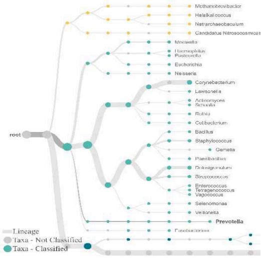
Aug 30, 2022 Version 1

Characterization of the Archaeome, Bacteriome and Eukaryome in Nasopharyngeal Swabs V.1



DOI
dx.doi.org/10.17504/protocols.io.bp2l61ok1vqe/v1

Carolin Baehren[1], Anton Pembaur[2], Patrick P. Weil[2], Frank Schult[3], Stefan Wirth[3], Jan Postberg[2], Malik Aydin[4]

[1]Laboratory of Experimental Pediatric Pneumology and Allergology, Center for Biomedical Education and Research, School of Life Sciences (ZBAF), Faculty of Health, Witten/Herdecke University, 58455 Witten, Germany;
[2]Clinical Molecular Genetics and Epigenetics, Faculty of Health, Center for Biomedical Education & Re-search (ZBAF), Helios University Hospital Wuppertal, Witten/Herdecke University, Alfred-Herrhausen-Str. 50, 58448 Witten, Germany;
[3]Center for Child and Adolescent Medicine, Center for Clinical and Translational Research (CCTR), Helios University Hospital Wuppertal, Witten/Herdecke University, 42283 Wuppertal, Germany;
[4]Laboratory of Experimental Pediatric Pneumology and Allergology, Center for Biomedical Education and Research, School of Life Sciences (ZBAF), Faculty of Health, Witten/Herdecke University, 58455 Witten, Germany, Center for Child and Adolescent Medicine, Center for Clinical and Translational Research (CCTR), Helios University Hospital Wuppertal, Witten/Herdecke University, 42283 Wuppertal, Germany

Anton Pembaur

OPEN ACCESS

DOI: dx.doi.org/10.17504/protocols.io.bp2l61ok1vqe/v1

Protocol Citation: Carolin Baehren, Anton Pembaur, Patrick P. Weil, Frank Schult, Stefan Wirth, Jan Postberg, Malik Aydin 2022. Characterization of the Archaeome, Bacteriome and Eukaryome in Nasopharyngeal Swabs. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bp2l61ok1vqe/v1

Manuscript citation:
Carolin Baehren, Anton Pembaur, Patrick P. Weil, Frank Schult, Stefan Wirth, Jan Postberg, and Malik Aydin 2022 Characterization of the Archaeome, Bacteriome and Eukaryome in Nasopharyngeal Swabs

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

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

Created: August 12, 2022

Last Modified: August 30, 2022

Protocol Integer ID: 68548

Keywords: strong focus on archaea research, archaea research, archaea, characterization of the archaeome, archaeome, eukaryome in nasopharyngeal swab, microbiome, sequencing technique, nasopharyngeal swab specimen, use of nasopharyngeal swab specimen, oxford nanopore, nasopharyngeal swab, bacteriome, minion oxford nanopore, eukaryote, pathogenicity in human disease, respiratory disease, sequencing platform, nanopore technology, pathogenicity, eukaryome, respiratory disorder, pathogenesis in term, pathogenesi

Abstract

This protocol describes the Characterization of the Archaeome, Bacteriome and Eukaryome in Nasopharyngeal Swabs by sequencing with nanopore technology.

For a long time, archaea were under-represented in the literature, and less is known about their pathogenicity in human diseases. Using conventional methods, the cultivability particularly of archaea is challenging and they are still classified as the 'dark matter' of the microbiome. The evolution of advanced sequencing techniques in the twenty-first century, a strong focus on archaea research is interestingly observed. However, the influence on disease course or even pathogenesis in terms of respiratory disorders remain unexplored. Thus, more attention has to be paid on the characterization of the archaeome with the goal of translation into clinical contexts. Considering this important issues lacking good methodological reports in the literature, we evaluated previously developed primer sets and sequencing platforms. With these useful hints, we share potential alternative procedures with the aim how to increase the quality of research on archaeome and eukaryotes. The use of nasopharyngeal swab specimens derived from a cohort suffering from respiratory diseases enable to study translational aspects on disease course and eventually pathogenesis. The optimization of 'pre-sequencing' steps, starting from the DNA isolation, amplification, right choice of sequencing platforms e.g., MinION Oxford Nanopore rule some important traces to a high-qualitative in-depth sequencing success. However, those descriptive data significantly contribute to optimize existing archaic models with the aim to exploit translational approaches *ex vivo*.

