal control used is laboratory specific and the volume used is dependent on its specific characteristics and lab optimisation prior to use
- Enter the sample details in the table below

	Sample identifier	Sample input volume (ul)
	Positive control	
	Negative control	

Table: Sample details and input volume

The following consumables are required for this step:

Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf Catalog #**0030108051

Viral Multiplex Reference 11/242 NIBSC **National Institute of Biological Standards and Control (NIBSC) Catalog #**11/242

The following equipment is required for this step:

Equipment

NUCLISENS easyMag	NAME
Nucleic acid extraction	TYPE
bioMerieux	BRAND
N/A	SKU
https://www.biomerieux-usa.com/clinical/nuclisens-easymag	LINK

Equipment

Centrifuge	NAME
Benchtop Centrifuge	TYPE
Eppendorf	BRAND
5405000441	SKU
https://online-shop.eppendorf.us/US-en/Centrifugation-44533/Centrifuges-44534/Centrifuge-5425-PF-243560.html	LINK
Any benchtop centrifuge will suffice	SPECIFICATIONS

0.2 Remove CSF samples from -70 °C storage and defrost On ice . Label the sample elution tubes (1 for RNA and 1 for total nucleic acids for each sample) and easyMAG lysis buffer tubes while the samples defrost. 15m

0.3 Briefly centrifuge the samples. 5m

280 x g, 4°C, 00:00:30

0.4 Add 200-1000 µL of sample to a easyMAG lysis buffer tube. Similarly add 200-1000 µL of the positive control to a lysis buffer tube. CAUTION: Avoid sample cross-contamination. 5m

Viral Multiplex Reference 11/242 NIBSC **National Institute of Biological Standards and Control (NIBSC) Catalog #**11/242

0.5 Add 10 µL of internal control to each sample in lysis buffer (including the positive control and negative control lysis buffer tube which contains no sample). CAUTION: Avoid sample cross-contamination. 5m

0.6 Incubate at Room temperature for 00:10:00 . Program the sample order on the easyMAG instrument and load the required disposable plastics. Set the elution volume to 35 µL . 10m

0.7 Add 50 µL of easyMAG magnetic silica to each sample and internal control containing lysis buffer tube. 5m

0.8 Add the contents of the lysis buffer tube to the easyMAG processing wells and start the instrument. 40m

0.9 Store 20 µL of extracted total nucleic acids in a labelled tube and store at -20 °C until ready to proceed to library preparation. The remaining 15 µL of 5m

extracted total nucleic acids is used generate cDNA from the RNA fraction of extracted total nucleic acids.

cDNA from purified RNA

4h 10m

1 Perform RNA purification from a fraction of the extracted total nucleic acids using the RNA Clean & Concentrator™-5 **Zymo Research Catalog #**R1015 1h 10m

The following consumables are required for this step:

Ethanol (100%, Molecular Biology Grade) **Fisher Scientific Catalog #**BP2818500

Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf Catalog #**0030108051

The following equipment is required for this step:

Equipment

Centrifuge

NAME

Benchtop Centrifuge

TYPE

Eppendorf

BRAND

5405000441

SKU

https://online-shop.eppendorf.us/US-en/Centrifugation-44533/Centrifuges-44534/Centrifuge-5425-PF-243560.html

LINK

Any benchtop centrifuge will suffice

SPECIFICATIONS

Prior to starting:

- Add 96 mL of 100% ethanol to the 24 mL RNA wash buffer concentrate.
- Reconstitute the lyophilised DNAse I with nuclease-free water to a concentration of 1 U/ul or defrost an aliquot of previously reconstituted DNAse I.

