Genomic surveillance of SARS-CoV-2 in Belgium

Report of the National Reference Laboratory (UZ Leuven & KU Leuven)

Situation update – 6th of April 2021
(report 2021_21)

Executive summary

12.866 Belgian sequences of SARS-CoV-2 are publicly available on GISAID.
Since the 1st of January 2021, 9.527 unbiased positive samples were sequenced in the context of baseline surveillance.

For baseline surveillance samples collected during the last two weeks,
- B.1.1.7 (20/501Y.V1) represented 84,4%
- B.1.351 (20H/501Y.V2) represented 3,5%
- P.1 (20J/501Y.V3) represented 1,7%

The decrease in relative frequency of B.1.351 and P.1 are to be associated with the current surge of B.1.1.7 (third wave) and some delay in sequencing activity reported by participating laboratories. Therefore, these changes are to be therefore considered as temporary artefacts.

In this report, we propose a methodology for the definition, reporting and investigation of emerging variants.

Authors (National Reference Laboratory – UZ Leuven and KU Leuven):
Guy Baele, Lize Cuypers, Piet Maes, Els Keyaerts, Elke Wollants, Marc Van Ranst, Emmanuel André.

With the collaboration of the laboratories of UCL, ULB, UMons, UNamur, ULiège, UGent, UAntwerpen, Jessa ZH, AZ Delta, AZ Klina, IPG, AZ St Lucas Gent, OLV Aalst, Briant network, ZNA, AZ St Jan Brugge, and UZ Leuven/KU Leuven.

Previous reports can be downloaded using the following link:
https://www.uzleuven.be/nl/laboratoriumgeneeskunde/genomic-surveillance-sars-cov-2-belgium

Table of content

1. Baseline surveillance

Since the end of 2020, the list of variants of concern (VOCs) and variants under investigation has grown regularly, and we expect that this list will continue to increase as a consequence of both the upscaling of genomic surveillance around the world and the increased selective pressures exerted by the combination of partial herd immunity and stepwise vaccination rollout.

	Cases detected in Belgium 6/4/2021 (30/03/2021)	Epidemiological situation in Belgium (6/4/2021)	Regions with active circulation
B.1.1.7 (20/501Y.V1)	5002 (3909)	Dominant lineage	All regions
B.1.351 (20H/501Y.V2)	649 (495)	Emerging	Southern African region
P.1 (20J/501Y.V3)	212 (131)	Emerging	Latin America
B.1.427	1 (1)	Sporadic	Northern America
P.2	2 (2)	Sporadic	Latin America
B.1.526	0 (0)	Unreported	Northern and Latin America
B.1.525	21 (15)	Sporadic (increasing)	Western Africa
B.1.214.2	323 (254)	Emerging	Europe
A.2.5	0 (0)	Unreported	Central America

Table 1: Updated list of internationally recognized variants of concern (red) and variants of interest (orange) and number of sequenced strains in Belgium as reported in GISAID.

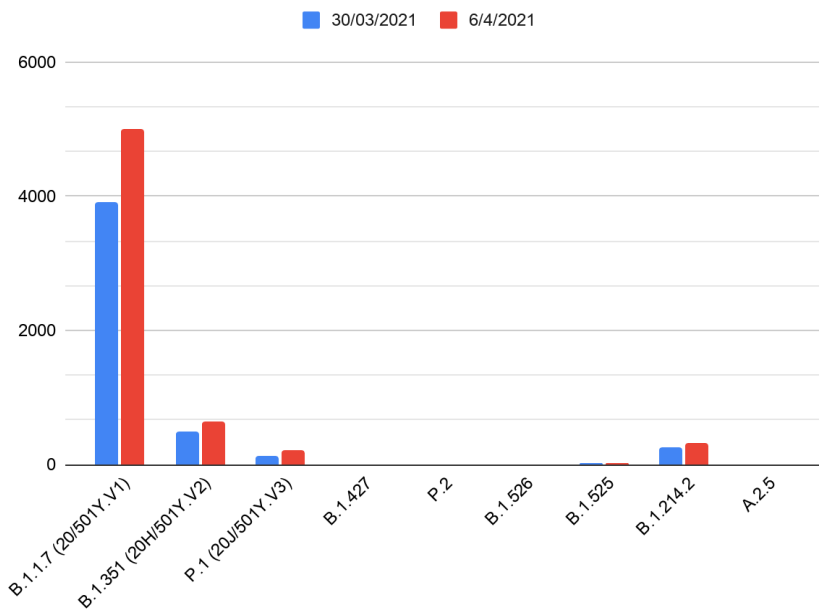


Figure 1: *Number and evolution of Belgian sequences available on GISAID for variants of concern and variants of interest.*

2. Monitoring of VOCs in Belgium

After a constant rise in proportion starting from January 2021, most new SARS-CoV-2 infections in Belgium are currently associated with a variant of concern (VOC), principally B.1.1.7 (501Y.V1). This phenomenon had not translated into a significant rise of cases until recently. We are currently observing an increase in the number of infections and hospitalisations, which can be directly related to the spread and dominance of 501Y.V1, a more transmissible and more virulent variant compared to historical circulating strains.