Materials

DNA isolation:

QIAmp DNA-Mini Kit by Qiagen	(Qiagen 51304)

PCR:

Q5 HotStart High Fidelity 2x MM	(NEB M0492)

Library Preparation + Sequencing:

NEBNext FFPE DNA Repair Mix	(NEB M6630)
NEBNext Ultra II End Repair/dA-Tailing Module	(NEB E7546)
NEBNext Quick Ligation Module	(NEB E6056)
Agencourt AMPure XP	(Beckman Coulter A63880)
Long Amp Tag Polymerase MM	(NEB M0323)
PCR Barcoding Expansion 1-96	(ONT EXP-PBC096)
Ligation Sequencing Kit	(ONT SQK-LSK110)

DNA Isolation

2h

1 DNA was isolated from nasal swabs with amies medium, using approximately 1000 µL .

For the Isolation, the QIAmp DNA-Mini Kit by Qiagen was used, following the QiAamp tissue protocol from the 'QIAamp DNA Mini and Blood Mini Handbook 05/2016'

1.1 Centrifugation of sample at 7500 rpm, 00:10:00 until pellet formation.

10m

1.2 Resuspension of pellet in 180 µL ATL

5m

1.3 Adding 20 µL Proteinase K , vortexing,

Incubate at 56 °C until complete lysis. Occasionally vortexing.

10m

1.4 Brief Centrifugation of the sample.

2m

1.5 Adding 200 µL Buffer AL , pulse-vortex afterwards 00:00:15 , Incubation at 70 °C for 00:10:00 . Brief centrifugation of the sample.

15m

1.6 Adding 200 µL Ethanol (96-100%) , and mix by pulse-vortexing for 00:00:15 . Afterwards, short centrifugation of the sample.

6m

1.7 Transfer mixture (including precipitate) to the QIAamp Mini spin column. CAVE: without wetting the rim. Centrifugation: 8000 rpm, 00:01:00

Replace the QIAamp Mini spin column, use a clean 2 ml collection tube, Discard tube with the filtrate.

5m

1.8 Adding 500 µL Buffer AW1 CAVE: without wetting the rim.

Centrifuge: 8000 rpm, 00:01:00 .

Replace the QIAamp Mini spin column, use a clean 2 ml collection tube, Discard tube with the filtrate.

5m

1.9 Add 500 µL Buffer AW2 to the QIAamp Mini spin column without wetting the rim. Closing of the column, Centrifugation:
14000 rpm, 00:03:00 , Centrifugation at full speed

8m

1.10 Replace QIAamp Mini spin column with a new 2 ml collection tube. Discard tube with the filtrate. Centrifugation: 14000 rpm, 00:01:00 , Centrifugation at full speed

2m

1.11 Placing QIAamp Mini spin column in a new 1.5 ml microcentrifuge tube. Discard tube with the filtrate.
Add 100 µL AE . Incubation at room temperature (00:05:00), centrifugation:
8000 rpm, 00:01:00

7m

1.12 Repeat Step 1.11: Add the flowthrough of the previous step to the Mini spin column and incubation at room temperature for 00:05:00 , centrifugation
8000 rpm, 00:01:00 at .

