

CORONA ONDER DE LOEP (30)

Krantenkoppen

In *De Volkskrant* van 21 februari jl. schrijft Rosanne Hertzberger halverwege haar column:

B.1.1.7 maakt pas op de plaats.

Als dit de kop van haar bijdrage zou zijn geweest, zouden velen onder ons waarschijnlijk geen enkel idee hebben gehad waarover het stukje zou gaan. Maar dat was dus niet het geval. Het opschrift van deze column luidde:

Als onzichtbare lettertjes veranderen in een virus.

Een dergelijke aankondiging zou de lezer nieuwsgierig kunnen maken naar de overgang van 'onzichtbare lettertjes' in een virus.

Als je hier opheldering over wil raad ik je aan verder te lezen. Hertzberger waagt zich niet aan een nadere uitleg.

En de drager van codenaam B.1.1.7? Wel, zijn aanwezigheid in onze samenleving is nog steeds de belangrijkste reden voor de slechts geringe versoepeling van de ons opgelegde vrijheidsbeperkingen.

Op 18 februari kopte het dagblad *TROUW*:

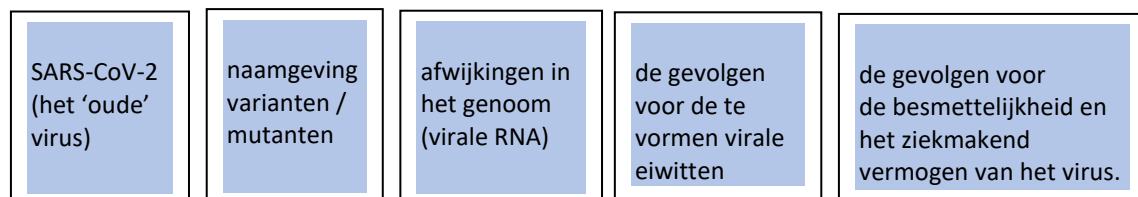
E484K heeft een zorgelijke list verzonnen.



Zou je, zonder de begeleidende tekst te lezen, weten wie of wat E484K is? Gaat hierachter soms een Russische spion schuil of een undercoveragent in de Amsterdamse onderwereld? Ik denk dat de lezer van dit artikel eerst de tekst moet doornemen voordat hij achter de ware identiteit komt. Maar het moet gezegd, ook ná het lezen van het tamelijk omvangrijke verhaal weet je niet *waarom* de oorspronkelijker naamgever specifiek gekozen heeft voor de letters E en K en voor de cijfercombinatie 484!

Ik heb me voorgenomen hierin duidelijkheid te verschaffen en weet vrijwel zeker dat je deze en verwante codes kunt ontcijferen, nadat je de afleveringen 30 en 31 hebt gelezen. Daarvoor is wél een aantal stappen nodig.

Uitgaande van het oorspronkelijke virus zal ik bij de uitleg van deze geheimtaal dít pad van links naar rechts volgen:



De voorspelling is dat de coronapandemie op een goed moment zal uitdoven, maar ook dat de mensheid daarna niet zomaar bevrijd is van het coronavirus en zijn mutanten. Mogelijk dienen bepaalde bevolkingsgroepen jaarlijks tegen deze virussen te worden gevaccineerd, volgens een programma dat te vergelijken is met dat van de griepvaccinatie. Het is dan ook niet zo gek om wat meer kennis op te doen over deze infectie dan die je uit de media kunt halen.

Mijn advies is dan ook: vooral doorlezen!

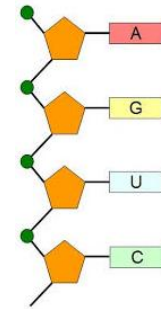
De nucleotiden-volgorde

Je zult het al wel vermoeden hebben: de codenamen slaan op mutanten van SARS-CoV-2.

Oorzaak van mutaties zijn wijzigingen in het genetische materiaal van het virus, dus in dit geval afwijkingen in het RNA-molecuul.

Voor de verklaring van de code B.1.1.7 én de daarbij passende mutaties die zich in het enkelstrengs (+)RNA voordoen is het verhelderend het genoom van het virus onder de microscoop te leggen.