For baseline surveillance samples collected during the last two weeks,

- B.1.1.7 (20/501Y.V1) represented 84,4%
- B.1.351 (20H/501Y.V2) represented 3,5%
- P.1 (20J/501Y.V3) represented 1,7%

The decrease in relative frequency of B.1.351 and P.1 are to be associated with the current surge of B.1.1.7 (third wave), and are therefore to be considered as a temporary artefact.

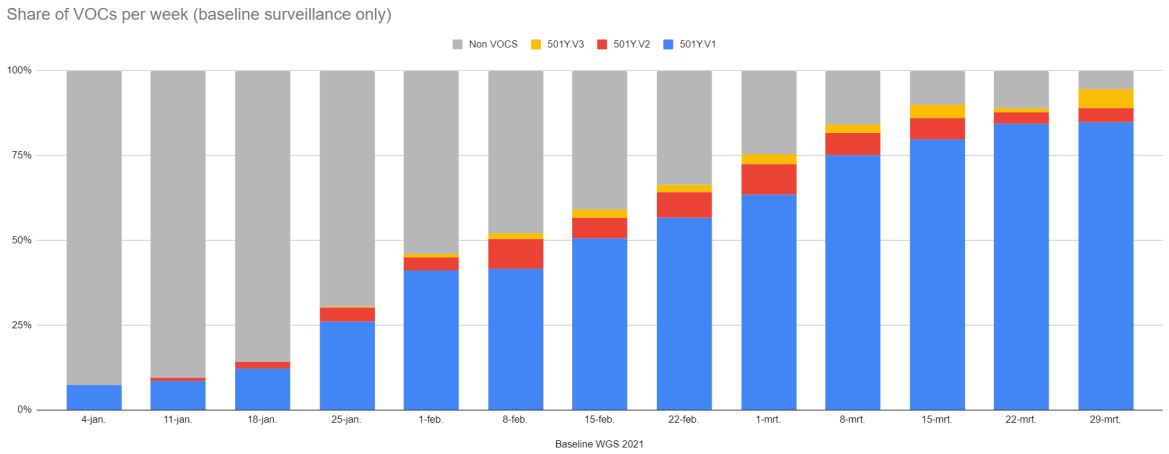


Figure 2: *Share of VOCs circulating in Belgium as measured through baseline WGS tests performed per sampling date since week 1 of 2021. Colour code: Non-VOCs (grey), 501Y.V1 (blue), 501Y.V2 (red) and 501Y.V3 (yellow).*

2.1. Update with regard to the circulation of P.1 (20J/501Y.V3) in Belgium

In comparison to most western European countries, Belgium (212) and Italy (391) reported a relatively high number of P.1 sequences on GISAID.

As illustrated in Figure 3, this lineage has been observed in all provinces since the start of the year, and most cases have been found in and around major cities.