7m

1.13 Concentration measurement with nanophotometer or qubit.

5m

PCR

2h 30m

2 **PCR**
Archaea: Nested PCR
Eukaryotes: single PCR

2.1 **Primer selection for archaea and eukaryotes**

A	B	C	D
Nr.	**Name**	**Primer Name**	**Sequence (5′ ◊ 3′)**
1	344F	S-D-Arch-0344-fw	5′-acggggygcagcaggcgcga-3′
2	1041R	S-D-Arch-1041-rev	5′-ggccatgcaccwcctctc-3′
3	519F	Arch-519F-Tag	5′-tttctgttggtgctgatattgccagcmgccgcggtaa-3′
4	786R	Arch-786R-Tag	5′-acttgcctgtcgctctatcctcggactacvsgggtatctaat-3′

A	B	C	D
5	563F	Euk-563F-Tag	5′-tttctgttggtgctgatattgcgccagcavcygcggtaay-3′
6	1132R	Euk-1132R-Tag	5′-acttgcctgtcgctctatcttcccgtcaatthcttyaart-3′

2.2 **1st PCR Mix:**

10m

- 8 µL nuclease free water
- 12.5 µL Q5 Polymerase
- 2 µL Primer Mix
- 2.5 µL DNA-Template

2.3 **PCR-Run. 1**

53m

Primer pair Arch-344-F-1041R / Eck.563F-1132Rtag

Heated Lid: 110 C

Denaturation 95 °C 00:03:00

Cycles (30):

Denaturation 95 °C 00:00:30

Annealing 55 °C 00:00:30

Elongation 72 °C 00:00:30

End Cycle

Final Elongation 65 °C 00:05:00

2.4 **2nd PCR (archaea only) Mix:**

10m

- 9.5 µL nuclease free water
- 12.5 µL Q5 Polymerase
- 2 µL Primer-Mix
- 1 µL DNA-Template

2.5 **PCR-Run.** 2 (nested)

50m

Primer pair Arch-519-F-786Rtag

Heated Lid: 110 C

Denaturation 95 °C 00:03:00

Cycles (28):

Denaturation 95 °C 00:00:30

Annealing 55 °C 00:00:30

Elongation 72 °C 00:00:30

End Cycle

Final Elongation 65 °C 00:05:00

2.6 **Gel Electrophoresis** 1h 30m

Check, if the wanted sequences were amplified. Ether through classic gel electrophoresis or through microcapillary gel electrophoresis.

Library preparation + sequencing: 10m

3 **Library Preparation + Sequencing:**

- 1st Purification
- PCR preparation
- 2nd Purification
- Concentration measurement

3.1 **1st Purification:** 10m

Add 36 µL Beats AMPure XP and apply an external magnetic field for 00:05:00 . Afterwards discard fluid supernatant.

3.2 Add 150 µL Ethanol 70% and discard fluid supernatant. 3m

3.3 Add another 150 µL Ethanol 70% . Afterwards discard the fluid supernatant and dry tube with open lid 8m

3.4 Resuspend pellet in 15 µL nuclease free water 5m

3.5 **PCR-preparation: Mix** 10m

12.5 µL Long Amp Tag Polymerase MM
2 µL sample
9.5 µL nuclease free water
1 µL Barcode

3.6

Heated Lid: 110 C 30m

Denaturation 95 °C 00:03:00

Cycles (18):

Denaturation 95 °C 00:00:15

Annealing 62 °C 00:00:15

Elongation 65 °C 00:00:45

End Cycle

Final Elongation 65 °C 00:05:00

3.7 2nd Purification

3.8 Add 36 µL Beats and apply an external magnetic field for 00:05:00 . Afterwards discard fluid supernatant 10m

3.9 Add 150 µL Ethanol 70% and discard fluid supernatant. 3m

3.10 Add another 150 µL Ethanol 70% . Afterwards discard the fluid supernatant and dry tube with open lid 8m

3.11 Resuspend pellet in 15 µL nuclease free water 5m

3.12 Concentration measurement with nanophotometer 5m

3.13 **Library preparation:**

We used a modified version of the PCR barcoding (96) genomic DNA (SQK-LSK109) protocol by Nanopore.
Quantify the barcoded library using a nanophotometer and pool all barcoded libraries in the desired ratios in a 1.5 ml DNA LoBind Eppendorf tube.

