

Document de poziție

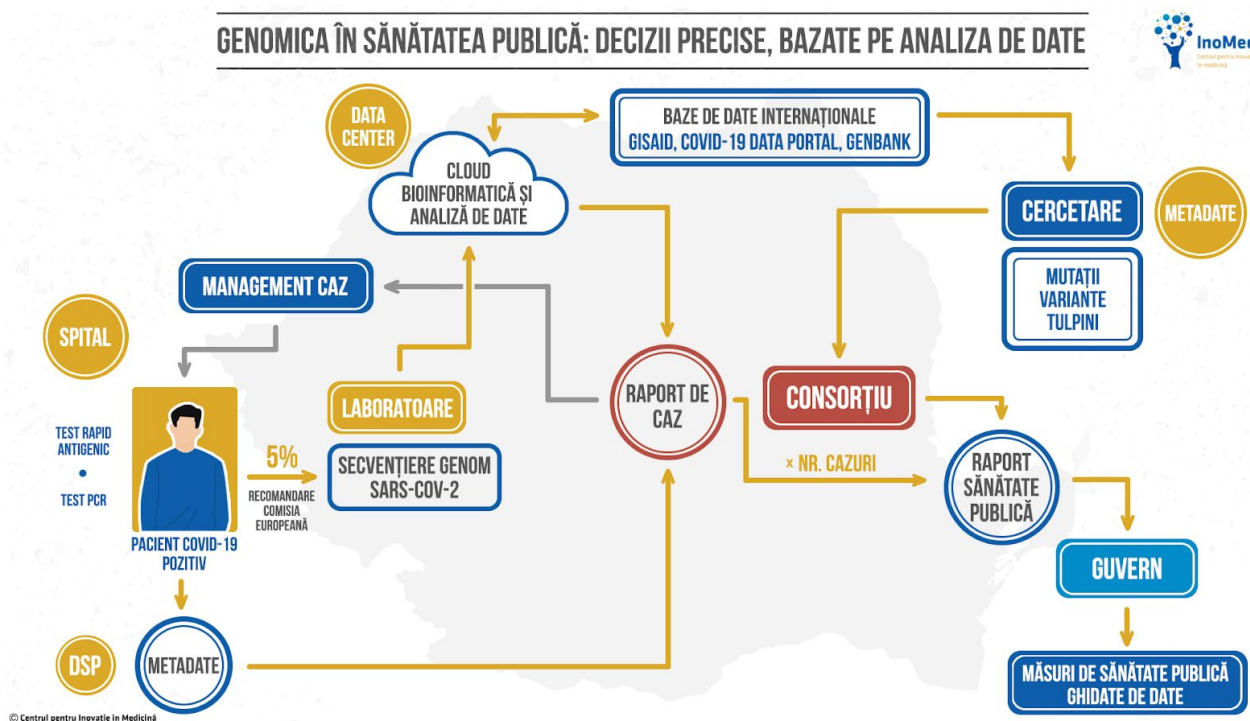
Supravegherea prin secvențiere genomică a virusului SARS-CoV-2 - propunere de înființare a Centrului Național de Date și Analiză Genomică pentru Sănătatea Publică

Sumar executiv

- Pe lângă vaccinare, pentru controlul pandemiei trebuie utilizate sisteme avansate de supraveghere a circulației virusului și de identificare în timp real a modificărilor acestuia, prin secvențiere genomică pe scară largă
- În anul 2021, au fost comunicate la CNSCBT 499 secvențieri genomice, finalizate în săptămânile 53/2020-07/2021, respectiv până la data de 21 februarie 2021 inclusiv, dintre care 136 au fost confirmate cu VOC 202012/01 (linia genetică B.1.1.7).
- Comisia Europeană recomandă ca Statele Membre să ajungă la o capacitate de secvențiere de cel puțin 5% dintre probele pozitive Covid-19. Conform recomandărilor Comisiei Europene, România ar trebui să realizeze între 5.000 și 10.000 de secvențieri pe lună.
- Recomandăm implementarea unui sistem de supraveghere prin secvențiere genomică a virusului SARS-CoV-2, la nivel național, coordonat de un **Centru Național de Date și Analiză Genomică pentru Sănătatea Publică**.
- Setarea **Centrului Național de Date și Analiză Genomică pentru Sănătatea Publică** este în linie cu solicitările Comisiei Europene privitoare la tranziția digitală din cadrul Mecanismului de Redresare și Reziliență al Comisiei Europene și cu recomandările European Semester (*“Consiliul RECOMANDĂ ca, în 2020 și în 2021,România să întreprindă acțiuni astfel încât: Să direcționeze cu prioritate investițiile către tranziția ecologică și digitală, în special către infrastructura de servicii digitale”*).
- Recomandăm ca implementarea și guvernarea **Centrului Național de Date și Analiză Genomică pentru Sănătatea Publică** să fie realizată de un **Consortiu** la care să

participe: institute de cercetare cu experiență în domeniul secvențierii genomice a SARS-CoV-2 (Institutul Cantacuzino, Institutul Matei Balș), Rețeaua de Genetică Medicală, una sau mai multe Universități de Medicină și Farmacie, organizații neguvernamentale relevante (Societatea Română de Bioinformatică, Societatea Română de Genetică Medicală, Centrul pentru Inovație în Medicină), entități private relevante (furnizori de servicii medicale, furnizori de soluții de secvențiere, furnizori de soluții ICT)

- **Recomandăm** implicarea în proiect, încă de la început, a unor experți din consorțiile similare din Danemarca și Marea Britanie
- Design-ul sistemului de supraveghere genomică pentru SARS-CoV-2 este în imaginea de mai jos:



Context

Genomul coronavirusului SARS-CoV-2 a fost secvențiat într-o lună de la raportarea primului caz de boală COVID-19¹. Comunitatea științifică internațională răspuns cu o viteză fără precedent în cazul urgenței de sănătate publică reprezentată de noua tulpină de coronavirus. Acest răspuns a fost posibil datorită tehnologiilor de secvențiere masivă paralelă, instrumentelor digitale, căilor de comunicare a datelor aproape în timp real, dar și deschiderii către colaborare și împărtășire a datelor.

SARS-CoV-2 face parte din categoria virusurilor ARN, care au o tendință de a acumula erori pe măsură ce se replică. SARS-CoV-2 acumulează în medie 2 mutații pe lună. Nu toate variantele virale care au apărut în cursul anului 2020 au avut o semnificație biologică, însă pe măsură ce virusul evoluează și determină infecții răspândite la nivel mondial pot apărea mutații cu un impact epidemiologic. Există în prezent 3 variante virale care sunt în atenția autorităților internaționale: B.1.1.7 (Marea Britanie), B.1.351 (Africa de Sud) și P1 (Brazilia) - care pun probleme din punctul de vedere al controlului pandemiei, ca urmare a transmisibilității crescute și a interferenței cu răspunsul imun obținut după administrarea vaccinurilor.

După aproximativ 10 luni în care virusul a evoluat fără a acumula mutații semnificative (care ar putea sugera o schimbare de fenotip și ar putea avea impact epidemiologic), se observă o schimbare a tiparului evolutiv. Proteina spike, și mai ales domeniul de legare de receptor (RBD), reprezintă o regiune esențială pentru evoluția virală și cu impact asupra răspunsului imun. Până în octombrie 2020 s-a observat un număr mic de modificări la nivelul genomului viral, în regiunile care codifică aceste structuri. Totuși, cele 3 variante menționate anterior au acumulat mai multe mutații, într-un timp scurt, comparativ cu alte virusuri.

Apariția a 3 noi variante care au caracteristici similare sugerează posibilitatea unei *evoluții convergente*. Virusurile se adaptează la populația umană, în moduri asemănătoare, în diferite părți ale lumii. Există posibilitatea de apariție a unor variante care vor scădea semnificativ eficacitatea vaccinurilor autorizate, dacă nu se introduc măsuri de sănătate publică pentru reducerea transmiterii infecției².

În acest context, pentru controlul pandemiei trebuie utilizate sisteme avansate de supraveghere a circulației virusului și de identificare în timp real a modificărilor acestuia, prin secvențiere genomică pe scară largă.

¹ <https://virological.org/t/novel-2019-coronavirus-genome/319>

² https://www.sciencedirect.com/science/article/pii/S0092867421000878?dgcid=rss_sd_all

Analizarea genomului SARS-CoV-2 permite obținerea unei imagini de rezoluție înaltă, cu multiple aplicații, de la dezvoltarea de teste de diagnostic și vaccinuri, și până la informarea în timp real, asupra măsurilor de sănătate publică.

