Instituto de Microbiología
Universidad San Francisco de Quito

Quito, 12th March 2020

SARS-CoV2 genome sequencing from the first patient identified with COVID-19 in Quito, Ecuador

Sully Márquez[1], Belén Prado-Vivar[1,2], Juan José Guadalupe[4], Manuel Jibaja[5], Milton Tobar[5], Verónica Barragán[1], Patricio Rojas-Silva[1], Gabriel Trueba[1], Michelle Grunauer[3], Paúl Cárdenas[1,2*]

[1]Universidad San Francisco de Quito, COCIBA, Instituto de Microbiología
[2]Universidad San Francisco de Quito, Centro de Bioinformática
[3]Universidad San Francisco de Quito, COCSA
[4]Universidad San Francisco de Quito, COCIBA, Laboratorio de Biotecnología Vegetal
[5]Unidad de Cuidados Intensivos, Hospital Eugenio Espejo, Quito

Contact: pacardenas@usfq.edu.ec, tel.: +59322971700 ext 1517

A broncho-alveolar lavage (BAL) sample was received from a 57-year-old patient of Dutch origin who presented symptoms in the Sucumbios province and then transferred to Quito. The Ecuadorian Ministry of Public Health and INSPI officially reported the positivity for SAR-CoV2 using the standard RT-PCR approach in public media. The BAL sample was collected on 03/10/2020 in DNA / RNA Shield (Zymo) to ensure virus inactivation and to conserve the genetic material and transported immediately at 4ºC in a sealed container with all the biosecurity and containment measures recommended by the CDC of the USA (https://www.fda.gov/media/134922/download).

The genetic material of the sample was extracted in a biosafety type 2 chamber with HEPA filters in the Virology Laboratory of the IM of the USFQ. The QIAamp® Viral RNA Extraction Kit from Qiagen was used. Retranscription of RNA to cDNA was carried out using a superscript II platinum One Step PCR retrotranscriptase (invitrogen). The quality and concentration of the genetic material was quantified using QuBit obtaining 2.46 ng / ul of total cDNA. cDNA sequencing was performed according to the RNA Viral Metagenomics MinION One-Pot Sequencing Protocol of Public Health England, Genomics Lab at the USFQ Bioinformatics Center.

Bioinformatic analysis used the PoreChop algorithm to assign the sequences to their genetic barcodes, and the Kaiju and Kraken platforms to assign the taxonomy to the sequences found.

The metagenomic analysis found a total of 206,111 DNA sequences with 43,603,091 bases, of which 0.3% corresponded to Coronavirus sequences, 83% were not assigned and 17% were identified as SARS coronaviruses (Figure 1).

