Sign In Get started

Record the first discovery of the novel coronavirus

Wenzhong Tang
Jan 30, 2020 · 8 min read

original edition

Original author: Wenjun Sill dog

Liner notes to work on December 26, 2019, and I still browsed the automatic interpretation results of mNGS pathogenic microorganisms of this day as usual. If there is no problem, I will start a day's research and development work.

Surprisingly, it was found that a sample reported a sensitive pathogen - the SARS coronavirus, with dozens of sequences, and this sample only had such a meaningful pathogen. If it was an ordinary virus, it was a quite reliable result. As soon as I was worried, I quickly checked the detailed analysis data in the background and found that the similarity was not very high, only about 94.3% (this is related to the threshold of card similarity, which is mainly determined by the sequence of the virus family, and there are several possibilities to think of...). There are several differences in the genomes of different strains of SARS. 2. RNA viruses are prone to mutations, 17 years after the SRAS event, and the mutation is relatively large. 3. Microsporidiosis of mixed species, etc. In order to confirm the reliability of the results, a detailed analysis has begun.

Fortunately, I have encountered several similar sensitive pathogen confirmation and analysis work before, and the leader has discussed with me several times whether I can do an analysis process for automatic mining and reporting of such pathogens. Automatic mining and reporting is preliminary; we where doing other projects with higher weights and higher priority. When the version comes out, this sample is just useful. I named it. Compared with the analysis process used in daily production, it has an additional suffix: "Exploration Edition", which contains almost all sequenced viral genomes.

The analysis results of the Discovery Edition suggest that this pathogen is most similar to bat SARS like coronavirus, with an overall similarity of about 87% and SARS of about 81%. The number of sequences on the comparison has increased from dozens to more than 500. In addition, five contigs have been assembled, with a total of more than 1,200 bp. At this time, the results of automatic mining and reporting have been well analyzed in detail. During the analysis, confidential discussions with the interpreter and the leader also began.

Further analysis, whether it is taking the original sequence to the NCBI nt library blastn or the assembled sequence to blastn, the results are the most similar to bat SARS like coronavirus, but the overall similarity is only about 81%. The similarity of our bank protein blastn contrast is similar, so the results are kept suspicious.

The classification information of viruses has always been chaotic, and the classification rules are not uniform. Some are based on typing genes (such as rhyniviruses), while others do not have clear typing genes, and are classified by other methods. I have not investigated the classification methods of the SARS like and SARS belts. On NCBI Taxonomy, the SARS like belt belongs to the SARS like belt, and the SARS belt belongs to the SARS belt. I have not investigated the literature and I don't have much data to hand. Based on the similarity of the genome, I subconsciously believe that since bat SARS like is a subordinate classification of SARS, the unknown virus detected is at least one like bat SARS 1. The parallel species of bat, that is, a new bat SARS like (not into SARS and Bat SARS like are classified through several unannounced promises).

Thousands of coronavirus genomes were further analyzed carpetsel, and the two most similar genomes, bat SL-CoZC45 and bat SL-CoZXC21, were finally found in terms of similarity. genomes, bat SL-CoVZC21 After the sequence, many articles also analyze these two).

(The picture was sent to the group only in the afternoon)

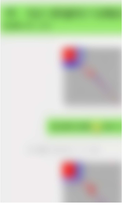
This information alone is not enough. At least look at the evolutionary information, so I began to do evolutionary tree analysis.

After downloading the genomes of all coronaviruses, 160 coronavirus genes (basically containing the coronaviruses of all known animals) were screened out through mass filtering, clustering and other analysis. The preliminary evolutionary tree was constructed using the genome sequence based on the average similarity of the whole genome (by the topic, I personally think that in species evolution analysis, it is more accurate and reasonable to build an evolutionary tree based on the average similarity of the whole gene than an evolutionary tree based on a certain gene. After all, it is impossible to use a single gene to represent the whole genome in terms of its nature, function, etc. In addition, the sequence of assembly in this note is also very short, and there is no complete gene sequence. As a result, it is not surprising. It is still the closest to Bat SARS like coronavirus, and it is also on the branch of SARS.

(The picture was not released until evening)

Note: In large piece of text in the upper left corner is SARS like bags, one is in red, one is in blue (SARS like, and the blue bag is the bat SARS like), and the other is in green (SARS), and the other is in purple (other coronaviruses). Language and symbols are all in English, so it is not a mixed language.

