Fatal Sabiá virus infection in a yellow fever suspected case, São Paulo, 2019

28 January 2020

Disclaimer
We are currently preparing a manuscript with the data reported here. We ask that you communicate with us if you wish to publish results that use this data in a journal. If you have any other questions relating to data availability, please also contact us directly.

Summary

- We have detected Sabiá virus (SABV) genome fragments from two patients suspected for yellow fever virus infection with hemorrhagic fever: Patient 19, a previously undetected case, and Patient 17, a recently reported case. SABV was identified in the patients using an untargeted metagenomic approach;

- Genetic analysis of recovered genome sequences shows that SABV strains are genetically diverse (≥10% nucleotide distance) at the nucleoprotein gene compared to each other, and compared to the reference virus sequence from 1990;

- The broad and previously unreported genetic diversity in SABV species revealed by this report poses challenges for existing molecular diagnostic tools and indicates the need to re-evaluate retrospective and prospective cases of hemorrhagic fever with negative yellow fever;

- In-house PCR and nested PCR detection protocols were designed to test SABV strains and all contacts from the most recent patient presenting with SABV; all contacts tested negative.

Laboratory investigation using metagenomic sequencing

Due to a recent case of arenavirus reported in Sao Paulo (Patient 17), the Institute of Tropical Medicine - Faculty of Medicine of the University of São Paulo (IMT-FMUSP) conducted a detailed investigation on yellow fever negative samples initially included in the SOFFA study. Of the 68 patients included in the study, a total of 8 patients were PCR negative for YFV and were analysed using an unbiased metagenomic approach. Patients' samples were collected in EDTA tubes. Plasma was separated in two aliquots and centrifuged at 10,000 g for 5 minutes. RNA was extracted from 200 uL of the sample supernatant using Qiagen RNA mini kits in a P3 Laboratory at IMT-FMUSP.

Sequencing was conducted using the MinION platform from Oxford Nanopore Technologies (ONT) at the IMT-FMUSP as described in CADDE Protocols with small modifications, specifically the metagenomic amplicon fragments were prepared for sequencing using the genomic sequencing kit LSK109 without barcoding. Each sample was subjected to two full technical replicates (RNA extraction, cDNA synthesis, PCR, library generation) before running on separate, fresh ONT flowcells (one per replicate) to help rule out the possibility of between-sample or barcode contamination.

Two patients yielded reads related to arenavirus: Patients 17 and 19. SABV was detected in all four flowcells. The genome of SABV virus consists of two negative-stranded RNA molecules, L and S, of about 7.5 kb and 3.5 kb in length, respectively. Thus, reads were mapped against the SABV

reference genome (L and S segments, Accession numbers NC_006317.1 and NC_006313.1) with minimap2 with parameters '-a' (SAM output) and '-x map-ont' (nanopore mapping mode), '-w1' and '-k11' (increased sensitivity). Sequencing and alignment statistics for each sample are summarized in **Table 1.**

Table 1. Sequencing coverage statistics for the two SABV patients analysed in this report.

Strain	Genome segment	Sequence coverage >1X (%)	Sequence coverage ≥ 5X (%)	Mean coverage depth (X)
Patient 17	S	100	100	407
	L	91.6	80.1	117
Patient 19	S	74.24	57.7	7.48
	L	34	14.2	2.57

Genetic diversity of SABV challenges commonly used diagnostic tools

The high within-species nucleotide diversity of Sabia virus reported here may challenge commonly used lab diagnostic tools to detect SABV strains circulating in Brazil. To investigate whether PCR primer nucleotide sequences from existing assays would detect the new sequences from the 2019 cases in Sao Paulo state, we summarized the matches and mismatches between SABV strains (reference NC_006317 and new sequences) and existing PCR primer sequences from Bowen et al. that is typically used in public health laboratories in Brazil (**Figure 2**).

Alignment of the primers from this assay indicates that the new Patient 17 strain is more divergent to the 1696R primer than the SABV reference sequence genome (**Figure 2**). The presence of two additional mismatch sites to this primer raises the possibility that currently used primers may not be able amplify newly identified SABV strains. Unfortunately, we currently lack the genomic sequence of Patient 19 at these sites (please note that we are currently working to generate more a more complete genome sequence for Patient 19).