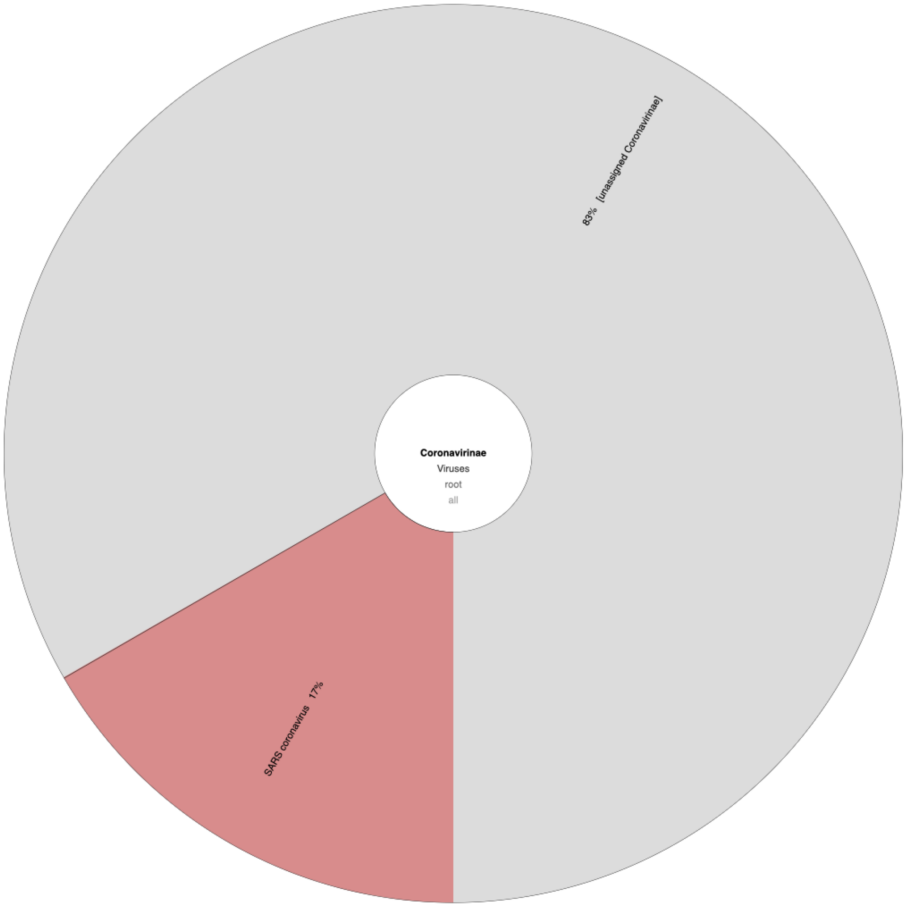

Figure 1. Krona chart summarizing the percentage of sequences assigned to Coronavirus in the metagenome.

The sequences assigned to Coronavirus were extracted and mapped using the Wuhan-Hu-1 reference genome (GenBank accession number MN908947). A sequence similarity of 99.68% was found with this sequence with 100% query coverage. Additionally, it was performed a phylogenetic tree with the sequences found and the reference strains used in GenBank NCBI. The phylogenetic alignment grouped the query sequence with the ORF1AB segment of the virus polyprotein (it encodes replication genes). It was determined that the virus genome is grouped with those of the strains previously described in the USA and China (Figure 2).