Het genoom van de coronavirussen, dus het complete genetische materiaal, behoort tot de grootste RNA-strengen onder de RNA-virussen. Het omvat bijna 30.000 basen. Op de tekening rechts zien we een héél klein stukje van een denkbeeldig genoom, met links de suiker (ribose; lichtbruine vijfhoeken) en de fosfaatgroep (groene bolletjes) en aan de rechterkant de basen arginine (A), guanine (G), uracil (U) en cytosine (C). De combinatie: **base + ribose + fosfaatgroep** noemen we een nucleotide (afl. 4 en 5). We zien hier dus vier aan elkaar geknoopte nucleotiden.



Tijdens de replicatie van het virale RNA in de menselijke cel worden per 24 uur vele duizenden kopieën van dit nucleïnezuur gevormd. We lazen eerder dat hierbij een enkele keer verkeerd wordt afgelezen, zodat er een fout binnensluipt in de structuur van het RNA: in plaats van bijvoorbeeld adenine (A) wordt bij toeval cytosine (C) ingebouwd. Het gevolg kan zijn dat dit foute RNA de overhand krijgt en als variant (mutant) een eigen leven gaat leiden. Als dat gebeurt kan het zijn dat dit nieuwe virus besmettelijker of dodelijker is dan het moeder-virus. Maar ook is er de mogelijkheid dat de variant geen enkele betekenis voor onze gezondheid meer heeft. Alleen als het nieuwgevormde virusdeeltje in evolutionair opzicht een voorsprong heeft op de virusdeeltjes in zijn omgeving kan het gaan domineren en de andere verdringen (*'the survival of the fittest'*). Die voorsprong kan zich uiten in een doelmatiger replicatie en/of het vermogen beter zijn gastheer te besmetten.

Als we wat meer zouden willen weten van de mutanten van het oorspronkelijk SARS-CoV-2, dan dienen we ons dus in eerste instantie bezig te houden met mogelijke afwijkingen in het virale RNA-molecuul.



Tussen haakjes: het oorspronkelijke SARS-CoV-2, zoals dit in december 2019 in China de pandemie in gang zette, wordt in de literatuur met de algemene term *'wild-type'* aangeduid. We hebben dus bij virussen te maken met een wild-type en zijn mutanten (varianten).

Al vanaf het begin van de coronapandemie houden virologen zich bezig met onderzoek dat erop gericht is mutaties op te sporen. Hiertoe worden bij corona geïnfecteerde personen in veel landen monsters (neusvocht) ten behoeve van dergelijke studies afgenomen. De techniek om het genoom te analyseren en de nucleotiden-volgorde (de nucleotiden-sequentie) in kaart te brengen wordt al vele jaren routinematig toegepast en vormt dus geen enkel probleem bij dergelijk onderzoek.

Chaos in de naamgeving

In december 2020 werd voor het eerst in ZO-Engeland (Kent) bij een aantal corona-positieve patiënten een dominante, agressieve variant van het oude, vertrouwde SARS-CoV-2 aangetroffen. Hij kreeg in eerste instantie naar de streek de naam Kent-variant. Al snel werd hij ook in andere delen van het VK ontdekt. Wetenschappers deelden deze mutant in een apart fylogenetisch cluster in.¹ Wat is dat? In zijn algemeenheid geldt dat de virussen van zo'n cluster dezelfde genetische kenmerken hebben, maar dat die kenmerken net iets anders zijn dan die van de inmiddels bekende virusstam(men).

Deze Kent-variant werd aanvankelijk aangeduid als VUI-202012/01 (VUI = *Variant Under Investigation*), maar, mede gezien de snelle verspreiding, weldra genaamd VOC-202012/01 (VOC = *Variant of Concern*). Ook was de code 20I/501Y.V1 in gebruik. In ons land werd het in het dagelijks taalgebruik de Britse of Engelse variant.

¹ Het bijvoeglijke "fylogenetisch" komt van het zelfstandig naamwoord fylogenie = leer der afstamming.

Tegen een dergelijke naamgeving wordt door sommige auteurs bezwaar gemaakt. Het is namelijk niet helemaal uitgesloten dat VOC-202012/01 op een kwade dag uit een ander land naar het VK werd overgebracht. Bovendien zou deze naamgeving (Engelse variant) ook enig verwijt kunnen inhouden richting de Engelsen.