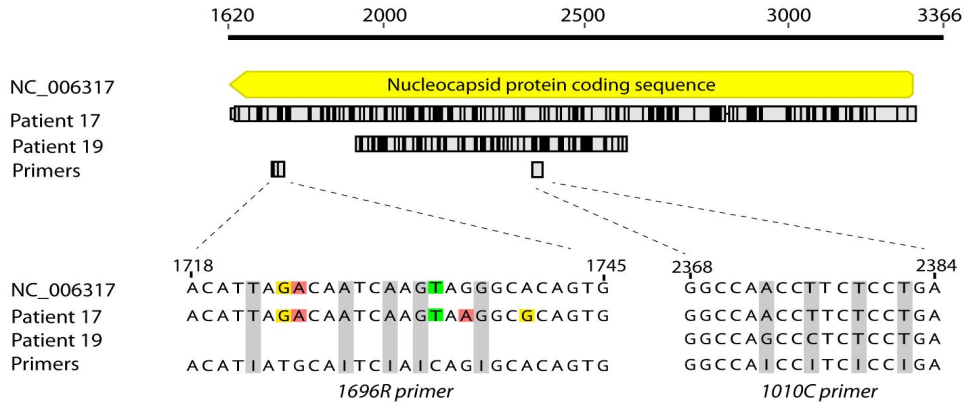

Figure 2. Comparison of matches and mismatches between SABV and existing diagnostic tools. Genomic locations are given relative to the start of the S segment of the SABV reference genome (NC_006317). Upper bars show the available sequence of the new strains relative to the SABV reference strain and location of primers, with variants colored in black. Lower part shows the primer and SABV sequences at primer binding sites. Colors here indicate mismatches to the primers, and pale grey bars indicate inosine sites.

PCR primers to detect circulating SABV strains

SABV is highly infectious and lethal and may spread through aerosols. Person-to-person transmission can occur with direct contact with the blood or other body fluids of infected individuals. Airborne transmission of arenaviruses has been reported and contact with infected medical material can also be associated with transmission. In an attempt to design an effective PCR approach to investigate virus infection and possible contamination to close contacts of Patient 17 and Patient 19, we next designed and evaluated the performance of six new PCR primers to detect the virus strains recovered here. A list of the primer sequences and additional details is shown in **Table 2**.

Table 2. Description of the primer sets used for Conventional PCR validation. Tm = primer melting temperature.

Primer sets	Name	Sequence 5' - 3'	N° of bases	Amplicon Length	Tm
1	Pos_29F	GTCACGCTTAAATCTTTGATTGC	23	350 bp	55°C
	Pos_381R	ACAGACACCTCAAGACACCA	20		
2	Pos_1642_2F	AGTGTGAAGAGCAATGCATCAG	22	200 bp	52°C
	Pos_1842R	AATCTGGTCAGATTTTGGTCACC	23		
3	All_3F	TCACGCTTAAATCTTTGATTGCG	23	121 bp	65°C
	All_3R	GCAGCAATYAAGCTCACTGC	20		
4	Pos_2F	CGCACCGGGGATCCTAGGC	19	380 bp	55°C
	Pos_381R	ACAGACACCTCAAGACACCA	20		
5	All_2F	GCAGTGAGCTTRATTGCTGC	20	723 bp	60°C
	All_2R	YACTGCATCAGTTATTGTCC	20		
6	Pos_171F	AACCTGTGGAAGAGTGGCCT	20	200 bp	52°C
	Pos_381R	ACAGACACCTCAAGACACCA	20		

All primer sets were ordered on Jan 21, 2020. Upon arrival on Jan 23, 2020, primers were pre-diluted in 1X TE buffer (Sigma) to 100 µM. The extracted RNA from Patient 17, Patient 19 and an additional negative control was converted into cDNA using Random Hexamers (Invitrogen) and ProtoScript II First Strand cDNA Synthesis Kit (New England BioLabs) following instructions. The PCR reactions were prepared according to the conditions described in **Tables 3 and 4**, changing the Tm according to the set of primers used.

Table 3. PCR performing reagents and conditions for the new SABV PCR detection protocols.

Component	**Volume in 25 uL reaction**
5x Q5 reaction buffer	5 uL
10 mM dNTPs	0.5 uL
Q5 DNA polymerase	0.25 uL
Nuclease-free water	Up to 25 uL
Primers 10 uL	1.25 uL forward primer 1.25 uL reverse primer
cDNA	2.5 uL

Table 4. PCR cycling conditions for the new SABV PCR detection protocols.

Temperature	**Time**	**No. cycles**
98 ^{0}C	1 min	1x
98 ^{0}C	30 s	35 x
Variable	30 s	
72 ^{0}C	1 min	
72 ^{0}C	2 min	1x
4 ^{0}C	Hold	

We next prepared a gel using the E-gel EX Agarose 2% (Invitrogen). To load the PCR products, 5 µl of the products were mixed with 15 µl of NFW and 20 µl of E-gel Sizing DNA ladder (Invitrogen) was used in the first band. The gel was placed into the E-gel equipment and the run was performed until the bands were distinguishable by transillumination. The specific bands of the correct size were observed according to expected amplicon lengths to all the reactions, except for the set 4 (Patient 19) in which no bands were detected (**not shown**). No bands were detected in the negative controls, indicating that there was no contamination.