Strangely, this unknown virus gathers into a relatively independence branch with bat SL-CoZC45 and bat SL-CoVZC21, while the other bat SARS like coronaviruses are in the other two branches. This phenomenon seems to mean that there is something wrong with the classification of the literature, bat I looked at the source, and the method is no different from others. Interpret the classification of the literature and think it's right for the time being.) This is also one of the bases for some later experts to judge this unknown virus as the novel coronavirus.



Screenshot of some evolutionary trees

The Boss and Bodhisattva that the patient is seriously ill and urgent to test. On the morning of the 27th, the preliminary results of the analysis of the 26th were held an emergency meeting with several leaders and decided to continue in-depth analysis, delay the issuance of reports, and share data with the Institute of Pathogens of the Chinese Academy of Medical Sciences for analysis.

Later, an in-depth analysis was carried out from the genetic level (mTalk, X, N and other genes), and there was no significant discovery. The main reason was that the number of sequences detected was small and the coverage rate was too low. They were incomplete genes and could not do anything.

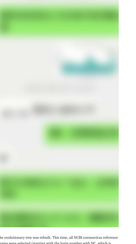
At noon, the retested supplementary data has been prepared for analysis. Retesting can carry out technical repeatability verification, avoid false yang caused by pollution of some unknown factors, ensure that the sample does have the pathogen, and more data can be analyzed. If a complete genome can be assembled, the analysis results will be more reliable and more in-depth analysis can be done.

The next day (2019 12.27), after the data came out early in the morning, the assembled analysis was carried out quickly, and finally the nearly complete genome sequence was assembled. The data is also shared with other in-depth analysis by the Pathogen Institute of the Chinese Academy of Medical Sciences. The number of sequences this time has increased from more than 500 to more than 470,000!

Because time is limited and there are other research and development projects to be done, there is no detailed gap explication for the assembly results to obtain the complete genome. In addition, the data has been shared with the pathogen, and they will do this, as they have not assembled a complete genome sequence. The existing assembly results can meet most of the analysis needs.

Later, I also continued to do some in-depth analysis.

The reply sequence is evenly distributed, and there is no obvious bias. The sequencing depth is mainly concentrated in the 100-fold volume, and the depth reaches 1000x, indicating that there is no problem with assembly, the sequencing is also quite good, and the unknown pathogen also exists in a complete genome.



The evolutionary tree was rebuilt. This time, all MCM coronavirus reference strains were selected (starting with the login number with NC, which is officially considered the more credible), plus several places that analyzed the latest source the day before.

The results of the evolutionary tree are basically the same as the previous day.

Genome Linear analysis, ORF annotation, etc. show that the unknown coronavirus is typical BetaCoV (orf1ab, S, M, N, E, etc.). The lighter color in the rainbow graph is the 3 protein region, which is the gene with the greatest difference.

After comparing with the 7 CNS verification target sequences of SARS coronavirus, it was found that the sequence of the unknown virus similarity is only about 90%. This key is that there are also several variations in the primer sequence. It is speculated that the detection kit of SARS cannot detect this unknown pathogen (the same is true of experiments in many places later). :)

In addition, many other detailed in-depth analyses of genes, proteins, etc. have been done, and will not be repeated.

So far, the analysis can basically confirm that there is indeed a virus in the patient's sample, but this virus is not very similar to all viruses with known genome information. It may be a new virus similar to that SARS the coronavirus.

(Skim looking back at what I said at that time, it's really a prophecy I'm sorry to the people of the whole country and blame me, a crow's mouth)

This is not necessarily SARS. The teleconsumer and pachyrgenity are unknown. At the same time, I am aware of the potential seriousness of the problem. The laboratory has been thoroughly cleaned up and disinfected, and the laboratory personnel have been isolated. The laboratory equipment of the experimental operation have carried out relevant monitoring.

Next is the question of how to report it. Reporting directly may scare the doctor. What's more, it may be a new virus. Reporting a mistake is a major problem. There is some necessary information that needs to be checked first.

Of course, the first thing that came to mind was the history of wildlife exposure. At that time, the information obtained was that the patient had returned to his hometown, and it could not be ruled out that he had been in contact with a bat or bitten by a bat.

