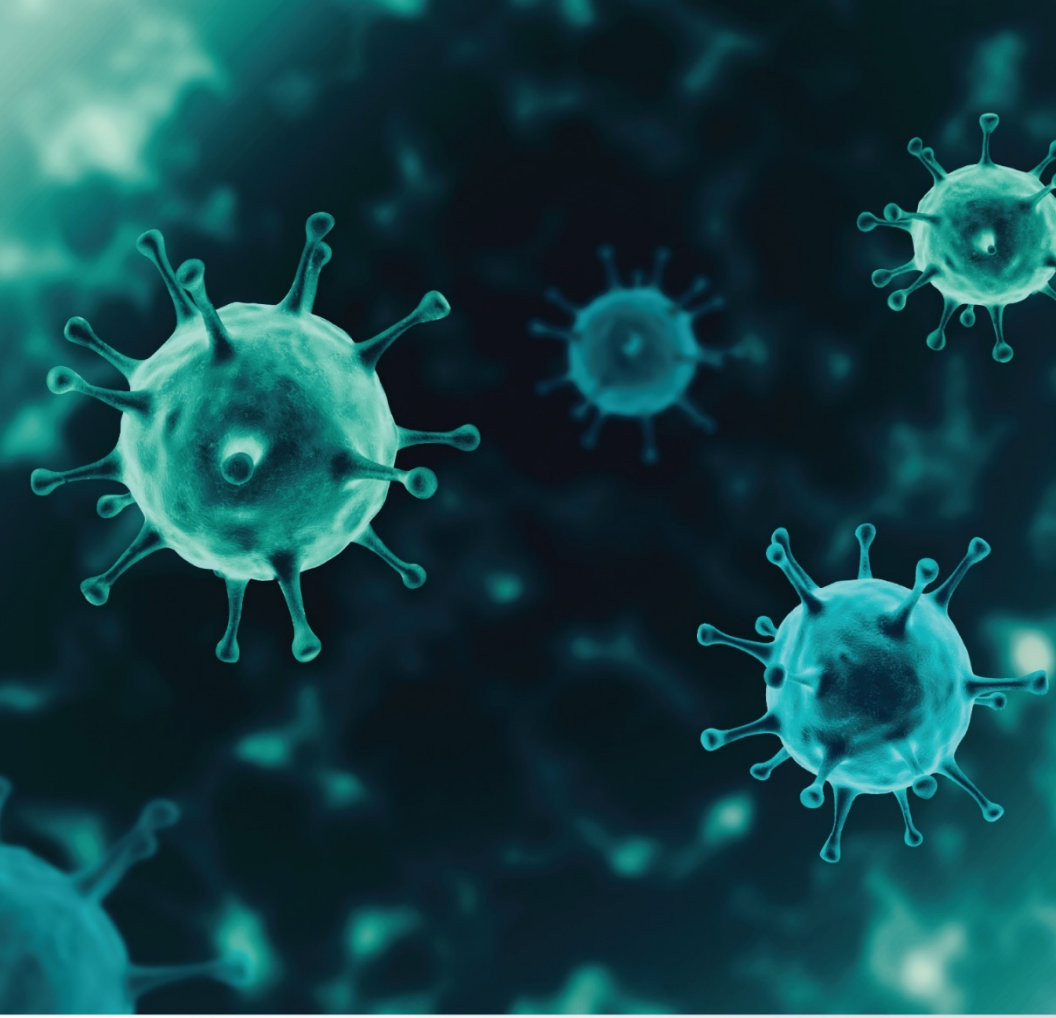

Andreas Scherer, Ph.D.

GENETIC ANALYSIS OF THE COVID-19 VIRUS AND OTHER PATHOGENS

GOLDEN HELIX
Enabling Precision Medicine

Copyright © 2020 Andreas Scherer, Ph.D.
All rights reserved.
ISBN: 978-0-9986882-8-2

Table of Contents

Preface

The COVID-19 pandemic is reaching historic proportions. We are dealing with an infectious disease that is caused by a novel coronavirus we discovered just a few months ago. Since then, it has brought healthcare systems to the brink, it altered how we work, it changed how we socialize, and it impacted in a major way the world economy.