We herinneren ons dat Donald Trump met nadruk sprak van 'het Chinese virus,' met als duidelijk waarneembare ondertoon een sneer naar de bestuurders van de Volksrepubliek China. Inmiddels kennen we de Zuid-Afrikaanse en Braziliaanse variant en nog een aantal andere. Er valt veel voor te zeggen om de naam van een virus niet te laten verwijzen naar een bepaald land of streek. Nog steeds spreken we van 'De Spaanse griep' die in 1918 wereldwijd tientallen miljoenen slachtoffers maakte. De ontstaanswijze van deze destijds verwoestende influenza-pandemie had echter niets met Spanje van doen.



De PANGOLIN nomenclatuur

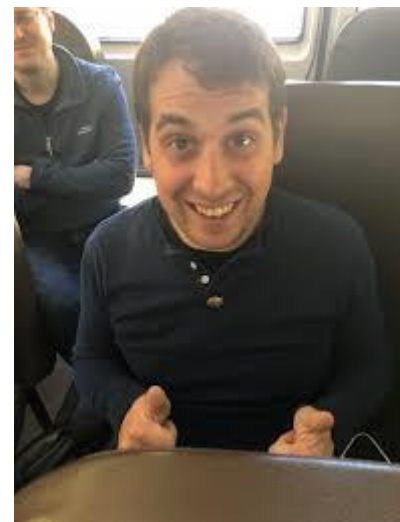
De namen VOC-202012/01 en Britse/Engelse variant werden in wetenschappelijke kringen dan ook grotendeels in de ban gedaan. Er was behoefte aan een nieuwe wijze van naamgeving met als basis een logisch indelingssysteem. Het werd de hoogste tijd om orde in de chaos te scheppen.

Andrew Rambaut, een Britse evolutiebioloog, nam de uitdaging aan. Hij ontwikkelde in het voorjaar van 2020 een fylogenetisch raamwerk (een virusstamboom) ten behoeve van de benoeming en classificatie van het inmiddels grote aantal afstammelingen van het oorspronkelijke virus SARS-CoV-2.

Halverwege 2020 kon worden vastgesteld dat sinds het begin van de coronapandemie al talrijke nieuwe genomsequenties waren ontstaan. Maar er bestond dus geen goed systeem om aan al die varianten van SARS-CoV-2 volgens een logisch model een naam te geven. Ook was het lastig ze volgens goed gedefinieerde lijnen te ordenen. Rambaut had behoefte aan een coherent raamwerk dat logische naamgeving mogelijk zou maken.

De evolutiebioloog vond het de hoogste tijd om tot een werkbaar systeem te komen, een systeem dat moest passen in de genetische diversiteit van het moment. Hij ontwikkelde een nieuwe nomenclatuur die past in een consistent fylogenetisch raamwerk. Het staat bekend onder de namen PANGOLIN lijn en PANGOLIN nomenclatuur.²

Wij buigen ons over de systematiek die Rambaut hierbij volgde. Houd hierbij in gedachte dat de volgorde (sequentie) en identiteit van de bouwstenen (de nucleotiden) in het genoom (RNA) belangrijke kenmerken zijn op grond waarvan hij een (nieuwe) variant van SARS-CoV-2 een naam geeft en indeelt.

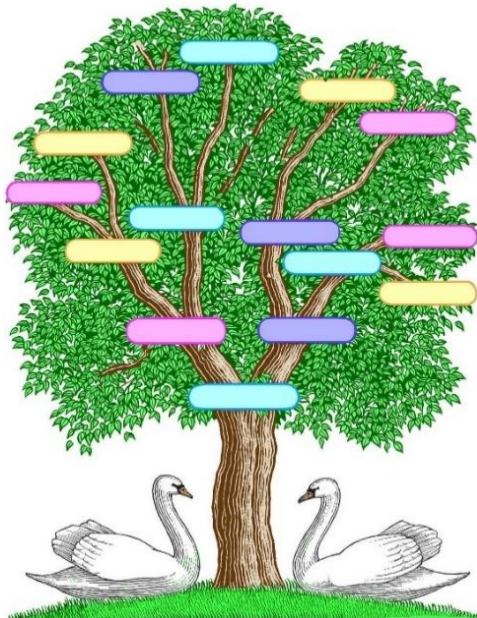


Andrew Rambaut

Omdat helemaal in het begin van de uitbraak van SARS-CoV-2 de nucleotiden-sequentie van het oorspronkelijke virus reeds diverse keren was vastgesteld, lag het voor de hand om dat patroon als uitgangspunt te nemen bij de classificatie van de soorten die er inmiddels uit waren voortgekomen.