3.14 Prepare 1 µg pooled barcoded libraries in 47 µL nuclease free water . 10m

3.15 **DNA repair and end-prep**

3.16 Thaw DNA CS (DCS) at RT, spin down, mix by pipetting, and place on ice. 5m

3.17 Prepare the NEBNext FFPE DNA Repair Mix and NEBNext Ultra II End repair / dA-tailing Module reagents in accordance with manufacturer's instructions, and place on ice. 5m

3.18 In a 0.2 ml thin-walled PCR tube, mix the following: 10m

- 1 µL DNA CS
- 47 µL DNA
- 3.5 µL NEBNext FFPE DNA Repair Buffer
- 2 µL NEBNext FFPE DNA Repair Mix
- 3.5 µL Ultra II End-prep reaction buffer
- 3 µL Ultra II End-prep enzyme mix

Mix gently by flicking the tube, and spin down.

3.19 Using a thermal cycler, incubate at 20 °C for 00:05:00 and 65 °C for 00:05:00 10m

3.20 **AMPure XP bead clean-up**

3.21 Resuspend the AMPure XP beads by vortexing.
Transfer the DNA sample to a clean 1.5 ml Eppendorf DNA LoBind tube. 5m

3.22 Add 60 µL of resuspended AMPure XP beads to the end-prep reaction and mix by flicking the tube. 5m

3.23 Incubate on a Hula mixer (rotator mixer) for 00:05:00 at room temperature. 5m

3.24 Prepare 500 µL of fresh 70% ethanol in Nuclease-free water. 5m

3.25 Spin down the sample and pellet on a magnet until eluate is clear and colourless. Keep the tube on the magnet, and pipette off the supernatant. 10m

3.26 Keep the tube on the magnet and wash the beads with 200 µL freshly prepared 70% ethanol without disturbing the pellet. Remove the ethanol using a pipette and discard. 3m

3.27 Repeat the previous step. 3m

3.28 Spin down and place the tube back on the magnet. Pipette off any residual ethanol. Allow to dry for 00:00:30 , but do not dry the pellet to the point of cracking. 5m

3.29 Remove the tube from the magnetic rack and resuspend the pellet in 61 µL nuclease-free water
Incubate for 00:02:00 at RT. 5m

3.30 Pellet the beads on a magnet until the eluate is clear and colourless. 2m

3.31 Remove and retain 61 µL of eluate into a clean 1.5 ml Eppendorf DNA LoBind tube. 3m

3.32 Take forward the repaired and end-prepped DNA into the adapter ligation step. However, at this point it is also possible to store the sample at 4 °C overnight.

3.33 **Adapter ligation and clean-up** (PCR barcoding (96) genomic DNA (SQK-LSK109) protocol by Nanopore)

Although the recommended 3rd party ligase is supplied with its own buffer, the ligation efficiency of Adapter Mix (AMX) is higher when using Ligation Buffer supplied within the Ligation Sequencing Kit.

3.34 Spin down the Adapter Mix (AMX) and Quick T4 Ligase, and place on ice. 1m

3.35 Thaw Ligation Buffer (LNB) at RT, spin down and mix by pipetting.

Due to viscosity, vortexing this buffer is ineffective. Place on ice immediately after thawing and mixing. 1m

3.36 Thaw the Elution Buffer (EB) at RT, mix by vortexing, spin down and place on ice. 5m

3.37 To retain DNA fragments of < 3 KB, thaw one tube of Short Fragment Buffer (SFB) at RT, mix by vortexing, spin down and place on ice.

3.38 In a 1.5 ml Eppendorf DNA LoBind tube, mix in the following order: 10m

- 60 µL DNA sample from the previous step
- 25 µL Ligation Buffer (LNB)
- 10 µL NEBNext Quick T4 DNA Ligase
- 5 µL Adapter Mix (AMX)

Mix gently by flicking the tube, and spin down.

3.39 Incubate the reaction for 00:10:00 at RT. If you have omitted the AMPure purification step after DNA repair and end-prep, do not incubate the reaction for longer than 00:10:00 . 10m

3.40 Resuspend the AMPure XP beads by vortexing. Add 40 µL of resuspended AMPure XP beads to the reaction and mix by flicking the tube. 5m

3.41 Incubate on a Hula mixer (rotator mixer) for 00:05:00 at RT. 5m

3.42 Spin down the sample and pellet on a magnet. Keep the tube on the magnet, and pipette off the supernatant. 10m

3.43 Wash the beads by adding 250 µL Short Fragment Buffer (SFB) . Flick the beads to resuspend, spin down, then return the tube to the magnetic rack and allow the beads to pellet. Remove the supernatant using a pipette and discard. 5m