While 5 of the 6 sets of primers tested here were able to capture SABV in Patient 17 (the most recent case), we were not able to detect SABV in Patient 19 (oldest cases from early Dec 2019). [Noteworthy, to date no known contact of Patient 19 developed symptoms.] This is probably due to the low viremia of SABV in Patient 19 as suggested by the low fraction of SABV reads (0.12% in Patient 17 versus 0.002% for Patient 19) (**Table 2**) and the high within-species diversity of arenaviruses.

We then selected primer sets 1 and 3 and tested all patients who had direct contact with Patient 17 or with his blood using primer sets 1 and 3 (**Table 2**) at the Hospital das Clínicas and at IMT-FMUSP. Reassuringly, all close contacts of Patient 17 (and two negative controls) tested negative for SABV while the positive control tested positive (**not shown**).

Nested PCR to detect SABV in samples with low viremia

We next designed two nested PCR primer sets in an attempt to detect the virus in Patient 19 (oldest cases from early Dec 2019) and the virus in the Patient 17 (newest case) in the same reaction. A nested PCR approach is a more sensitive reaction because it employs a second stage of amplification with a pair of internal primers (inner) to those used in the first stage (outer). A list of the primer sequences is shown in **Table 5.**

Table 5. Description of the primer sets used for nested PCR validation. Tm = primer melting temperature.

Primer sets	Name	Sequence 5' - 3'	N° of bases	Amplicon Length	Tm
1	Pos_29F_outer	GTCACGCTTAAATCTTTGATTGC	23	350 bp	55°C
	Pos_381R_outer	ACAGACACCTCAAGACACCA	20		
	Pos_171F_inner	AACCTGTGGAAGAGTGGCCT	20	200 bp	52°C
	Pos_381R_inner	ACAGACACCTCAAGACACCA	20		
2	S_outer_1_f	TCAGTGCAGGGACAGATCCA	20	494 bp	69°C
	S_outer_1_r	TCCCTGAGAAGAGGGCTCAG	20		
	S_inner_1_f	TACAACCCCTGGAGACCTCA	20	112 bp	68°C
	S_inner_1_r	TCAGGAGGTGTGTACCTGGG	20		

For this reaction, the extracted RNA from Patient 19 and an additional negative control was converted into cDNA as described above. Conventional PCR approach and the PCR reactions were prepared according to the conditions described in **Tables 6 and 7**, changing the Tm according to the set of primers used.

Table 6. Nested PCR performing reagents and conditions for the new detection protocols. Note that the reaction is done twice, once for the outer PCR reaction, and then for the inner PCR reaction.

Component	**Volume in 25 uL reaction**
5x Q5 reaction buffer	5 uL
10 mM dNTPs	0.5 uL
Q5 DNA polymerase	0.25 uL
Nuclease-free water	Up to 25 uL
Primers 10 uL	1 uL forward primer and 1 uL reverse primer
cDNA	2.5 uL

Table 7. PCR cycling conditions for the SABV Nested PCR detection protocol. Note that the reaction is done twice, once for the outer PCR reaction, and then for the inner PCR reaction.

Temperature	**Time**	**No. cycles**
98 ^{0}C	1 min	1x
98 ^{0}C	30 s	35 x
Variable	30 s	
72 ^{0}C	1 min	
72 ^{0}C	2 min	1x
4 ^{0}C	Hold	

Reassuringly, our nested PCR allowed the amplification of the sample from Patient 19 even with a low viremia. A new gel was prepared as before, and the specific bands of the correct size were observed according to the Table 8 (Amplicon Length) to the reactions (not shown).

Conclusions

Our metagenomic approach was able to rapidly identify the third naturally occurring case of Sabiá virus in a patient with yellow fever-like symptoms. We found that the currently used PCR tools may not be able to detect circulating SABV strains due to the virus' genetic diversity in primer binding sites. Using our metagenomic data, we then designed a PCR scheme to test Patient's 17 close contacts for SABV. All contacts tested negative. We also designed a more sensitive nested PCR protocol that successfully detected SABV in a low virus titer sample from Patient 19. The framework described here, from metagenomic testing to virus sequencing, development and validation of standard and nested PCR laboratory diagnostic tools for divergent strains and variable viremia, was conducted in a total of 10 days.