1.1 Prepare DNAse I master mix as follows: 5m

	Component	Volume per sample (ul)	Total volume (ul) (13)
	DNA digestion buffer	5	65
	DNAse I enzyme (1 U/ul)	5	65
	Nuclease-free water	25	325
	Total	35	455

Table: DNAse I master mix

1.2 Add 35 µL of DNAse I master mix to each sample tube with 15 µL of total nucleic acids and mix my inverting gently. 5m

1.3 Incubate at Room temperature for 00:15:00 . Label a Zymo-Spin IC column and an elution tube for each sample and set aside. 15m

1.4 Add 100 µL RNA binding buffer to each sample 5m

1.5 Add 150 µL of 100% ethanol to each sample and mix by inverting 5m

1.6 Transfer the sample to its labelled Zymo-Spin IC column in a collection tube and centrifuge. Discard the flow-through. 5m

10000-16000 x g, 00:00:30

1.7 Add 400 µL RNA Prep Buffer to each column and centrifuge. Discard the flow-through. 5m

10000-16000 x g, 00:00:30

1.8 Add 700 µL RNA Wash Buffer to each column and centrifuge. Discard the flow-through. 5m

10000-16000 x g, 00:00:30

1.9 Add 400 µL RNA Wash Buffer to each column and centrifuge. Transfer the column to a labelled elution tube. 10m

10000-16000 x g, 00:01:00

1.10 Add 12 µL of DNAse/RNAse-Free Water directly to the column matrix and centrifuge. 10m

10000-16000 x g, 00:01:00

2 Perform first-strand synthesis on the purified RNA using the 1h

SuperScript™ IV First-Strand Synthesis System **Thermo Fisher Scientific Catalog #**18091050

kit.

The following consumables are required for this step:

96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504

The following equipment is required for this step:

Equipment

SimpliAmp Thermal Cycler	NAME
PCR	TYPE
Applied Biosystems	BRAND
A24811	SKU
https://www.thermofisher.com/order/catalog/product/A24811	LINK
Any standard PCR thermocycler will suffice	SPECIFICATIONS

2.1 Prepare priming master mix as follows On ice : 5m

	Component	Volume per sample (ul)	Total volume (ul) (13)
	50ng/ul random hexamers	1	13
	10mM dNTP mix	1	13
	Total	2	26

Table: Priming master mix

2.2 Combine the priming master mix with 11 µL of purified RNA 5m

2.3 Incubate the priming reaction at 65 °C for 00:05:00 and then put the reaction On ice for 00:01:00 . Set the thermocycler lid to 105 °C 10m

2.4 5m

Prepare the 1st strand synthesis master mix as follows On ice :

	Component	Volume per sample (ul)	Total volume (ul) (13)
	5X SSIV Buffer	4	52
	100mM DTT	1	13
	Ribonuclease inhibitor	1	13
	Superscript IV RT (200U/ul)	1	13
	Total	7	91

Table: 1st strand master mix

2.5 Add the 13 µL of RNA with annealed primers to 7 µL of 1st strand master mix 5m
On ice

2.6 Incubate the reaction in a thermocycler with the lid set to 105 °C as follows: 30m

	Temperature (degrees Celsius)	Time (minutes)
	23	10
	55	10
	80	10
	4	Hold

Table: 1st strand thermocycler parameters

3 Perform 2nd strand synthesis on the 1st strand synthesis product using the 1h 15m
Second strand synthesis kit **Invitrogen - Thermo Fisher Catalog #**A48571 kit.

The following consumables are required for this step:

96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504

The following equipment is required for this step:

Equipment

SimpliAmp Thermal Cycler	NAME
PCR	TYPE
Applied Biosystems	BRAND
A24811	SKU
https://www.thermofisher.com/order/catalog/product/A24811	LINK
Any standard PCR thermocycler will suffice	SPECIFICATIONS

3.1 Prepare the 2nd strand cDNA synthesis master mix as follows On ice : 5m

	Component	Volume per sample (ul)	Total volume (ul) (13)
	Nuclease-free water	55	715
	5X second strand reaction mix	20	260
	Second strand enzyme mix	5	65
	Total	80	1040

Table: 2nd strand synthesis master mix

3.2 Add 20 µL of 1st strand synthesis product to 80 µL of 2nd strand synthesis master mix 5m

3.3 Incubate the reaction at 16 °C for 01:00:00 . Set the thermocycler lid to 40 °C . 1h

3.4 Stop the reaction by adding 6 µL of [M] 0.5 Mass Percent EDTA, pH 8 5m

4 Purify the cDNA using Agencourt AMPure XP **Beckman Coulter Catalog #**A63880 beads. 45m

The following consumables are required for this step:

