Control experiment phase 2 - unmasked: How the Chinese scientists cheated in the authoritative study on SARS-CoV-2

Corona_Facts on Telegram • April 03, 2022

When scientists bend data to the max and others follow them

A few days ago we published our phase 1 control experiment on our Corona_Fakten channel, in which we discussed the results commissioned in the laboratory on the so-called cytopathic effect.

The results obtained clearly show that this effect**Not**–as claimed by virologists since 1954 - may be claimed as proof of the presence of a virus, butthe experimental setup itself leads to the effect, too

if**no**allegedly infected material is added.

Those who are not yet familiar with these control experiments can study them in detail under [14].

Today we present you the sequel: control experiment phase 2.

It is quite common for a single study to act as the basis on whose findings all other scientists then conduct their research, and SARS-CoV-2 is no different.

However, it is not difficult to imagine the far-reaching consequences if this very study, which in this case tells everyone else worldwide what the alleged disease-causing virus should look like genetically, has generated false and faulty data with an anti-scientific approach.

It should be obvious that a huge avalanche of faulty subsequent sciences will automatically be unleashed. With the well-known consequences that we have been able to experience live in the last two years.

While with the measles virus it was the study by John F. Enders from 1954 that became the "mother" of all other studies, with SARS-CoV-2 we see the scientific study by Prof. Zhang et. al., which was published in "Nature" and bears the following title:

"A new coronavirus associated with human respiratory disease in China" [1]

We now have the study by Prof. Zhang et. al. dismantled and had to admit to us with shock that the Chinese scientists not only let through significant defects, no, they were not above themselves even to "cheat"! What else is this if it's not a scam?

- Half of all published sequences have been blacked out so that the original nucleotide sequence can no longer be extracted, making it impossible for any scientist to trace them. [13]
- No virus was isolated, not even centrifuged, let alone sedimented.

- The assembly, i.e. the construction of the genome of SARS-CoV-2, which became the template for all scientists worldwide, is with the published sequences of the Chinese**not right** **preducible.**
- There is no evidence that the claimed SARS-CoV-2 genome was constructed from pure viral RNA.
- The Chinese authors were subtly influenced to the effect that they only concentrated on the search for possible respiratory tract pathogens due to knowledge of the medical history and symptoms of the patient under consideration.
- The Chinese received two completely different outcomes. The longest contiguous sequence found by overlap using the program "**Megahit" was 30,474 nucleotides**, while**the other program "Trinity"** from the same record a longest contig of**generated 11,760 nucleotides.** Conversely, Trintiy produced significantly more coherent sequence pieces, namely 1,329,960 pieces, than Megahit (384,096). This is from a scientific point of view, **particularly critical under the aspect of reproducibility.**

- It is the sequence data made available to the public by the Chinese**not the original, raw sequence data.**
- It was**no causal cause**described for the present disease because the authors could not discover it. The finding thus**only represents an extremely weak correlation, since only ONE patient was examined!**
- The sequences probably also contain those of human origin, more precisely: ribosomal RNA (if the virologists' databases are correct). This means that not all ribosomal RNA was removed as part of the RNA depletion.

- The coverage of the target genome with the connected short sequences (PCR primers) shows a**very uneven distribution.**With an unbiased sequencing, one would expect that each nucleotide can be covered with the sequences found with a similar number of times. Under the simplified

Assuming a binomial distribution one would expect the coverage **usually lies in the corridor shown.**the**significant fluctuation**the cover raises the question of whether certain sequences that were not contained in the original sample were accidentally created during the sequencing step.

- **There were no controlx pexperiments through Gled.**

Some examples of essential control experiments that virologists failed to carry out:

- *The control experiments, which rule out that the cytopathic effect, which is claimed by virologists as virus detection, does not come about through the experimental setup itself. Some studies claim they did a check, but it was never documented, making it impervious to scrutiny.*

- *The other control experiment, resulting from scientific logic, is the intensive one using the developed PCR method (real-time RT-PCR) with clinical samples from people with**other diseases**than those attributed to the virus and based on**samples from healthy people**,**Animals**and**Plant**to check whether their samples also turn out to be "positive" tested.*

- *Conduct control experiments to rule out*
  - that even with human/microbial RNA from a lung lavage of a healthy person,

  - a person with another lung disease,

  - a person who tested negative for SARS-CoV-2,

  - or from such RNA from reserve samples from the time when the SARS-CoV-2 virus was still unknown,
  Gexactly the Gdead A ufaddition Ga virus genome from short RNA fragments mö Gis!

- *Control experiments regarding genome sequencing from the "infected cell culture" to rule out that other virus genomes can also be assembled "de novo" or by alignment with other reference genomes.*
- *Control experiments to rule out that the genome of the target virus "de novo" or by alignment from the**negative control culture**was assembled.*

At this point, as a true scientist, you actually gasp, because it was once again confirmed that all the rules of science have been broken and disregarded.

Can anyone explain to me why no scientist basing their publication on the Chinese work has noticed the massive flaws? Or were they simply ignored with unprecedented audacity? Or was there an incentive to look away?

True to the motto: The main thing is that I take the laurels, because I am the first!

A new coronavirus linked to human respiratory disease in China -
Was a "novel" and "disease-causing" virus found in Wuhan?

Part A: How was the sequence for the "novel" corona virus SARS-CoV-2 determined?

Part B: Critical consideration of the methods and conclusions

introduction G

In the current pandemic, physicians, primarily virologists, occupy a prominent position. Your assessments are hardly questioned in our society and are decisively used to justify the measures taken. This approach does not appear to be adequate, especially considering the enormous impact on social life. If we

As a society, if we open up a significant space for technocratic thinking and acting to shape social life, we must also have the necessary knowledge and the corresponding methodological skills on a broad basis, especially within the ranks of decision-makers. Otherwise belief comes before knowledge. A problem that is currently manifesting itself dramatically and appears to be highly toxic. It is now time for a great many people to prove themselves mature, to embark on a research trip and to trust in their own abilities. It's not about distrust, it's about questioning. We cannot attain or expect an unequivocally absolute truth, should it exist anywhere. Findings must be constantly questioned, checked and dogmas uncovered. The engine of scientific work is found in these important attributes. We should always be open and grateful for new ideas, arguments and interpretations of the supposedly familiar, because these enrich our lives like a stimulating elixir.

Doubting the known can be bolder than exploring the unknown. alexander v Humboldt

This article deals in detail with the scientific publication "A new coronavirus associated with human respiratory disease in China" [1], in which the genetic sequence of the alleged "novel" corona virus SARS-CoV-2 was proposed for the first time. This work was published on February 03, 2020 and was originally published in the scientific journal Nature. The consideration is in two parts. In the first part, the methodical procedure for the construction of the alleged SARS-CoV-2 virus genome is explained. Then, in the second part, there follows a critical examination of the present publication in a scientific sense. Methodological weaknesses, critical questions and control options are pointed out,

Part A How was the sequence for the "novel" corona virus SARS-CoV-2 determined?