It has mobilized a global response trying to defeat it. There is no good answer currently available. The goal is to reach a sufficiently high level of immunization in the global population and to develop treatment options. In the meantime, we have to be efficient in diagnosing infections, isolating COVID-19 cases and study this virus by understanding its subtypes, its epidemiology, its routes of transmission, and its clinical manifestation.

In this eBook, I summarize to the best of my knowledge our current understanding of COVID-19 and the virus SARS-Cov-2 causing it. Most of the papers that I used to prepare this eBook were published since the beginning of this year. The body of knowledge is quickly expanding and evolving. Next-Generation Sequencing (NGS) can deliver significant insights in this process. This eBook outlines how NGS can be used and how Golden Helix software solutions can be utilized in the process.

I'd like to thank Gabe Rudy for his contributions and feedback. Also, I'd like to thank Delaina Hawkins for her support in publishing this eBook.

Andreas Scherer, Ph.D.

April 2020
Bozeman, Montana

Chapter 1: Introduction

At the end of 2019, a virus appeared somewhere in the Chinese city of Wuhan. It caused cold and flu-like symptoms in most, but also pneumonia and death in a few. It proved to be highly contagious. The disease it caused would soon be coined COVID-19, standing for coronavirus disease of 2019. It quickly emerged as a global phenomenon mobilizing resources in every country to defeat it.

At the time of this writing, we are in the midst of a global pandemic. COVID-19 has arrived in many countries: Asia, Europe, and Latin-America. There are cases reported in Australia, the Middle East, Africa, and Canada. The United States was hit especially hard, trying to contain the exponential spread in a country that is based on individual freedom and liberty. As the country is adopting, based on federal and state guidance, social distancing, we are facing the reality that the reported case numbers are climbing undeterred (see fig. 1).

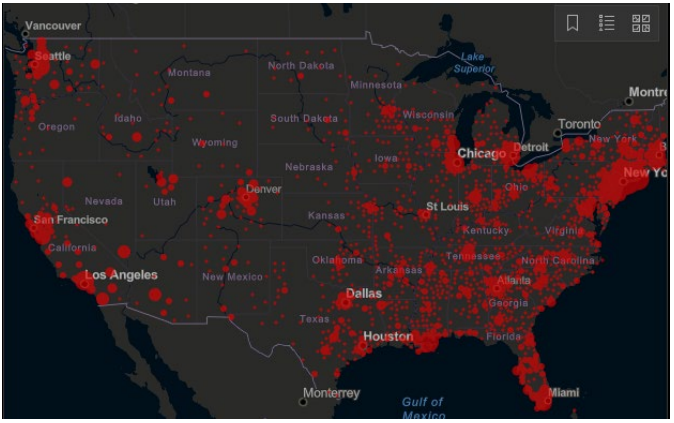

Fig 1: COVID-19 infections in the US

As I am writing this sentence, there have been 159,184 confirmed cases of the coronavirus and the death toll stands at 2,953 here in the US. This will obviously be out of date by the time you are reading this paragraph. We are dealing with exponentially growing numbers. There are estimations for the US, that put the number of infections in the tens of millions and the number of deaths in the hundreds of thousands. The Johns Hopkins University has developed a website, the Coronavirus Resource Center, that gives up-to-date information on this pandemic with updated information multiple times a day.

There are other informative resources, that help to quantify the spread of the virus. Dr. Edward Parker, from the Vaccine Centre at the London School of Hygiene & Tropical Medicine, is maintaining a website that allows you to visualize current trends with his COVID-19 tracker (see fig. 2).