3.44 Repeat the previous step.

5m

3.45 Spin down and place the tube back on the magnet. Pipette off any residual supernatant. Allow to dry for 00:00:30 but do not dry the pellet to the point of cracking. 3m

3.46 Remove the tube from the magnetic rack and resuspend the pellet in 15 µL Elution Buffer (EB). Spin down and incubate for 00:10:00 at RT. 10m

3.47 Pellet the beads on a magnet until the eluate is clear and colourless. 5m

3.48 Remove and retain 15 µL of eluate containing the DNA library into a clean 1.5 ml Eppendorf DNA LoBind tube. 2m

3.49 Quantify 1 µL of eluted sample using a Qubit fluorometer. The prepared library is used for loading into the flow cell. Store the library on ice until ready to load. 1m

3.50 The prepared library is used for loading into the flow cell. Store the library on ice until ready to load.

3.51 **Priming and loading the SpotON flow cell**

3.52 Thaw the Sequencing Buffer (SQB), Loading Beads (LB), Flush Tether (FLT) and one tube of Flush Buffer (FB) at RT. 1m

3.53 Mix the Sequencing Buffer (SQB), Flush Tether (FLT) and Flush Buffer (FB) tubes by vortexing and spin down at RT. 1m

3.54 Open the MinION Mk1B lid and slide the flow cell under the clip. Slide the priming port cover clockwise to open the priming port. 1m

⚠ Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µl, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.

3.55 After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles (a few µl): 1m

Set a P1000 pipette to 200 µl
Insert the tip into the priming port
Turn the wheel until the dial shows 220-230 µl, or until you can see a small volume of buffer entering the pipette tip

3.56 To prepare the flow cell priming mix, add 30 µL of thawed and mixed Flush Tether (FLT) directly to the tube of thawed and mixed Flush Buffer (FB), and mix by vortexing at RT. 1m

3.57 Load 800 µL of the priming mix mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Wait for 00:05:00 . During this time, prepare the library for loading by following the steps below. 6m

3.58 Thoroughly mix the contents of the Loading Beads (LB) by pipetting because it contains a suspension of beads which settle very quickly. It is vital that they are mixed immediately before use! 1m

3.59 In a new tube, prepare the library for loading as follows: 2m
37.5 µL Sequencing Buffer (SQB)
25.5 µL Loading Beads (LB) , mixed immediately before use
12 µL DNA library

3.60 Complete the flow cell priming through Gently lifting the SpotON sample port cover to make the SpotON sample port accessible. Load 200 µL of the priming mix into the flow cell via the priming port (not the SpotON sample port), avoiding the introduction of air bubbles. 1m

3.61 Mix the prepared library gently by pipetting up and down just prior to loading. 1m

3.62 Add 75 µL of sample to the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop flows into the port before adding the next. 1m

3.63 Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port and replace the lid. 1m

3.64 If you using a MinION Mk1C turn basecalling while sequencing on.

1m

3.65 **Ending the experiment**

After your sequencing experiment is complete, if you would like to reuse the flow cell, please follow the Wash Kit instructions and store the washed flow cell at 2-8°C, OR

Follow the returns procedure by washing out the flow cell ready to send back to Oxford Nanopore.

Bioinformatics:

1d

4

4.1 If you were unable to basecall in real time, perform the basecalling now using the Guppy basecaller (newest version).

https://community.nanoporetech.com/docs/prepare/library_prep_protocols/Guppy-protocol/v/gpb_2003_v1_revag_14dec2018/guppy-software-overview

4.2 Now, using the resulting .fastq files, run the WIMP workflow from the Epi2Me software.

(https://nanoporetech.com/resource-centre/epi2me-wimp-workflow-quantitative-real-time-species-identification-metagenomic)

4.3 If the graphical output from the WIMP workflow is not sufficient for your analysis, you can download the results in a .csv dataset. Due to the size of this dataset, further analyses may be performed by creating an SQL database.