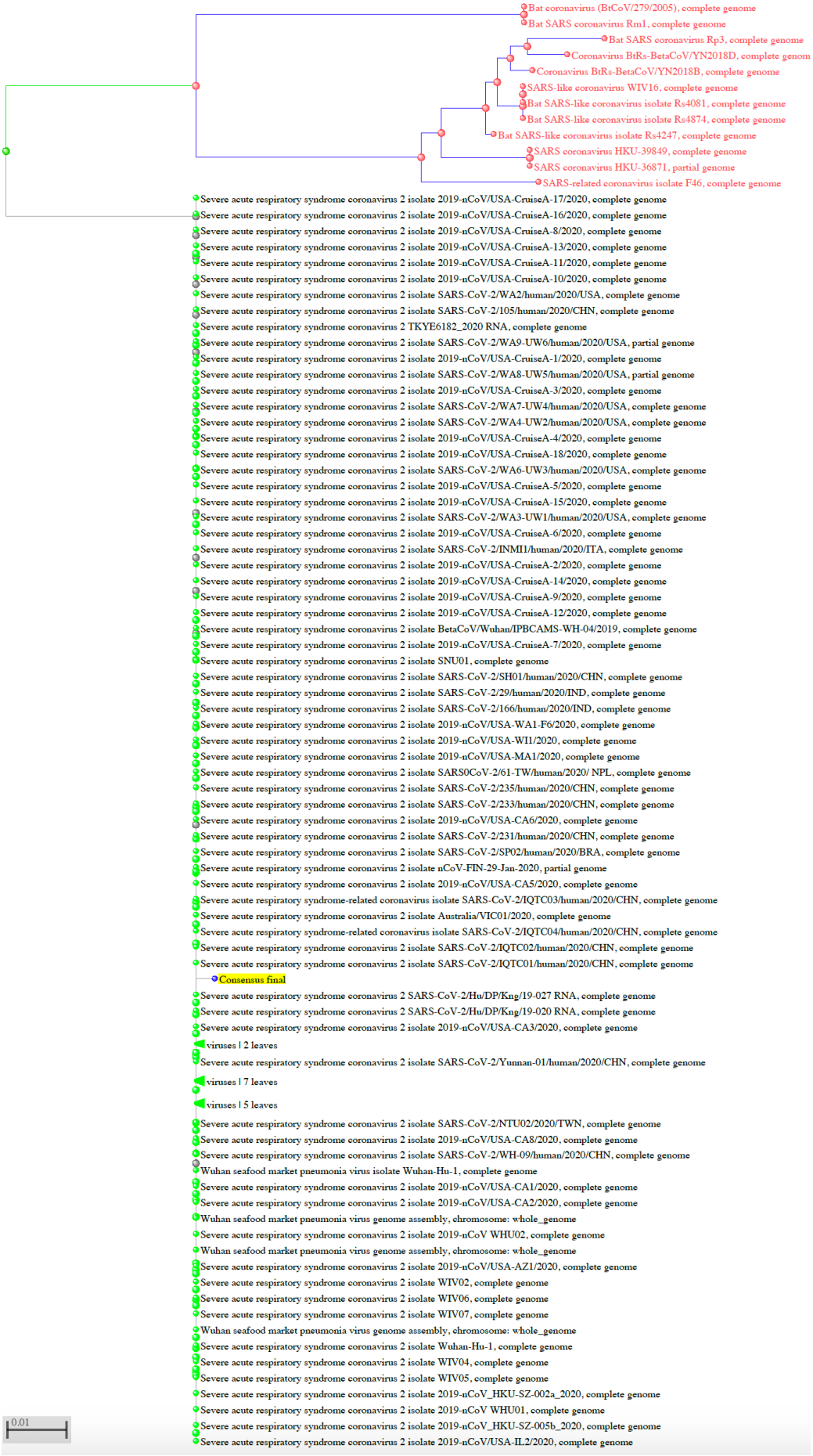
Bat coronavirus (BtCoV/279/2005), complete genome
Bat SARS coronavirus Rm1, complete genome
Bat SARS coronavirus Rp3, complete genome
Coronavirus BtRs-BetaCoV/YN2018D, complete genom
Coronavirus BtRs-BetaCoV/YN2018B, complete genome
SARS-like coronavirus WIV16, complete genome
Bat SARS-like coronavirus isolate Rs4081, complete genome
Bat SARS-like coronavirus isolate Rs4874, complete genome
Bat SARS-like coronavirus isolate Rs4247, complete genome
SARS coronavirus HKU-39849, complete genome
SARS coronavirus HKU-36871, partial genome
SARS-related coronavirus isolate F46, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-17/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-16/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-8/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-13/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-11/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-10/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA2/human/2020/USA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/105/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 TKYE6182_2020 RNA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA9-UW6/human/2020/USA, partial genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-1/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA8-UW5/human/2020/USA, partial genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-3/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA7-UW4/human/2020/USA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA4-UW2/human/2020/USA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-4/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-18/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA6-UW3/human/2020/USA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-5/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-15/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WA3-UW1/human/2020/USA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-6/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/INMI1/human/2020/ITA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-2/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-14/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-9/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-12/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate BetaCoV/Wuhan/IPBCAMS-WH-04/2019, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-7/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SNU01, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/SH01/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/29/human/2020/IND, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/166/human/2020/IND, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-WA1-F6/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-WI1/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-MA1/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS0CoV-2/61-TW/human/2020/ NPL, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/235/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/233/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA6/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/231/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/SP02/human/2020/BRA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate nCoV-FIN-29-Jan-2020, partial genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA5/2020, complete genome
Severe acute respiratory syndrome-related coronavirus isolate SARS-CoV-2/IQTC03/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate Australia/VIC01/2020, complete genome
Severe acute respiratory syndrome-related coronavirus isolate SARS-CoV-2/IQTC04/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/IQTC02/human/2020/CHN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/IQTC01/human/2020/CHN, complete genome
Consensus final
Severe acute respiratory syndrome coronavirus 2 SARS-CoV-2/Hu/DP/Kng/19-027 RNA, complete genome
Severe acute respiratory syndrome coronavirus 2 SARS-CoV-2/Hu/DP/Kng/19-020 RNA, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA3/2020, complete genome
viruses | 2 leaves
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/Yunnan-01/human/2020/CHN, complete genome
viruses | 7 leaves
viruses | 5 leaves
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/NTU02/2020/TWN, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA8/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/WH-09/human/2020/CHN, complete genome
Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA1/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA2/2020, complete genome
Wuhan seafood market pneumonia virus genome assembly, chromosome: whole_genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV WHU02, complete genome
Wuhan seafood market pneumonia virus genome assembly, chromosome: whole_genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-AZ1/2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate WIV02, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate WIV06, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate WIV07, complete genome
Wuhan seafood market pneumonia virus genome assembly, chromosome: whole_genome
Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate WIV04, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate WIV05, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-002a_2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV WHU01, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-005b_2020, complete genome
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-IL2/2020, complete genome
0.01

Figure 2. Maximum likelihood phylogenetic tree using the closest 100 BLASTx results with the HEE patient sequence (highlighted in yellow).