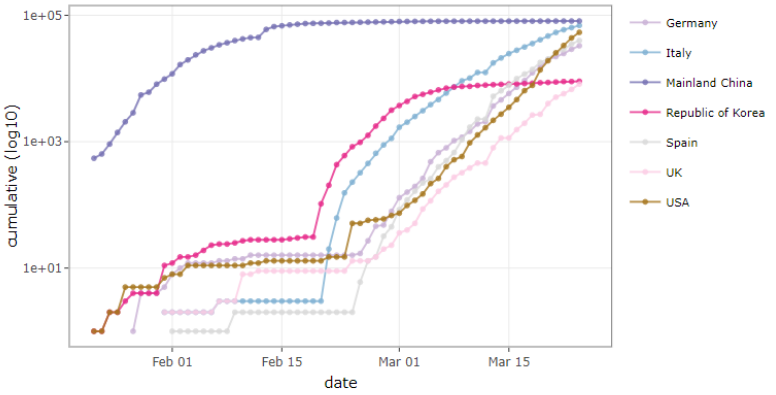

Fig 2: Cumulative (log10) cases in Germany, Italy, Mainland China, Republic of Korea, Spain, UK, and US

In addition, the website gives information about other recent outbreaks, such as the epidemic of Severe Acute Respiratory Syndrome (SARS) in 2003, the 2009 swine flu pandemic, and the 2014 Ebola outbreak.

The virus behind COVID-19 is called SARS-CoV-2. It is a pathogen that has unique characteristics turning it into a threat to our lives and the global economy. According to Fang et al. (2020), the preliminary estimate of R0, which indicates the expected number of cases directly produced by one person in a population susceptible to infection, for COVID-19 is 2.2 - 3.7 (see also Qun Li et al., 2020). It could be shown that it is able persist for days on uncleaned surfaces.

As it stands, this new coronavirus is one of the topmost lethal pathogens we have witnessed. Here is a list of other noteworthy, very lethal viruses (see Harding 2020).

Marburg-Virus: This virus was identified in 1967 when a number of lab technicians were exposed to infected monkeys from Uganda. It causes hemorrhagic fever that can lead to shock, organ failure, and ultimately death with a mortality rate of 25%-80%.

Ebola Virus: It was simultaneously observed in Sudan as well as in the Republic of Congo in 1976. Its mortality rate varies and can reach up to 71%.

Rabies: This disease severely affects the brain and is lethal if left untreated. The development of vaccines has tremendously helped to limit the impact of this virus globally.

Human Immunodeficiency Virus (HIV): There are two species of Lentivirus, a retrovirus, that are capable of infecting humans. It's a sexually transmitted infection that occurs via contact with contaminated blood, semen, vaginal fluids, and breast milk. HIV infects cells of the human immune systems, such as helper T cells, macrophages, and dendric cells. It persists as a problem. Just in the US it is estimated that about 1.1 million people had HIV at the end of 2016.

Dengue: In the 1950s the Dengue Virus was first seen in South-East Asia, especially in the Philippines and Thailand. Since then, it spread further in tropical and subtropical regions. The Dengue virus infects up to 100 million people every year with a mortality rate of 2.5% with Ebola like symptoms if left untreated.

SARS-CoV: This virus causes acute respiratory syndrome. It was first observed in the Guangdong providence in China. It spread to 26 countries. It has a high mortality rate of 9.6% but only effected about 8,000 people.

MERS-CoV: The MERS virus belongs to the family of coronaviruses that causes respiratory syndrome. It has a high mortality rate of up to 40%.

Infectious diseases significantly impact our quality of life, our health, and ultimately our ability to survive. Hence, we need to understand how they function and how we can protect ourselves from negative consequences. The next chapter dives deeper into what we currently know about COVID-19 and the virus causing it (SARS-Cov-2). Chapter 3 reviews in what ways Next-Generation Sequencing (NGS) can be used to study this virus

and to potentially diagnose it. Lastly, in Chapter 4 we review all capabilities of the Golden Helix software stack vis a vis the NGS-based use cases.