² PANGOLIN staat voor **Phylogenetic Assignment of Named Global Outbreak Lineages**; ook komen we het gebruik van PANGO in plaats van PANGOLIN tegen.

Rambaut wees bij de constructie van het raamwerk deze initiële sequentie (*root-sequence*) dan ook als referentiekader aan. Bovendien nam hij alleen de virussen in beschouwing die aanleiding gaven tot een aanzienlijke verspreiding in een bepaalde regio en die tegelijkertijd ook daadwerkelijk ziekte veroorzaakten. Een voorbeeld hiervan is het coronavirus dat de omvangrijke epidemie in Noord-Italië (febr. 2020) veroorzaakte en die zich razendsnel over Europa verspreidde.



Dit virus voldeed aan de criteria van opname in zijn systeem en moest er dus een plaats in krijgen. Rambaut stelde voor om de naam van de belangrijkste oer-stam(men) met een enkele letter te laten beginnen en vervolgens de (zij)takken van de stamboom aan te duiden met cijfers die achter de letter moesten worden geplaatst. De wortel van de stamboom vertegenwoordigt als het ware het denkbeeldige oer-virus. Bovengronds verschijnen uit de wortel twee dikke stammen, met de namen 'stam A' en 'stam B' (in de Engelse taal: *lineage A* en *lineage B*). Deze twee 'uitlopers' onderscheiden zich van elkaar in slechts kleine verschillen in de nucleotiden. De virussen die tot cluster A of cluster B behoren zijn de SARS-CoV-2 virussen die zich in China in december 2019 en januari 2020 vertoonden. In korte tijd ontwikkelden zich allerlei nieuwe varianten (mutanten) als (zij)takken van de moeder-stammen A en B. Alle codenamen moesten dus in principe of met de letter A of B beginnen.

De zijtakken van A en B

In mei 2020 bestudeerde Rambaut bijna 28.000 complete genomen van het SARS-CoV-2 die inmiddels in een database waren ondergebracht. Hij definieerde nieuwe SARS-CoV-2 takken die ofwel uit stam A of uit stam B ontsproten waren (bijv. tak A.1, tak A.2, tak B.1).

Een verdere onderverdeling wordt noodzakelijk geacht als een variant van SARS-CoV-2, die bijv. in feite tot de soort B.1 behoort, niet meer volledig aan de vereiste criteria van B.1 voldoet. De nieuwe tak krijgt dan de naam B.1.1. Een zijtak van B.1.1 wordt dan bijvoorbeeld B.1.1.2.

In de zomer van 2020 werden aan de hand van dit classificatiemodel door Rambaut en zijn medewerkers 81 takken gedefinieerd. Zes daarvan konden worden afgeleid van A (aangegeven met A1-A6) en nog twee als onderverdeling van A1. Ook konden toen zestien takken worden benoemd die uit stam B ontsproten. Het zijn er inmiddels véél meer.

De Engelse variant diende dus een eigen naam te krijgen. Het werd, gezien de fylogenetische kenmerken, **B.1.1.7**.³



Op 5 maart 2020 werd in **Brazilië** voor de eerste keer een stam waargenomen die (later) met de codenaam B.1.1.28 een plaats kreeg in de PANGOLIN stamboom.

Ruim een maand later ontstond uit deze stam een variant die aanvankelijk code **B.1.1.28.1** kreeg toebedeeld. Maar aangezien de uitbreiding de codenaam wel erg lang maakte, besloot men tot de codenaam **P.1**.⁴

Dus: de Braziliaanse variant waar we in de krant over lezen mag voor ons die naam hebben, maar bezit in de wereld van virologen en epidemiologen de codenaam P.1.

