

# Anreicherungs- workflow zum Coronavirus- Nachweis mit NGS-Systemen von Illumina

Äußerst sensitive Erkennung  
und Charakterisierung gängiger  
Atemwegsviren, einschließlich  
Coronavirus-Stämmen



## Viren-Surveillance mit NGS

Virusinfektionen sind weltweit ein schwerwiegendes Gesundheitsproblem und es treten ständig neue Varianten von Infektionskrankheiten auf. Der Ausbruch des neuartigen Coronavirus (SARS-CoV-2), das sich 2019 vom chinesischen Wuhan aus schnell in zahlreiche Länder ausbreitete, ist ein wichtiges Beispiel für die potenziellen Risiken globaler Viruspanidemien. Coronaviren (CoV) sind eine große Familie von Viren, die Menschen infizieren und milde bis schwerwiegende Atemwegserkrankungen verursachen können, u. a. das MERS-CoV (Middle East Respiratory Syndrome, nahöstliches Atemwegssyndrom) und das SARS-CoV (Severe Acute Respiratory Syndrome, schweres akutes respiratorisches Syndrom). Die Weltgesundheitsorganisation (WHO) bezeichnete die mit SARS-CoV-2 assoziierte Erkrankung (COVID-19) als gesundheitliche Notlage von internationaler Tragweite, was die Notwendigkeit eines schnellen, genauen Virennachweises unterstreicht.<sup>1</sup>

NGS (Next-Generation Sequencing, Sequenzierung der nächsten Generation) bietet eine effektive, neuartige Möglichkeit für das Screening von Proben sowie zum Nachweis von Viren ohne vorherige Kenntnisse des Infektionserregers.<sup>2</sup> Bei der Target-Anreicherung handelt es sich um eine Sequenzierungsmethode, bei der genomische Regionen von Interesse durch Hybridisierung von Nukleinsäuren an targetspezifische biotinylierte Sonden erfasst werden. Die Hybrid-Capture-Target-Anreicherung ermöglicht einen äußerst sensitiven Nachweis ohne die hohe Read-Tiefe, die bei der Shotgun-Metagenomik-Sequenzierung erforderlich ist. Darüber hinaus ermöglicht der NGS-Workflow für die Target-Anreicherung nahezu vollständige Sequenzdaten von Targets und damit Anwendungen wie Variantenanalysen für den Virennachweis oder die Viren-Surveillance.<sup>3</sup> Im Vergleich zu anderen gezielten Sequenzierungsmethoden wie der Amplikonsequenzierung ermöglicht die Anreicherung durch Hybrid-Capture ein deutlich größeres Sondenpanel mit umfassenderem Profiling der Target-Regionen.

Zudem sind die für Hybrid-Capture-Protokolle verwendeten Oligosonden auch in hochgradig mutierten Regionen wirksam, sodass auch Viren erfasst werden können, die sich schnell verändern (z. B. RNA-Viren).

## Virenanreicherungsworkflow

Dieser Anwendungshinweis erläutert einen optimierten Workflow für den Coronavirus-Nachweis und die entsprechende Analyse mit dem Illumina DNA Prep with Enrichment Library Preparation Kit, Virus-Targeting-Panels, bewährter Illumina-Sequenzierung und vereinfachter Datenanalyse (Abbildung 1). Dieser Workflow ist für die Anreicherung viraler DNA- und RNA-Targets aus der Gesamtnukleinsäureextraktion vorgesehen. Zu Beginn des Protokolls steht die Extraktion von Nukleinsäureproben und die anschließende reverse Transkription extrahierter RNA in doppelsträngige cDNA. Die Bibliotheksvorbereitung erfolgt mit Illumina DNA Prep with Enrichment mit IDT for Illumina DNA/RNA UD Indexes. In diesem Protokoll werden DNA und cDNA einer Tagmentierung, Aufreinigung und Voranreicherungsamplifizierung unterzogen. Nach der Amplifikation können bis zu 12 Proben für eine Anreicherungsreaktion gepoolt werden. Hierbei kommt ein Oligo-Panel zum Einsatz, das auf virale Sequenzen abzielt. Im Anschluss an die Sondenhybridisierung erfolgen Sondenerfassung, Anreicherungsamplifikation, Quantifizierung und Sequenzierung.

