# Strukturanalyse mit maschinellem Lernen enthüllt Taktiken des SARS-CoV-2-Virus

Categories : Anlagen & Komponenten, Messtechnik & Analytik

Date : 20. September 2021

Bei der Fähigkeit des Virus spielen die Proteine des SARS-Cov-2-Virus eine Schlüsselrolle, die menschliche Immunabwehr auszutricksen und sich in Patientenzellen zu vermehren. Ein internationales Forschungsteam unter Beteiligung der Technischen Universität München (TUM) hat nun den bislang umfassendsten und detailreichsten Überblick aller weltweit verfügbaren 3D-Strukturen der Virusproteine zusammengetragen. Bei der Auswertung mit Methoden der künstlichen Intelligenz kamen überraschende Erkenntnisse zutage.
Wie gelingt es dem SARS-CoV-2-Virus, sich der Immunabwehr zu entziehen und sich in den Zellen von Patienten zu replizieren? Um diese Frage zu klären, hat ein internationales Forschungsteam den bislang umfassendsten Überblick aller bisher verfügbaren Analysen der exakten dreidimensionalen Form der SARS-CoV-2-Proteine  darunter auch das bekannte Spike-Protein  zusammengetragen.
Um diese Übersicht zu zusammenzustellen, verwendete das Team Hochdurchsatz-Maschinenlernen. Dieser Ansatz ermöglicht es, basierend auf Analysen verwandter Proteine, strukturelle Zustände von Coronavirus-Proteinen vorherzusagen. Inzwischen besteht die Datenbank aus 2.060 3D-Modellen mit atomarer Auflösung. Auf der Website Aquaria-COVID (https://aquaria.ws/covid) sind alle Strukturmodelle frei verfügbar.

> Dies bietet einen beispiellosen Detailreichtum, der Forschenden helfen wird, die molekularen Mechanismen der COVID-19-Infektion besser zu verstehen und Therapien zur Bekämpfung der Pandemie zu entwickeln, beispielsweise indem sie potenzielle neue Angriffspunkte für zukünftige Behandlungen oder Impfstoffe identifizieren.
> - Burkhard Rost, Inhaber des Lehrstuhls für Bioinformatik an der TU München

## Die strukturelle Landkarte erschließt das zusammengetragene Wissen

In einem zweiten Teil der Studie wurde ein komplementärer Ansatz verwendet, der als Human-in-the-Loop Machine Learning bekannt ist. Hier wurde eine neuartige, visuelle Schnittstelle erzeugt, die alles zusammenfasst, was derzeit über die dreidimensionale Form von SARS-CoV-2-Proteinen bekannt ist  und was nicht.
Forschende können die visuelle Schnittstelle auch als Navigationshilfe verwenden, um für spezifische Forschungsfragen geeignete Strukturmodelle zu finden. Die Arbeit mit den Modellen hat bereits einige wichtige Hinweise darauf gegeben, wie Coronaviren es schaffen, das Kommando in unseren Zellen übernehmen.

## Wie Coronaviren das Kommando in unseren Zellen übernehmen

Mithilfe der Machine-Learning-Algorithmen identifizierte das Team drei Coronavirus-Proteine (NSP3, NSP13 und NSP16), die menschliche Proteine nachahmen und den Wirtszellen erfolgreich vorgaukeln, sie seien körpereigene Proteine, die im besten Interesse der Zelle arbeiten.
Die Modellierung ergab auch fünf Coronavirus-Proteine (NSP1, NSP3, Spike-Glykoprotein, Hüllprotein und ORF9b-Protein), die Prozesse in menschlichen Zellen zweckentfremden oder stören. Auf diese Weise gelingt es dem Virus, die Kontrolle zu übernehmen, seinen Lebenszyklus zu vervollständigen und sich auszubreiten.

## Verstehen, wie das Virus funktioniert  und wie man es stoppt

Bei der Analyse dieser Strukturmodelle haben wir außerdem neue Hinweise darauf gefunden, wie das Virus sein eigenes Genom kopiert  das ist der zentrale Prozess, der es dem Virus ermöglicht, sich bei Infizierten schnell auszubreiten, sagt Burkhard Rost. Die Erkenntnisse aus unserer Studie bringen uns näher an das Verständnis, wie das Virus funktioniert und was wir tun können, um es zu stoppen.
Je länger das Virus zirkuliert, desto größer ist die Gefahr, dass es mutiert und neue Varianten wie den Delta-Stamm bildet, sagt Sean O Donoghue, Erstautor der Studie und Professor am Garvan Institute in Sydney. Unsere Ressource wird Forschern helfen zu verstehen, wie sich neue Virusstämme voneinander unterscheiden  ein Puzzleteil, von dem wir hoffen, dass es beim Bekämpfen neu auftauchender Varianten hilft.