Ethanol (100%, Molecular Biology Grade) **Fisher Scientific Catalog #**BP2818500
Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf Catalog #**0030108051
Nuclease-free Water

The following equipment is required for this step:

Equipment	
Magnetic Stand-96	NAME
ThermoFisher Scientific	BRAND
AM10027	SKU
https://www.thermofisher.com/order/catalog/product/AM10027	LINK

Prior to starting:

- Prepare 80% ethanol
- Label the purified cDNA tubes
- Bring the Agencourt AMPure XP **Beckman Coulter Catalog #**A63880 beads to Room temperature and resuspend by vortexing.

4.1 Add 180 µL (1.8X) of Agencourt AMPure XP **Beckman Coulter Catalog #**A63880 to 100 µL of 2nd strand synthesis product from 4.4 and mix well by pipetting up and down 10 times. 5m

4.2 Incubate for 00:05:00 at Room temperature 5m

4.3 Place the sample plate on a magnetic stand and allow the solution to clear (00:05:00) 5m

4.4 Discard the supernatant without disturbing the beads 2m

4.5 Add 200 µL of 80% ethanol to each well with the sample plate on the magnetic stand and incubate at room temperature for 00:00:30 . 2m

4.6 Discard the supernatant without disturbing the beads 2m

4.7 Repeat the wash steps once. Use a P20 pipette to remove residual ethanol from each well 5m

4.8 Air dry the beads (± 00:02:00) but do not allow the surface to crack. 2m

4.9 Remove the plate from the magnetic rack. Elute the DNA from the beads by adding 25 µL nuclease-free water to the beads. Mix well by pipetting up and down 10 times. 5m

4.10 Incubate for 00:02:00 at Room temperature 2m

4.11 Place the sample plate on a magnetic stand and allow the solution to clear (00:05:00) 5m

4.12 Remove 22 µL of the supernatant and transfer to a clean nuclease-free tube 5m

Pre-Sequencing QC PCR 40m

5 Perform a QC real-time PCR using the 40m

iTaq Universal SYBR® Green One-step kit **Bio-Rad Laboratories Catalog #**1725150

kit. The primers should be specific to the internal control used. This protocol used Tobacco Mild Green Mosaic virus as an internal control with the following primers diluted as per the iTaq kit package insert:

	Primer	Sequence
	Forward	5'-GGATATGTCTAAGTCTGTTGC-3'
	Reverse	5'-CAGACAACTCGGGTGCG-3'

Ellis MD, Hoak JM, Ellis BW, Brown JA, Sit TL, Wilkinson CA, Reed TD, Welbaum GE. Quantitative real-time PCR analysis of individual flue-cured tobacco seeds and seedlings reveals seed transmission of tobacco mosaic virus. Phytopathology. 2020 Jan 19;110(1):194-205.

The following consumables are required for this step:

96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504

The following equipment is required for this step:

Equipment

CFX96 Touch Real-Time PCR	NAME
qPCR	TYPE
Bio-Rad	BRAND
#1855195	SKU
https://www.bio-rad.com/en-no/sku/1855195-cfx96-touch-real-time-pcr-detection-system?ID=1855195	LINK

5.1 Prepare the master mix as follows On ice 5m

	Component	Volume per sample (ul)	Total (ul) (13)
	iTaq Universal SYBR reaction mix (2X)	10	130
	iScript RT	0.25	3.25
	Nuclease-free water	2.75	35.75
	Forward primer	1	13
	Reverse primer	1	13
	Total volume	15	195

Table: Spiked internal control PCR master mix. Sample volume per reaction is 5ul

5.2 Add 5 µL of post 2nd strand synthesis clean-up product to 15 µL of master mix 5m

5.3 Run the following PCR program on a qPCR machine reading the SYBR green fluorophore 30m

	Temperature (degrees Celsius)	Time (seconds)	Cycles
	95	60	1
	95	10	40
	58	20 (+Read)	

Table: Spiked internal control PCR thermocycler parameters. The internal control spike-in should be to a target Ct value of 30±2. The kit reverse transcriptase step is skipped.