## Methoden

### Probenvorbereitung

Um die Leistung des Panels für die Viren-Target-Anreicherung zu demonstrieren, wurde eine Probe mit dem deaktivierten CoV-Stamm (OC43) von Microbiologics (QC Sets and Panels: Helix Elite, Katalog-Nr. 8217) analysiert (Tabelle 1). Die Coronavirus-Kulturprobe wurde mit dem QIAGEN QIAmp Viral Mini Kit (QIAGEN,

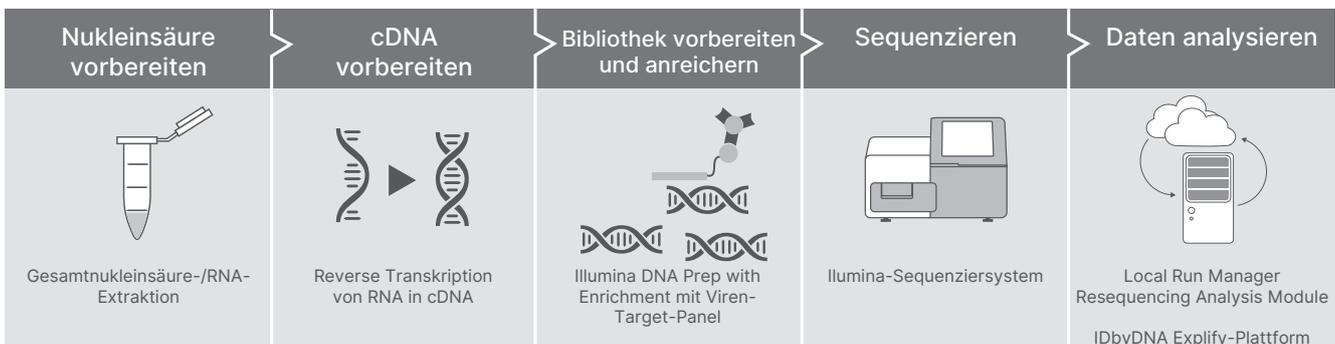


Abbildung 1: Anreicherungsworkflow für den Nachweis von Coronaviren: Der optimierte NGS-Workflow für den Nachweis von Coronaviren umfasst Probenvorbereitung, Bibliotheksvorbereitung, Zielanreicherung, Sequenzierung und Datenanalyse.

Katalog-Nr. 52904) in einer BSL2-Laborumgebung extrahiert. 150 ng extrahierte RNA wurden mit zwei verschiedenen Workflows in cDNA transkribiert: dem für Illumina TruSeq RNA-Reagenzien und dem für das Thermo Fisher Scientific Maxima H Minus Double-Stranded cDNA Synthesis Kit (Thermo Fisher Scientific, Katalog-Nr. K2561). Um Sequenzierungsergebnisse von Patientenproben zu simulieren, wurde die Virusprobe zusätzlich per Spike-in in 95%ige Universal Human Reference(UHR)-Hintergrund-RNA (Agilent Technologies, Katalog-Nr. 74000) gegeben und mit den oben beschriebenen Methoden einer reversen Transkription in cDNA unterzogen (Tabelle 1).

Zusätzlich wurde ein Pool mit vier respiratorischen RNA-Viren und einem DNA-Virus (Tabelle 2) mit dem QIAGEN AllPrep PowerViral DNA/RNA Kit (QIAGEN, Katalog-Nr. 28000-50) in einer BSL2-Laborumgebung extrahiert. Das maximale Volumen der gesamten Nukleinsäure wurde einer reversen Transkription in cDNA (ohne DNase-Behandlung) unterzogen. Hierbei kam ein Workflow zum Einsatz, der von dem für Illumina TruSeq RNA-Reagenzien abgeleitet war. Um Sequenzierungsergebnisse von Patientenproben mit mehreren Pathogenen zu simulieren, wurde die Virenpoolprobe zusätzlich per Spike-in in 95%ige Universal Human Reference(UHR)-Hintergrund-RNA (Agilent Technologies, Katalog-Nr. 74000) gegeben und einer reversen Transkription in cDNA unterzogen (Tabelle 1).

