Michael Worobey **@MichaelWorobey** *Sun Aug 08 21:33:16 +0000 2021*

SARS-CoV-2/COVID-19 in Italy in September 2019: the most important finding yet on the origin of the pandemic*.

(*or an error with big consequences.)

A thread. 1/24

https://t.co/ATG2ifRbB2

The study, led by Dr. Elisabetta Tanzi, also includes heavy-hitters of molecular evolution **@sergeilkp** and Sudhir Kumar. I greatly admire both but respectfully disagree with their conclusions here and feel it is important to explain why. 2/

Dr. Tanzi led an earlier study claiming to find evidence of SARS-CoV-2 in a boy in Northern Italy who presented with measles symptoms in Nov 2019. 3/

https://t.co/K2QV86e5XH

Although that paper claims that her lab was "designated SARS-CoV-2 RNA-free", my correspondence with her indicates that that was not the case. A positive control from a local patient was used, presumably to develop the in-house nested PCR approach used. 4/ **https://t.co/TZl36jJ0M8**

Dear Mike,
thank you very much for your interest in our study.
Here are the answers to your questions (see below). Have a good day
Elisabetta

Prof.ssa Elisabetta Tanzi
Dipartimento di Scienze Biomediche per la Salute
Università degli Studi di Milano
Via Pascal 36, 20133 Milano
Tel 0250315139 – Fax 0250315120
e-mail: elisabetta.tanzi@unimi.it

Il 24/01/2021 19:42, Worobey, Michael - (worobey) ha scritto:

> Dear Dr. Tanzi,
>
> I hope you are well!
>
> I was hoping you might be able to clarify a few points raised in your recent EID paper, "Evidence of SARS-CoV-2 RNA in an Oropharyngeal Swab Specimen, Milan, Italy, Early December 2019".
>
> You describe your lab as "designated free of SARS-CoV-2 RNA". Yet you also mention the use of positive controls, and the development of an in-house hemi-nested PCR assay.

The positive control was used for the first time in our lab for this study. No SARS-positive samples (except the one detected in the study) were ever tested in our laboratory.

> 1. What did you use as your positive control?

The positive control was RNA isolated from a positive patient given to us from a local hospital.

> 2. I assume that when developing your PCR assay you checked that it worked by successfully amplifying your positive control. Is that correct?

The developed PCR amplified the positive control as expected.

> 3. You mention that, perhaps due to thawing, that sequencing of longer genomic regions was prevented. Which other genomic regions did you try unsuccessfully to amplify and sequence?

Unfortunately, the material left from this patient was very few and the viral load was so low that the PCR became positive only after the second amplification round (nested). The amplification of longer regions was therefore not possible.

> 4. Have you tried to sequence the complete genome of the virus using a tiling approach like the ARTIC Protocol of the Swift Biosciences kit? These approaches use multiplex PCR of fragments that are shorter than the one you report amplifying successfully, thus may work to amplify other genomic regions.

We have considered applying such protocols but, unfortunately, the few available microliters were used to detect the virus and there is nothing left for further tests.

> 5. Would you be willing to share an aliquot of the specimen you reported to be positive wit my lab so that we could try to replicate your findings and attempt complete genome sequencing the the virus?

As explained above, unfortunately, there is no RNA left.

Nested PCR is super-sensitive. It involves a first-round amplification of a PCR-product, then a second-round amplification of a subregion of that amplified PCR product. 5/

It is also *highly* prone to false positives: massive amounts of first-round product can contaminate later PCR reactions unless strict protocols are in place.

Especially true when a positive control is amplified before or at the same time you screen an "unknown" sample. 6/

Let's turn to the current study.

Again, an in-house nested PCR assay was used - to detect <500-nucleotide regions coding for (1) NsP3, (2) RdRp, (3)

Spike, (4) another Spike region.

Although not mentioned, I assume a positive control was used the validate the assay. 7/

The authors looked at hospitalized patients in Italy with measles-like rashes who were measles and rubella Dx-negative.

They report SARS-CoV-2 RNA in 11/44 cases from the "pre-pandemic" period, Aug 2019-Feb 2020. The earliest in Sept 12, 2019, plus 6 in Oct, 1 in Nov. 8/

(They also report SARS2 RNA in 2/12 such cases from Mar 2020-Mar 2021 who had also tested negative for SARS2.)