Library Preparation

6h 40m

6

1h 17m

Prepare the tagmentation reaction from the

Illumina DNA Prep (M) Tagmentation (24 Samples) **Illumina, Inc. Catalog #**20018704

kit.

The following consumables are required for this step:

96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504

The following equipment is required for this step:

Equipment

SimpliAmp Thermal Cycler	NAME
PCR	TYPE
Applied Biosystems	BRAND
A24811	SKU
https://www.thermofisher.com/order/catalog/product/A24811	LINK
Any standard PCR thermocycler will suffice	SPECIFICATIONS

Equipment

Magnetic Stand-96	NAME
ThermoFisher Scientific	BRAND
AM10027	SKU
https://www.thermofisher.com/order/catalog/product/AM10027	LINK

Prior to starting:

- Bring the bead-linked transposomes (BLT), tagmentation buffer 1 (TB1), tagmentation stop buffer (TSB) and tangent wash buffer (TWB) to Room temperature and vortex to mix

6.1 Pool 15 µL of extracted total nucleic acids (from initial nucleic acid purification step) and 15 µL of purified cDNA (prepared from purified and isolated RNA) in a 96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504 5m

6.2 Prepare the tagmentation master mix as follows: 5m

	Component	Volume per sample (ul)	Total (ul) (12)
	Bead-linked transposomes (BLT)	11	132
	Tagmentation buffer 1 (TB1)	11	132
	Total	22	264

Table: Tagmentation master mix. Vortex to mix.

6.3 Add 20 µL of master mix to 30 µL of pooled DNA and mix by pipetting up and down 10 times 5m

6.4 Run the following program on a thermocycler with the lid set to 100 °C 15m

	Step	Temperature (degrees Celsius)	Time (minutes)
	Tagmentation	55	15
	Hold	10	Hold

Table: Tagmentation thermocycler parameters.

6.5 Add 10 µL of tagmentation stop buffer to each reaction and pipette up and down slowly to resuspend 5m

6.6 Run the following program on a thermocycler with the lid set to 100 °C 15m

	Step	Temperature (degrees Celsius)	Time (minutes)
	Incubation	37	15

	Step	Temperature (degrees Celsius)	Time (minutes)
	Hold	10	Hold

Table: Tagmentation stop thermocycler parameters.

6.7 Place the sample plate on a magnetic stand and allow it to clear 00:03:00 3m

6.8 Remove and discard the supernatant 2m

6.9 Wash by removing the plate from the magnetic stand and adding 100 µL of tagment wash buffer directly onto the beads, pipetting up and down slowly to resuspend 5m

6.10 Place the plate on the magnetic stand and wait for the solution to clear 00:03:00 3m

6.11 Remove and discard the supernatant 2m

6.12 Repeat the wash step 10m

6.13 Remove the plate from the magnetic stand and add 100 µL of tagment wash buffer directly onto the beads and resuspend. Cover the plate and place on the magnetic stand until the amplification step. 2m

7 Amplify the tagmented DNA using the Illumina Nextera DNA CD indexes (24 indexes) **Illumina, Inc.** and the Illumina DNA Prep (M) Tagmentation (24 Samples) **Illumina, Inc. Catalog #**20018704 kit. 2h 31m

The following consumables are required for this step and steps immediately following:

Ethanol (100%, Molecular Biology Grade) **Fisher Scientific Catalog #**BP2818500

96 well LoBind PCR plates Semi-skirted **Eppendorf Catalog #**0030129504

Nuclease-free Water

Qubit 1X dsDNA HS Assay Kit **Thermo Fisher Scientific Catalog #**Q33230

The following equipment is required for this step:

Equipment

SimpliAmp Thermal Cycler	NAME
PCR	TYPE
Applied Biosystems	BRAND
A24811	SKU
https://www.thermofisher.com/order/catalog/product/A24811	LINK
Any standard PCR thermocycler will suffice	SPECIFICATIONS