From Gon Gsla Ge

A 41-year-old patient admitted to Wuhan Central Hospital on December 26, 2019 was examined. He was a worker at the local fish market which, according to epidemiological studies, has been implicated as the original outbreak site of the "novel" virus. The patient was admitted with a temperature above 37.5°C, chest tightness, cough, pain and general weakness. Preliminary etiological investigations ruled out, among other things, the presence of influenza viruses. Other common respiratory pathogens, such as human adenoviruses, were also denied. After the patient's condition did not improve even after administration of antibiotic and antiviral medication, was transferred to the intensive care unit and later to another hospital in Wuhan for further treatment. In order to determine possible pathogens in connection with the present disease, bronchoalveolar lavage fluid (BALF) was taken from the patient and then sequenced meta-transcriptomically.

The methods part

The scientific publications within virology generally include a methods section in which the procedure is described in more detail. For example, the methodical procedure and the sequencing technologies used are described here and information on the bioinformatic protocols is included.

data re pplace G

As already mentioned,**exactly one person was examined.**The experiments carried out and the evaluation of the results**were not done blindly.**All scientists involved had complete information, in particular they knew the clinical picture of the patient.

Patient information and sample collection

One patient was examinedwith an acute onset of fever (temperature above 37.5 °C), cough and chest tightness, who presented at the Central Hospital in Wuhan, China. As part of the admission was taken from the patient bronchoalveolar

Lavage liquid (BALF) is removed and preserved at -80 °C until further processing. All patient information, such as examination and laboratory data, was in the form of clinical records.

The patients probe and the RNA se quence G

First, the total RNA (ribonucleic acid) was extracted from the BALF. Commercial extraction kits are used for this. These kits are generally used in accordance with the manufacturer's instructions. After extraction, the quantity and quality of the RNA solution was assessed before subsequent sequencing. However, before the RNA can be sequenced, a so-called RNA library has to be built up. Here, the existing RNA is technically prepared for the sequencing step. For example, the RNA must be transcribed into cDNA using reverse transcription and cut into short fragments. In the present case, during

library building the**ribosomal RNA**removed according to the manufacturer's instructions. The ribosomal ribonucleic acid (rRNA) is found in the human ribosomes, the site of protein biosynthesis. Before the actual sequencing, the short fragments are amplified using the polymerase chain reaction (PCR) so that the sequencing machine used can detect them and determine the nucleotide sequence. The RNA library was sequenced on the Illumina MiniSeq platform in Shanghai.

Info

RNA (Ribonukleinsäure) besteht aus einem Phosphat, dem Zucker Ribose und den vier Basen *Adenin (A)*, *Cytosin (C)*, *Guanin (G)* und *Uracil (U)*.

DNA *(Desoxyribonukleinsäure)* besteht ebenfalls aus einem Phosphat, dem Zucker *Desoxyribose* und ebenfalls aus vier Basen, wobei anstatt *Uracil (U)*, die Base *Thymin (T)* vorkommt. Man spricht bei einem derartigen Molekül von einem Nukleotid und bezeichnet diese ebenfalls mit den Buchstaben *A*, *C*, *G*, *T* und *U*.

Eine **DNA-Sequenz** ist eine Abfolge von Nukleotiden und wird im Computer in Form von Buchstabenfolgen gespeichert.

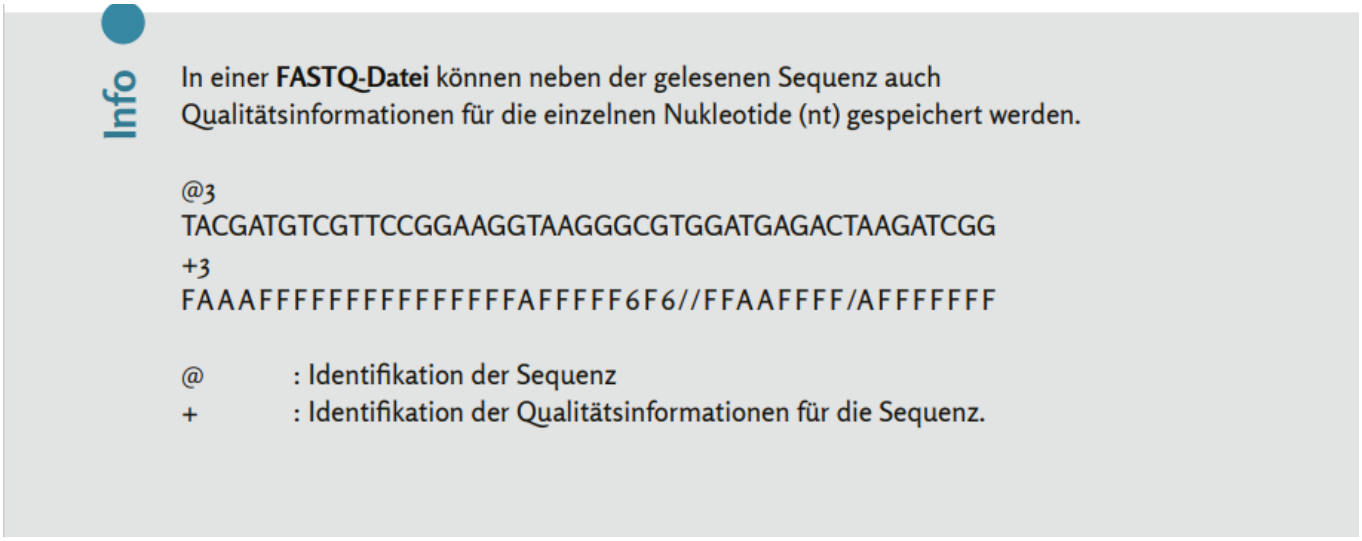
Die ersten acht Nukleotide des „Spike-Proteins“ lauten: ATGTTTGT [2].

source [2]

Bioinformatic Processing Gthe se quenzlesun Gen and virus identification

Sequencing according to the protocol described above and subsequent adaptation and quality control using the Trimmomatic program resulted in a total of 56,565,928 short sequences (reads) with a length of about 150 nucleotides. As explained above, the short sequences can be described with strings of letters consisting of A, T, C, G for the four nucleotides.

The strings of letters are stored in ordinary text files. In the present case, the short readings were stored in FASTQ format [3], which contains additional quality information on the individual nucleotides. From this quality information for each individual nucleotide within a short sequence read, the probability can be determined that the respective nucleotide was correctly registered by the sequencing machine.



One tries now from the short sequential readings*(56,565,928 reads x 150 nt long)* to determine longer contiguous nucleotide sequences. For this purpose, overlaps between the short sequences are sought.