Tabelle 1: Zusammensetzung der zu analysierenden Virusproben

Probe	Zusammensetzung <sup>a</sup>	Referenzgenom
CoVOC43	150 ng CoV OC43-RNA	AY391777.1 Humanes Coronavirus OC43
CoVOC43_95UHR	7,5 ng CoV OC43-RNA und 142,5 ng UHR-RNA	AY391777.1 Humanes Coronavirus OC43
ViralPool	Gesamtnukleinsäure des Viruspools mit maximalem Volumen (8,5 µl)	Siehe Tabelle 2
ViralPool_95UHR	5 % Gesamtnukleinsäure im Viruspool und 95 % UHR-RNA nach Volumen	Siehe Tabelle 2

a. Die empfohlene Mindestmenge der Zugabe von RNA/Gesamtnukleinsäure für die reverse Transkription beträgt 10 ng. Optimale Ergebnisse werden erzielt, wenn die reverse Transkription an frisch extrahierten Nukleinsäureproben durchgeführt wird.

Tabelle 2: Zusammensetzung des Virenpools

Virus	Typ der Nukleinsäure	Referenzgenom
Influenza-A-Virus (H1N1)	RNA	Influenza A/ Michigan/45/2015
Influenza-B-Virus	RNA	Influenza B/ Colorado/06/2017
Humanes Parainfluenzavirus 3	RNA	NC_001796.2
RSV B9320	RNA	AY353550.1
Adenovirus 7	DNA	AY594255.1

### Bibliotheksvorbereitung

Sequenzierungsfähige Bibliotheken wurden mit Illumina DNA Prep with Enrichment (Illumina, Katalog-Nr. 20025524) und IDT for Illumina DNA UD Indexes (Illumina, Katalog-Nr. 20027213)\* unter Verwendung von cDNA aus der CoV-Probe (CoVOC43), der Virenpoolprobe (ViralPool) und beiden Spike-in-Proben (CoVOC43\_95UHR und ViralPool\_95UHR) vorbereitet. Die für die Tagmentierung empfohlene Gesamtzubemenge DNA beträgt 10–1.000 ng.

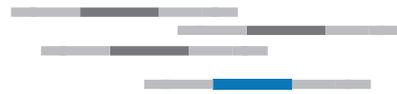
Nach der Amplifikation wurden die Proben zur Anreicherung nach Probenotyp (CoVOC43, CoVOC43\_95UHR, ViralPool und ViralPool\_95HR) in separate Pools aufgeteilt. Anreicherungsreaktionen wurden mit dem Respiratory Virus Oligos Panel (Illumina, Katalog-Nr. 20042472) durchgeführt. Dieses enthält ca. 7.800 Sonden für den Nachweis von Atemwegsviren, aktuellen Grippestämmen und SARS-CoV-2 sowie Humansonden zur Positivkontrolle (Tabelle 3). Die Bibliotheken wurden im Anschluss an die Anreicherung (Abbildung 2) quantifiziert, gepoolt und zur Sequenzierung in das MiSeq™ System geladen.

\* Katalog-Nr. 20027213 ist veraltet und wurde durch Illumina DNA/RNA UD Indexes Set A, Tagmentation (96 indexes, 96 samples), Katalog-Nr. 20091654 ersetzt.