The data contains the
- filename of the .fastq file
- Read ID → is the unique primary key, wich enables to identify the read and therefore the sequence
- Run ID
- exit_status (of the read from the WIMP workflow)
- barcode
- taxID (every phylogenetic rank of each species has its own ID, with these IDs the lineage is composed
- name (of the organism)
- score

  - lineage

4.4 Python scripts

While working on this project, a few Python scripts may be useful, depending on analysis you want to perform.

This script we used to split large files into smaller ones:

```
#IMPORTANT: this script must be started from the same file
directory as your input file!

filecounter=0
filelinecounter=0

inputfilename="file_i_want_to_split.txt"  #set the correct name of
the file, you want to #split.
filename=inputfilename.split(".")[0]
file_lines= open(inputfilename, 'r').readlines()
print(len(file_lines))
while filelinecounter
outputfilename=filename+"_"+str(filecounter).zfill(3)+".txt"  #set
#the correct ending for your file here
        print(outputfilename)

        while filelinecounter
outfile.write(file_lines[filelinecounter])
                filelinecounter=filelinecounter+1
        else: filecounter=filecounter+1
```

This script was used, to append the lenght of each analysed read (or with small changes the whole sequence) to the .csv table:

```
#IMPORTANT: this script must be started from the same file
directory as your input file!
# This script, the .fastq files from the run you want to analyse
and the WIMP.csv file must be in the same directory!


inputfilename="WIMP_inputfile.csv" #change the inputfile here

import os
from multiprocessing import Pool
import concurrent.futures #imports the multithreading library
import shutil
from pathlib import Path
filecounter=0
filelinecounter=0
i=1

# Define a function for the thread
def search_fasta(WIMP_inputline):
        WIMP_inputline=WIMP_inputline.rstrip()
        fastqfilename=WIMP_inputline.split("-",2)[0]+".fastq"
        #print(str(fastqfilename))
        readID=WIMP_inputline.split(",",3)[1]
        #print(str(readID))
        fqfile=open(fastqfilename, 'r').readlines()
        #print("fqfile is open")
        #print(str(fqfile[0]))
        fqcounter=0
        found= False
        while found == False:
                fqreadID= fqfile[fqcounter*4].split()[0][1:37]
                #print(str(fqreadID))
                if (readID == fqreadID):
                        readlenght=len(fqfile[fqcounter*4+1])  #
if you want to get the sequence instead of the lenght, remove the
len() function.
                        #print(str(readlenght))
                        found=True
                else: fqcounter=fqcounter+1
        completeline=WIMP_inputline+","+str(readlenght)+"\n"
        #print("Thread")
        return completeline

if __name__ == "__main__":
```

```
        dirname = os.path.join("C:/WIMPlenght_tmp")
        os.mkdir(dirname)

        filename=inputfilename.split(".")[0]
        print(filename)
        file_lines= open(inputfilename, 'r').readlines()
        print(len(file_lines))
        while filelinecounter
throughputfilename=filename+"_"+str(filecounter).zfill(6)+".csv"
                print(throughputfilename)
                while filelinecounter
outfile=open(dirname+"/"+throughputfilename, 'a')
                        print(filelinecounter)
                        outfile.write(file_lines[filelinecounter])
                        filelinecounter=filelinecounter+1
                        #print(filelinecounter)
                else:
                        filecounter=filecounter+1
                        print("Filenumber: ", filecounter)

        print("tmpfiles complete")
        outputfilename=inputfilename.split(".")
[0]+"_Output_WIMP&Seqlenght.csv"
        print(outputfilename)

        while i < filecounter:

                tmpfilename=filename+"_"+str(i).zfill(6)+".csv"
                WIMP_lines =
open(os.path.join(dirname+"/"+tmpfilename), 'r').readlines()
#opens the tmp WIMP outputfile and creates a list with each line
as one item in the list
                p=Pool()
                with open(outputfilename, 'a') as outfile:
                        result=p.map(search_fasta, WIMP_lines)
                        p.close()
                        p.join()
                        #print(result)
                        for f in result:
                                #print(f)
                                outfile.write(f)
                i=i+1
                print(i)
        else:
                print("task complete")
```

```
                shutil.rmtree(dirname)
                print("tmpfiles deleted")
```