Due to the large number of short readings, around 56 million here, bioinformatic programs and powerful computers are required for this work step.

The longer nucleotide sequences calculated by overlap are called contig. The word contig comes from the English word contiguous, which means connected. The following illustration illustrates this process.

Info

Zielsequenz: Zweifel ist der Motor wissenschaftlichen Arbeitens.

Lesungen: Zweif st der Mo nschaftlich eitens.
Zweifel ist Motor sens lichen Arbeite

Überlappung: Zweif st der Mot nschaftlich eitens.
eifel ist Motor sens lichen Arbeite

Contigs: Zweifel ist der Motor senschaftlichen Arbeitens.

Man beachte, dass die beiden gefundenen Contigs nicht die gesamte Zielsequenz abdecken. In den Publikationen ist in der Regel nicht angegeben, wie diese Lücken gefüllt werden. In diesem Fall wird sich nach den Termini der deutschen Sprache gerichtet und die Lücke mit der Silbe „wis“ gefüllt, obwohl diese in den Daten nicht enthalten war.

In the publication considered here, the time-consuming contiger creation was carried out with the two so-called assemblers Megahit and Trinity in versions v.1.1.3 and v.2.5.1.

To obtain longer contigs from the millions of short sequence readings, the default settings were used. Prior to this, however, possible sequences of human origin were removed using the human reference genome (human release 32, GRCh38.p13) [4]. After this step, 23,712,657 non-humans remained*(known to date)* short sequences. The table below provides an overview of the two assemblies.

Software	Anzahl Contigs	Kürzestes Contig	Längstes Contig
Megahit (v.1.1.3)	384.096	200 nt	30.474 nt
Trinity (v.2.5.1)	1.329.960	201 nt	11.760 nt

Table 1: Summary of results for contiger creation with Megahit and Trinity.

The comparison of the contigs with the known sequences in the relevant nucleotide and protein databases showed that the longest contigs (Megahit: 30,474 nt and Trinity 11,760 nt) were very similar to the SARS-like bat virus SL-CoVZC45, MG772933. According to the authors, the longer of the two contigs, with a length of 30,474 nucleotides, covered almost the entire virus genome and was used for the primer design for the so-called subsequent PCR confirmation and determination of the genome termini.

The organization of the viral genome was done by sequence alignment using two representative betacoronaviruses, namely SARS-CoV Tor2, AY274119 and bat SL-CoVZC45, MG772933. The former is associated with humans and the latter with bats.

data disposal Gability

The sequence readings were filed in the NCBI Sequence Read Archive (SRA) database under BioProject accession number PRJNA603194. The complete genome sequence of WHCV (today SARS-CoV-2) was deposited in GenBank under the identifier MN908947.

control attempts

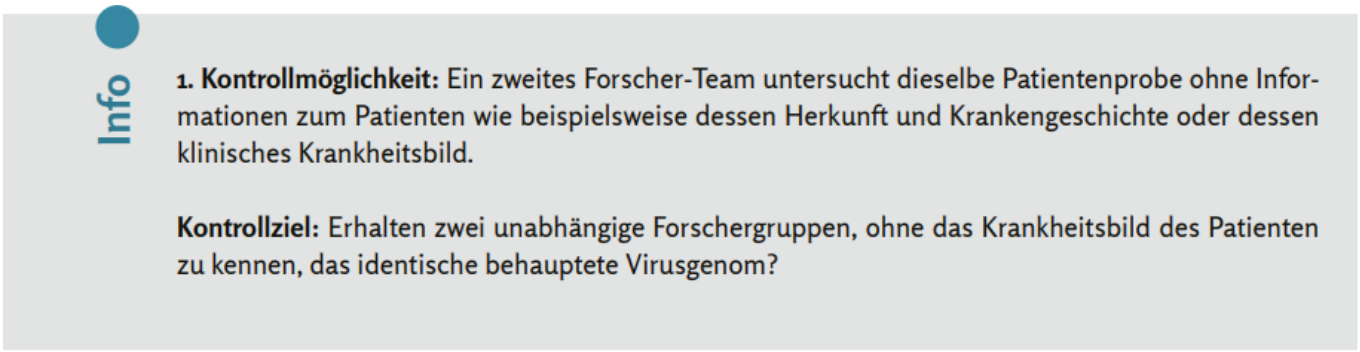
In the present publication "A new coronavirus associated with human respiratory disease in China" [1]**no control tests are documented.**

Part B Critical consideration of the methods and conclusions

After the procedure for determining the virus genome of the alleged "novel" and "disease-causing" virus SARS-CoV-2 has been examined in detail, this section provides a critical examination of the individual work steps. Questioning scientific knowledge is the elixir of living science.

annotation Gen: data re pplace G

The scientists involved were not blind and therefore had knowledge of the medical history and symptoms of the patient under consideration. The authors probably did**influenced in this way**to look for possible respiratory pathogens. Already at this point there would be a simple control option.



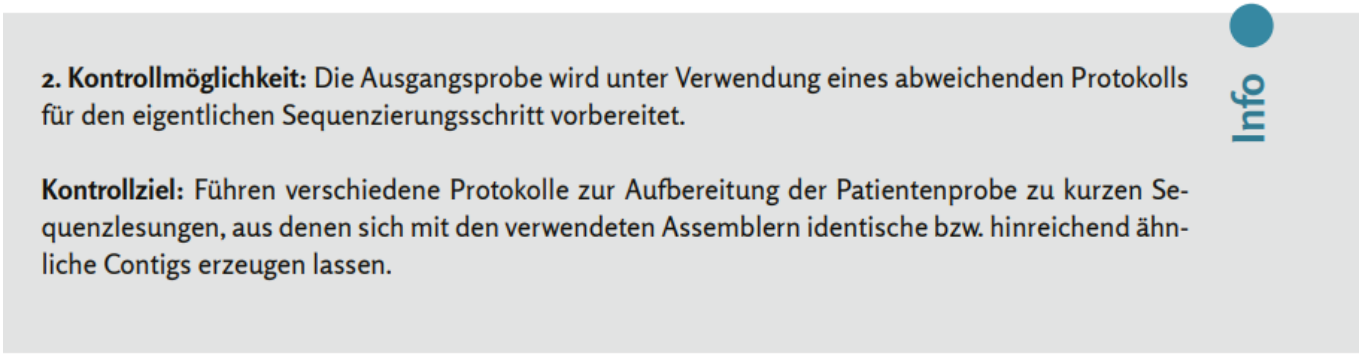
---

annotation Gen: The patients probe and the RNA se quence G

As in section *"The patient sample and RNA sequencing"*was explained in more detail, a few technical steps are necessary before the actual sequencing, i.e. the determination of the nucleotide sequence of short fragments, can be carried out.