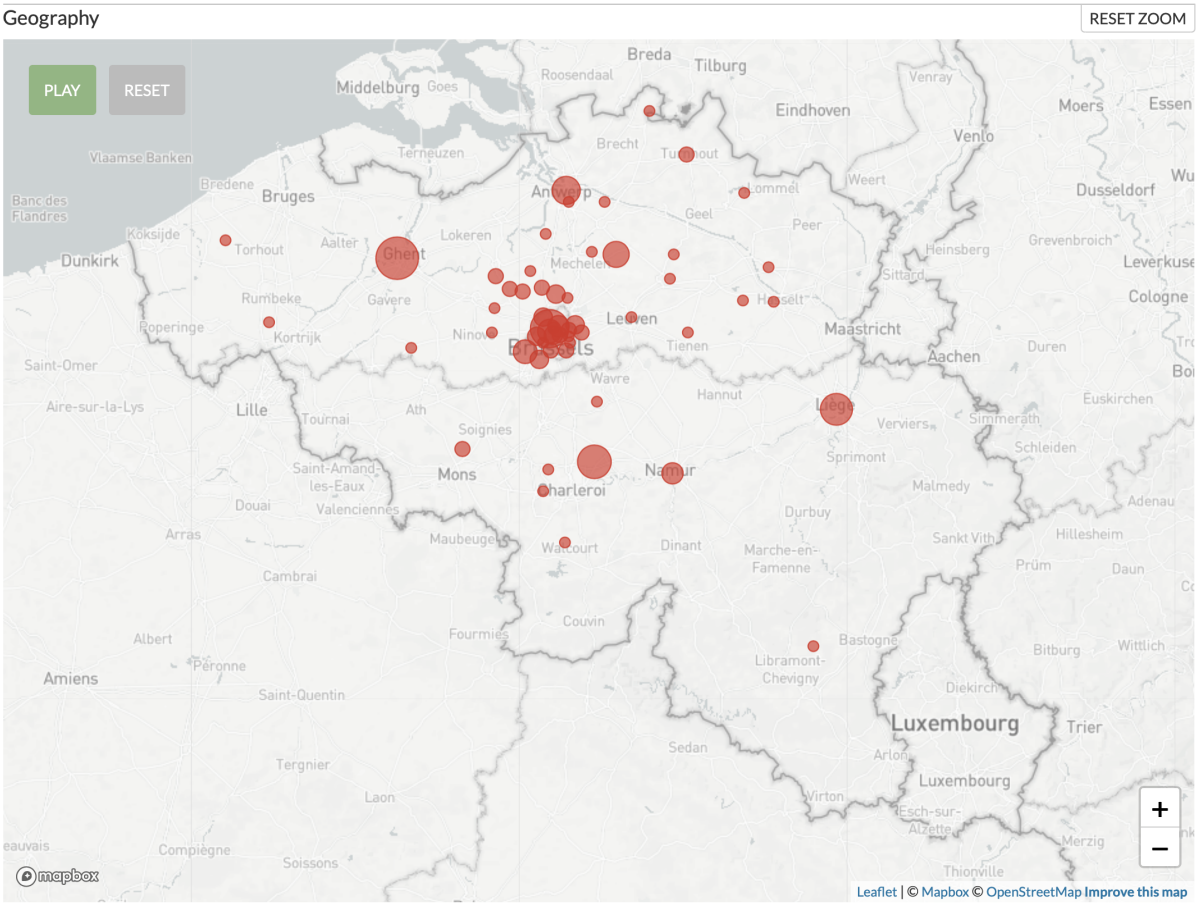


Figure 3: Dispersion of P.1 cases in Belgium since 1/1/2021

The presence of P.1 in Belgium is associated with several parallel introductions in early 2021, but most of the cases are from a large predominantly Belgian clade of cases, suggesting that the P.1 progression in the country is the direct consequence of one unique and initially undocumented introduction event (returning traveller untested or which did not respect quarantine). A second smaller, but yet persistent cluster is observed in the region of Antwerp, Ranst, Merchtem and Mechelen. A few isolated strains are still being analysed, and may be related to a successful outbreak containment or an artefact caused by sequences not meeting all quality standards.

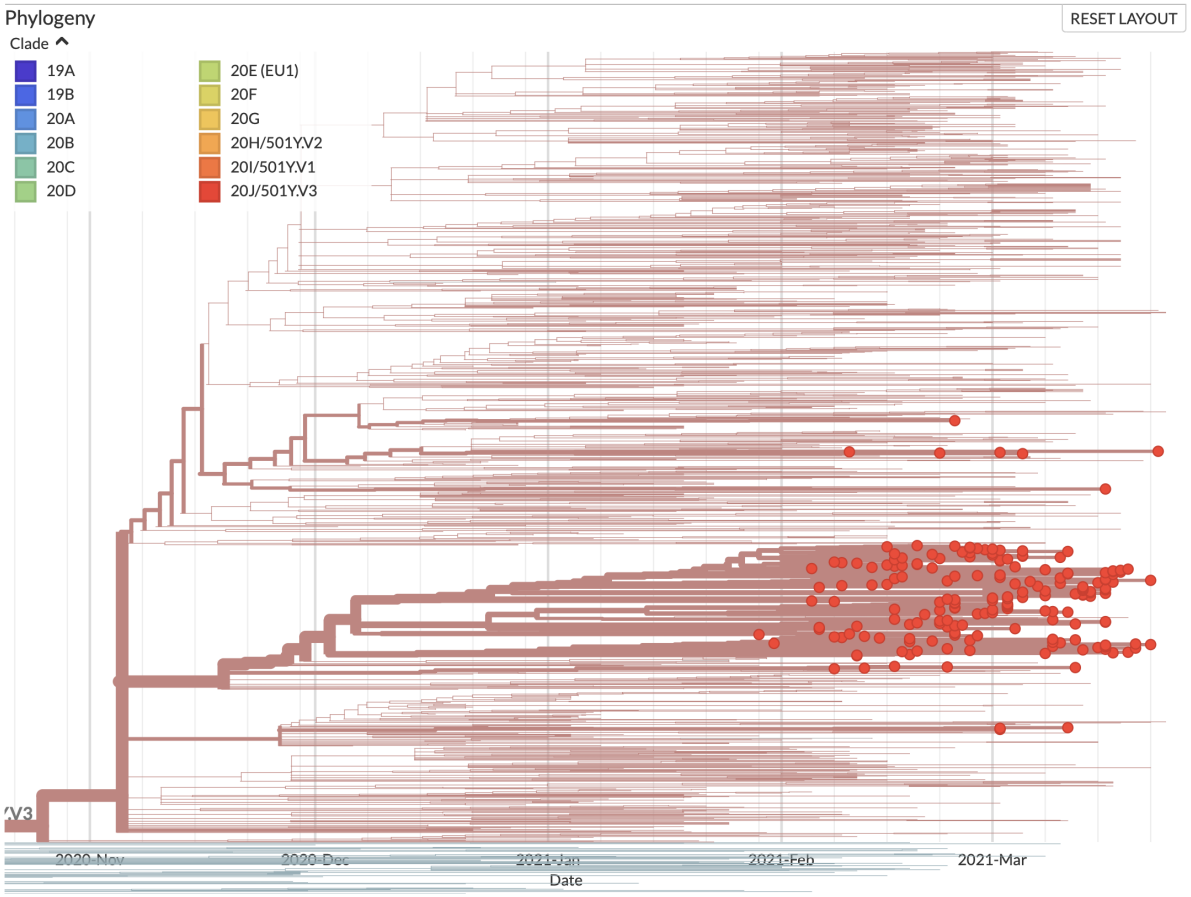


Figure 4: Clustering of P.1. sequences originating from Belgium suggest the presence of one large cluster, a second - but small (5) - cluster and a few isolated sequences.