Equipment

Centrifuge	NAME
Benchtop Centrifuge	TYPE
Eppendorf	BRAND
5405000441	SKU
https://online-shop.eppendorf.us/US-en/Centrifugation-44533/Centrifuges-44534/Centrifuge-5425-PF-243560.html	LINK
Any benchtop centrifuge will suffice	SPECIFICATIONS

Equipment

Plate centrifuge	NAME
Centrifuge	TYPE
Any	BRAND
N/A	SKU

Equipment

Magnetic Stand-96	NAME
ThermoFisher Scientific	BRAND
AM10027	SKU
https://www.thermofisher.com/order/catalog/product/AM10027	LINK

Equipment

Qubit Fluorometer	NAME
Fluorometer	TYPE
Invitrogen	BRAND
Q33238	SKU
https://www.thermofisher.com/order/catalog/product/Q33238#/Q33238	LINK

Prior to starting this step and and steps immediately following:

- Thaw and bring to Room temperature the index adaptors. Thaw the Enhanced PCR mix On ice .
- Bring the sample purification beads (SPB) to Room temperature
- Thaw the resuspension buffer (RSB) and allow it to come to Room temperature
- Prepare 80% ethanol

	Sample identifier	Sample indexes

	Sample identifier	Sample indexes
	Positive control	
	Negative control	

Table: Sample identifier and barcoding details.

7.1 Prepare amplification master mix as follows: 5m

	Component	Volume (ul)	Total (ul) (12)
	Enhanced PCR mix (EPM)	22	264
	Nuclease-free water	22	264
	Total	44	528

Table: Amplification master mix.

7.2 Vortex and centrifuge 280 x g, 00:00:10 1m

7.3 With the plate/tube from the post-tagmentation clean-up, discard the supernatant and remove from the magnetic stand 2m

7.4 Add 40 µL of PCR master mix directly onto the beads of each sample and pipette up and down to ensure the beads are fully suspended before sealing the plate and 5m

centrifuge 280 x g, 00:00:10

7.5 Add adaptors to each sample – 5 µL i7 adaptor and 5 µL i5 adaptor and record the used combination – and mix by pipetting up and down 10 times and centrifuge 280 x g, 00:00:10 10m

7.6 Run the following PCR program with the lid set to 100 °C 50m

	Temperature (degrees Celsius)	Time (seconds)	Cycles
	68	180	1
	98	180	1
	98	45	12
	62	30	
	68	120	
	68	60	1
	10	Hold	Hold

Table: Amplification thermocycler parameters. The cycle number can be adjusted based on input DNA.

7.7 Centrifuge the samples 280 x g, 00:01:00 1m

7.8 Place the plate on a magnetic stand and allow the liquid to clear 00:05:00 5m

7.9 Transfer 45 µL supernatant from each well into a new plate 2m

7.10 Vortex the sample purification beads and add 81 µL to each well containing supernatant and pipette 10 times to mix 5m

7.11 Incubate the plate for 00:05:00 at Room temperature 5m

7.12 Place plate on a magnetic stand and wait for the liquid to clear 00:05:00 — 5m

7.13 Remove and discard the supernatant — 2m

7.14 Wash by adding 200 µL of freshly prepared 80% ethanol to each well without mixing — 5m

7.15 Incubate for 00:00:30 — 30s

7.16 Remove and discard the supernatant — 5m

7.17 Repeat the wash step once — 10m 30s

7.18 Use a small volume pipettor remove residual ethanol — 2m

7.19 Air-dry for up to 00:01:00 but do not let the beads crack — 1m

7.20 Add 32 µL of resuspension buffer (RSB) to the beads and resuspend — 5m

7.21 Incubate at room temperature for 00:02:00 — 2m

7.22 Place on a magnetic stand and allow the liquid to clear (00:02:00) — 2m

7.23 Transfer 30 µL of supernatant to a new plate — 5m

7.24 Quantify the libraries using a Qubit fluorometer using:
Qubit 1X dsDNA HS Assay Kit **Thermo Fisher Scientific** **Catalog** #Q33230 — 15m

	Sample identifier	Concentration (ng/ul)
	Positive control	
	Negative control	

Table: Sample identifier and library concentration after clean-up 1.