For example, commercial kits - for extracting total RNA from the sample - for removing ribosomal RNA from the sample - for converting RNA into cDNA - for fragmenting the RNA or cDNA into short pieces or - for amplification (PCR) of the cDNA in order to be able to determine the nucleotide sequence.

Commercial technical equipment is used to test the quantity and quality of the RNA. These are complex processes that scientists can only control to a very limited extent. This leads directly to a further possibility of control,**which is not documented in the scientific literature.**



2. Kontrollmöglichkeit: Die Ausgangsprobe wird unter Verwendung eines abweichenden Protokolls für den eigentlichen Sequenzierungsschritt vorbereitet.

Kontrollziel: Führen verschiedene Protokolle zur Aufbereitung der Patientenprobe zu kurzen Sequenzlesungen, aus denen sich mit den verwendeten Assemblern identische bzw. hinreichend ähnliche Contigs erzeugen lassen.

Info

It is thus stated that the sequencing of the RNA from the existing bronchoalveolar lavage fluid requires a highly complex protocol. The work steps carried out are difficult to control. It is therefore questionable to what extent the reproducibility required for scientific purposes can be guaranteed without knowledge of a reference genome.

---

annotation Gen: Bioinformatic processing Gthe se q uenzlesun Gen and virus identification

First of all, there is a clear difference between the results of the two assemblers Megahit and Trinitiy.**The discrepancy between the two results is enormous.** Thus, the longest contiguous sequence found by overlap was with**Megahit, 30,474 nucleotides**, while**Trinity**from the same record a longest contig of 1**generated 1,760 nucleotides.**Conversely, Trintiy produced significantly more coherent sequence pieces, namely 1,329,960 pieces, than Megahit (384,096).

Here the question arises immediately Ge:

What would the proposed sequence for the alleged SARS-CoV-2 virus have been if the Megahit software and the corresponding reference genomes (bat-like SARS viruses and previously known human-associated SARS viruses) had not existed?

Another important Ga woman Ge is:

Why is the length of the claimed coronavirus SARS-CoV-2 only 29,903

nucleotides[2] and thus**571 nucleotides fewer than the longest contig, which is 30,474 nucleotides?**

The big difference between the two software programs in terms of maximum contig length is most notable. There do not seem to be any recognized regularity criteria to be followed for the algorithms used to find overlaps.

From a scientific point of view, in particular from the point of view of reproducibility, this must be assessed particularly critically. Furthermore, the previous analysis, especially with regard to the possible random errors during the sequencing step or the finding of nucleotide sequences that may not exist in reality, shows that the calculated sequences or claimed virus genomes can only be models from elementary logical considerations. It is therefore not permissible to speak of a true virus genome or a true virus sequence. The naming of this important fact and possible methods for model validation are not found in the scientific literature or only in passing.

In *"Choice of assemblers has a critical impact on de novo assembly of SARS-CoV-2 genome and characterizing variants"*[5], the authors show that when reconstructing the SARS-CoV-2 genome from several hundred samples**the choice of assembler plays a significant role.**Further, these showed that at least 9% of the variants between Megahit and metaSPAdes (another assembler) are unique to the assembly methods. The scientists conclude from this that their analyzes show what crucial role the used
assembly methods for the construction of SARS-CoV-2 genomes from short reads to the characterization of so-called variants.

of this observation Gtherefore includes fol Gend Mrs Ge to:

What are virus variants or virus mutations? Biological Reality or a Bioinformatic Artifact?

In summary, assembly results show large differences. The present publication mentions that the longest contiguous sequence computed with Megahit (30,474 nt),

almost the entire viral genome.**Based on this sequence, PCR primers were designed.**In Supplementary Table 8. *"PCR primers used in this study"*, [Table 8] lists 52 primers for genome amplification. These evenly cover the entire claimed virus genome, as shown in the graphic below.