I also doubted whether the artificial virus-related staff inadvertently infected it. After all, the collective infection of Beurdfoun some time ago was caused by incomplete sterilization in a factory.

If there is no more information, we still need to communicate with the doctor to obtain more information. At the same time, we know that there is this unknown virus in the sample now for examination, and other things will be left to the doctor for investigation and handling.

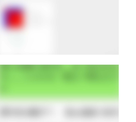
I had communicated with the doctor before noon, and the patient was quarantined.

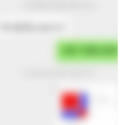
Because I didn't know more information, the patient was also quarantined. The virus is not a new SARS. Thinking about it, it should not be SARS like, the information is unknown, so the nervous mood is also a little relaxed. However, because the patient is seriously ill, it should not be underestimated, and he is still communicating closely with the hospital. During this period, some in-depth analysis will also continue. On the 27th and 28th, I also had some communication and communication with colleagues with the hospital and the disease control personnel by phone. On the 29th and 30th, they also went to Wuhan to report and communicate the matter face to face with the leaders of the hospital and the Center for Disease Control and Prevention, including all our analysis results and the analysis results of the unknown virus sent by colleagues from other units at night. Everything is under intense, confidential and rigorous investigation (at this time, the hospital and disease control personnel already knew that there were many similar patients. After communicating the test results, we had already started emergency treatment, but I didn't know it).

I thought this matter would pass quickly. After all, except for this patient, I didn't seem to hear any other patients infected, but by December 30, I heard that there were still many patients with similar symptoms, and my nerves suddenly tightened again. In particular, about the aftermath of the 360s, a "friend" may have detected the same virus in the sample of another patient, but they directly sent a report that the SARS coronavirus was detected and the laboratory was immediately closed. At the same time, the relevant units also issued an announcement of "unexplained pneumonia", and the relevant rumors also began to spread on Weibo in the early morning of the 31st.

What really makes me nervous again is that friends dared sequence for us to analyze. When I analyzed it, it was indeed the same virus! The first idea in the subconscious is that "this virus is contagious"! Maybe it's really a new type of SARS!

I got the sequence of friends late at night on the 30th for analysis.





Two unknown viruses come together, with more than 99% similarity.

Blue is the sequence of the family synthesized by bases and is compared and contrasted with SARS, the vertical lines with SARS has mutated into other, and it is completely different from the other two. Red is the mutated sequence of the virus, and it is consistent with SARS.

If the two or two are directly compared, all sequences can be compared, with a similarity of 99.6%. No matter how conservative the region is, the similarity between different species is almost impossible to be so high, and the whole genome coverage exceeds 20%. There is no doubt that it is the same virus!

At this time, I was both very nervous and excited. The tension is that this unknown virus may be as terrible as SARS, the excitement is that we have discovered and confirmed the pathogen early through mNGS technology, and indeed helped the epidemic prevention work. I quickly sorted through prevention and control measures before it spreads widely!

I was in a complicated mood and called with emotion, and then I posted a circle of friends that few people could see:

On the morning of December 31, rumors about SARS on Weibo began to spread widely. I've been waiting to see how the official responds. In the afternoon, the official announcement only said it was "unexplained pneumonia", and the "unexplained pneumonia" in Wuhan has "more than twelve cases, including 7 severe cases. After seeing this news, I felt that things were not going well, and I guessed that the transmission capacity of the virus was not low. However, the official announcement is that "no obvious human to human transmission phenomenon has been found". There is not much data held in the early stage, and the situation is difficult to judge, not to mention the new virus. In order to stabilize social emotions and avoid unnecessary panic among the people, such a notice is actually understandable.

At this time, the expert group has begun to intervene, and the "national team" such as Wuhan Virus Institute also began to analyze and identify. They know more information, there were more samples and cases, and there were more consultations and more professional, so I didn't do much in-depth analysis in the future, waiting for the official results.

At first, I was quite confident in the country's ability to deal with these major public health emergencies. After all, I have been through SARS, the flu and other events. In addition, the Beijing plague incident more than two months ago was also reported by the official department. After reporting on the same day, they formed a joint emergency response mechanism and quickly controlled the spread of the plague within a month. The next morning, they saw the news, and no new infections were found in the future. The prevention and control was very good. , mNGS has made great contributions to the plague incident, thinking that it can also play a great role in the prevention and control of this unknown virus.