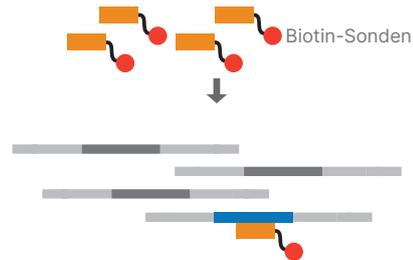
Tabelle 3: Viren, auf die das Respiratory Virus Oligos Panel abzielt

Humanes Coronavirus 229E	Humanes Parainfluenzavirus 1
Humanes Coronavirus NL63	Humanes Parainfluenzavirus 2
Humanes Coronavirus OC43	Humanes Parainfluenzavirus 3
Humanes Coronavirus HKU1	Humanes Parainfluenzavirus 4a
SARS-CoV-2	Humanes Metapneumovirus (CAN97-83)
Humanes Adenovirus B1	Respiratorisches Synzytial-Virus (Typ A)
Humanes Adenovirus C2	Humanes respiratorisches Synzytial-Virus 9320 (Typ B)
Humanes Adenovirus E4	Influenza-A-Virus (A/Puerto Rico/8/1934(H1N1))
Humanes Bocavirus 1 (Primaten-Bocaparvovirus-1-Isolat st2)	Influenza-A-Virus (A/Korea/426/1968(H2N2))
Humanes Bocavirus 2c PK-Isolat PK-5510	Influenza-A-Virus (A/New York/392/2004(H3N2))
Humanes Bocavirus 3	Influenza-A-Virus (A/Gans/Guangdong/1/1996(H5N1))
Humanes Bocavirus 4 NI-Stamm HBoV4-NI-385	Influenza-A-Virus (A/Zhejiang/DTID-ZJU01/2013(H7N9))
KI-Polyomavirus Stockholm 60	Influenza-A-Virus (A/Hongkong/1073/99(H9N2))
WU-Polyomavirus	Influenza-A-Virus (A/Texas/50/2012(H3N2))
Humanes Parechovirus Typ 1	PicoBank/HPeV1/a Influenza-A-Virus (A/Michigan/45/2015(H1N1))
Humanes Parechovirus 6	Influenza-B-Virus (B/Lee/1940)
Humanes Rhinovirus A89	Influenza-B-Virus (B/Wisconsin/01/2010)
Humanes Rhinovirus C (Stamm 024)	Influenza-B-Virus (B/Brisbane/60/2008)
Humanes Rhinovirus B14	Influenza-B-Virus (B/Colorado/06/2017)
Humanes Enterovirus C104-Stamm: AK11	Influenza-B-Virus (B/Washington/02/2019)
Humanes Enterovirus C109-Isolat NICA08-4327	Humankontrollgene

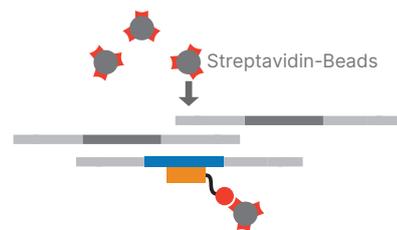
A. Illumina DNA Prep with Enrichment-Bibliotheken poolen



B. Biotinylierte Sonden an Zielregionen hybridisieren



C. Erfassung mithilfe von Streptavidin-Beads



D. Aus Beads eluieren

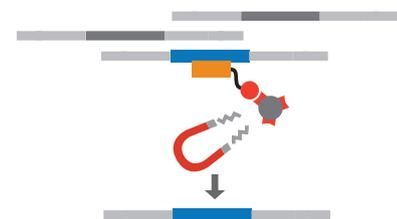


Abbildung 2: Anreicherungsschemie: Vorbereitete Bibliotheken werden mit einem Viren-Target-Panel angereichert. Anschließend folgen Amplifikation, Quantifizierung und Sequenzierung.