La nivel european, ECDC consideră că riscul de transmitere a noilor variante SARS-CoV-2 este **înalt/foarte înalt**, iar Comisia Europeană recomandă ca Statele Membre să ajungă la o capacitate de secvențiere de cel puțin 5% dintre probele pozitive Covid-19. La începutul anului 2021, un singur Stat Membru UE testa mai mult de 1% dintre probele pozitive, în timp ce alte state nu secvențiau suficient sau nu secvențiau deloc. În acest moment, modelul de bune practici în Uniunea Europeană este Danemarca (secvențiază toate cazurile pozitive), iar pe plan mondial este Marea Britanie (deține jumătate dintre genomurile secvențiate, la nivel mondial).

Un raport emis de ECDC pe 16 februarie 2021 cuprinde o analiză a capacității statelor membre UE de a detecta noile variante SARS-CoV-2. În cele mai multe state, capacitatea de secvențiere cu scopul identificării variantelor SARS-CoV-2 este sub standardul impus de Comisia Europeană, respectiv secvențiere a 5-10% din cazurile pozitive. Majoritatea statelor UE secvențiază sub 1% din probele pozitive. Nevoile specifice identificate în multe dintre țările evaluate includ capacitatea de secvențiere, existența protocoalelor și în mod special, componenta de analiză bioinformatică³.

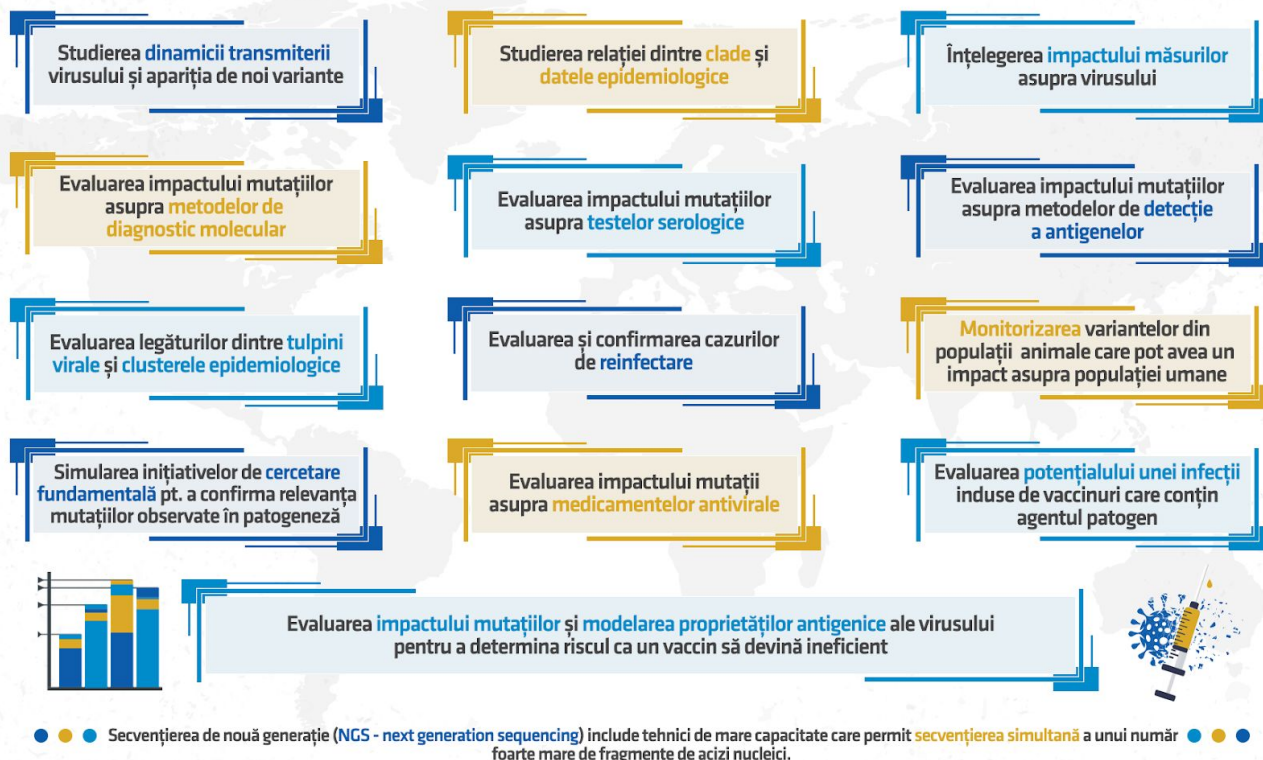
Organizația Mondială a Sănătății și Centrul European pentru Prevenție și Controlul Bolilor recomandă intensificarea monitorizării evoluției genomului viral și împărtășirea datelor obținute prin secvențiere⁴. Pe baza datelor disponibile, se așteaptă o creștere a numărului de cazuri de infecție COVID-19 din cauza răspândirii accelerate a noilor variante caracterizate de o transmitere crescută. Rata transmiterii și impactul asupra statelor membre va depinde de modul în care sunt aplicate măsurile non-farmacologice și complianța la aceste măsuri. Mesajul principal emis de ECDC este pregătirea pentru o creștere rapidă a cazurilor care va impune introducerea unor măsuri stricte de control pentru a proteja capacitatea sistemelor de sănătate, alături de accelerarea campaniilor de vaccinare.

³<https://www.ecdc.europa.eu/en/publications-data/detection-and-characterisation-capability-and-capacity-sars-cov-2-variants>

⁴ <https://www.who.int/publications/i/item/9789240018440>

Obiectivele secvențierii genomului SARS-CoV-2

OBIECTIVELE SECVENȚIERII GENOMULUI SARS-COV-2/ ADAPTAT DUPĂ ECDC



Fezabilitatea secvențierii întregului genom viral în timp real (real time WGS) necesită accesul la platforme de secvențiere, accesul la metadate, dar și o strategie pentru analiza rapidă a datelor și care să implice costuri scăzute. Pentru ca monitorizarea pandemiei folosind date genomice să fie eficientă, strategiile trebuie să aibă o bună acoperire globală, să fie standardizate și integrată în programe naționale de prevenție a pandemiilor.

Organizația Mondială a Sănătății a publicat un document care cuprinde recomandări pentru maximizarea impactului inițiativelor pentru secvențierea genomului SARS-CoV-2 în strategiile de sănătate publică.

Indiferent de câte genomuri SARS-CoV-2 sunt generate, acestea vor avea un impact pozitiv asupra deciziilor de sănătate publică dacă există strategii pentru producerea și comunicarea unor rezultate utilizabile, într-un timp rezonabil. Construirea unei rețele globale reziliente poate maximiza impactul secvențierii în domeniul sănătății publice, nu doar pentru SARS-CoV-2, dar și pentru alți agenți patogeni care ar putea apărea. Investițiile pe termen lung sunt necesare pentru a crește capacitatea de analiză bioinformatică și filogenetică, două componente care în prezent au rămas în urma altor obiective.

Obiective ale secvențierii genomului SARS-Cov-2 din perspectiva optimizării strategiilor de sănătate publică, conform OMS:

- **Identificarea și caracterizarea SARS-CoV-2 și dezvoltarea de măsuri de combatere a răspândirii infecției**
- **Monitorizarea transmiterii și răspândirii geografice**
 - Studiarea răspândirii geografice și a reintroducerii între populații
 - Evaluarea dovezilor asupra căilor de transmitere sau clustere
 - Cuantificarea perioadelor de transmitere și urmărirea parametrilor epidemiologici în timp – R0
 - Monitorizarea mediului înconjurător
 - Evaluarea posibilelor cazuri de reinfecție
- **Monitorizarea evoluției SARS-CoV-2**
 - Evaluarea structurată a posibilelor mutații relevante. Prin secvențiere genomică se pot identifica modificări genetice care pot determina schimbări de fenotip (transmisibilitate, virulență)
 - Monitorizarea impactului evoluției SARS-CoV-2 asupra măsurilor pentru combaterea pandemiei (eficacitatea vaccinurilor, acuratețea testelor moleculare, eficacitatea terapierilor antivirale)
 - Monitorizarea evoluției SARS-CoV-2 la animale

OMS recomandă ca deciziile asupra scopurilor secvențierii genomului viral să se ia în cadrul unor **echipe multidisciplinare**. Aceste decizii vor ghida aspectele tehnice legate de metodele folosite pentru secvențiere și alegerea probelor. Pentru majoritatea tipurilor de obiective urmărite, atât datele obținute prin secvențiere cât și metadatele sunt importante.