³ We verdiepen ons niet in alle details van het systeem; zo maken we ons niet druk over het feit dat het cijfer achter B.1.1 een 7 is. Daar zullen Rambaut en medewerkers hun argumenten voor hebben.

⁴ Inmiddels bestaat ook stam P.2.

Op 23 december 2020 werd een tot dan toe onbekende variant van SARS-CoV-2 aangetroffen bij twee mensen die van **Zuid-Afrika** naar het VK reisden. Het kreeg bij die gelegenheid de naam 501Y.V2 en ook 20H501Y.V2. In de weken die volgden verspreidde het virus zich over een aantal landen in Afrika, Europa en Amerika. Volgens de PANGOLIN nomenclatuur werd de Zuid-Afrikaanse variant: **B.1.351**. Het is dus eveneens een aftakking van B.1.



In ons land werd op 7 januari jl. bekend gemaakt dat bij nader onderzoek van het coronavirus waarmee iemand al op 22 december 2020 besmet was, de Zuid-Afrikaanse variant was.

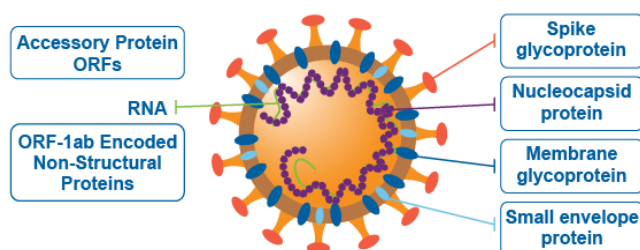
Er zijn ook nog een paar andere classificatiesystemen in gebruik. Zo bestaat de zogenaamde GISAID nomenclatuur. In dit systeem draagt SARS-CoV-2 de naam hCoV-19.

Ook is er de Nextstrain nomenclatuur. Voorbeelden hiervan zijn de codenaam 20I/501Y.V1 voor de Britse en 20H/501Y.V2 voor de Zuid-Afrikaanse variant.

Het genoom onder de microscoop

Het genoom (het RNA) van SARS-CoV-2 is een groot molecuul, opgebouwd uit nucleotiden. Nucleotiden zijn onderdelen van genen. Een gen bestaat uit meerdere codons. Dat zijn aaneenschakelingen van drie nucleotiden. Een codon (triplet) bepaalt tijdens de vorming van een peptide (eiwit) in het ribosoom op welk moment welk aminozuur aan de groeiende ketting moet worden geregen. Die boodschap (dat recept) wordt door het messenger RNA aan het ribosoom overgebracht (afl. 7).

Om de situatie goed voor ogen te hebben, halen we 'de plattegrond' van het coronavirus er nog even bij. Mutaties in het RNA van SARS-CoV-2 hebben vooral een verstrend effect op de synthese van de virale eiwitten. Over welke eiwitten hebben we het dan ook al weer?



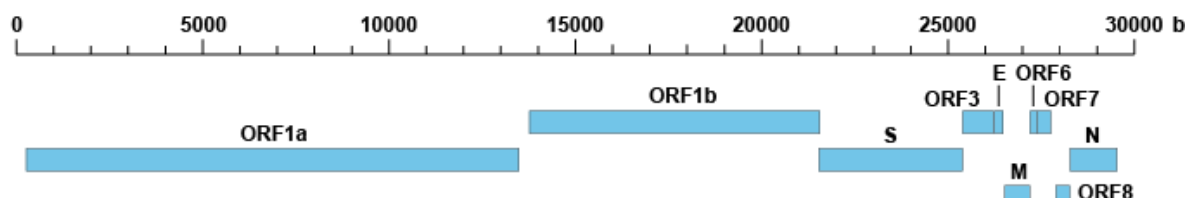
We herinneren het ons wellicht nog.

We onderscheiden:

- ◆ niet-structurele eiwitten; dat zijn eiwitten die niet bijdragen aan de structuur van het deeltje; het zijn vooral virale enzymen;
- ◆ structurele eiwitten; hiertoe behoren het envelopeiwit (E-eiwit), het membraaneiwit (M-eiwit), het nucleocapside-eiwit (N-eiwit) en in het bijzonder het spijkereiwit (S-eiwit).