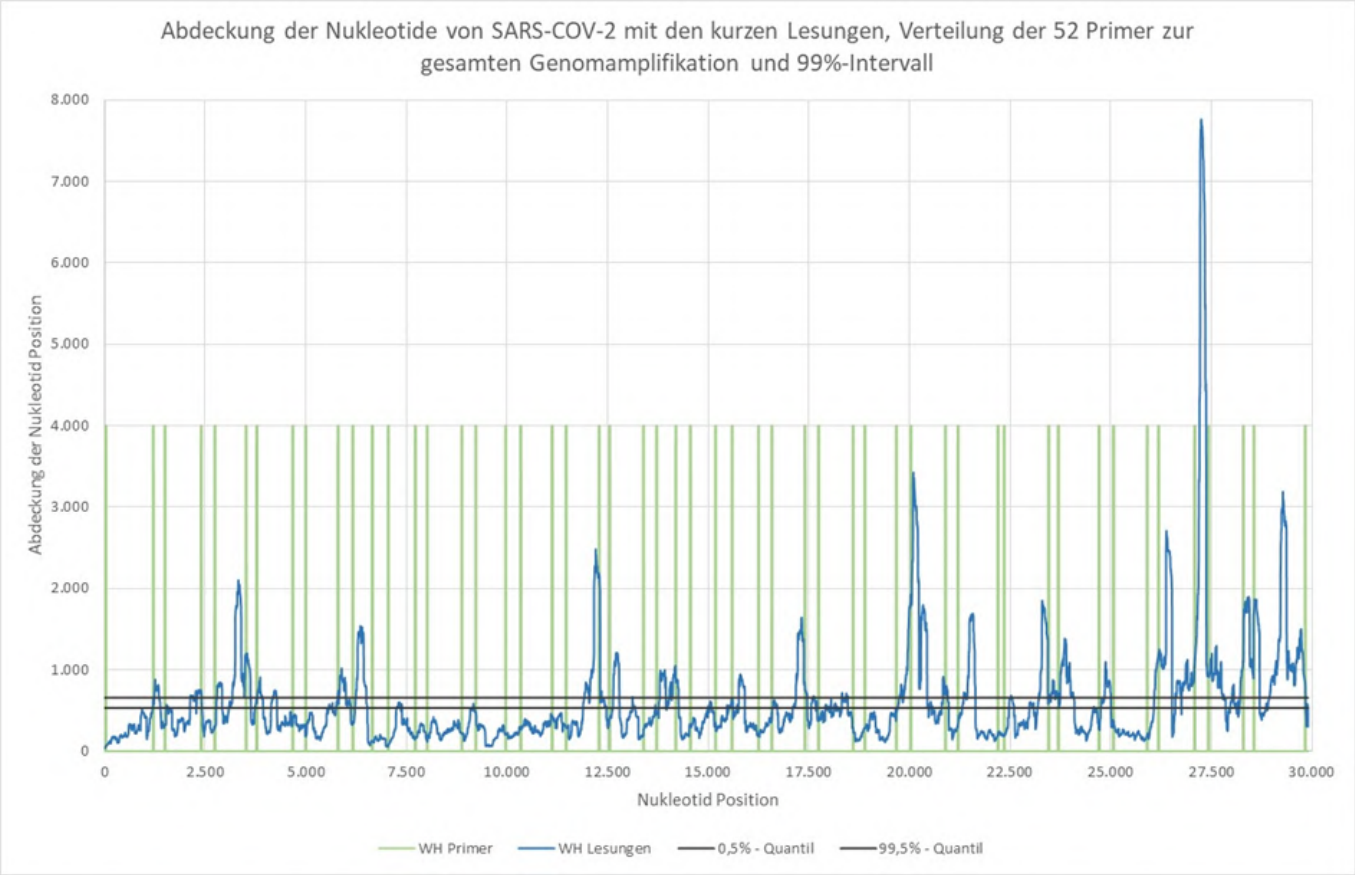


Figure 1: Coverage of SARS-CoV-2 nucleotides with short reads, distribution of the 52 primers for whole genome amplification and 99% interval. The nucleotides were covered with Bowtie2 [6] and Samtools [7].

From the present publication it is not clear exactly how the PCR confirmation or the determination of the genome terms was carried out. the**illustration 1**but shows a particularly high coverage of the nucleotides of the claimed viral genome in the vicinity of the listed primers. The strongly varying depth of coverage of the individual nucleotide positions is also striking.

Assuming that all 29,903 nucleotide positions in SARS-CoV-2 associated readings occur with equal probability, the coverage for each nucleotide position should be between the two horizontal gray lines with 99% probability.

This is not the case for about 90% of the nucleotide positions (see blue lines), which is quite remarkable. A priori one would expect that if sufficient virus RNA is present in the sample and a sufficient number of sequence fragments are read, a homogeneous coverage of the nucleotides within the virus genome is achieved.

The question also arises as to why a second, specific primer-based confirmation and identification step is necessary to clarify the genome termini. This leads to the assumption that that step leads to the distortion clearly visible in the graph.

Furthermore, the impression could arise that the use of the many specific PCR primers for the claimed virus genome and the amplification with an unscientifically high**Number of 35 cycles "connection sequences" arise that do not exist in reality.**

This assumption is supported by the sequencing of the entire RNA of a commercial and uninfected culture of human cells. The necessary amplification of the cDNA fragments was carried out by polymerase chain reaction (PCR) with 14 cycles and exclusively using random hexamers. So short primers with 6 random and unspecific nucleotides, which should lead to a relatively homogeneous coverage of the calculated contigs through the associated sequence readings.

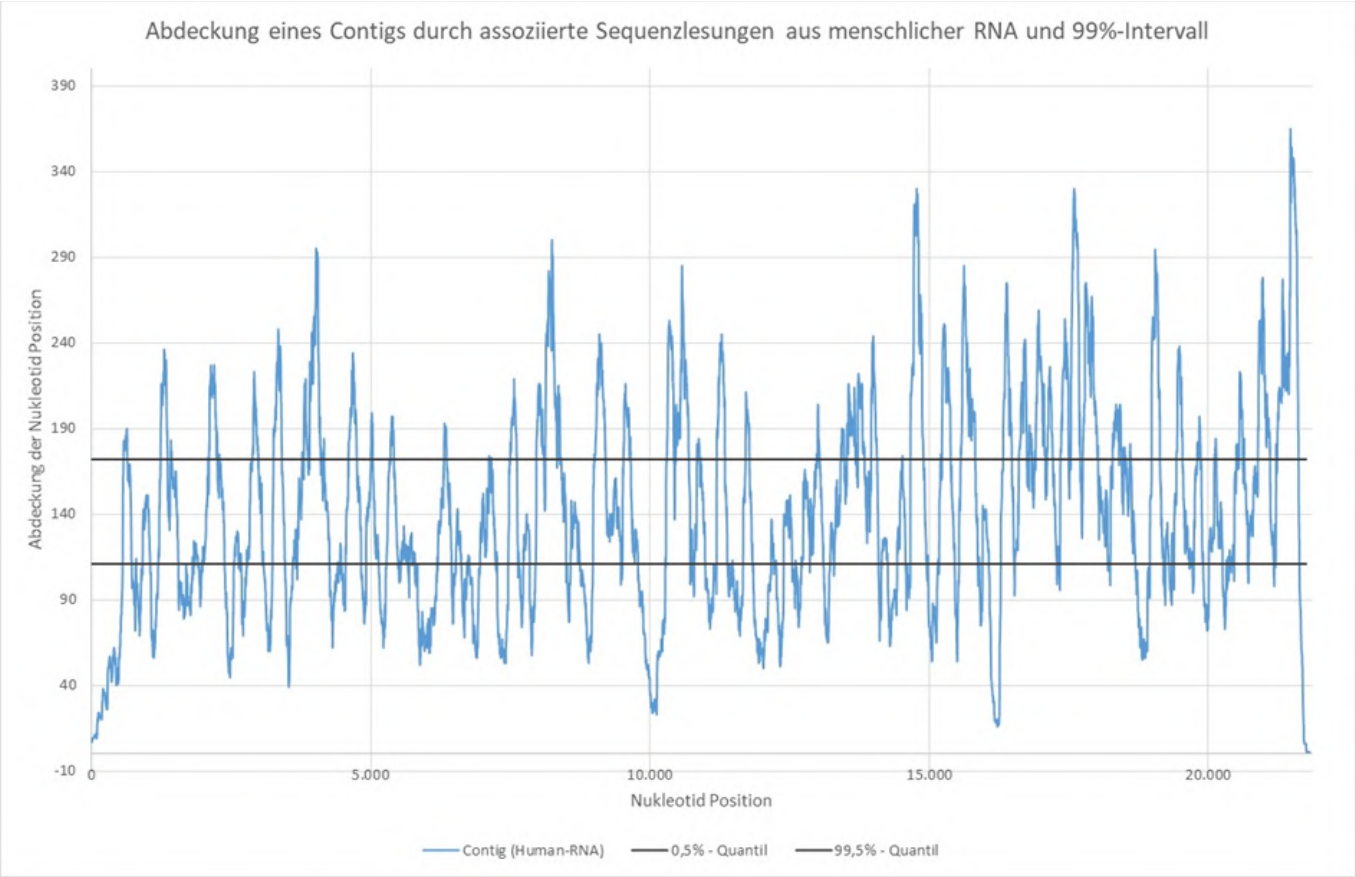


Figure 2: Coverage of a contig (21,814 nt) by its associated human RNA sequence reads and 99% interval. The nucleotides were covered with Bowtie2 [6] and Samtools [7] determined.

In contrast to Figure 1, Figure 2 shows a significantly more homogeneous coverage of the 21,814 nucleotide positions, although here too the coverage for many nucleotide positions is outside the 99% interval. However, the deviations here are rather undistorted below and above the 99% interval.