Multe analize se bazează pe capacitatea de a compara secvențele virale locale cu diversitatea genomului viral de la nivel global. În acest sens, **se recomandă împărtășirea datelor** la nivelul resurselor internaționale precum GISAID și GenBank. **Alegerea probelor pentru secvențiere** depinde de întrebarea la care se dorește un răspuns și de context. Trebuie luate în considerare și elementele logistice, cum ar fi modul de transport al materialului și cum se poate realiza optim extragerea ARN-ului viral și secvențiere, fără a pune în pericol materialul genetic.

Prioritizarea obiectivelor în funcție de contextul local

Atunci când resursele dedicate proiectelor de secvențiere sunt limitate, obiectivele trebuie să vizeze acele activități care pot avea un potențial clinic sau impact asupra deciziilor de sănătate

publică, activități care să poată fi sustenabile. În acest caz, secvențierea genomului SARS-CoV-2 poate fi prioritizată în următoarele situații, conform recomandărilor **OMS și ECDC**:⁵⁶

- Izolarea și secvențierea genomului viral de la persoane vaccinate pentru SARS-CoV-2, dar care se infectează ulterior în ciuda unui răspuns imun optim la vaccin
- Situații de risc precum comunitățile în care există animale susceptibile pentru infecția cu SARS-CoV-2
- Cazurile pacienților imunocompromiși, care rămân pozitivi pentru perioade lungi de timp, mai ales atunci când primesc tratament cu anticorpi în contextul infecției cu SARS-CoV-2
- Atunci când există o creștere neașteptată a transmisibilității sau virulenței SARS-CoV-2
- Atunci când există suspiciunea unei modificări a performanței unor teste de diagnostic (teste moleculare sau care folosesc antigene/anticorpi)

Genomul SARS-CoV-2

La doar o lună de la apariția primelor cazuri de boală COVID-19, genomul SARS-CoV-2 a fost secvențiat. Prin comparație, cazul epidemiei SARS din 2003, secvențierea genomului a necesitat 6 luni, iar în cazul epidemiei de gripă porcină din 2009 au fost necesare 2 luni și jumătate⁷. Dezvoltarea genomicii are deja un impact semnificativ asupra epidemiologiei, virusologiei, bolilor infecțioase și sănătății publice. Sute de mii de genomuri au fost analizate până în prezent prin secvențiere masivă paralelă. De exemplu, baza de date GISAID (The Global Initiative on Sharing All Influenza Data), o resursă publică ce urmărește evoluția virusului și răspândirea la nivel mondial, a înregistrat peste 528.000 genomuri SARS-CoV-2⁸, până la data de 14 februarie 2021.

SARS-CoV-2 are genomul organizat sub forma unei molecule lungi de acid ribonucleic – ARN. SARS-CoV-2 face parte din genul betacoronavirus împreună cu SARS-CoV și MERS-CoV. Coronavirusurile au cele mai mari genomuri dintre toate virusurile din familia ARN. Cea mai mare parte din genom (două treimi) codifică gene pentru polimeraza virală și alte proteine nestructurale (ORF1a, ORF1b). O treime din genom corespunde unor proteine structurale – proteina spike (S), proteine care fac parte din anvelopa virală (E), membrana (M), nucleocapsida (N) și alte proteine helper. Virusul pătrunde în celulele umane prin legarea de receptorului ACE2, cu ajutorul proteinei spike (S), ce reprezintă o structură antigenică. Modificările de la nivelul genei care codifică proteina spike ar putea avea un impact semnificativ asupra eficacității vaccinurilor, imunității naturale și a terapiei bazate pe anticorpi.

Datele preliminare asupra ratei de evoluție a SARS-CoV-2 sunt aproape de o medie de 1×10^{-3} substituții pe an, similară cu rata de evoluție a altor virusuri ARN. Rata de mutații a coronavirusurilor este mai scăzută decât a altor virusuri ARN deoarece au capacitatea de a corecta o parte dintre aceste mutații. SARS-CoV-2 înregistrează, în medie, o substituție la două săptămâni.

⁵ <https://apps.who.int/iris/bitstream/handle/10665/338480/9789240018440-eng.pdf>

⁶ <https://www.ecdc.europa.eu/sites/default/files/documents/Sequencing-of-SARS-CoV-2-first-update.pdf>

⁷ <https://www.illumina.com/company/news-center/blog/2020-in-genomics.html>

⁸ <https://www.gisaid.org/>

Pe 14 decembrie 2020, autoritățile din Marea Britanie au raportat către Organizația Mondială a Sănătății (OMS) **identificarea unei noi variante SARS-CoV-2** - SARS-CoV-2 VOC 202012/01 (Variant of Concern, year 2020, month 12, variant 01), *cluster filogenetic B.1.1.7*. Aceasta se caracterizează prin 23 de mutații, 17 dintre acestea sunt mutații care se traduc prin modificarea proteinelor virale (6 mutații sinonime, 14 non-sinonime și 3 deleții). La nivelul proteinei Spike sunt înregistrate 8 mutații. **Mutația N501Y** alterează un aminoacid prezent la nivelul domeniului de legare a receptorului (RBD) din proteina Spike și **se asociază cu o afinitate de legare crescută față de receptorul de pe suprafața celulelor gazdă (ACE2)**. **Deleția din poziția 69/70 poate afecta performanța testelor de diagnosticare (PCR)** care utilizează gena proteinei S ca țintă și de aceea este recomandată utilizarea testelor PCR care folosesc mai multe ținte.

Pe 26 ianuarie 2021, Public Health England a raportat identificarea mutației E484K, la nivelul genei care codifică proteina S. Datele provin în urma unei analize care a inclus peste 200.000 de genomuri secvențiate în cadrul proiectului COG-UK⁹. Se estimează că administrarea de anticorpi monoclonali sau anticorpii dezvoltati în urma vaccinării au o eficacitate mai mică în neutralizarea variantelor virale cu această mutație¹⁰.

Datele colectate până în prezent sugerează o transmisibilitate mai ridicată între contactii B.1.1.7. O variantă mai contagioasă a SARS-CoV-2 ar putea avea un impact semnificativ asupra controlului pandemiei (un impact chiar mai mare decât ar avea o variantă asociată cu forme mai severe). **NERVTAG** (New and Emerging Respiratory Virus Threats Advisory Group) a publicat, la începutul anului, date care arată că există posibilitatea ca această variantă să fie asociată cu o creștere a mortalității¹¹. De asemenea, date NERVTAG publicate la începutul lunii februarie 2021, indică faptul că B.1.1.7 este dominantă în toate regiunile Regatului Unit și estimează că sunt necesare 7-8 săptămâni pentru ca incidența cazurilor să crească de la 10% la 80%. La 2 februarie 2021 se estima că în cele mai multe țări europene varianta B.1.1.7 reprezintă aproximativ 5-10% din cazurile pozitive.

B1.1.7 nu este singura variantă SARS-CoV-2 identificată în ultimele luni. O diversificare a SARS-CoV-2 a fost observată la nivel global, fiind determinată de procesele de evoluție și adaptare. Alte variante identificate și care au semnificație din punctul de vedere al sănătății publice sunt:

- **20C/501Y.V2** este o altă variantă care a fost prima dată detectată la nivelul provinciei Eastern Cape din Africa de Sud și raportată la 18 decembrie 2020. Aceasta, de asemenea, se poate lega mai ușor de celulele umane și e asociată cu o creștere a ratei de transmitere a infecției. Prezintă mutația N501Y, la fel ca varianta din Regatul Unit, dar și alte mutații de interes, localizate la nivelul domeniului de legare de receptor din gena care codifică proteina spike (K417N și E484K). Studiile sugerează o posibilă modificare a imunogenității din cauza mutației E484K. Două studii recente indică faptul că serul convalescent care

⁹https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957504/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5_England.pdf

¹⁰ <https://science.sciencemag.org/content/371/6527/329>

¹¹https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955239/NERVTAG_paper_on_variant_of_concern_VOC_B.1.1.7.pdf

conține anticorpi neutralizanți împotriva SARS-CoV-2 a prezentat o **putere de neutralizare substanțial redusă contra variantei 501Y.V2**. Până la 19 ianuarie s-au identificat cazuri de infecție cu varianta 501Y.V2 în 10 țări europene.

- **20J/501Y.V3** este o variantă detectată pe 6 ianuarie, în Japonia¹², de la persoane întoarse din regiunea Amazonas din Brazilia. Aceasta prezintă 12 mutații la nivelul regiunii spike, printre care N501Y și E484K.

