

Studie ermittelt Sars-CoV-2-Variante mit Lücke im Erbgut

Analysesoftware konnte fehlende Gen-Bausteine bislang nicht feststellen

Automatisierte Gen-Analysen von Sars-CoV-2-Proben übersehen durchgängig, wenn im Erbgut des Virus durch Mutationen Gen-Abschnitte verschwunden sind. Das zeigt eine neue Studie von Forschenden des Centrums für Biotechnologie (CeBiTec) der Universität Bielefeld und des Evangelischen Klinikums Bethel, eine der Trägerkliniken des Universitätsklinikums OWL.

Analysesoftware konnte fehlende Gen-Bausteine bislang nicht feststellen. Das Studienteam konnte in seinen analysierten Proben nachweisen, dass ein großer Teil des Genabschnitts ORF8 fehlte. Von dieser Genregion wird vermutet, dass sie dazu beiträgt, die Abwehrreaktion im menschlichen Körper zu verzögern. Fehlt sie, besteht die Chance, dass das Virus weniger pathogen wird, also weniger schwerwiegende Krankheiten verursacht. Das Forschungsteam hat seine Studienergebnisse in dem Journal „Viruses“ veröffentlicht.

„Wenn in Krankenhäusern wie auch in Schulen mit Routinekontrollen erfasst wird, ob sich Personen mit dem Virus infiziert haben, geht es hauptsächlich darum, weitere Ansteckungen einzudämmen“, sagt der Studienleiter Professor Dr. Jörn Kalinowski, Genetiker am CeBiTec. Werden die Proben mit den verbreiteten PCR-Tests in Laboren analysiert, soll nicht nur festgestellt werden, ob eine Infektion vorliegt. Liegt ein Infektionsfall vor, wird auch untersucht, um welche Variante des Virus es sich handelt. „Dafür reicht es aus, einzelne charakteristische Genabschnitte zu identifizieren, die für die verbreiteten Virusvarianten typisch sind.“ Aktuell kommt bei solchen Analysen europaweit meist dasselbe Ergebnis heraus: die Deltavariante – sie ist weitaus ansteckender als andere Varianten.

„Weil nur wenige Genabschnitte nötig sind, um eine gängige Virusvariante zu erkennen, nehmen Labore es in der Regel in Kauf, wenn andere Genabschnitte nicht identifiziert werden können“, sagt Kalinowski. Ein Grund dafür, dass das Genom nicht komplett bestimmt werden kann, ist zum Beispiel die mangelhafte Aufbereitung der Probe. Außerdem kommt es häufig vor, dass die Analysesoftware einzelne Nukleotide – die Gen-Bausteine des Viren-Erbguts – nicht optimal erkennt. Statt der Buchstaben A, T, G und C, die genutzt werden, um die Gen-Bausteine des Virenerbguts wiederzugeben, schreibt die Software dann den Buchstaben N in die Gensequenz.

Software dokumentiert fehlende Genabschnitte bisher missverständlich

Das Bielefelder Forschungsteam hat in seiner Studie herausgefunden, dass dieser Behelf zu einem weitreichenden Problem führen kann. „Durch Mutationen können Varianten von Sars-CoV-2 entstehen, denen längere Genabschnitte fehlen“, sagt der Bioinformatiker Professor Dr. Alexander Sczyrba vom CeBiTec, Co-Autor der Studie. „Wir haben festgestellt, dass die häufig eingesetzte Standardsoftware auch dann Platzhalter in die Gensequenz einträgt, wenn ein ganzer Genabschnitt gar nicht vorhanden ist.“ Dann steht reihenweise der Buchstabe N in der Gensequenz.

„Das ist ein systematischer Fehler“, sagt Jörn Kalinowski. „Denn eine solche Lücke im Genom ist ein wichtiger Anhaltspunkt, wenn es um die künftige Gefährdung durch das Coronavirus geht.“ Ist eine Lücke in einem Genabschnitt, verschwinden auch Eigenschaften, die das betroffene Gen gespeichert hatte. Bei der Vermehrung des Virus werden die Eigenschaften nicht mehr weitergegeben. „Hinzu

kommt, dass an solchen Leerstellen im Genom auch keine Mutationen mehr entstehen können, die das Virus für den Menschen gefährlicher machen.“ Solche fehlenden Genabschnitte können laut Kalinowski mit dazu führen, dass sich Sars-CoV-2 an die Menschen als ihre Wirte anpasst. So wird das Virus dann zwar infektiöser, zugleich aber ungefährlicher. „Das Virus würde damit endemisch werden, tritt also in verschiedenen Regionen regelmäßig auf, so wie es auch bei anderen, lange bekannten Coronaviren der Fall ist, die bei uns heutzutage lediglich harmlose Erkältungskrankheiten verursachen.“