### Sequenzierung

Vorbereitete Bibliotheken können auf jedem Gerät von Illumina sequenziert werden. Die Tischsysteme iSeq™ 100, MiniSeq™ und MiSeq sind aufgrund der geringen Read-Anforderungen gut für diese Proben geeignet. In dieser Studie wurden aus Virenproben vorbereitete Bibliotheken gemäß dem MiSeq System Denature and Dilute Libraries Guide (Handbuch zum Denaturieren und Verdünnen von Bibliotheken für das MiSeq System) (Dokument-Nr. 15039740 v10) denaturiert sowie auf eine Endladekonzentration von 10 pM verdünnt und anschließend auf dem MiSeq System mit einer Read-Länge

von 2 × 151 bp mit MiSeq v3-Reagenzien (Illumina, Katalog-Nr. MS-102-3001) sequenziert. Außerdem wurden vor der Sondenhybridisierung Aliquots sämtlicher Bibliotheken sequenziert, um die relative Veränderung der Anreicherungsreaktion zu bestimmen.

Der Virustiter, die RNA-Qualität und die Anzahl der Reads pro Probe wirken sich auf die Anzahl der erhaltenen viruspezifischen Reads und die Coverage des viralen Genoms aus. Im Allgemeinen werden für diesen Workflow 500.000 Reads je Probe empfohlen. Hierbei handelt es sich jedoch nur um einen empfohlenen Ausgangspunkt.

### Datenanalyse

Für die automatische Geräteanalyse wurden Sequenzierungsläufe in Local Run Manager mit dem Resequencing Analysis Module konfiguriert. Dieses Modul ermöglicht die Eingabe sämtlicher Laufinformationen sowie die Auswahl des Referenzgenoms für das nachfolgende Sequenz-Alignment. Anwender können Referenzgenome direkt auf das Gerät hochladen. Dies ermöglicht eine einfache Anpassung. Die Analyse wird nach Abschluss der Sequenzierung automatisch gestartet, sodass Anwender die Ergebnisse so schnell wie möglich auswerten können. Das Resequencing-Modul liefert Daten zu Alignment, Coverage und kleinen Varianten sowie FASTQ-, BAM- und VCF-Dateien, die in anderen Datenanalysepipelines verwendet werden können (Abbildung 3).



Abbildung 3: Anreicherungsanalyse-Workflow: Sequenzierungsläufe werden in Local Run Manager konfiguriert. Nach Abschluss des Laufs wird die Datenanalyse automatisch gestartet. Gegebenenfalls ist auch eine zusätzliche Analyse mit Anwendungen von Drittanbietern möglich.

Bei der IDbyDNA Explify-Plattform handelt es sich um eine anwenderfreundliche Lösung für die ausführliche Datenanalyse mit zuverlässiger Datenqualitätssicherung (QC, Quality Control), standardisierter Ergebnisauswertung, sorgfältig kuratierten Datenbanken und anwendungsspezifischer Befunderstellung. Die Datenanalyse basiert auf k-mers und Alignment-Schritten, einschließlich des Virennachweises auf Proteinebene, was die Fähigkeit zur Bestimmung neuer und stark divergenter Viren erhöht.

Auf die Plattform kann über BaseSpace™ Sequence Hub zugegriffen werden.

## Ergebnisse

Im Anschluss an die Vorbereitung und Sequenzierung der Bibliotheken wurden die einzelnen Proben mit dem Local Run Manager Resequencing Analysis Module (v2.5.56.11) auf das Coronavirus-Referenzgenom OC43 aligniert, wie im [Handbuch zum Workflow für das Local Run Manager Resequencing Analysis Module](#) erläutert.

### Reverse Transkription und Nachweis von CoV mit dem Local Run Manager Resequencing Module

Im Rahmen dieser Studie wurde die cDNA für die Bibliotheksvorbereitung anhand der reversen Transkription mit zwei unterschiedlichen Kits generiert – einem handelsüblichen Kit von Thermo Fisher Scientific und einem Workflow, der aus dem für Illumina TruSeq RNA-Reagenzien abgeleitet wurde (Einzelheiten unter „Methoden“). Für CoVOC43-Proben wurde bei der Illumina-Methode mit reverser Transkription ein Alignment auf das Referenzgenom von über 99 % erzielt und von über 97 % bei der Methode von Thermo Fisher Scientific (Abbildung 4). Das Alignment der Spike-in-Proben lag unabhängig von der Methode für die reverse Transkription bei über 64 % und bei über 19 % für das Respiratory Virus Oligos Panel (Abbildung 4).