După aproximativ 10 luni în care virusul a evoluat fără a acumula mutații semnificative (care ar putea sugera o schimbare de fenotip și ar putea avea impact epidemiologic), se observă o schimbare a tiparului evolutiv. Proteina spike, și mai ales domeniul de legare de receptor (RBD), reprezintă o regiune esențială pentru evoluția virală și cu impact asupra răspunsului imun.

Metode de secvențiere

Pandemia COVID-19 este prima pentru care tehnologiile de secvențiere de nouă generație au fost disponibile încă de la început. Strategiile de secvențiere pentru SARS-CoV-2 includ **metode țintite**, care se bazează pe cunoașterea genomului și **metode metagenomice**, care nu necesită cunoașterea genomului. Înainte de realizarea investițiilor într-o strategie, trebuie să fie evaluate cerințele impuse de diferite tehnologii (resurse umane competente, infrastructură, durata până la obținerea rezultatelor, costuri, procesarea datelor, capacitatea și acuratețea). Numărul de probe care trebuie analizate depinde de obiectivul programului de secvențiere.

Tehnologiile de tip Sanger pot fi utilizate în cazul secvențierii fragmentelor individuale (de până la 1000 perechi de baze), în reacții separate. Secvențierea întregului genom ar necesita cel puțin 30 de ampliconi individuali să fie secvențiați separat pentru fiecare probă. De aceea, **secvențierea Sanger** este utilă pentru secvențierea unor fragmente scurte, de exemplu, să acopere posibilele lacune care pot apărea în urma secvențierii masive paralele sau pentru studierea diversității virusului la nivelul unor regiuni scurte (zona de legare a primerilor, în cazul în care un test de diagnostic sugerează anomalii).

Platformele pentru secvențiere masivă paralelă sunt mai potrivite pentru secvențierea de rutină a întregului genom. Platformele frecvent folosite în prezent includ cele produse de Illumina, IonTorrent și Oxford Nanopore. Acestea permit secvențierea concomitentă a unui număr mai mare de probe. Alegerea platformei depinde de obiectivului programului de secvențiere. Un timp rapid de obținere a rezultatelor este util atunci când obiectivul ține de aplicații clinice, în timp ce o acuratețe ridicată este esențială pentru a studia diversitatea evoluției virului la nivelul unei gazde.

Secvențierea de mare capacitate de tip **shotgun metagenomics** permite identificarea și genotiparea tuturor tipurilor de microorganisme dintr-o probă fără a cunoaște anterior ce tipuri de microorganisme sunt prezente. Această metodă permite inclusiv identificarea unor co-infecții (virală/bacteriană), ceea ce poate informa deciziile terapeutice și prognosticul pacientului. Înțelegerea materialului genetic viral permite identificarea modului în care evoluează un virus pentru a monitoriza modul de transmitere și acuratețea testelor de diagnostic sau eficacitatea terapiilor. Analiza transcriptomului uman (setul de ARN de la nivelul epiteliului nazal inflammat și

¹² <https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-en.html>

celulelor imune) poate indica modul în care sistemul imun reacționează în urma infecției. În acest fel se pot descoperi mecanismele prin care SARS-CoV-2 infectează celulele umane și de ce anumite persoane fac o formă de boală mai severă.

mNGS nu este indicată în cazul unor probe cu încărcătură virală redusă. În aceste cazuri se preferă metodele țintite (amplicon/capture-based) prin care se pot detecta tulpinile noului coronavirus cu o acuratețe similară cu secvențierea de tip shotgun¹³.

Pentru asigurarea implementării unui program sustenabil de secvențiere trebuie luate în considerare punctele slabe și provocările alături de oportunitățile pe care fiecare tehnologie le-ar aduce. Țările care nu au implementat încă programe care folosesc tehnologii de secvențiere de mare capacitate trebuie să ia în considerare dovezile din literatura de specialitate. În acest, ECDC a publicat o listă de recomandări¹⁴.

Testarea PCR care se realizează în prezent doar determină prezența virusului, nu oferă informații despre genom. Atunci când resursele sunt limitate pentru secvențierea la scară largă, testarea PCR poate permite detecția noilor variante. De exemplu, **Applied Biosystems TaqPath COVID-19 assay** (ThermoFisher) este un test PCR care conduce la rezultate particulare atunci când varianta virală prezintă deleția 69/70. Lipsa amplificării genei care codifică proteina S permite evaluarea frecvenței B.1.1.7 și selectarea probelor destinate secvențierii. Însă, și alte variante SARS-Cov-2 prezintă această deleție, de aceea TaqPath nu permite stabilirea diagnosticului de infecție cu B.1.1.7.

- ***Extragerea materialului genetic***

Obținerea unor date de calitate în urma secvențierii depinde de extragerea unei cantități suficiente de material genetic viral (ARN), de calitate, pentru a fi prelucrat. Acolo unde sunt disponibile mai multe tipuri de probe, se recomandă selectarea probei care are cea mai mare încărcătură virală și un nivel scăzut de contaminare cu materiale bacteriene/umane. Dacă este fezabil, izolatele destinate secvențierii ar trebui selectate din probe pozitive care au fost deja procesate de un laborator de diagnostic molecular. Distribuirea optimă a resurselor previne duplicarea probelor procesate și poate economisi resurse umane și materiale. Este importantă stabilirea unui sistem de identificare a probelor dacă acestea migrează între laboratoare.

Modul de stocare a ARN-ului viral este important pentru obținerea unor date de calitate. Trebuie să se mențină temperatura optimă între etapa prelevării probei și analiză.

- ***Selectarea probelor***

Strategiile pentru obținerea probelor vor depinde de obiectivele urmărite. Pentru studierea **transmiterii nosocomiale** sau evaluarea rezultatelor **anchetelor epidemiologice** este importantă obținerea unui număr mai mare de probe din clusterul epidemiologic de interes, precum și probe care nu fac parte din acesta, pentru a observa dacă acestea sunt conectate.

¹³ <https://www.nature.com/articles/d42473-020-00120-0>

¹⁴ <https://www.ecdc.europa.eu/en/publications-data/sequencing-sars-cov-2>

Analizele de filodinamică, în schimb, pot fi realizate pe un număr mai redus de probe, alese aleatoriu. Acestea pot fi influențate negativ dacă probele nu sunt alese aleatoriu din totalul cazurilor confirmate. Modelele filodinamice pornesc de la ideea că secvențele sunt colectate uniform, ceea ce nu se poate obține dacă acestea sunt prelevate ca urmare la fenomenului de contact-tracing. Pentru analizele de filodinamică, genomurile virale ar trebui să fie secvențiate în raport cu adevărata incidență a cazurilor.

Dacă la nivelul unei anumite regiuni există o acoperire bună în ceea ce privește diagnosticul, un subset aleatoriu de probe pozitive, rămase în urma diagnosticului molecular, ar putea fi destinate secvențierii. Cu toate acestea, în anumite țări, probele pozitive nu vor reflecta distribuția reală a răspândirii infecției din cauza lacunelor care există în materie de resurse sau accesibilitatea dintre locații. Dacă raportarea cazurilor suspecte este mai reprezentativă decât raportarea cazurilor confirmate, se recomandă selectarea unor probe din diferite locații și de la momente diferite în raport cu numărul de cazuri suspecte.

Numărul de probe destinate secvențierii depinde de resursele disponibile, de capacitatea de monitorizare și nivelul de detecție. Strategia de selectare a probelor trebuie adaptată, de asemenea, în funcție de capacitatea locală/națională. În acest sens, există un ghid publicat de ECDC¹⁵ pentru stabilirea numărului necesar de genomuri secvențiate, raportat la unitatea de timp și zona geografică pentru a obține detectarea și cuantificarea variantelor virale dintre alte tipuri de virusuri circulante. Conform evaluării de risc privind noile variante virale, din 21 ianuarie 2021, ECDC recomandă un număr minim de 500 de probe secvențiate aleatoriu, în fiecare săptămână.

