Twitter Thread by Theo Sanderson

Theo Sanderson
@theosanderson

1/ Today has brought a lot of news about 'variants', as we discover community transmission of B.1.351 (SA variant) in the UK.

But some of the people who study these genomes have increasingly started focusing on a different concept: 'constellations', which may be more useful.

B.1.1.7 / 20I/501Y.V1 (first found in UK)	B.1.351 / 20H/501Y.V2 (first found in SA)	20J/501Y.V3 (first found in Brazil)	20B/S.484K (first found in Brazil)
S: D 614 G	S: D 614 G	S: D 614 G	S: D 614 G
	S: E 484 K	S: E 484 K	S: E 484 K
S: N 501 Y	S: N 501 Y	S: N 501 Y	
	S: L 18 F	S: L 18 F	
	S: K 417 N	S: K 417 T	

2/ We're actually all using 'variant' wrong anyway – we're using it to refer to a 'lineage'. A lineage is a set of viruses descended from a common ancestor. All B.1.1.7 (UK variant) viruses descend from an infection that probably occurred in August or September.

3/ But what makes this lineage practically important is that it carries a set of mutations which make it more transmissible.

I.e. we don't really care about the ancestry in itself.

4/ For example, one of the mutations that we think is important to B.1.1.7's faster transmission is called N501Y.

5/ It turns out that a lot of the lineages identified as important recently, in Brazil, South Africa, and the UK, have mutations such as N501Y in common (figure adapted from https://t.co/lpJeW0GGnU)

B.1.1.7 / 20I/501Y.V1 (first found in UK)	B.1.351 / 20H/501Y.V2 (first found in SA)	20J/501Y.V3 (first found in Brazil)	20B/S.484K (first found in Brazil)
S: D 614 G	S: D 614 G	S: D 614 G	S: D 614 G
	S: E 484 K	S: E 484 K	S: E 484 K
S: N 501 Y	S: N 501 Y	S: N 501 Y	
	S: L 18 F	S: L 18 F	
	S: K 417 N	S: K 417 T	

6/ You can see in the table that the lineage originating in South Africa, and those originating in Brazil, have another mutation, called E484K, in common.

7/ Today PHE's Technical Briefing 5 reveals that a small number of B.1.1.7 (UK lineage) genomes have recently been found which newly carry the E484K mutation, which had previously been seen in these South African and Brazilian lineages.

Detection of E484K mutation in B.1.1.7 VOC 202012/01

The COG-UK dataset (total sequences 214,159) was analysed on 26/01/2021. The spike protein mutation E484K (found in VOC 202012/02 B1.351 and VOC 202101/02 P1) has been detected in 11 B1.1.7 sequences. Preliminary information suggests more than one acquisition event.

8/ This shows why thinking about constellations is useful. A 'constellation' is just a set of shared mutations carried by viruses, irrespective of ancestry.

9/ Should we treat B.1.1.7+E484K cases as less concerning than viruses descending from those that originated in South Africa?

10/ We don't know the answer to that one way or another. We don't know to what extent E484K is the key mutation to driving the properties of a lineage, and we don't know how this mutation might interact with other mutations B.1.1.7 carries.

11/ But increasingly, it's going to be important to focus on the mutations that viruses carry, and understanding what effect these have, rather than simply focusing on their ancestry.

12/ From everything we have seen across the continents, E484K seems significant, and we should be very alert to its emergence in another lineage with high transmissibility. (There's every indication that the public health authorities are alert).

13/ Here's some more on constellations: https://t.co/dOzmEG5Vju

> There is a lot of discussion at the moment about naming SARS-CoV-2 variants and coming up with a standardised naming system. There was some discussion of this last week at the WHO @edwardcholmes @EvolveDotZoo @mvankerkhove @JeremyFarrar @firefoxx66
>
> — Andrew Rambaut \U0001f9a0\U0001f9ec\U0001f332\U0001f52e\U0001f926\u200d\u2642\ufe0f (@arambaut) January 15, 2021

(I regret saying the use of 'variant' was 'wrong' above, it's all very fuzzy!)