Conclusion: It was possible to sequence the orf1ab segment of the SARS-CoV-2 virus genome, which is grouped with strains from the USA and China using a metagenomic approach. The sequence didn´t clustered with the European strains (which could be caused by the sequence region alignment). In order to identify a precise genotype, the entire virus genome must be sequenced (we are currently performing the sequencing using the ARTIC project database).

Additional information: bacterial sequences related to the patient's respiratory microbiota were identified by metagenomics (Figure 3).

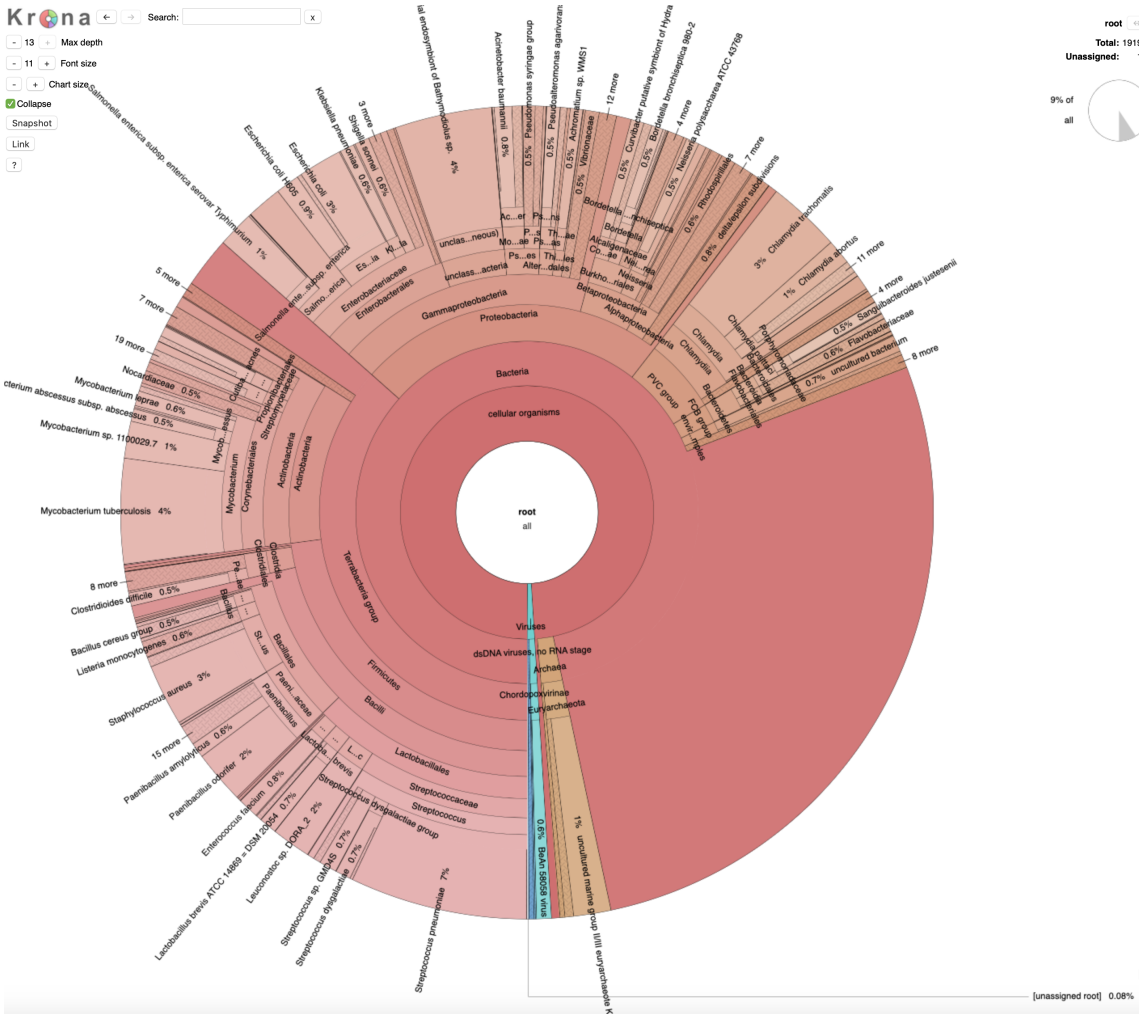

Figure 3. Krona chart summarizing the percentage of sequences assigned to bacteria and eukaryotes from the patient's respiratory microbiota.