Pentru statele membre care au capacitate limitată de a realiza secvențierea genomului SARS-CoV-2 se recomandă monitorizarea modificărilor bruște în ceea ce privește ratele de transmitere sau severitatea bolii. De asemenea, pregătirea laboratoarelor trebuie să reprezinte o prioritate. În acest sens, ECDC recomandă¹⁶:

- Implementarea unor strategii de pre-screening pentru variantele care sunt în atenția autorităților internaționale (mutație N501Y, del69-70)
- Asigurarea resurselor umane și materiale pentru a controla numărul crescut de cereri pentru detecția și caracterizarea probelor SARS-CoV-2
- Creșterea capacității de secvențiere prin utilizarea tuturor resurselor disponibile la nivelul statelor membre sau solicitarea asistenței ECDC

Conform raportului ECDC publicat pe 16 februarie, privind capacitatea de secvențiere la nivelul UE, în multe state membre, timpul de obținere a rezultatelor testelor RT-PCR pentru pre-screening este mai mare de 48 de ore. Scopul realizării acestor teste care pot ridica suspiciunea apariției unei noi variante includ: reducerea cazurilor infecției cu variante necunoscute de import, pentru detecția precoce a cazurilor comunitare sau pentru evaluare focarelor.

¹⁵ <https://www.ecdc.europa.eu/en/publications-data/sequencing-sars-cov-2>

¹⁶ <https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-variants-concern-eueea-first-update>

CELE MAI UTILIZATE PLATFORME PENTRU SECVENȚIEREA ȘI ANALIZA SARS-COV-2



	Avantaje 	Limitări 	Timpul de lucru 	Capacitatea de secvențiere 	Valabilitatea și costurile
Secvențiere Sanger	Accesibilitate la scală largă Ușor de utilizat Cost eficient dacă sunt utilizate puține ținte	Capacitatea de secvențiere este foarte mică Ampliconii trebuie să fie amplificați și secvențiați individual Costuri ridicate pentru analiza genomului întreg Nu poate fi utilizat pentru analiza metagenomurilor	Câteva ore	100 kB - 2 Mb într-un singur ciclu de utilizare	Valabilitate pe scală largă Costuri reduse pentru a utiliza puține ținte
Illumina (ex: iSeq, MiniSeq, MiSeq, NextSeq, HiSeq, NovaSeq)	Randament foarte crescut Acuratețe foarte mare iSeq este portabil Metodele de manipulare a datelor sunt foarte bine stabilite	Cu excepția dispozitivului Illumina iSeq, costurile sunt ridicate pentru achiziție și mentenanță, comparativ cu alte platforme Capacitatea maximă de citire este 2 x 300 bp	10 - 55 ore depinde de instrumentul utilizat	1,2 - 6000 Gb depinde de instrumentul utilizat	Costuri ridicate de mentenanță Costuri moderate de funcționare
Oxford Nanopore Technologies (Flongle, MinION, GridION, PromethION)	Dispozitiv portabil Date în timp real Costuri de mentenanță scăzute Secvențierea poate fi orpită în momentul în care se strâng suficiente date Capacitate de citire mare	Limitări la utilizarea homopolimerilor Rata de eroare/citire este de ~5%, așadar utilizarea instrumentelor adecvate este critică pentru obținerea secvențierii de precizie Momentan nu este potrivită pentru determinarea variației la nivelul gazdei	Datele sunt disponibile imediat Timpul de lucru poate fi monitorizat, iar dispozitivul lăsat să funcționeze pentru zile întregi dacă este nevoie	Variază de la < 2 Gb pentru Flongle, până la 220 Gb pentru PromethION Pot fi folosite până la 48 de celule pentru PromethION	Costuri reduse pentru pornire Costuri moderate de funcționare
Ion Torrent	Timp rapid de obținere a rezultatelor odată ce a început secvențierea	Limitări la utilizarea homopolimerilor Costuri ridicate la achiziție Capacitate de citire în jur de 400 bp	2 ore - 1 zi depinde de cipul și dispozitivul utilizate	30 Mb - 50 Gb depinde de cipul și dispozitivul utilizate	Costuri moderate

Recomandări ale Comisiei Europene

Comisia Europeană¹⁷ a publicat pe 19 ianuarie un document care cuprinde o serie de acțiuni necesare pentru a intensifica lupta împotriva pandemiei COVID-19. Acestea vizează recomandări pentru statele membre, Centrul European de Prevenire și Control al Bolilor (ECDC), Agenția Europeană pentru Medicamente (EMA).

- **Toate statele membre UE trebuie să ajungă la o capacitate de secvențiere de cel puțin 5% din probele pozitive. Este de preferat să se ajungă la 10%, iar rezultatele trebuie să fie obținute rapid.**
- ECDC poate suplimenta capacitatea de secvențiere cu până la 18.000 de probe. Până acum, doar două State Membre au folosit această opțiune și încă două urmează a primi ajutorul ECDC. Mai ales pentru țările în care această capacitate nu există, **utilizarea resurselor ECDC este o urgență.**
- Comisia Europeană va sprijini statele membre pentru creșterea capacității de secvențiere prin mobilizarea fondurilor pentru a achiziționa echipamente și a sprijini Rețeaua de Laboratoare OMS. Echipamentele existente în instituțiile de cercetare naționale ar trebui folosite la capacitate maximă. De asemenea, Comisia Europeană va susține inițiativele pentru procesarea și analiza datelor genomice.

¹⁷ https://ec.europa.eu/info/sites/info/files/communication-united-front-beat-covid-19_en.pdf

La finalul lunii octombrie 2020, Comisia Europeană încuraja transferul de date și cunoștințe despre Covid-19, în timp real, la nivelul UE¹⁸. În discursul președintelui Comisiei Europene, Ursula von der Leyen, se regăsește și îndemnul de a partaja datele pe platforma ECDC și în baze de date internaționale, ca GISAIID, menționată mai sus. Experiența cu SARS-CoV-2 a dovedit că nu doar datele culese în mod direct sunt importante, ci și metadatele. Aceeași recomandare vine și din partea Organizației Mondiale a Sănătății¹⁹.

Partajarea rapidă a datelor provenite din secvențierea genomică (GSD), împreună cu metadatele epidemiologice și clinice relevante anonimizate, vor maximiza impactul secvențierii genomice în răspunsul sănătății publice la COVID-19 (public health genomics).

ECDC recomandă ca statele membre să monitorizeze în continuare modificările apărute brusc ale ratelor de transmitere sau de severitate a bolii, ca parte a procesului de identificare și evaluare a impactului variantelor. Autoritățile naționale de sănătate publică ar trebui să anunțe cazurile cu noile variante cunoscute, precum și cu orice alte variante noi de SARS-CoV-2 de potențială îngrijorare, prin intermediul sistemului de avertizare timpurie (EWRS - Early Warning and Response System) și al sistemului european de supraveghere (TESSy - The European Surveillance System)²⁰.

Partajarea și interpretarea datelor – nomenclaturi, instrumente și baze de date utilizate la nivel internațional

O componentă majoră a supravegherii bazate pe secvențierea agentului patogen este utilizarea unei nomenclaturi potrivite pentru datele obținute în urma secvențierii, bazată pe înrudirea genetică a secvențelor. Acest lucru simplifică comunicarea între diferiții actori din domeniul epidemiologiei moleculare și permite tabelarea simplificată a datelor genomice pentru integrarea cu analiza epidemiologică standard. Au fost implementate mai multe nomenclaturi pentru SARS-CoV-2:

- Nomenclatura GISAIID, folosind termenul hCoV-19 pentru virus (www.gisaid.org)
- Nomenclatura Nextstrain (<https://nextstrain.org/ncov>)
- Nomenclatura liniei genetice descrise de Andrew Rambaut și colab. (<https://cov-lineages.org>)

În timp ce nomenclaturile de tip cladă (o ramură a unui arbore filogenetic conținând toate organismele descendente dintr-un strămoș comun) utilizate de Nextstrain și GISAIID vizează

¹⁸ https://ec.europa.eu/commission/presscorner/detail/en/statement_20_2015

¹⁹ <https://apps.who.int/iris/handle/10665/338483>

²⁰ <https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-sars-cov-2-variants-eueea>

clasificarea la scară largă a diversității virale circulante, nomenclaturile de tip descendențe (<https://cov-lineages.org>) sunt destinate să evidențieze focarele din anumite regiuni geografice.

Secvențele de consens, complete (canonice, obținute în urma alinierii a mai multe secvențe și identificării unor modele comune) ar trebui să fie partajate în baza de date EpiCov a GISAID (www.gisaid.org) pentru a permite analiza filogenetică la nivel mondial. Datele brute, pe de altă parte, pot fi încărcate în portalul de date COVID-19 (www.covid19dataportal.org) pentru a le pune la dispoziția comunității de cercetare globale. Atât GISAID, cât și ENA / SRA pot fi legate la Sistemul european de supraveghere (TESSy). Dacă se constată că mutațiile virusului au dus la un rezultat fals negativ al RT-PCR, această informație trebuie raportată rapid către ECDC și către Biroul regional al OMS în Europa.