8 **Additional clean-up:** Libraries quantified as >10ng/ul will undergo a double-sided clean-up 27m

8.1 Add 25 µL of cleaned library to a new plate (double-sided clean-up plate 1) 5m

8.2 Add 60 µL of nuclease-free water 5m

8.3 Add 45 µL of SPB 5m

8.4 Incubate at Room temperature for 00:05:00 5m

8.5 Place the plate on a magnetic stand and allow the liquid to clear. During this time, to a second plate (double-sided clean-up plate 2) add 15 µL of undiluted SPB to each well to be used. 5m

8.6 Transfer 125 µL of supernatant from double-sided clean-up plate 1 to the predispensed 15 µL of SPB in double-sided clean-up plate 2. Discard double-sided 2m

clean-up plate 1.

9 **Additional clean-up:** Libraries quantified as <10ng/ul will undergo a single-sided clean-up 15m

9.1 Add 25 µL of libary to a new plate (corresponding wells in the double-sided clean-up plate 2 from may be used) 5m

9.2 Add 60 µL of nuclease-free water 5m

9.3 Add 60 µL of SPB 5m

10 **Finalisation of libraries** 50m

10.1 Incubate the plate(s) (double-sided clean-up plate 2) at Room temperature for 00:05:00 5m

10.2 Place on a magnetic stand and wait for the liquid to clear (00:02:00) 2m

10.3 Remove and discard the supernatant 5m

10.4 Wash by adding 200 µL of freshly prepared 80% ethanol to each well without mixing 5m

10.5 Incubate for 00:00:30 30s

10.6 Remove and discard the supernatant 5m

10.7 Repeat the wash step once 10m 30s

10.8 Using a small volume pipettor remove residual ethanol 2m

10.9 Air-dry for up to 00:01:00 but do not let the beads crack 1m

10.10 Add 30 µL of resuspension buffer (RSB) to the beads and resuspend 5m

10.11 Incubate at room temperature for 00:02:00 2m

10.12 Place the plate on a magnetic stand and wait for the liquid to clear (00:02:00) 2m

10.13 Transfer 28 µL of supernatant to a nuclease-free tube 5m

11 Quantification and QC of libraries 1h 20m

11.1 Quantify the libraries using a Qubit fluorometer using: 15m

Qubit 1X dsDNA HS Assay Kit **Thermo Fisher Scientific Catalog #**Q33230

	Sample identifier	Concentration (ng/ul)
	Positive control	
	Negative control	

Table: Sample identifier and library concentration after final clean-up.

11.2 (Optionally) evaluate the library mean fragment size using a TapeScreen instrument. If this is not done, use a default mean fragment size of 600bp. 1h

11.3 Calculate the Molar concentration of the library using the following formula: 5m

$$Molarconcentration = concentration(ng/ul) * 10^6/660 * avgerage fragmentsize(bp)$$

Sequencing 2d 9h 6m

12 Prepare the quantified and quality controlled libraries for loading and load the Sequencing instrument. 2d 9h 6m

The following consumables are required for this step:

- Sodium Hydroxide 1M solution
- Safe-Lock Tubes 1.5 ml PCR clean DNA LoBind **Eppendorf Catalog #**0030108051
- Nuclease-free Water
- MiSeq v3 Sequencing Reagents (600 cycles) **Illumina, Inc. Catalog #**MS-102-3003
- PhiX Control v3 **Illumina, Inc. Catalog #**FC-110-3001

The following equipment is required for this step:

Equipment

MiSeq

NAME

Sequencer

TYPE

illumina

BRAND

SY-410-1003

SKU

https://www.illumina.com/systems/sequencing-platforms/miseq/order-miseq.html

LINK

Equipment

Centrifuge

NAME

Benchtop Centrifuge

TYPE

Eppendorf

BRAND

5405000441

SKU

https://online-shop.eppendorf.us/US-en/Centrifugation-44533/Centrifuges-44534/Centrifuge-5425-PF-243560.html