3. Proposed consensus methodology for investigating emerging variants

3.1. Rationale

During the first year of the SARS-CoV-2 pandemic, an explosion of the number of cases in an immunologically naïve population led to a rapid divergent evolution of the virus, translated in the emergence of multiple viral lineages around the world. International travel associated with uncontrolled transmission has subsequently contributed to a rapid mixing of lineages and a relative competition favorable to the most infectious variants. These more infectious strains are characterized by the accumulation of shared mutations leading to better stability[1] or an optimized interaction between the virus and the humans cells it infects[2]. Importantly, these few optimisations have been reported to be associated with increased disease severity, in particular for the lineages B.1.1.7[3] and P.1[4].

Along the recent months, characterized by a continued pandemic and the progressive rollout of vaccination, the proportion of the world population presenting immunity after natural infection or vaccination is constantly increasing. Because this build-up of herd immunity is unequal and slow, it is also by definition incomplete. Until the moment when each region of the world will reach a vaccination rate of at least 50%[5], we will be in an in-between zone defined by partial herd immunity, a status which generates a very important selective pressure leading to the natural emergence and selection of viral mutations harbouring epitope modifications. This phenomenon is illustrated by the continuous emergence of variants harbouring a high number of diverse mutations, deletions and insertions in the spike protein of the virus. In a general context characterized by a societal trend towards relaxation measures and increased immunity, it is expected that epitope changes leading to relative immune escape mechanisms will become a key driver of epidemiologically successful viral lineages[6]. This phenomenon will be relatively independent of genetic evolution associated with transmissibility of disease severity patterns. In the future, it is therefore not excluded that less virulent strains become dominant and therefore could be considered as a complement to vaccination strategies, as their circulation will entertain herd immunity achieved through vaccination. Nevertheless, before such hypotheses could eventually be transposed into public health interventions, it is important that we develop at the international level a common approach towards characterizing emerging variants.

The paragraphs below represent a first attempt of consensus at the national level, before it would be presented at the international level. Anticipating on the continuous emergence of novel variants around the world, we hereunder provide guidance on how to improve our global reactivity with regard to this phenomenon and its intrinsic risks.

---

[1] Jun Zhang, Yongfei Cai, Tianshu Xiao, Jianming Lu, Hanqin Peng, Sarah M. Sterling, Richard M. Walsh Jr., Sophia Rits-Volloch, Haisun Zhu, Alec N. Woosley, Wei Yang, Piotr Sliz, Bing Chen. Structural impact on SARS-CoV-2 spike protein by D614G substitution. DOI: 10.1126/science.abf2303

[2] https://virological.org/t/mutations-arising-in-sars-cov-2-spike-on-sustained-human-to-human-transmission-and-human-to-animal-passage/578/2

[3] https://www.nature.com/articles/s41586-021-03426-1

[4] https://www.medrxiv.org/content/10.1101/2021.03.24.21254046v1

[5] https://www.nature.com/articles/s41577-021-00531-0

[6] https://www.nature.com/articles/s41586-021-03471-w

4.2. Genomic surveillance setup

The aim of the Belgian genomic surveillance initiative is to be able to detect the introduction and spread of emerging variants with a detection limit close to 1%. To achieve this goal, our surveillance is based on two main strategies: baseline surveillance and active surveillance programs.

Baseline surveillance in Belgium is based on the systematic analysis of 10 (up to 20%) of positive samples collected from a network of over 25 geographically dispersed clinical laboratories. These laboratories represent the different pillars of the national testing strategy, including community screening, active case finding, symptomatic cases and severe infections. Sample collection by sentinel laboratories is unbiased, although the positive results must have a sufficiently high viral load (>10^4 copies/ml). A high viral load allows to ensure the technical yield of whole genome sequencing and the temporal representativity, as low viral loads can be associated with residual signals observed several weeks after infection.