Resursele disponibile pentru raportarea și interpretarea datelor obținute în urma secvențierii

GISAID²¹: În 2006, ca urmare a fatalităților cauzate de gripa aviară, accesului limitat la secvențele virusului H5N1, reticenței națiunilor de a împărtăși datele obținute public (erau partajate la nivel național – EMBL, DDBJ, GenBank), a fost clară nevoia unei baze de date sigure, în care autorii datelor să se simtă protejați și să poată primi acreditarea meritată. Astfel, în 2008, a fost lansată GISAID, o bază de date publică, dar care necesită autentificarea utilizatorilor - „acces liber”, dar nu public.

Nextstrain²² este un proiect open-source care își propune valorificarea și integrarea potențialului științific și de sănătate publică al datelor genomice provenite din secvențierea SARS-CoV-2i. Nextstrain oferă o vizualizare actualizată continuu a datelor disponibile public, alături de instrumente de analiză. Echipa Nextstrain analizează datele genomice SARS-CoV-2 la nivel global și oferă contextul cu ajutorul instrumentelor vizuale. Mai mult, există un instrument, Nextclade, care permite utilizatorului să compare secvența SARS-CoV-2 obținută cu aceea de referință și să își încadreze filogenetic secvența în structura arborescentă a noului coronavirus. Nextstrain este deosebit de utilă în contextul în care, de exemplu, existau 180 de mii de genomuri virale complete la nivel internațional (noiembrie 2020), creșterea era exponențială, iar probabilitatea ca multe să fie asemănătoare, dar și să fie obținute altele la fel în continuare este foarte mare.

PANGOLIN²³ (Phylogenetic Assignment of Named Global Outbreak Lineages – Analiza filogenetică a filiațiilor globale): a fost dezvoltat pentru a pune în aplicare nomenclatura dinamică a descendențelor genetice (lineages, linii genetice) SARS-CoV-2, cunoscută sub numele de nomenclatura Pango. Permite utilizatorului să atribuie secvența genomică SARS-CoV-2 cu cea

²¹ <https://www.gisaid.org/about-us/mission/>

²² <https://nextstrain.org/sars-cov-2/>

²³ <https://cov-lineages.org/>

mai probabilă descendență (descendența Pango) secvențelor de interogare SARS-CoV-2. Programul Pangolin oferă și algoritmul pangoLEARN pentru identificarea descendențelor, bazat pe crearea structurii arborescente de clasificare a descendențelor SARS-CoV-2 utilizând machine learning. PangoLEARN a fost antrenat utilizând aproximativ 60.000 de secvențe SARS-CoV-2 din GISAID.

GenBank²⁴ este baza de date pentru secvențiere genetică a NIH (National Institutes of Health din Statele Unite), o colecție adnotată a tuturor secvențelor de ADN disponibile publicului. GenBank face parte din International Nucleotide Sequence Database Collaboration - INSDC, care cuprinde DNA DataBank din Japonia (DDBJ), European Nucleotide Archive (ENA) și GenBank a NCBI (National Center for Biotechnology Information). Aceste trei organizații fac schimb zilnic de date.

COVID-19 Data Portal²⁵: o platformă europeană pentru împărtășirea și analizarea datelor provenite din secvențiere. Comisia Europeană și Institutul European de Bioinformatică EMBL (EMBL-EBI), împreună cu statele membre ale UE și parteneri de cercetare, cum ar fi ELIXIR, operează o platformă europeană dedicată datelor COVID-19 care permite colectarea rapidă și partajarea datelor disponibile de cercetare din diferite surse pentru comunitățile de cercetare europene și globale. Acest efort comun reprezintă un pilon prioritar pentru realizarea obiectivelor Cloudului European de Științe (EOSC - European Open Science Cloud).

Baza de date este împărțită în mai multe subcategorii, dintre care:

- Viral Sequences, care cuprinde atât secvențe raw, cât și secvențe canonice ale SARS-CoV-2 (288.428 de înregistrări)
- Host Sequences, care cuprinde secvențe ale gazdelor SARS-CoV-2, umane și nu numai (10.377 înregistrări).
- Expression - date cuprinzând expresia genetică și proteică a genelor umane implicate în infectarea celulelor gazdă (87 înregistrări).
- Proteine - date despre proteinele SARS-CoV-2 și receptorii asociați (1.045 înregistrări)

Rolul metadatelor

Supravegherea la nivel global a secvențelor genetice SARS-CoV-2 și a metadatelor conexe contribuie la un răspuns optim, particular schimbărilor pe care patogenul le poate suferi și potrivit pentru provocările impuse de acest mutații. **Datele genomice fără metadata nu au valoare pentru influențarea deciziilor de sănătate publică.** Pentru a obține metadata complete, este necesară urmărirea răspândirii geografice a SARS-CoV-2 în timp și obținerea certitudinii că mutațiile care apar ar putea influența semnificativ patogenitatea, transmiterea sau contra măsurile (vaccinurile, tratamentul și diagnosticul)²⁶.

²⁴ <https://www.ncbi.nlm.nih.gov/genbank/>

²⁵ <https://www.covid19dataportal.org/>

²⁶ <https://apps.who.int/iris/handle/10665/338483>

Identificarea și cuantificarea perioadelor de transmitere reprezintă un alt nivel de detaliu important. Odată ce s-a atins un grad înalt de diversitate într-o filiație, rata de substituții poate fi estimată. În acest caz, diversitatea dintre două probe, pentru care se cunoaște data prelevării, poate fi utilizată pentru estimarea TMRC (the most recent common ancestor - cel mai recent strămoș comun). Durata estimată a circulației poate permite realizarea unor predicții asupra posibilelor date în care apare un caz clinic, cu o acuratețe de săptămâni sau luni. Abordările filogenetice care utilizează „ceasul molecular” sunt folositoare pentru identificarea posibilelor cazuri de transmitere nedetectată și pentru estimarea datelor posibile ale unor evenimente care trec neobservate. Dacă există metadate asupra locației de unde a fost prelevată proba, secvențierea genomului SARS-CoV-2 poate indica dacă infecția este rezultat al transmiterii locale sau dacă e de import.

Datele obținute în urma secvențierii pot fi partajate în mai multe formate, însă calitatea datelor trebuie evaluată cu atenție înainte de partajare, iar laboratoarele ar trebui să contacteze platformele de partajare a datelor pentru consultare.

Orice date partajate ar trebui să protejeze anonimatul pacientului - datele brute care conțin „human reads” (amprenta genomică a gazdei de la care a fost secvențiat genomul viral²⁷) trebuie să fie filtrate pentru a păstra numai GSD non-uman (adică viral) înainte de partajare.

Partajarea metadatelor conexe, cum ar fi data colectării eșantionului sau locația aproximativă de eșantionare, este necesară pentru a permite utilizarea GSD în multe aplicații filogenetice. Cu toate acestea, metadatele care pot fi partajate fără a compromite anonimatul pacientului ar trebui puternic luate în considerare. În acest sens, pacientul trebuie să fie informat despre colectarea metadatelor, care este utilitatea lor în tabloul global și care sunt beneficiile individuale dacă își dă consimțământul.

Tipul de metadate necesare datelor genomice

Metadatele ar trebui să includă întotdeauna cel puțin data și locația colectării mostrei, dar metadate suplimentare vor spori considerabil utilitatea secvențierii în diverse aplicații sau pentru a lua decizii pentru controlul pandemiei. Prin urmare, acolo unde este posibil, metadatele ar trebui să includă date referitoare la tipul eșantionului, modul în care a fost obținută secvența, legături cu alți viruși, istoricul de călătorie al pacientului și informații demografice sau clinice.

Deoarece SARS-CoV-2 a apărut abia recent la oameni, diversitatea genetică a virusului rămâne relativ limitată și, prin urmare, secvențele complete sunt importante pentru a captura cât mai multe site-uri filogenetice. În cazul în care secvențierea completă nu este posibilă, pot fi generate secvențe parțiale. Genomurile SARS-CoV-2 parțiale sunt încă valoroase și ar trebui să fie de asemenea publicate în bazele de date internaționale.

Consensul organizațiilor internaționale de sănătate publică și supraveghere este de a partaja datele și metadatele obținute la nivel internațional, oricât de simpliste ar părea ele spre deosebire de cele publicate de națiuni cu un sistem de secvențiere genomică bine pus la punct, ca Marea

²⁷ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7478626/>

Britanie, deoarece este nevoie să se cunoască situația epidemiologică la nivel internațional pentru a putea adopta cele mai bune măsuri de control al pandemiei.