LINK

Any benchtop centrifuge will suffice

SPECIFICATIONS

Prior to starting this step and and steps immediately following:

- Defrost the PhiX control
- Ensure the MiSeq instrument doesn't require maintenance or a weekly wash
- Remove the Illumina Miseq v3 reagent cartridge from -20 °C storage and defrost in a water bath approximately 01:00:00 before its intended use. Alternatively place it at 4 °C overnight prior to its intended use

- Defrost buffer HT-1 On ice and keep at 4 °C until use
- Prepare a fresh solution of [M] 0.2 Molar NaOH
- Create the sample sheet using the appropriate software that will be used for the MiSeq sequencing run (this will indicate the number of cycles, the library preparation parameters and the sample names and indexes for use by the instrument to set the run parameters and demultiplex the output data)

12.1 Dilute the prepared libraries to [M] 4 nanomolar (nM) using nuclease-free water and pool them in a single LoBind Eppendorf tube 10m

12.2 Combine the diluted library and [M] 0.2 Molar NaOH solution in a LoBind Eppendorf tube to denature the libraries as follows: 2m

	Component	Volume (ul)
	4nM pooled libraries	5
	0.2M NaOH	5
	Total volume	10

Table: Library denaturation.

12.3 Vortex briefly and spin-down the denaturing libraries and incubate for 00:05:00 at Room temperature 5m

12.4 Stop the reaction by adding 990 µL of chilled HT-1 buffer 2m

12.5 Invert several times and spin-down the solution 2m

12.6 Repeat the denaturation and stopping of a [M] 4 nanomolar (nM) PhiX library 21m

12.7 The denatured library concentration is [M] 20 picomolar (pM) . Remove 100 µL from the library and add 100 µL of the [M] 20 picomolar (pM) denatured and 2m

stopped PhiX library

12.8 Further dilute the library (with added PhiX) to a [M] 10 picomolar (pM) concentration and keep the reaction On ice until the library is loaded into the MiSeq reagent cartridge 2m

12.9 Follow the prompts on the MiSeq screen to load the flow cell, wash buffer and reagent cartridge and load 600 µL of [M] 10 picomolar (pM) pooled libary solution (with PhiX) prior to loading the cartridge. 20m

12.10 Start the sequencer 2d 8h

Bioinformatics

13 The bioinformatics tools needed to perform the complete analysis are available at: https://github.com/RuanMarais/UCT_metagenomics. Alternatively, basic analysis can be performed as described below.

14 The analysis described below was performed on a computer with a 32-core Intel(R) Xeon(R) CPU and 124Gb of RAM. Analysis was done locally and the computer was running Linux (Ubuntu distribution).

14.1 Install conda, instructions are available at: https://docs.conda.io/projects/conda/en/latest/user-guide/install/linux.html

14.2 To install the required packages run the following commands:

Command

Installs biobakery_workflows (Linux Ubuntu)

Install biobakery

```
conda install -c biobakery biobakery_workflows
```

Command

Biobakery dependency (Linux Ubuntu)

Install bio bakery dependency

```
conda install tbb=2020.2
```

Command

Download kneaddata database (Linux Ubuntu)

Biobakery database

```
kneaddata_database --download human_genome bowtie2
$PATH_TO_WORKING_DIRECTORY
```

The kneaddata database can also be directly dowloaded from:
http://huttenhower.sph.harvard.edu/kneadData_databases

Command

Install kraken2 (Linux Ubuntu)

```
conda install -c bioconda kraken2
```

Command

Install SPAdes (Linux Ubuntu)

Install spades

```
conda install -c bioconda spades
```

Command

Install DIAMOND (Linux Ubuntu)

```
conda install -c bioconda diamond
```

Command

Install krakentools (Linux Ubuntu)

```
conda install -c bioconda krakentools
```

Command

Install PEAR (Linux Ubuntu)

```
conda install -c bioconda pear
```

14.3 Additional dependencies:

Kraken2 databases can be constructed de novo or are available as pre-constructed databases at:
https://benlangmead.github.io/aws-indexes/k2

Download the standard database and the EuPathDB48 database and extract into the working directory where you will analyse the sequencing data.