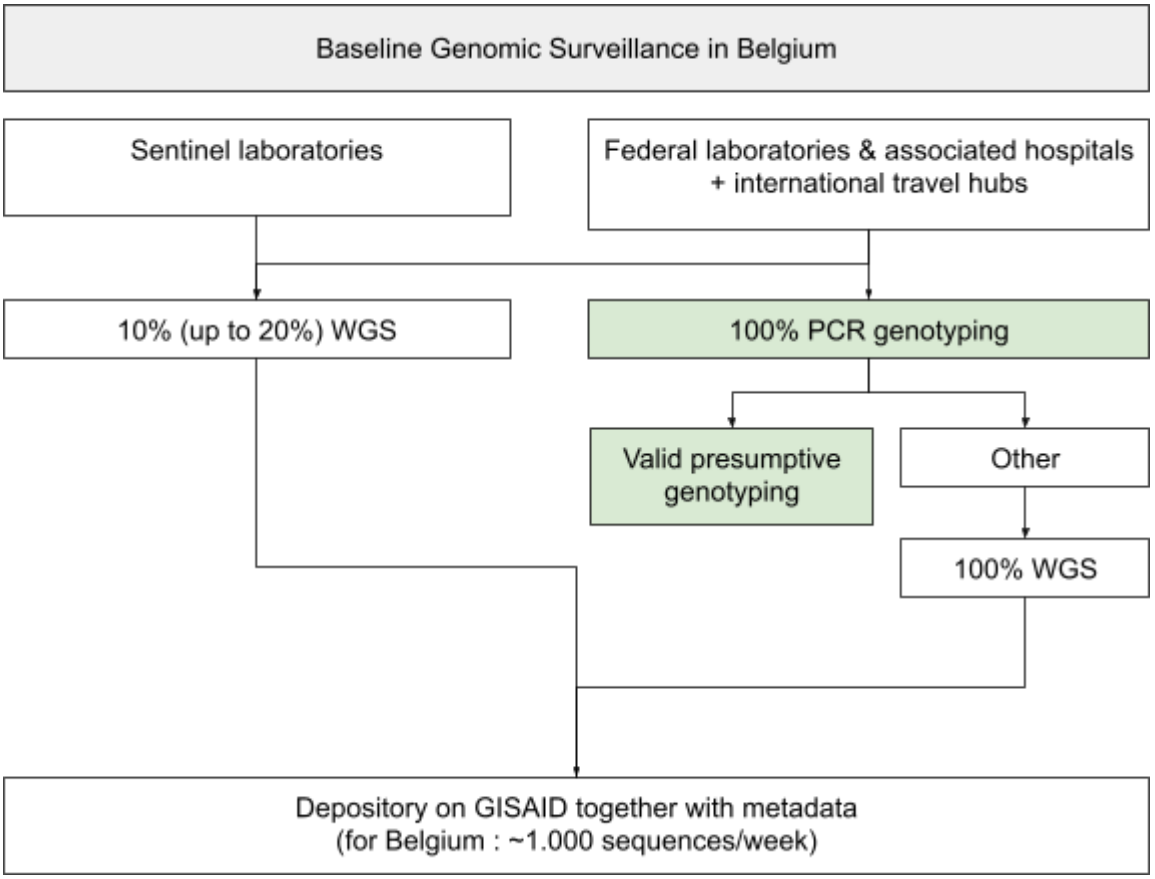


Figure 1: ***Baseline genomic surveillance in Belgium.***

In complement to whole genome sequencing, Belgium has recently introduced a complementary PCR approach aiming to detect key molecular markers of the dominant and emerging lineages currently detected in Belgium. As illustrated in Figure 1, this approach aims to increase the detection limit for lineages present at a rate of approximately 1%. This approach is progressively being implemented in the 8 federal platform laboratories and their associated hospitals. Reports of this PCR-based surveillance are to be included in the weekly reports of the National Reference Laboratory and are not uploaded on GISAID.

	S:H69 del		S:K417N	S:K417T	S:E484K	S:N501Y
B.1.1.7 (20/501Y.V1)	Yes		wt	wt	wt	Yes
B.1.351 (20H/501Y.V2)	wt		Yes	wt	Yes	Yes
P.1 (20J/501Y.V3)	wt		wt	Yes	Yes	Yes
B.1.525	Yes		wt	wt	Yes	wt

Table 2: ***Current design of the VOC PCR used in Belgium.*** *The present table is used for the presumptive genotyping of samples. All markers should be present in order to have a valid result. All other results should be investigated with whole genome sequencing. wt: wild type.*

In complement to baseline surveillance (unbiased sampling), Belgium has introduced a complementary active surveillance arm based on whole genome sequencing. The principle of this arm is to voluntarily introduce a bias in the selection of positive samples for a selected number of indications associated with an increased risk of emergence of lineages of concern or epitope switch. Under normal circumstances, this active surveillance arm should not represent more than 20% of all genomes deposited on GISAID, and should not create a significant bias in general trends with regard to persistent and dominant lineages. The advantage of this approach is that very early introductions of variants and/or emergence of new mutants will be documented and archived for further investigation.

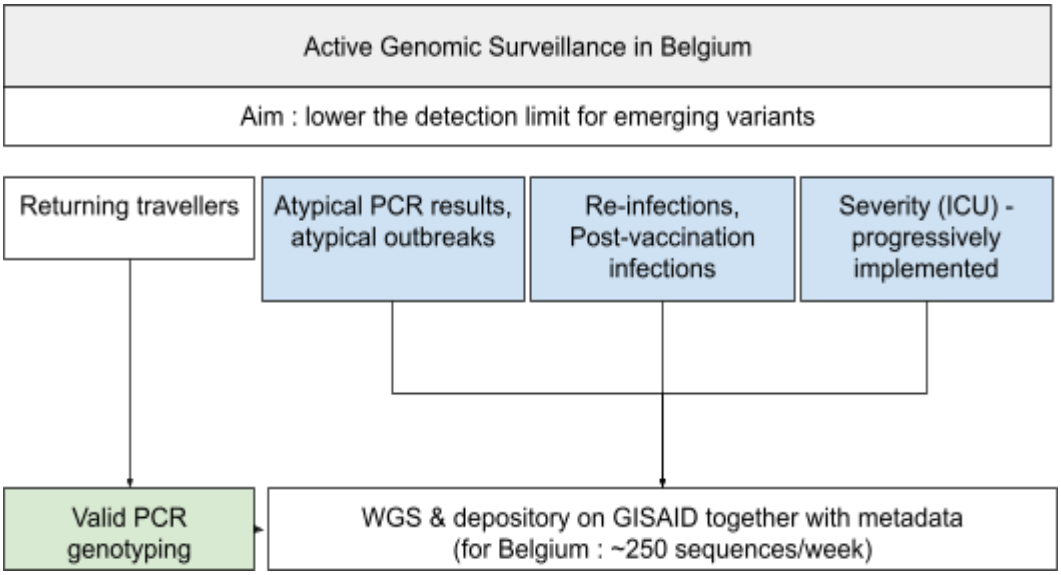


Figure 2: ***Active surveillance in Belgium.*** *The aim of active surveillance strategies is to lower the detection limit for emerging variants and variants harboring immune-escape mechanisms or higher severity patterns.*