Informațiile clinice și demografice, cum ar fi vârsta, sexul, prezența comorbidităților, severitatea bolii și rezultatul tratamentului, precum și legăturile cu alte secvențe din baza de date, sunt încurajate atunci când astfel de informații nu riscă să ducă la identificarea pacientului. Cu cât mai multe genomuri virale sunt secvențiate, cu atât mai ușor este de păstrat anonimatul gazdei.

Un consens internațional cu privire la formatele specifice pentru metadate (cum ar fi data colectării) ar permite ca datele obținute în urma secvențierii genomice provenite de la mai multe laboratoare, diferite ca flux de lucru și raportare, să fie compilate rapid în seturi de date mai mari și să se reducă ambiguitatea. Unele baze de date, inclusiv GISAID, plasează deja restricții de format pentru anumite câmpuri.

Anonimatul pacientului este pus în pericol la începutul răspândirii patogenului, de obicei, sau în cazul în care laboratorul sau statul secvențiază foarte puțin. În cazul în care nu este posibil să fie partajate metadatele complete fără a risca confidențialitatea pacientului, datele care sunt cele mai utile pentru studii globale ar trebui să fie partajate preferențial. De exemplu, locația eșantionării, data și istoricul călătoriilor sunt mai utile pentru studiile filodinamice decât vârsta sau sexul pacientului.

Evoluția variantei B.1.1.7 pe teritoriul României

Pe 8 ianuarie, România a raportat prima varianta B.1.1.7, identificată la o persoană care nu avea istoric de călătorie internațională. În anul 2021, au fost comunicate la CNSCBT 499 secvențieri genomice, finalizate în săptămânile 53/2020-07/2021, respectiv până la data de 21 februarie 2021 inclusiv, dintre care 136 au fost confirmate cu VOC 202012/01 (linia genetică B.1.1.7). Aceasta corespunde unei rate de confirmare de 27,3%.

Călătoria în Regatul Unit a fost menționată pentru numai 5 cazuri (4%). Varianta SARS-CoV-2 linia genetică B.1.1.7 s-a identificat în jumătate din județele țării. Analiza epidemiologică asupra primelor 136 de cazuri realizată de INSP asupra indică faptul că există o probabilitate mare ca B.1.1.7 să se răspândească pe întreg teritoriul României în următoarea perioadă.

Universitatea „Ștefan cel Mare” din Suceava a secvențiat, până acum, peste 150 de probe provenind de la pacienți diagnosticați pozitiv cu infecție cu SARS-CoV-2. La 25 ianuarie 2021, Universitatea raporta un singur caz B1.1.17 dintr-un lot de 20 de probe secvențiate²⁸. La data de 11 februarie au fost raportate 14 cazuri noi de infectare cu B.1.1.7, dintr-un lot de 47 de probe, provenite din 4 județe, secvențiate în cadrul laboratorului de Metagenomică și Biologie Moleculară al Universității „Ștefan cel Mare” din Suceava (USV)[2]^{29,30}. Conform unui comunicat de presă emis

²⁸ <https://usv.ro/un-caz-cu-noua-varianta-britanica-a-tulpinii-sars-cov-2-b-1-1-7-identificat-la-suceava-in-urma-secventierilor-efectuate-la-universitatea-stefan-cel-mare-din-suceava/>

²⁹ <https://usv.ro/category/stiri/>

³⁰ <https://usv.ro/un-caz-cu-noua-varianta-britanica-a-tulpinii-sars-cov-2-b-1-1-7-identificat-la-suceava-in-urma-secventierilor-efectuate-la-universitatea-stefan-cel-mare-din-suceava/>

de universitate, creșterea semnificativă a numărului de cazuri noi și distribuția geografică a acestora indică o răspândire accentuată a noii tulpini într-o mare parte din teritoriul României.

Aceeași observație asupra creșterii incidenței variantei B.1.1.7 s-a desprins în urma unei analize realizate de compania MedLife, care a secvențiat 404 probe până la 11 februarie, toate rezultatele fiind raportate către Direcția de Sănătate Publică și Ministerul Sănătății. Sistemul Medical MedLife a început în luna aprilie 2020 primul și cel mai amplu studiu din Europa Centrală și de Est privind gradul de imunizare naturală a populației pentru Covid-19.

În luna decembrie 2020, o analiză realizată de Medlife, care a inclus 60 de genomuri virale secvențiate, indica absența variantei B.1.1.7. Ulterior, au fost secvențiate încă 79 de genomuri SARS-CoV-2, din probe recoltate în perioada 13-21 ianuarie 2021, provenite din mai multe zone ale țării și s-a constatat că 10 dintre acestea erau caracterizate de prezența variantei B.1.1.7. Pe baza acestui lot, **prevalența a fost de 12,6%**. Cea mai recentă analiză, care a inclus un lot de 90 de probe recoltate în intervalul 25 – 29 ianuarie 2021, din mai multe orașe ale țării indică o **prevalență de 26,6%** (alte 24 de probe B.1.1.7).

Conform recomandărilor Comisiei Europene, România ar trebui să realizeze între 5.000 și 10.000 de secvențieri pe lună.

Modele internaționale de secvențiere și analiză a datelor genomice pentru SARS-CoV-2

La finalul anului 2020, GISAID a cuprins peste 360.000 de genomuri SARS-CoV-2 din peste 140 de țări. Dar cele mai multe dintre acestea au introdus un număr mic de genomuri. Două excepții sunt Regatul Unit, care a introdus 45% dintre secvențe și Danemarca (7%). În noiembrie 2020, COG-UK a atins un număr record de 100.000 de genomuri virale secvențiate după 9 luni de eforturi susținute. Două luni mai târziu, numărul s-a dublat.

- **Marea Britanie secvențiază aproape jumătate dintre genomurile raportate la nivel global**

Pe 14 decembrie, autoritățile din Marea Britanie au raportat către Organizația Mondială a Sănătății (OMS) **identificarea unei noi variante SARS-CoV-2** (*SARS-CoV-2 VUI 202012/01, cluster filogenetic B.1.1.7*) prin secvențiere genomică. Primul Ministru al Marii Britanii, Boris Johnson, a sporit restricțiile în perioada Crăciunului și a solicitat autorităților din alte țări europene să interzică călătoriile din Marea Britanie.

Secvențierea SARS-CoV-2 se desfășoară în Marea Britanie la o rată mult superioară celorlalte state, iar rezultatele informează în timp real deciziile de sănătate publică. Până la data de 2 februarie 2020 au fost secvențiate 214.039 de genomuri virale. În decembrie 2020, COG-UK secvențiază între 8000 și 10.000 de genomuri SARS-CoV-2 pe săptămână.

Identificarea și caracterizarea genomică rapidă și eficientă a noii variante SARS-CoV-2 au fost posibile datorită lansării, în martie 2020, a celui mai mare proiect din lume de secvențiere a SARS-CoV-2 – Consorțiul COVID-19 Genomics UK (COG-UK). COG-UK are scopul de a asigura secvențierea pe scară largă și rapidă a genomurilor virale cu scopul de a susține centrele NHS și guvernul Regatului Unit. COG-UK este dezvoltat printr-un parteneriat între cele 4 agenții de sănătate publică din cadrul NHS, Wellcome Sanger Institute și 12 parteneri din sfera academică ce asigură capacitatea de secvențiere și analiză. Proiectul a beneficiat de o finanțare de 20 milioane de lire sterline.