14.4 Additional dependencies:

Download the Trimmomatic (Trimmomatic-0.39) binary file from:
usadellab.org/cms/?page=trimmomatic

and move it (after extraction) to bin as follows:

Command

Move trimmomatic binary file to bin (Linux Ubuntu)

Move trimmomatic to bin

```
sudo mv $PATH_AFTER_DOWNLOAD /bin
```

14.5 Additional dependencies:

Download the {Reviewed (Swiss-Prot)} fasta file to generate the DIAMOND database from:
https://www.uniprot.org/help/downloads

14.6 Analyse the raw sequencing reads as follows:

1) Human read, contaminant and low-complexity read removal with kneaddata:

Command

kneaddata - human, contaminant and low complexity read removal (Linux Ubuntu)

Run kneaddata

```
kneaddata --input $PAIRED_END_RAW_READ_1 --input
$PAIRED_END_RAW_READ_2 --reference-db
Homo_sapiens_hg37_and_human_contamination_Bowtie2_v0.1 --output
~/$WORKING_DIRECTORY/ --threads 30 --trimmomatic /bin/Trimmomatic-
0.39/
```

Review the kneaddata paired reads in FastQC and make any required modifications to the trimming algorithm if necessary.

2) Taxonomic classification using kraken2:

Command

Run kraken2 taxonomic classifier (Linux Ubuntu)

Run kraken2

```
kraken2 --db $STANDARD_DATABASE_PATH_OR_EUPATHDB48 --threads 30 --
report $NAME_OF_REPORT_WITH_DBNAME_DATACLEANSTRATEGY_DATE_SAMPLE --
paired $KNEADDATA_PAIRED_READ_1 $KNEADDATA_PAIRED_READ_2 --output
$KRAKENFILE
```

This command is run once for each kraken2 database used. Thus, for this analysis the command twice for the standard and EuPathDB48 databases.

3) Extract unclassified reads using krakentools:

Command

Extract unclassified reads using kraken tools (Linux Ubuntu)

Extract unclassified reads

```
extract_kraken_reads.py -k $KRAKENFILE -t 0 –fastq-output -s
$KNEADDATA_PAIRED_READ_1 -s2 $KNEADDATA_PAIRED_READ_2 -o
$UNMATCHED_FORWARD -o2 $UNMATCHED_REVERSE
```

The reads not classified by kraken2 are extracted for further analysis.

4) Generate contigs from unclassified reads using SPAdes:

Command

Generate contigs using SPAdes (Linux Ubuntu)

SPAdes

```
spades.py --meta -1 $UNMATCHED_FORWARD -2 $UNMATCHED_REVERSE -o
$SPADES_OUTPUT_DIRECTORY
```

5) Generate merged paired-end reads if no contigs are generated:

Command

Assemble paired-end reads (Linux Ubuntu)

```
pear -f $UNMATCHED_FORWARD -r $UNMATCHED_REVERSE -o $OUTPUT_DIRECTORY
```

6) Generate the DIAMOND database as follows:

Command

Generate DIAMOND database (Linux Ubuntu)

Generate DIAMOND database

```
diamond makedb --in $SWISS_PROT.fasta -d $DIAMOND_DATABASE_NAME
```

7) Run DIAMOND on generated contigs and/or assembled unclassified reads (if no contigs are generated):

Command

Run DIAMOND (Linux Ubuntu)

```
diamond blastx -d $DIAMOND_DATABASE_NAME -q
$CONTIGS_OR_ASSEMBLED_READS.fasta -o $DATABASE_MATCHES.tsv
```

14.7 Data visualisation can be done using Pavian. Installation and usage instructions are available here:
https://github.com/fbreitwieser/pavian