Kenntnis der Virusvariante hilft bei Rekonstruktion von Infektionsketten

Entdeckt haben die Forschenden die Gen-Lücke im Virus durch Analysen von Viren-Proben aus dem Evangelischen Klinikum Bethel (EvKB). Dort werden medizinisches Personal und Patient*innen kontinuierlich auf Infektionen mit dem Coronavirus getestet. Seit April 2020 wertet die Arbeitsgruppe von Kalinowski Proben aus dem Klinikum aus. Für die Analyse erhält sie Extrakte, in denen die Erbsubstanz des Virus enthalten ist. Sie stammen aus Abstrichen positiv getesteter Patient*innen. „Die detaillierte Analyse der Proben ermöglicht uns, Infektionsketten zu rekonstruieren, wenn Infektionen auftreten“, sagt die Oberärztin Dr. Christiane Scherer, Leiterin der Mikrobiologie im EvKB und ebenfalls Co-Autorin der Studie. Auf dem Höhepunkt der zweiten Infektionswelle stellten Scherer und ihr Team im Januar und Februar 2021 ein Infektionscluster fest. Die Infektionen wurden durch aufwendige Screenings und Kontaktverfolgung eingedämmt. Die Virusvariante B.1.1.294 konnte sich auf den Stationen nicht weiterverbreiten.

Die CeBiTec-Analysen bestätigten, dass die Abschottungsmaßnahmen auf den betroffenen Stationen erfolgreich waren. „Wir konnten das so genau sagen, weil wir eine Besonderheit der Virusvariante entdeckt haben: In ihrem genetischen Code fehlen 168 Nukleotide“, berichtet Jörn Kalinowski. Die Gen-Bausteine fehlten in der Genregion „Open Reading Frame 8“ (ORF8). Die genetischen Informationen sind mutmaßlich mit dafür verantwortlich, dass es dem Virus gelingt, die Immunreaktion von Infizierten zu verzögern.

Funktion in Analysesoftware ergänzt

Nachweisen konnten die Wissenschaftler*innen die fehlenden Nukleotide, weil sie – anders als beim Standard-PCR-Test üblich – zusätzlich Nanoporen-Sequenzierung einsetzen. Mit den Spezialgeräten lassen sich längere Genabschnitte bestimmen als mit den üblichen Sequenziermaschinen. Außerdem ergänzten die Forschenden eine frei verfügbare Software zur Gen-Analyse um eine Funktion, die fehlende Nukleotide in Gensequenzen korrekt erkennt und kennzeichnet. „Nur dadurch konnten wir feststellen, dass innerhalb der Genregion ORF8 ein Stück verschwunden war“, sagt Kalinowski.

„Durch diese Analyse ließ sich nicht nur das Cluster in unserem Klinikum feststellen“, sagt Christiane Scherer. „Wir konnten auch absichern, dass die Virusvariante bei uns in eine Sackgasse geraten ist und sich nach der Eindämmung niemand mehr damit angesteckt hat.“

Stammbaum zeigt verwandte Varianten der untersuchten Virus-Mutante

Die CeBiTec-Forschenden wollten wissen, woher die Virus-Variante mit dem verlorenen Genabschnitt stammt. Dafür entnahmen sie Rohdaten aus der zentralen Datenbank der Coronaviren-Varianten und werteten sie mit ihrer eigens entwickelten Software aus. „Wir konnten so ermitteln, wo weitere Vorgänger der von uns untersuchten Variante aufgetreten sind und wo ähnliche Varianten vorkamen“, erklärt Alexander Sczyrba. Das Ergebnis veranschaulichen die Wissenschaftler*innen in einem Stammbaum von Sars-CoV-2. Darin ist zum Beispiel zu sehen, dass ein Vorläufer der Virusvariante vorher in Dänemark entdeckt wurde.

Die CeBiTec-Forschenden wollen auch anderen Wissenschaftler*innen ermöglichen, fehlende Genabschnitte in Sars-CoV-2-Varianten präzise zu ermitteln. Dafür stellen sie ihre Weiterentwicklung der Analysesoftware samt des Quellcodes auf einer einschlägigen Plattform zum Download zur Verfügung.

Um zu klären, welche Funktionen einzelne Gene des Virus haben und wie es sich entwickelt, sei es wichtig, nach weiteren Varianten mit gelöschten Genabschnitten suchen zu können, sagt Jörn Kalinowski. „Doch dafür müssen bundesweit alle Rohdaten von analysierten Corona-Proben verfügbar gemacht werden. Rigide Datenschutzbestimmungen verhindern das momentan leider“, beklagt er.

Originalveröffentlichung:

David Brandt, Marina Simunovic, Tobias Busche, Markus Haak, Peter Belmann, Sebastian Jünemann, Tizian Schulz, Levin Joe Klages, Svenja Vinke, Michael Beckstette, Ehmke Pohl, Christiane Scherer, Alexander Sczyrba, Jörn Kalinowski: Multiple Occurrences of a 168-Nucleotide Deletion in SARS-CoV-2 ORF8, Unnoticed by Standard Amplicon Sequencing and Variant Calling Pipelines. *Viruses*, <https://doi.org/10.3390/v13091870>, erschienen am 18. September 2021.