The media began to refine rumors. The Beijing News first released "SARS rumors", and the People's Daily and other words were slightly euphemistic, "I'm not sure it was SARS", and the last sight "rumormongers" people were arrested. After seeing these news, somehow, I suddenly felt a little disappointed. What has not been determined or controversial in science is directly a rumor? At first, there was a feeling that the arguments in online public opinion were quite reasonable, and the government's response was more or less unpalatable situation. The high profile refining of rumors in the media will interfere with the characterization of this virus science, and excessively optimistic judgments will make the public lose vigilance and take defensive measures. The subsequent developments have once again confirmed my worries.

After the incident exploded, some friends also come to ask me if I know it, especially friends in Wuhan. After telling them that it must be kept strictly confidential (after all, someone has been arrested), I still revealed a little message to them so that they must pay attention to protection.

This dialogue may also underestimated why SARS is refined and named is the novel coronavirus. After all, the genome similarity is only about 80% (after digression, many similarities are only about 80% of different species of the same type of coronavirus). SARS has brought us too much trauma, the public I was extremely panicked about it.

The similarity of different genes of novel coronavirus and SARS varies from 75% to 94%, especially the E gene, which is related to human cell receptors (ACE2), with only 75% similarity. So it was SARS, which was regarded as a basis. However, in Shi Zhengli's latest article, through their method analysis, the novel coronavirus also belongs to SARS-CoV (SARS-related coronavirus).

Everyone knows the subsequent development of the incident. Here are a few more questions.

Why have we analyzed that this unknown virus is a coronavirus similar to SARS in two days, and all the analysis results have been reported. The official announcement will not be issued until January 7 to clarify that pneumonia is caused by the novel coronavirus?

In fact, the week we do is only to analyze and identify whether there is such a virus in the sample sent for testing, but whether pneumonia is caused by this virus, we have not analyzed it, nor can we analyze it. Detection of a virus does not mean that pneumonia is caused by the virus. Normally, officials need to strictly demonstrate such a major health incident, and there is a rigorous scientific demonstration process and a long cycle. Why is it that the government has to do it now only detect such a virus in multiple samples, but also to verify that pneumonia is caused by this virus, etc.

Separation, cultivation, verification, etc. are all very time-consuming, and they also need to be discussed by experts and reach a consensus. From many papers on the novel coronavirus published recently, it can also be seen that many sequencing data were completed in the first two or three days of January.

Another very general thing about this matter is that people who know the truth are so alienated, and people who do not know the truth are all kinds of "strong popular science" and "deep analysis", the companies that have experienced it dare not say anything, and other companies take advantage of the opportunity to market. This is the case from plague to novel coronavirus, hehe, it's so interesting.

When you want to talk about year-ender on the whole event, the biggest feeling is disappointment, sadness and anger. We have all heard that it is so timely. Why can't we control it now? Let the whole country enter the epidemic war? It's not more scientific or technical, but more thinking and media.

I used to be a very angry young man. At this point, I can't help but return to heaven. Passing confidence is the most important thing. After that, everyone is an elite. So our thoughts that things would develop like this, so they didn't bother to criticize many things. But I still want to say a few words.

There was not much problem with refining rumors of SARS and optimism promoted by the media in the earliest days. After all, the understanding of this virus was very limited. In dealing with these major public health events, the disease control system may adopt the rule of "strict internal and external isolation", with internal caution, strict verification and careful statements, and external communication with the media and public opinion mainly positive news. What's more, in this incident, how to explain to the public is obviously not something that the CDC system can make decisions alone.

Needless to say, the status of Wuhan's transportation hub was approaching the Spring Festival at that time, which was the peak season for returning, tourism, film and other service industries. Pessimistic propaganda will affect the economy, the stock market, and the mood of the entire consumer pool among the public, leading to serious consequences such as material hoarding and social chaos. If the virus really has no ability to spread, or because these measures have been taken to kill the spread of the virus, society does not see what harm the virus can bring, decision makers will inevitably be scolded for overreaction and making a fuss, and scientists have to take the blame at the time. On the contrary, if the virus is not controlled and spread, the "Wuhan Virus Institute" will be the first to blame, if the virus is strong, it will lead to the public's lack of awareness of prevention and control, and it will be difficult to carry out prevention and control work, which will eventually lead to the rapid spread of the virus and more serious consequences. Therefore, when you are a decision maker, you have to consider all aspects together. Social, economic and political aspects need to be balanced, which tests the ability and vision of policymakers and experts very much.