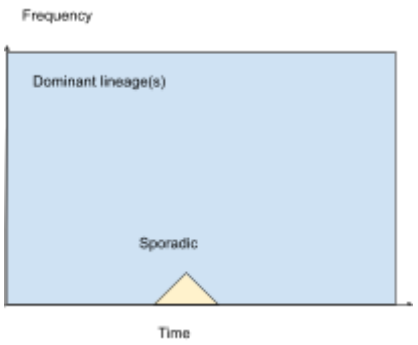
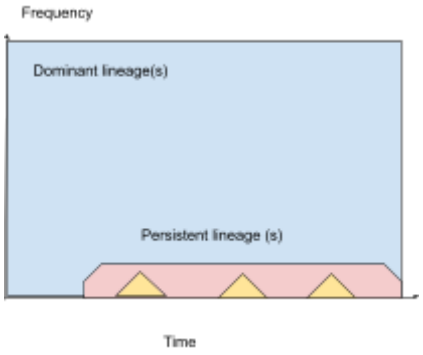
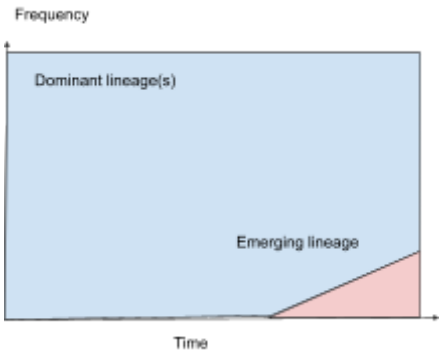
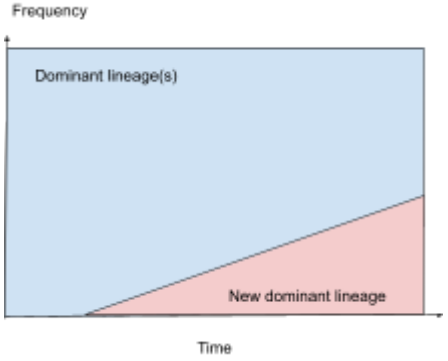
4.3. Definitions and risk evaluation associated with emerging variants

The level of risk associated with SARS-CoV-2 variants should be seen as the combination of their potential to disrupt epidemiological control and their epidemiological progression.
The table below highlights the level of risk taking into account these two factors.

	Dominant	Emerging	Persistent	Sporadic
Variant of concern with immune escape (eg. B.1.351 and P.1)				
Variant of concern without immune escape (eg. B.1.1.7)				
Variant of interest with potential immune escape (eg. B.1.525)				
Other variants				

Table 3: ***Estimation of the level of risk of variants*** *based on their intrinsic specificities (stable over the time) and their presence at a particular moment in time in a particular geographical entity. Green: no specific actions to be taken. Yellow: active monitoring and targeted interventions to be undertaken. Orange: active monitoring and reinforcement of control measures to be undertaken. Red: Additional general measures to be undertaken*

4.3. Proposed definitions for emerging variants

	Share of viral populations	Epidemiological definition
Level 1 Sporadic		No evidence of active community transmission Sporadic cases mainly associated with returning travellers (<1% at regional level)
Level 2 Persistent		Active transmission, outside the context of returning travellers. No sign of rapid increase in share of circulating viral populations (<1% at national level)
Level 3 Emerging		Active transmission, outside the context of returning travellers. Early signs of rapid increase in share of circulating viral populations (≥1% at national level or ≥10% at provincial/regional level)
Level 4: Dominant		Active transmission, outside the context of returning travellers. Confirmed rapid increase in share of circulating viral populations (≥5% at national level or ≥20% at provincial/regional level)

4.4. Proposed interventions for emerging variants

4.4.1. Level 1: Sporadic lineage

- Monitoring through usual genomic surveillance programs
- Upload of sequences and associated metadata on GISAID
- Regular literature review (transmissibility, immune escape, virulence)
- If lineage is not described: Pre-definition (genetic markers, phylogenetic analysis, epitope analysis)

4.4.2. Level 2: Persistent lineage

- Monitoring through usual genomic surveillance programs
- Upload of sequences and associated metadata on GISAID
- Regular literature review (transmissibility, immune escape, virulence)
- If lineage is not known: Submission of a new lineage (genetic markers, phylogenetic analysis, epitope analysis)

4.4.3. Level 3: Emerging lineage

- Monitoring through usual genomic surveillance programs
- Upload of sequences and associated metadata on GISAID
- If lineage is not known:
    - Submission of a new lineage
    - Phylogenetic analysis
    - Epitope analysis (3D modelling)
- Extensive literature review (international situation in any)
- Preliminary epidemiological assessment
    - Penetration
    - Transmissibility
    - Immune escape
    - Vaccine efficacy
- Preliminary communication to MoH, ECDC & WHO