The 21,814 nt contig shown in Figure 2 was calculated from the RNA of a commercial human and uninfected cell culture using the assembler Megahit. A comparison using Blastn with the entire nucleotide database revealed a high level of agreement (99.91%), over almost the entire contig length, with the mRNA sequence: Homo sapiens spectrin repeat containing nuclear envelope protein 2 (SYNE2), transcript variant 5 , mRNA, Access:**NM _182 91 4**.

In the associated publication entitled *"A novel SYNE2 mutation identified by*

Whole exome sequencing in a Korean family with Emery-Dreifuss muscular dystrophy"

[8th]

it says under results

*[... novel de novo pathogenic heterozygous missense mutation (**NM _182 91 4.2** : c.4858G > A; p.Ala1620Thr) of the SYNE2 gene, which had not been previously reported was identified by whole exome sequencing in the proband and by Sanger sequencing in his son. ...].*

In German:

"...a novel pathogenic heterozygous missense mutation NM_182914.2: c.4858G > A; p.Ala1620Thr) of the SYNE2 gene, which had not previously been reported, was identified by whole-exome sequencing in the subject and identified by Sanger sequencing in his son. ..."

An interesting observation in itself!

Irrespective of this observation, it can be stated that the mRNA sequence mentioned with the identifier**NM _182 91 4.2** could be assembled with readings of the RNA fragments of a non-"infected" cell culture of human origin, which were amplified using random and thus non-specific hexamers with only 14 cycles before the actual sequencing.

A PCR confirmation or determination using specific primers or the use of any reference sequences was not necessary. Therefore, the following natural questions arise.

- Why doesn't the longest contig already contain the entire SARS-CoV-2 viral genome sequence we were looking for?
- Why are a large number of specific primers and reference sequences required for the final sequence determination?
- Why is the longest contig (30,474 nt) longer than the claimed viral sequence of SARS-CoV-2 (29,903 nt)?

First of all, this contradicts naive logic. If already 30,474 contiguous

nucleotides are found, the "true" genome sequence should contain at least 30,474 nucleotides.

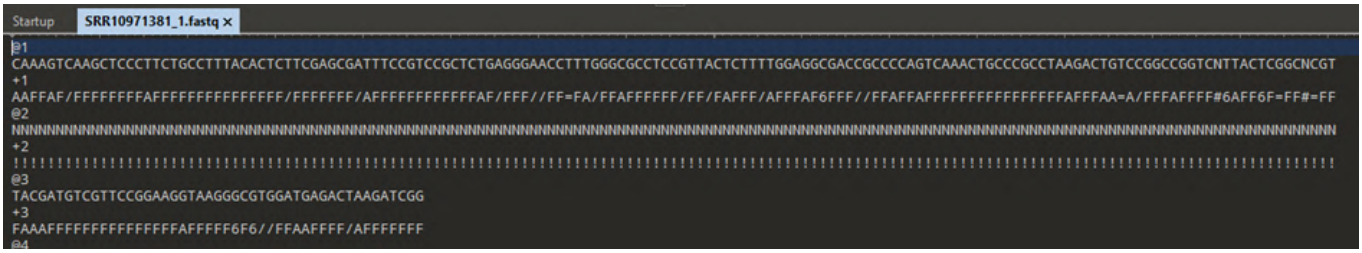
Because otherwise some of the overlaps identified by Megahit would be declared incorrect. There are no standards documented in the scientific literature for this. In the following, simple control experiments are to be presented, which only require basic knowledge of the Linux operating system, a certain period of familiarization with the bioinformatics used Computer programs and the published sequence data require.

Info

1. **Kontrollversuch:** Enthält der Datensatz unter der BioProject-Zugangsnummer PRJNA603194, 56.565.928 kurze Lesungen.
Protokoll:
Download der Sequenzdaten: fastq-dump --split-files --origfmt --gzip SRR10971381

He Gresult: The sequences can be created with the command *"fastq dump"* of the sratoolkit [9] package in FASTQ format and the paired end reads can be conveniently saved in two text files. Each of the two text files contains 28,282,964 short sequences with an average length of about 142 bp.

In total, both files contain 56,565,928 short sequences.**What is striking, however, is that a significant number of the sequences consist of N's, ie unknown nucleotides. [13]**These could be erroneous readings or subsequently overwritten, for example human sequence readings. In terms of scientific reproducibility, this approach would have to be viewed critically


Startup
SRR10971381_1.fastq
@1
CAAAGTCAAGCTCCCTTCTGCCTTTACACTCTTCGAGCGATTTCCGTCCGCTCTGAGGGAACCTTTGGGCGCCTCCGTTACTCTTTTGGAGGCGACCGCCCCAGTCAAACTGCCCGCCTAAGACTGTCCGGCCGGTCNTTACTCGGCNCGT
+1
AAFFAF/FFFFFFFFAFFFFFFFFFFFFFFF/FFFFFFF/AFFFFFFFFFFFFAF/FFF//FF=FA/FFAFFFFFF/FF/FAFFF/AFFFAF6FFF//FFAFFAFFFFFFFFFFFFFFFFFAFFFAA=A/FFFAFFFF#6AFF6F=FF#=FF
@2
NNNN

Info

2. Kontrollversuch: Lassen sich mit den veröffentlichten Sequenzen die Assemblies wiederholen? Können also insbesondere mit Megahit und Trinitiy die jeweils längsten Contigs reproduziert werden?
Protokoll:
Aufbereitung der Daten wie z.B. Adapterentfernung oder Eliminierung von Reads von schlechter Qualität mit der Software fastp [10]: fastp -i SRR10971381_1.fastq.gz -I SRR10971381_2.fastq.gz -o SRR10971381_1.fastq -O SRR10971381_2.fastq

Assembly mit Megahit v.1.2.9 [11] (Standardeinstellungen): megahit -1 SRR10971381_1.fastq -2 SRR10971381_2.fastq -o megahit_result

Assembly mit Trinity v.2.5.1 [12] (Standardeinstellungen): Trinity --seqType fa --left SRR10971381_1.fa --right SRR10971381_2.fa --CPU 6 --max_memory 20G

Sources: [10] [11] [12]

Remarks: First of all, it should be mentioned that version 1.1.3 of the Megahit software does not run on the computers used here. Therefore the current version v.1.2.9 was used.

The Trinity software does not run with FASTQ files. Therefore, Trinity was used with the setting – SeqType fa (FASTA). In contrast to FASTQ files, FASTA files only contain the sequences without quality information on the individual bases. These circumstances must be taken into account when analyzing the results.

Result: The table below provides a summary overview of the results.