Go right, plain, wrong, sinner through the ages.

There is no parallel universe in this world, and we can't know which decision is more correct or wrong. In the atmosphere of everything that likes to be publicized to optimism in China, and the existence of Tuhao psychology, you can imagine what decision makers made.

In the later stage of the incident, especially around January 12, I believe that experts have seen that things are developing in a bad direction (about 30% of the first 41 patients with confirmed cases have no history of contact with the seafood market, the number of medical staff infections keeps many places, and some have been initially positive for nucleic acid). However, the follow-up publicity is still too optimistic and artistic, and at the same time, it is still commonly refining rumors, and there is no sign of issuing an early warning. " It is not ruled out that limited human-to-human transmission, the risk of continuous human-to-human transmission is low, and it can be prevented and controlled". The wording is quite cautious and euphemistic, and it seems that it is still "controllable and safe" in the ordinary people know what "limited human-to-to-to human transmission" is? What is "continuous human-to human transmission"? Their understanding may be that they basically don't have human-to-human transmission, and naturally they are not so guard. I also popularized science in Moments.

From many papers published soon, it can also be seen that many experts have a deep understanding of the transmission ability of the virus very early. The more data they go on, the more data they have, and the clearer the prediction of the situation. I believe that experts have already given not only conservative suggestions, but also strong suggestions. Unfortunately, the force of some people or some media (after all, the publicity was too optimistic before, but now it hurts too much to slap in the face), ignore expert opinions and ignore the health of the people, they have to mislead optimism and tomed or not issuing early warnings, this is unmatchable. Forgive me. No matter who it is, the punishment that should be punished and the mistakes of office should be removed.

I need say more. Let's talk about some hope.

mNGS is indeed a good technology that plays a great role in the diagnosis of difficult and critical cases and the early monitoring and outbreak monitoring of each major public health events. At present, there are many manufacturers that provide mNGS pathogen detection services. It is recommended that medical institutions and disease control communication channels with some well-tech enterprises, so that they can respond to emergencies like this more quickly. In addition, mNGS enterprises can also establish an information sharing platform to share information in time when encountering these incidents to see if there is an outbreak (I know this is difficult. On the one hand, it is a business secret in the first place, and on the other hand, no one dares to report sensitive information for fear of being blamed. I think this matter is worth and deserves a cost between a trader and a doctor.)

I also hope that after experiencing this novel coronavirus incident, the country's ability to deal with major public health events will make great progress. The ability to deal with major public health events is also a manifestation of the modernization of the country. Pursuing an more harmful aspect" to minimize the public to the more prosperous and lower psychological expectations. As the saying goes. Without expectations, you will not be disappointed.

For me, this matter seems to be a year-end tears. I have learned everything. The personal and handed in a qualified answer, but has this answer placed the people's risk? For the first time, I personally participated in such a major public health event, made a little contribution, and learned and grown a lot.

As for us I know, it should be who first found this virus, because it was also after we reported the results that the CDC began to intervene. Unfortunately, I don't know the specific time, I only know that on January WEBSITE, THE SAMPLE COLLECTION TIME IS ALSO THE EARLIEST FOR US. Maybe other institutions have also detected this virus, but this is an unknown virus. There is no reference genome in the nucleic acid database. They may not have the ability to analyze and identify this genes.

Therefore, it should be the first time we have discovered this novel coronavirus. Record it.

I hope we can defeat the epidemic as soon as possible.

Come on, China! Come on, Wuhan!

2020.01.28

0 1

WRITTEN BY
More Less
Passionate about science, current affairs commentary, pursuit of equality and justice

Follow

Follow

coronavirus19
0 Comments

SHARE

Learn more
Medium is an open platform where 170 million readers come to find insightful and dynamic thinking. Here, expert and emerging writers throughout the world share their stories, ideas, and expertise. Find out more by reading, watching, listening.

Make Medium yours.
Follow all the topics you care about, follow your favorite writers, and curate their stories on your Topics page to find a personalized mix of stories.

Write a story on Medium.
Anyone can write on Medium. Your ideas matter, too. Share your stories in a distraction-free editor — everyone starts. Discover tips, learn how to write and find your first readers.

Medium

About Write Help Legal