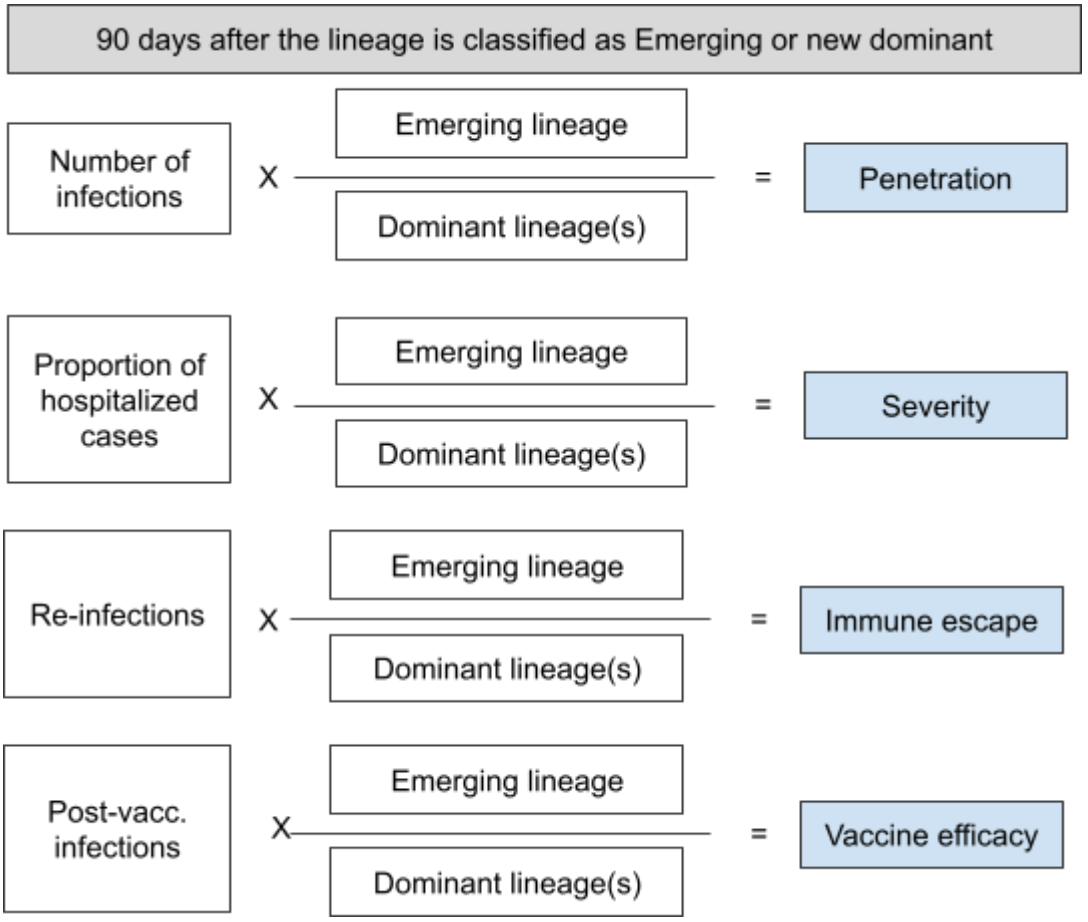
4.4.4. Level 4: New dominant lineage

- Monitoring through usual genomic surveillance programs
- Upload of sequences and associated metadata on GISAID
- Complementary analysis:
    - Updated Phylogenetic analysis
    - Antibody neutralisation assays
- Updated epidemiological assessment
    - Penetration
    - Transmissibility
    - Immune escape
    - Vaccine efficacy
- Communication to MoH, ECDC & WHO

4.5. Epidemiological assessment

An epidemiological assessment should be performed at the latest 3 months after the lineage has been classified as emerging. This assessment should thereafter be revised at the latest 3 months after the lineage has been classified as dominant. This assessment should be performed in collaboration with the National Reference Center and the National Public Health Institute (Sciensano), as the latter is the only organization to have access to individual databases on testing, travels, vaccination and hospitalisations.

The aim of this assessment is to evaluate the level of penetration, severity, immune escape and vaccine efficacy of an emerging lineage in comparison with the dominant lineages circulating in the country. This evaluation is considered as the baseline information for a formal risk assessment.



4.6. Submission of a new lineage

The current nomenclature system for SARS-CoV-2 lineages has been put in place by Rambaut *et al*. (2020) and is being used in the PANGOlin lineage assignment software that is linked to the GISAID platform. This system targets those lineages that contribute most to global transmission and genetic diversity, and is not intended to represent every evolutionary change in SARS-CoV-2 since these may eventually end up numbering in the thousands. Instead, the focus is on the genetic changes associated with **important epidemiological and biological events**. Hence, rather than naming every possible new lineage, classification should focus on those that have exhibited onward spread in the population, particularly those that have seeded an epidemic in a new location. Fortunately, because of the early sampling and genome sequencing of COVID-19 cases in China, especially in the Hubei province, it appears that the 'root sequence' of SARS-CoV-2 is known, making it a natural starting point (i.e., the reference sequence). Additionally, rather than maintaining a cumulative list of all lineages that have existed since the start of the pandemic, Rambaut et al. (2020) deemed it more prudent to mark lineages as 'active', 'unobserved' or 'inactive', as lineages that have not been seen for some time may re-emerge after a period of cryptic transmission in a region. A novel nomenclature system is currently being developed, in order to try to solve the confusion of using different naming systems and to improve communication with a non-expert audience (Callaway, 2021).