Acțiunile consorțiului se derulează pe baza a doi piloni principali:

- **Pilonul 1** (laboratoare regionale) – cuprinde laboratoarele de diagnostic din spitale care sunt concentrate pe testarea probelor de la pacienții internați.
- **Pilonul 2** (probe care ajung la Wellcome Sanger Institute) – cuprinde „superlaboratoarele” (*Lighthouse Lab*) care primesc probe de la persoane care nu sunt spitalizate precum și de la persoane instituționalizate

Pentru procesarea datelor, consorțiul a apelat la **MRC CLIMB (Cloud Infrastructure for Microbial Bioinformatics)**, un sistem de bioinformatică deja setat în Regatul Unit care a analizat un volum mare de date într-un timp rapid. Datele genomice sunt analizate împreună cu cele clinice și epidemiologice, ajutând la ghidarea politicilor și intervențiilor de sănătate publică. De asemenea, sunt realizate mai multe tipuri de studii:

- **GenOMICC (Genetics Of Mortality In Critical Care)** este un proiect de cercetare open-source care are ca scop înțelegerea factorilor genetici care determină forme severe de COVID-19. Proiectul își propune să secvențieze 100.000 de genomuri de la pacienții aflați în stare gravă. Datele genomice de la aceste persoane vor fi analizate alături de date legate de genomul viral din cadrul COG-UK. O analiză asupra genomurilor înscrise în GenOMICC, care a inclus 2.200 de pacienți internați în secțiile ATI, a identificat 5 gene asociate cu formele severe de boală.
- **COG-UK Hospital Onset Covid-19 Infection (HOCI)** este un studiu de fază III, prospectiv care evaluează avantajele secvențierii genomice pentru prevenția răspândirii SARS-CoV-2 la nivelul spitalelor din sistemul NHS. Studiul are două obiective: determinarea beneficiilor secvențierii rapide (sub 48h) a genomului viral pentru a identifica infecțiile nosocomiale COVID-19 și pentru a determina impactul secvențierii asupra reducerii focarelor COVID-19.

COG-UK a furnizat rezultate rapide asupra evoluției SARS-CoV-2, cu impact important la nivel local, regional, național dar și internațional. Acesta este o dovadă că decalajul dintre cercetare și implementarea rezultatelor pentru beneficiul cetățenilor poate fi redus semnificativ. COG-UK intră într-o nouă etapă, în care are obiectivul de a își dubla capacitatea de secvențiere, de a consolida infrastructura dezvoltată și va beneficia de un nou buget de 12,2 milioane de lire sterline din partea Department for Health and Social Care Testing Innovation Fund³¹.

³¹ https://www.cogconsortium.uk/news_item/what-next-for-cog-uk-following-new-funding-of-12-2-million/

- **Danemarca, model pentru Uniunea Europeană**

Pe 23 decembrie, în Danemarca au fost identificate 33 de cazuri de infecție COVID-19 cu varianta B.1.1.7 care a fost observată inițial în Regatul Unit. **Danish Covid-19 Genome Consortium (DCGC)**, un proiect danez, dedicat sprijinirii autorităților din domeniul sănătății publice pentru a monitoriza răspândirea infecției SARS-CoV-2, a fost înființat în martie 2020 prin colaborarea dintre Aalborg University, Statens Serum Institute, Hvidovre Hospital și Aalborg University Hospital. Până la data de 19 ianuarie 2020 au fost secvențiate peste 40.000 de probe, iar 28.000 au fost introduse în GISAID. În luna decembrie, 10% dintre probele pozitive erau secvențiate constant. Pe 25 decembrie, Danemarca a intrat în carantină națională, după ce s-a înregistrat un număr record de 3.692 cazuri pozitive. Durata carantinei a fost extinsă, având în vedere modul în care se răspândește pe teritoriul țării noua variantă SARS-CoV-2 (B.1.1.7). În ultima săptămână din 2020, noua variantă reprezenta 2,4% dintre cazurile COVID-19 pozitive. În prima săptămână din 2021, aceasta reprezenta 3,6%. La data de 11 februarie, aceasta a ajuns la 27%. În contextul unei creșteri rapide a incidenței cazurilor cu noua variantă, în Danemarca s-a luat decizia secvențierii genomice pentru fiecare caz pozitiv³².

- **Africa de Sud**

În iunie 2020, a fost lansată inițiativa Network for Genomic Surveillance in South Africa, pornind de la modelul COG-UK, cu scopul de a secvenția cel puțin 10.000 de genomuri SARS-CoV-2. Deși Africa de Sud nu a avut resursele necesare pentru a secvenția un număr mare de genomuri, au stabilit obiectivul de a prelucra un număr de 50-100 de probe aleatorii pe săptămână. Cu toate acestea, efortul a rezultat în identificarea, în luna noiembrie, în provincia Eastern Cape, a variantei 501Y.V2 (P2), care avea 8 mutații la nivelul genelor care codifică proteina spike. Aceasta s-a corelat cu o creștere importantă a cazurilor în regiuni – 90% dintre infecții erau cauzate de noua variantă. Rezultatele au fost împărtășite cu COG-UK și cu restul lumii. Identificarea noilor variante s-a tradus prin introducerea de restricții suplimentare. Modelul de supraveghere epidemiologică din Africa de Sud este reprezentativ pentru țările cu venituri mici și mijlocii, care au resurse limitate pentru proiectele de secvențiere³³.

Inițiativa din Africa de Sud a avut la bază 6 piloni esențiali:

- Secvențierea într-un interval de timp apropiat de colectarea probei
- O abordare de tip *platform-agnostic*
- Sprijinirea priorităților locale de sănătate publică
- Asigurarea împărtășirii datelor într-un proces rapid și etic
- Crearea sistemelor bioinformatic pentru procesarea și analiza locală a datelor
- Elaborarea rapidă de rapoarte pentru informarea autorităților

În plus, concomitent s-a investit în programe de formare pentru cercetătorii și personalul medical din Africa de Sud.

³²https://www.washingtonpost.com/world/europe/uk-variant-covid-denmark/2021/01/22/ddfaf420-5453-11eb-acc5-92d2819a1ccb_story.html

³³ <https://www.nature.com/articles/d41586-021-00065-4>

Propunere pentru PNRR

Centrul Național de Date și Analiză Genomică pentru Sănătatea Publică **- Sistem de supraveghere prin secvențiere genomică a virusului** **SARS-CoV-2**

În contextul în care România a secvențiat, de la începutul pandemiei, puțin peste 500 de genomuri SARS-CoV-2 și pentru a avea un control crescut asupra pandemiei, recomandăm implementarea unui sistem de supraveghere prin secvențiere genomică a virusului SARS-CoV-2, la nivel național, coordonat de un **Centru Național de Date și Analiză Genomică pentru Sănătatea Publică**.

Setarea **Centrului Național de Date și Analiză Genomică pentru Sănătatea Publică** este în linie cu solicitările Comisiei Europene privitoare la tranziția digitală din cadrul Mecanismului de Redresare și Reziliență al Comisiei Europene³⁴ și cu recomandările European Semester³⁵ (*“Consiliul RECOMANDĂ ca, în 2020 și în 2021,România să întreprindă acțiuni astfel încât: Să direcționeze cu prioritate investițiile către tranziția ecologică și digitală, în special către infrastructura de servicii digitale”*).

Recomandăm ca implementarea și governanța **Centrului Național de Date și Analiză Genomică pentru Sănătatea Publică** să fie realizată de un **Consortiu** la care să participe:

- institute de cercetare cu experiență în domeniul secvențierii genomice a SARS-CoV-2 (Institutul Cantacuzino, Institutul Matei Balș)
- Rețeaua de Genetică Medicală
- una sau mai multe Universități de Medicină și Farmacie
- organizații neguvernamentale relevante (Societatea Română de Bioinformatică, Societatea Română de Genetică Medicală, Centrul pentru Inovație în Medicină)
- entități private relevante (furnizori de servicii medicale, furnizori de soluții de secvențiere, furnizori de soluții ICT)
- cu rol de observator: reprezentanți ai unor consorții similare (Danemarca, Marea Britanie)

Recomandăm implicarea în proiect, încă de la început, a unor experți din consorțiile similare din Danemarca și Marea Britanie.

³⁴https://ec.europa.eu/info/business-economy-euro/recovery-coronavirus/recovery-and-resilience-facility_e

³⁵ <https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:52020DC0523&from=EN>

Obiectivele pentru anul 2021:

- Secvențierea genomului SARS-CoV-2 într-un interval scurt după ce se realizează colectarea probei (rolul important al Rețelei de Genetică Medicală)
- Maparea și utilizarea tuturor platformelor de secvențiere disponibile în prezent, cu condiția asigurării calității procesului (operatori publici și privați)
- Crearea infrastructurii de date (cloud) și a infrastructurii bioinformaticice necesare pentru procesarea datelor la nivel local și analiza datelor
- Împărtășirea rapidă a datelor obținute prin secvențiere cu bazele de date internaționale, respectând principiile etice
- Formarea specialiștilor necesari pentru implementarea și extinderea programului
- Elaborarea de rapoarte pentru autoritățile de sănătate publică, în mod continuu sau ori de câte ori sunt solicitate

Indicatori:

- Număr de genomuri secvențiate
- Număr de genomuri împărtășite cu bazele internaționale
- Număr specialiști care au accesat programele de training
- Rapoarte către autoritățile de sănătate publică
- Articole științifice publicate
- Apariții media cu scop de informare și educare (diseminare)