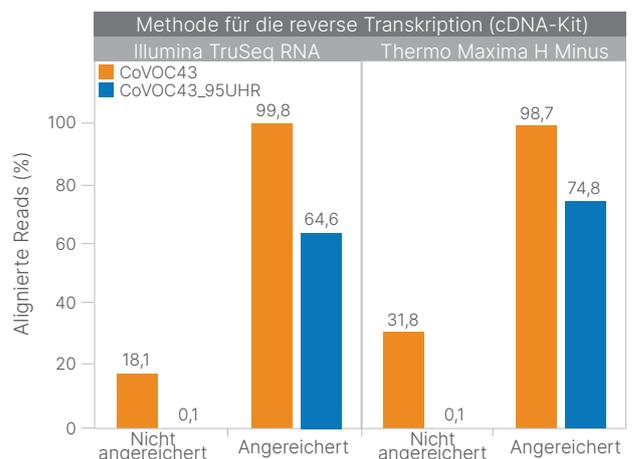


Abbildung 4: Äquivalente Leistung mit unterschiedlichen cDNA-Kits: Der Prozentsatz alignierter Reads bei Virenproben, für die die reverse Transkription mit den beiden unterschiedlichen cDNA-Kits durchgeführt wurde (siehe Abschnitt „Methoden“), wird vor (nicht angereichert) und nach der Anreicherung mit dem Respiratory Virus Oligos Panel (angereichert) dargestellt.

Als Maß für den Erfolg der Target-Capture- und Hybridisierungsreaktion wurden die prozentualen alignierten Reads aus Proben vor und nach der Anreicherung verglichen, um die relative Veränderung der Anreicherung zu berechnen. Die Ergebnisse zeigten unabhängig von der Methode für die reverse Transkription eine deutliche Steigerung der Anreicherung der analysierten CoV-Proben. Dies traf insbesondere auf die CoVOC43\_95UHR-Proben zu. Bei diesen zeigte sich für das Respiratory Virus Oligos Panel eine mehr als 640-fache Anreicherung (Tabelle 4).

Die mittlere Coverage bietet einen Überblick über die Coverage-Tiefe für sämtliche Basen im Referenzgenom und liefert einen Hinweis auf alle möglicherweise nicht sequenzierten Regionen. Die Anreicherung mit Viren-Target-Sonden erhöhte die mittlere Coverage der Virensequenz unabhängig von der Methode für die reverse Transkription bei allen Proben erheblich (Tabelle 5).

Tabelle 4: Metriken zur Anreicherung für das Respiratory Virus Oligo Panel

	Illumina TruSeq RNA		Thermo Maxima H Minus	
Probe	CoVOC43	CoVOC43_95UHR	CoVOC43	CoVOC43_95UHR
Relative Veränderung der Anreicherung	5,5×	646×	3,1×	748×

Tabelle 5: Mittlere Coverage mit dem Respiratory Virus Oligos Panel<sup>a</sup>

cDNA-Kit	CoVOC43		CoVOC43_95UHR	
	Nicht angereichert	Angereichert	Nicht angereichert	Angereichert
Illumina TruSeq RNA	427,1	3.283,5	2,6	2.437,3
Thermo Maxima H Minus	987,0	3.861,4	4,2	3.035,0

a. Die mittleren Coverage-Metriken wurden auf 1 Mio. Reads pro Probe normalisiert.

### Analyse mit der IDbyDNA Explify-Plattform

Das gespikete Coronavirus (Abbildung 5) und alle fünf Viren im Virenpool (Abbildung 6) wurden mit der Explify-Plattform nachgewiesen. Die Explify-Plattform zeichnet sich durch Konsensgenomsequenzen zu Viren, Coverage-Plots und die erwiesene Fähigkeit zum Nachweis von Koinfektionen mit anderen Viren, Bakterien, Pilzen oder Parasiten aus (Abbildung 5 und Abbildung 6).