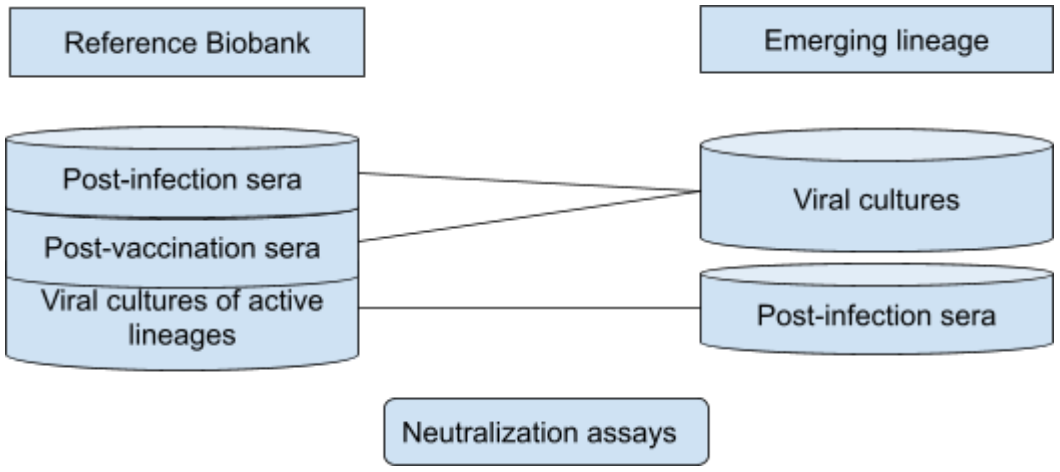
Rambaut et al. (2020) proposed lineage designations to meet a set of conditions: each descendant lineage should show phylogenetic evidence of emergence from an ancestral lineage into another geographically distinct population, implying substantial onward transmission in that population. In the case of a rapidly expanding global lineage, the recipient population may comprise multiple countries. In the case of large and populous countries, it may represent a new region or province. To show phylogenetic evidence, a new lineage must meet all of the following criteria: (a) it exhibits **one or more shared nucleotide differences from the ancestral lineage**; (b) it comprises at least five genomes with >95% of the genome sequenced; (c) genomes within the lineage exhibit at least one shared nucleotide change among them, which helps to focus attention only on lineages with evidence of ongoing transmission; and (d) a bootstrap (or posterior probability) value >70% for the lineage-defining node.

The three criteria (a, b and c) listed above can be assessed by evaluating the evolutionary relationships between SARS-CoV-2 genomes through phylogenetic and phylodynamic inference. Analysis of phylogenetic trees allows for conclusions to be drawn about epidemic and pandemic viruses that are important for public health (Martin et al., 2021). Phylodynamic analysis can provide insight into how a virus spreads spatially and temporally, and can also be used to infer the rate of viral spread through a host population and the basic reproduction number R_0, defined as the average number of infections generated by an infected host in a susceptible host population. Of particular interest is the study of viral adaptation through phylodynamic methods.

4.7. Neutralisation assays

Neutralizing antibodies appear after SARS-CoV-2 infection and vaccination and are maintained for several months. The emergence of SARS-CoV-2 variants has raised concerns about the breadth of neutralizing-antibody responses.

Comparing the neutralizing-antibody response of previously-infected and vaccinated individuals to emerging viral variants aims to determine how mutations within the spike protein are associated with virus neutralization[7].
Further, comparing the neutralizing-antibody response of individuals infected by an emerging viral variant against other circulating viral strains aims to estimate if this emerging variant will provide cross-protection against other strains.



[7] https://jamanetwork.com/journals/jama/fullarticle/2777898 , https://cdn.jamanetwork.com/ama/content_public/journal/jama/0/jld210021supp1_prod_1616113637.39581.pdf?Expires=1620451769&Signature=qczXqzToD78Jq~Cj5CvzgXqHDFGvxQ8OskCordVao~1jgKNHdsJZRkoOkPn3pa71-rs4Q~GMY8ZrmpJ7GLbFF3itmixoi8sT4JVqjRpfdXVLNX5Xufpc1G4ZoTqBsBLUgrX~1qS7wJoh~KgwZwR9PMM2HorJL1KqqpBRY-jzfXWkFq44rIEI-o60wXG3p4a50BfuWnN15KFy-QBgM-QKeoTLvXi6izmNUP5EJIb~GgyBuugbVprFWr~cZ63z0fnBjO760GezaASPxM1a3W2QEiA3eWz9P5yJwYQ5j4hCjKESyAYeaM-nEmQ9THoSRh25LQ3mhpC-l6DGJ6cWHR8-FQ__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

References

Rambaut, A., Holmes, E.C., O'Toole, Á. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5, 1403–1407 (2020). https://doi.org/10.1038/s41564-020-0770-5

Martin, M. A., Vaninsberghe, D., Koelle, K. Insights from SARS-CoV-2 sequences Science 29: 466-467 (2021). https://doi.org/10.1126/science.abf3995

Callaway, E. 'A bloody mess': Confusion reigns over naming of new COVID variants. Nature 589, 339 (2021). https://doi.org/10.1038/d41586-021-00097-w