Crucially, in each "pre-pandemic" case that yielded SARS2 RNA spanning the three mutations characterizing the spring 2020 Italian outbreak, they were present. 9/

This is the famous B.1 lineage, with D614G, which dominated the European outbreak then spread worldwide. These key mutations are:

C3037T

C14408T

A23403G/D614G 10/

In other words, instead of some ancestral genome, as might be expected for samples collected months before the earliest ones in Dec 2019 in Wuhan, they have highly derived genomes typical of what was circulating in Northern Italy in spring 2020. 11/

I have published a study, with Joel Wertheim, **@suchard_group**, **@LemeyLab** **@jepekar** and others that indicates that the B.1 variant jumped from China to Italy on or around Jan 28 2020. 12/

https://t.co/Nvyoq3lOrT

This helped lay to rest the idea that the Italian outbreak descended from a Jan-Feb 2020 one in Germany, which was extinguished through (guess what?) testing, tracing, quarantine, isolation.

The German lineage B virus was just *one* mutation different than the Italian one. 13/

It lacked C14408T.

We suspected that a B.1 progenitor with this mutation existed in China.

Sure enough, I *just* noticed that one does: hCoV-19/Zhejiang/HZ103/2020 EPI_ISL_422425.

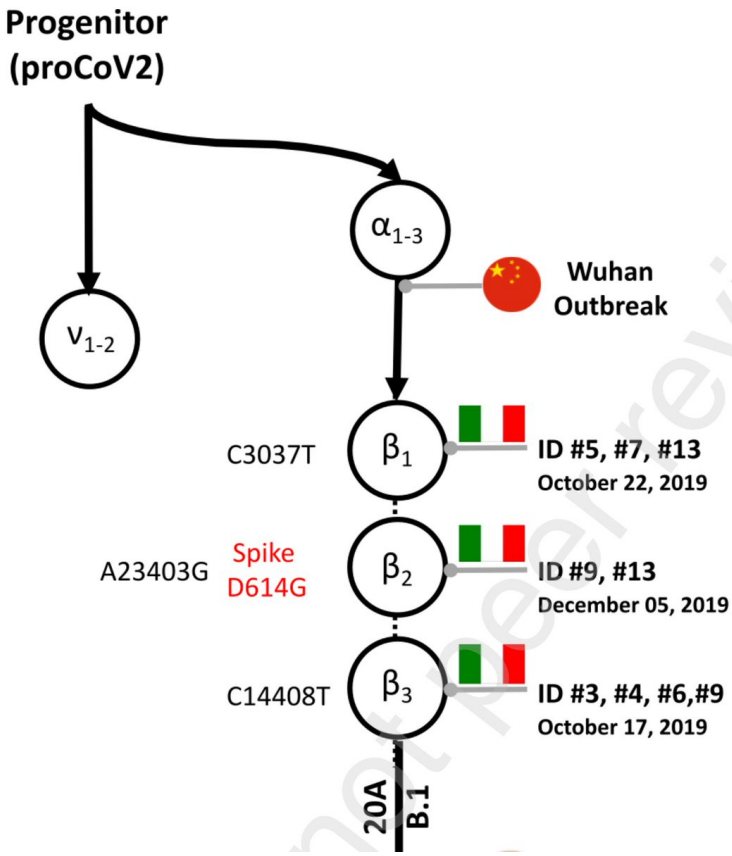
Indeed, B lineage viruses with the full progression to B.1 were present in PRC by Jan 2020... 14/

Just C3037T;

C3037T + A23403G/D614G; and

C3037T + A23403G/D614G + C14408T.

That makes Fig. 2 from Tanzi et al quite nonsensical. The figure portrays a stepwise evolution of the B.1 lineage in Italy, with B.1. present by Oct 2019. 15/ **https://t.co/BiDEi2HftM**



This would necessitate dispersal of each intermediate, plus B.1 itself, *back* to China from Italy. Or an independent progression of the identical 3 mutations in each country.

Both options strike me as *much* less likely than a nested-PCR false positive. 16/

Another important part of the study: none of the 11 pre-pandemic SARS2 RNA+ patients had detectable RNA using the CDC Real-Time RT-PCR assay.

What are the chances of hitting 11 true positives without a single one yielding a qRT-PCR positive result? 17/

The authors may be aware of the weirdness of this result, because they posit these early variants may bind the Spike protein more weakly than later variants.

But they are derived from viruses that were circulating in China that already possessed high affinity for hACE2. 18/

Sorry -- should read "these early variants' Spike protein may bind hACE2 more weakly than later variants."