Außerdem ließ sich mit der Explify-Plattform zeigen, dass die Anreicherungsdaten zu CoVOC43\_95UHR aus dem Respiratory Virus Oligos Panel eine Coverage liefern, die der von mit Shotgun-Metagenomik generierten Sequenzierungsdaten<sup>4</sup> gleichwertig ist (Abbildung 7). Bei der Anreicherung mit Virus-Target-Sonden-pools kommt es folglich im Vergleich zur Shotgun-Methode zu keiner Verzerrung der Coverage.

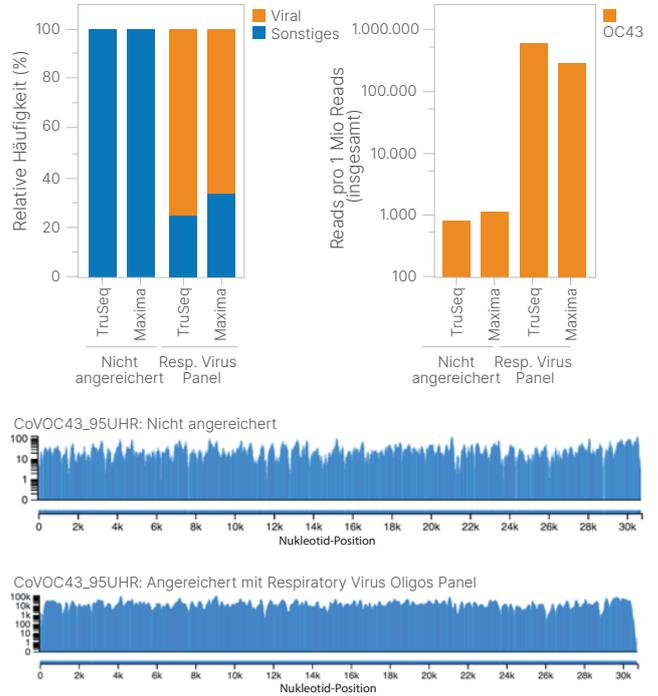
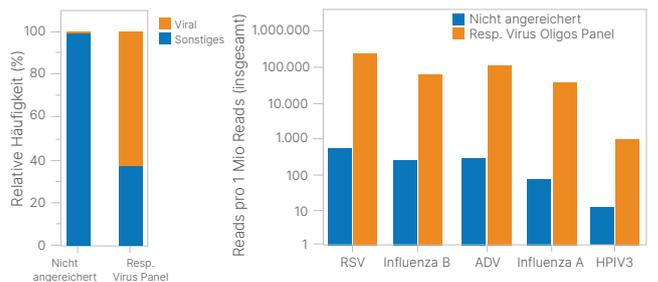
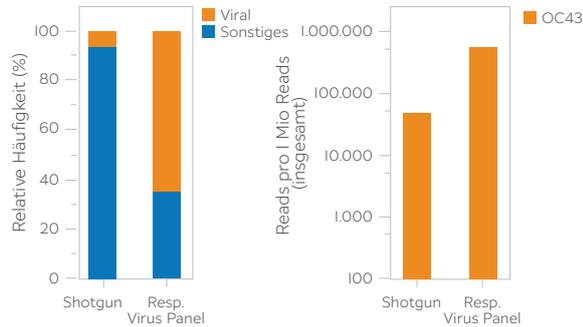


Abbildung 5: Nachweis von OC43: Der Nachweis von OC43 erfolgte unabhängig vom cDNA-Kit mit hoher genomischer Coverage. Die Daten wurden einem Downsampling auf 1 Mio. Reads pro Probe unterzogen.