Om inzichtelijk te maken waar zich in de RNA-streng mutaties (kunnen) voordoen, is het handig om het RNA van SARS-CoV-2 volgens een rechte lijn uit te rollen.

Als we dit doen, dan ontstaat de volgende schematische voorstelling.



Langs de RNA structuur, weergegeven met blauwgekleurde balken, is een meetlat gelegd. Hierop zijn van links naar rechts de posities van de nucleotiden in het genoom aangegeven. Het zijn er in totaal bijna 30.000. We kijken dus naar de samenstelling van het RNA van SARS-CoV-2, vanaf het begin (links) tot het einde (rechts) van de figuur. De getallen komen dus ook overeen met het aantal basen (A,C,G en U) die men in de nucleotiden tegenkomt. Het binnenkomen van één virusdeeltje in een van

onze cellen betekent dat een dergelijke streng van circa 30.000 nucleotiden aan het cytoplasma wordt afgegeven. Deze zetten de betrokken cel er vervolgens toe aan alle bouwelementen (voornamelijk eiwitstructuren) te maken die nodig zijn om, helemaal aan het einde van het replicatieproces, het nieuwe virusdeeltje te vormen. Aan dit proces werkt onze besmette cel dus driftig mee, terwijl die volgzzaamheid met grote waarschijnlijkheid zijn eigen ondergang inluidt.

De blauwe balken stellen de verschillende gebieden (domeinen) voor die men op grond van hun functie volgens dit patroon een plaats geeft. De verschillende genen die er deel van uitmaken coderen dus voor verschillende chemische structuren. ORF staat voor 'Open Reading Frame', een aanduiding die ik niet nader zal toelichten. Ze nemen de grootste plaats in en coderen voor de niet-structurele eiwitten van het virus en dragen de codes ORF1a, ORF1b, ORF3, ORF6, ORF7 en ORF8. Wij richten ons op de balken met de letters N, M, E en S. Domein N codeert voor de eiwitten van het nucleocapside, domein M codeert voor het membraaneiwit, domein E doet dat voor de productie van het envelopeiwit. Onze bijzondere belangstelling gaat echter uit naar domein S in het genoom van SARS-CoV-2. De genen van dit domein coderen voor de bestanddelen van de (uit)stekels, de spijkers en dus voor het complete S-eiwit.

We zien dat dit gebied de posities (ca) 21500 tot (ca) 25.300 omvat. We zijn vooral in dit gebied geïnteresseerd, omdat uitsluitend door middel van de spikes het coronavirus de cellen van de gastheer kan binnenkomen. Mochten zich mutaties in dit deel van het RNA van het virus voordoen, dan is de kans niet gering dat er afwijkingen in de spikes ontstaan. Deze zouden de penetratie van het virus in de cel kunnen bemoeilijken (gunstig) of juist vergemakkelijken (ongunstig). Ook zou het kunnen zijn dat een bepaalde mutatie in het S-eiwit de nieuwe variant de eigenschap meegeeft zich in meer of mindere mate te onttrekken aan ons (reeds getrainde) afweersysteem. En inderdaad, daar hebben we in de praktijk ernstig rekening mee te houden! Hierover later meer.

In de volgende aflevering zal ik het verband leggen tussen een mutatie in het genoom én de aminozuurvolgorde in het S-eiwit. Hiertoe zal ik het S-eiwit uitrollen tot een recht pad en je uitnodigen voor een alleszins boeiende wandeling. We zullen daarbij hier en daar onze voeten zetten op gemuteerde delen van het S-eiwit. Op onze tocht zullen we E484K tegenkomen, maar ook D614G en N501Y. Het is dus nog even afwachten wat dit alles te betekenen heeft.

Arijan Porsius
03-03-21



"Volgens mij is een van de grote fouten die de mensen maken wanneer ze proberen iets te begrijpen, dat ze zekerheid willen. De zoektocht naar kennis voedt zich niet met zekerheden: hij voedt zich met een radicale afwezigheid van zekerheden".

Een uitspraak van de Italiaanse natuurkundige Carlo Rovelli (geb. 1957) die zich specialiseerde in de kwantumfysica.