** The assembly without using the fastp software with the published sequences gives a total number of contigs of 29,463 and a longest contig of 29,802 bp.*

Software	Anzahl Contigs	Kürzestes Contig	Längstes Contig
Megahit (v.1.2.9)	28.459*	200 nt	29.802* nt
Trinity (v.2.5.1)	157.283	201 nt	29.875 nt

Table 2: Assembly results of the 2nd control experiment

The second control experiment shows an unexpected result. Both the total number of contigs and the lengths of the maximum contiguous sequences **differ significantly from the published results.**

Also noteworthy is the observation that a longer contiguous sequence can be calculated with Trinity than is given in publication [1]. Furthermore, the longest contigs agree almost completely with the published genome "MN908947", as shown by a Blastn query with the standard nucleotide database.

This means that the two longest contigs cannot be reproduced!

The published sequence data cannot therefore be the original, raw sequence data. It is noteworthy that the total number of sequences provided (56,565,928 pieces) corresponds to the information in the publication considered.

As described above, RNA sequences of human origin have been eliminated. The human reference genome (human release 32, GRCh38.p13) was used for this. In Wikipedia [4] it says

[... The human reference genome is derived from thirteen anonymous volunteers from Buffalo, New York. Donors were recruited by advertisement in The Buffalo News, on Sunday, March 23, 1997. ...].

In German:

"[The human reference genome was obtained from thirteen anonymous volunteers from Buffalo, New York. The donors were recruited through an advertisement in the Buffalo News on Sunday, March 23, 1997. ...]."

To what extent is the reference genome specific to humans and is it therefore suitable for the reliable detection of RNA sequences of human origin?**This question can be clarified with the next control experiment, which follows control experiment 2 in a natural way.**

Info

3. Kontrollversuch: Enthalten die im 2. Kontrollversuch mit Megahit (v.1.2.9) errechneten zusammenhängenden Sequenzen, solche menschlichen Ursprungs?

Protokoll:
Selektion der 25 längsten Contigs und Vergleich mit der Nukleotiddatenbank des NCBI

Blastn: blastn -db nt -remote -query **$Input.fasta** -out **$Ergebnis.txt** -outfmt „6 qseqid sseqid stitle pident length mismatch gapopen qstart qend sstart send evalue bitscore“

Result: The table below shows selected query hits. The longest contig calculated with Megahit (k141_11881) shows a perfect match with the published sequence for SARS-CoV-2 at a length of 29,801 nucleotides.**The two contigs (k141_18050 and k141_13168) show a high degree of identity with human sequences up to a length of about 6,000 nucleotides.**

Contig ID	Contig-Länge	Subjekt ID	Subjekt Titel	Übereinstimmung in %	Länge	Evalue	Bitscore
k141_11881	29.802	gi|1798174254|ref|NC_045512.2|	Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome	100.000	29.801	0.0	55.033
k141_18050	7.169	gi|2033711141|gb|CP068263.2|	Homo sapiens isolate CHM13 chromosome 15	98.322	5.842	0.0	10.218
k141_13168	5.924	gi|2033714914|gb|CP068260.2|	Homo sapiens isolate CHM13 chromosome 18	96.725	5.893	0.0	9.769
k141_18050	7.169	gi|2033716901|gb|CP068256.2|	Homo sapiens isolate CHM13 chromosome 22	98.136	5.848	0.0	10.163
k141_18050	7.169	gi|2033716899|gb|CP068257.2|	Homo sapiens isolate CHM13 chromosome 21	97.968	5.857	0.0	10.113
k141_18050	7.169	gi|555853|gb|U13369.1|HSU13369	Human ribosomal DNA complete repeating unit	98.100	5.843	0.0	10.142
k141_18050	7.169	gi|1021673812|gb|KX061889.1|	Nomascus leucogenys precursor 48S ribosomal RNA, 18S ribosomal RNA, 5.8S ribosomal RNA, and 28S ribosomal RNA genes, complete sequence	85.256	5.948	0.0	5.919
k141_18050	7.169	gi|1238280724|gb|KY962518.1|	Homo sapiens external transcribed spacer 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2, 28S ribosomal RNA gene, and external transcribed spacer, complete sequence	98.524	5.828	0.0	10.268
k141_18050	7.169	gi|36165|emb|X68195.1|	H.sapiens genomic DNA of ribosomal RNA intergenic spacer sequence	98.433	5.232	0.0	9.190
k141_18050	7.169	gi|1021673798|gb|KX061875.1|	Gorilla gorilla clone CH255-30A8 48S ribosomal RNA, 18S ribosomal RNA, 5.8S ribosomal RNA, and 28S ribosomal RNA genes, complete sequence	93.925	3.638	0.0	5.445
k141_18050	7.169	gi|1187954987|ref|NG_046806.1|	Homo sapiens origin of replication in ribosomal DNA intergenic spacer (LOC107403080) on chromosome 21	99.612	3.093	0.0	5.640
k141_18050	7.169	gi|1519313802|ref|NM_001256.6|	Homo sapiens cell division cycle 27 (CDC27), transcript variant 2, mRNA	90.490	2.061	0.0	2.699
k141_18050	7.169	gi|555853|gb|U13369.1|HSU13369	Human ribosomal DNA complete repeating unit	96.038	1.666	0.0	2.660

Table 3: Selected query hits comparing the 25 longest contigs from Control 2 and NCBI nucleotide database.

According to the subject descriptions, this is ribosomal and messenger RNA (rRNA and mRNA).**This would mean that not all human RNA sequences were completely removed from the original sequences, as the Chinese scientists claimed.** Finally, it must be noted that the supplementary tables 1 and 2 for the publication considered here show that the two longest contigs (Megahit: 30,474 nt and Trinity: 11,760 nt) each have a high nucleotide correspondence with the bat coronavirus bat SL-CoVZC45, MG772933 from

89.1% and 90.4%, respectively. However, it is not specified how many nucleotides match in total. It is quite possible that the match is only 10,000 nucleotides of the total of 30,474 nucleotides at 89.1% in the case of Megahit. That would mean 20,474 nucleotides had no significant match with bat SL-CoVZC45.

---

annotation G: control experiments

As already stated, no control experiments were documented in the publication considered here. It must therefore be assumed that no control tests were carried out.

Below are other obvious control options that could have secured the knowledge gained.

Info

3. Kontrollmöglichkeit: Kann das behauptete Virusgenom auch aus Proben weiterer Krankheitsfälle mit identischem Symptomenkomplex gewonnen werden, nicht aber aus Proben von Gesunden.
Kontrollziel: Erhärtung der Vermutung, dass das Vorhandensein der berechneten Sequenz kausal für ein ganz bestimmtes Symptomenkomplex (mit)verantwortlich zeichnet.