MD	Alert	Pathogen	Comment	Organism Name	Evidence	Type	% Coverage	ANI	Median Depth
				Human respiratory syncytial virus B <sup>9</sup>	13,658	RNA	100.0%	98.8%	22,444
				Influenza A virus (H1N1) <sup>9</sup>	10,950	RNA	99.3%	97.6%	2,337
				Influenza B virus <sup>9</sup>	3,687	RNA	99.2%	98.1%	8,457
				Human adenovirus 7 <sup>9</sup>	198	RNA	99.1%	99.7%	3,833
				Human adenovirus 7 <sup>9</sup>	198	DNA	99.1%	99.7%	3,832
				Human parainfluenza virus 3 <sup>9</sup>	3	RNA	85.0%	95.0%	15

Abbildung 6: Bestimmung von Viren-pools: Sämtliche im Virenpool enthaltenen RNA- und DNA-Virenstämme wurden nachgewiesen. Der Virenpool wurde nicht anhand von orthogonalen Verfahren verifiziert, sodass der Anteil der einzelnen Viren an der Gesamtvirolast nicht bekannt ist. Die Daten wurden einem Downsampling auf 1 Mio. Reads pro Probe unterzogen.



Die Verwendung eindeutiger dualer Indizes reduziert das Risiko eines entsprechenden Crossovers beim Multiplexing der Proben. Der anwenderfreundliche Workflow ermöglicht den Nachweis und die Charakterisierung von Krankheitserregern wie dem neuartigen SARS-CoV-2.

## Weitere Informationen

Virussequenzierungsmethoden

## Quellen

1. World Health Organization. COVID 19 Public Health Emergency of International Concern (PHEIC) Global research and innovation forum: towards a research roadmap. Veröffentlicht am 12. Februar 2020. Aufgerufen am 9. September 2023.
2. Bulcha B. Review on viral metagenomics and its future perspective in zoonotic and arboviral disease surveillance. *J Biol Agric Healthcare*. 2017;7(21):35–41.
3. Gaudin M, Desnues C. Hybrid Capture-Based Next Generation Sequencing and Its Application to Human Infectious Diseases. *Front Microbiol*. 2018;9:2924. Veröffentlicht am 27. November 2018. doi:10.3389/fmicb.2018.02924
4. Illumina. Comprehensive workflow for detecting coronavirus using Illumina benchtop systems. Veröffentlicht 2020. Aufgerufen am 29. August 2023.

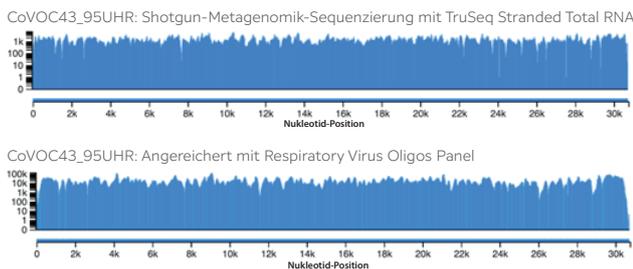


Abbildung 7: Mit der Shotgun-Methode gleichwertige Leistung: Die Anreicherung entsprach bei Leistung und Coverage den Ergebnissen der Shotgun-Metagenomiksequenzierung.

## Zusammenfassung

Der Bestimmung und die Charakterisierung neuartiger Viren ist im Bereich der öffentlichen Gesundheit von zentraler Bedeutung. NGS stellt hierbei eine leistungsstarke Methode zum breit gefächerten Nachweis bei der Bestimmung bekannter und neuartiger Viren dar. Der Einsatz von Illumina DNA Prep with Enrichment mit Panels, die auf virale Pathogene abzielen, liefert Forschern genomische Daten zum Nachweis von CoV und unterstützt Analyseanwendungen wie Genotypisierung und Variantenanalyse. Das offene Design ermöglicht die Bestimmung einer hohen Bandbreite an pathogenen Viren über alle Probentypen von Interesse hinweg.



+1 800 8094566 (USA, gebührenfrei) | +1 858 2024566 (Tel. außerhalb der USA)  
 techsupport@illumina.com | www.illumina.com

© 2023 Illumina, Inc. Alle Rechte vorbehalten. Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Inhaber. Spezifische Informationen zu Marken finden Sie unter [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).  
 M-GL-02146 DEU v1.0