4. Kontrollmöglichkeit: Kann das behauptete Virusgenom auch aus RNA-Quellen gewonnen werden, die sicher keine pathogenen „Viren" enthalten.
Kontrollziel: Sind die für die Assemblierung des behaupteten SARS-CoV-2 Genoms verwendeten RNA-Fragmente viralen Ursprungs?

---

summary Gand conclusion Gerun G

The patient sample from a case of illness in Wuhan became**all RNA extracted.** According to the authors, human RNA fragments were removed.

- Control experiment 3 has shown that in the published sequences for the alleged virus genome SARS-CoV-2 with the identification MN908947 there is a high probability that human ribosomal or messenger RNA

origin is included, provided the databases are properly maintained.

- The 2nd control experiment showed that the published results are not reproducible. Rather, the published sequences may not match the original sequences.

- According to the description in the present publication, it could not be assumed that almost the complete "virus genome" can be preserved with both Megahit and Trinity. This observation is of concern, as it can no longer be verified to what extent the 30,474 nt contig assembled with Megahit matches the bat coronavirus bat SL-CoVZC45 and to what extent it deviates from the final sequence proposal for SARS-CoV-2.

- No control experiments were documented. It can therefore be assumed that no control tests were carried out. We have outlined possible control experiments, but the list is not exhaustive.

- It is completely unclear a priori whether the disease of the patient under consideration is of "viral" origin.

- It could also be shown that the origin of the RNA fragments used to construct the alleged virus genome was not determined.
  It could simply be RNA of human origin.

- The question also arises as to whether it is even possible to clarify the origin with the methods used here. In short: are the sequences used “viral” in origin or not? This question inevitably leads to the concept of virus isolation, in the sense of the word "isolation". No (virus) particle was found in the present case.

The naive protocol for finding a possible (virus) particle is something like this:

. Discover “many” particles of the same morphology in the patient sample (here: bronchoalveolar lavage fluid). These should exist, since virologists have given the alleged SARS-CoV-2 virus particles a size in the range of around 100 nm

assign.

. Separate these particles from everything else.

. Draw several different samples of the found particles.

. Sequence all samples independently.

. Get (nearly) the same sequence in all cases.

Then, in the scientific sense, one would have described a particle characterized by its sequence and morphology. However, the proof of a possible pathogenicity or transmissibility of this particle would not be provided by far.

In conclusion, therefore, it should be noted that:

In the publication *"A new coronavirus associated with human respiratory disease in China"*[1] the existence of a "novel virus" or a novel "virus sequence" or a pathogen has not been demonstrated. In particular, no causal cause was found for the clinical symptom complex of the patient examined.

Our book series "The Eyewitnesses"

What we have worked out in more than 30 years of intensive research, independently financed control experiments, court cases won, personal exchange and correspondence with the world's leading virologists and institutions exceeds the imagination of most people.

Each book contains independent knowledge, which forces every state, politician, scientist and citizen into action.

Our promise: We PROVE that ALL claims of the existence of disease-causing viruses, pandemics, the theory of contagion and vaccines are based on error and pose a danger to life and limb.

Mrs Gen? Our contact person on Tele Gram is at your disposal GU.N G!

click@NotIsolate

With the purchase of one of the books "Die Zeitzeugen" you support our work and other book volumes at the same time. We thank you very much

Order now

-->The Eyewitnesses Volume 1.0 | [promotional video]

-->The Eyewitnesses Volume 1.1 | [promotional video]

Qsource directory

[1]A new coronavirus associated with human respiratory disease in China | Nature

*Nature*volume 579, pages 265–269 (2020)

[2]Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, co - Nucleotides - NCBI (nih.gov)

[3]FASTQ format. Aug 2021. URL: https://en.wikipedia.org/wiki/FASTQ_format

[4] Reference genome. Sep 2021

URL:https://en.wikipedia.org/wiki/Reference_genome

[5]Rashedul Islam et al. "Choice of assemblers has a critical impact on de novo assembly of SARS-CoV-2 genome and characterizing variants". In: Briefings in Bioinformatics 22.5 (2021). DOI: 10.1093/bib/bbab102.

Choice of assemblers has a critical impact on de novo assembly of SARS-CoV-2 genome and characterizing variants - PubMed (nih.gov)

[6]Bowtie 2. URL: Bowtie 2: fast and sensitive read alignment (sourceforge.net)

[7]samtools. samtools/samtools: Tools (written in C using htslib) for manipulating next-generation sequencing data. URL: GitHub - samtools/samtools:Tools (written in

C using htslib) for manipulating next-generation sequencing data

[8th]Sook Joung Lee et al. “A novel SYNE2 mutation identified by whole exome sequencing in a Korean family with Emery-Dreifuss muscular dystrophy”. In: Clinica Chimica Acta 506 (2020), pp. 50-54. DOI: 10.1016/j.cca.2020.03.021.

A novel SYNE2 mutation identified by whole exome sequencing in a Korean family with Emery-Dreifuss muscular dystrophy - ScienceDirect

[9]ncbi. ncbi/sra-tools: SRA Tools. URL: GitHub - ncbi/sra-tools: SRA Tools

[10]OpenGene. OpenGene/fastp: An ultra-fast all-in-one FASTQ preprocessor (QC/adapters/trimming/_ltering/splitting/merging...) URL: GitHub - OpenGene/fastp: An ultra-fast all-in-one FASTQ preprocessor (QC/adapters/trimming /filtering/splitting/merging...)

[11]voutcn. voutcn/megahit: Ultra-fast and me mory-efficient (meta-)genome assembler. URL: https://github.com/voutcn/megahit.

[12]Trinityrnaseq. trinityrnaseq/trinityrnaseq: Trinity RNA-Seq de novo transcriptome assembly. URL: https://github.com/trinityrnaseq/trinityrnaseq

[13]Instructions for downloading the sequences to remove the redacted sequences (replaced with "N")

1. Download the sequences via fastq-dump (console program for Windows or Linux)

fastq-dump --split-files --origfmt --gzip SRR10971381

2. If you only want to download a part (from reads 1 to 100, i.e. 100 reads:

fastq-dump --split-files -N 1 -X 100 -Z SRR10971381 > SRR10971381.fastq

3. Here is the official link https://trace.ncbi.nlm.nih.gov/Traces /sra/?run=SRR10971381

Generated the first 100 paired-end reads with the command under 2. Various reads are already "redacted" under these first 100

[14]Control experiment phase 1 - Several laboratories confirm the refutation of the Virology through the cytopathic effect | [Telegraph]

tabel

[Table 8]41586_2020_2008_MOESM1_ESM.pdf (springer.com)

Telegram main channel: https://t.me/Corona_Fakten

You can support us by buying our books

Order the first book here:"The Eyewitnesses Volume 1.0"

Order the second book here:"The Eyewitnesses Volume 1.1"

Contact person on Telegram for premium access:
Username: @NotIsolate

Below is a list of our most important items:

Corona_Fakten: List of the most important articles