Chapter 2: COVID-19 Key Facts

SARS-CoV-2 has just been recently discovered. The knowledge about this virus is fairly new and certain aspects of it are still under review or in-flux entirely as we learn more about this virus on a daily basis. Most of the papers that I cite for this eBook have been published in 2020. In this chapter, I summarize the state of our understanding of this virus based on an article written by Di Wu et al. (2020).

Epidemiology

The virus has rapidly spread from Wuhan to China's other areas and reached global proportions as it is now present on all continents except Antarctica. As of March 26, 2020, there over 510,000 confirmed cases and approximately 23,000 deaths globally. According to the European Centre for Disease Prevention and Control (ECDC), the latest daily risk assessment is moderate to high level. The case fatality rate of the currently reported cases in China is less than 4% which implies that so far this novel coronavirus does not seem to cause the high fatality rates previously observed for SARS-CoV and MERS-CoV. However, it has a higher R0 (2.2 -3.7) value than either of these viruses. SARS-CoV has an R0 of 0.67 – 1.23. MERS-CoV has an R0 of 0.29-0.8 (see Trilla, 2020).

Reservoir Hosts

Bats and other species can function as so-called reservoir hosts. They have played a critical role in transmitting various viruses, including Ebola. Cui et al. (2019) describe the origins of SARS-CoV and MERS-CoV likely to be in bats as there is a strong genetic overlap between the viruses extracted from bats and their human-transmissible versions. In fact, recent research showed

that SARS-CoV-2 is 96% identical at the whole-genome level to a bat coronavirus. Understanding the origins of a virus and when and how exactly the jump to humans occurred helps to understand and eventually control its spreading.

Route of Transmission

Currently, we can confirm that respiratory droplet transmission is the main route of transmission. There are a number of ways how these can transmit from human to human such as unclean surfaces (e.g. doorknobs), clothes, but also aerial droplets and contact. Asymptomatic infections can lead to a wider spread as the hosts are unaware of their ability to transmit. Recently, the new coronavirus was also found in the feces of confirmed patients in Wuhan, Shenzhen, and even in the US. This means that the virus is able to replicate in the digestive tract and opens up the possibility of fecal-oral transmission. At this moment in time, there is no confirmation that aerosol transmission of COVID-19 is possible, however this aspect is still being investigated. Also, neonatal infections, mother-to-child transmission, have been observed but need to be confirmed (see Fuk-Woo et al., 2020; Phelan et al., 2020; Jin et al., 2020; Shen et al., 2019; Zhu et al., 2020).

Clinical Manifestation

Let's look in further detail what we know about the clinical manifestation of this novel pathogen.

Incubation Period and Symptoms

A number of publications based on smaller enrollment numbers suggest an incubation period from 1 to up to 12 days with a mean of 5-7.5 days. In a larger study with 1,099 patients extracting data from laboratory-confirmed cases from 552 hospitals in 30 Chinese provinces, researchers reported that the estimated mean

incubation period of a SARS-CoV-2 infection was 4.0 days (Guan et al., 2020).

The same study (Guan et al., 2020) reported the following data. The median age of the patients was 47 years. 41.9% of the patients were female. 5% were admitted to the ICU. 2.3% underwent invasive mechanical ventilation and 1.4% died. Only 1.9% of the patients had a history of direct contact with wildlife. The most commons symptoms were the following:

- Fever: 43% on admission, 88.7% during hospitalization
- Cough: 67.8%
- Diarrhea: 3.8%

Coincidentally, the SARS-CoV-2 infected cases have symptoms like fever, fatigue, dry cough, dyspnea, etc., with or without nasal congestion, runny nose, or other upper respiratory symptoms. There are reports of loss of smell and taste in otherwise non-symptomatic cases.

Diagnosis

From a diagnostic standpoint here the available options:

Physical Examination

Some patients may not present any clinical noteworthy symptoms, despite being infected with the virus, except perhaps the loss of smell or taste. Patients in severe condition may have shortness of breath, moist rales in lungs, weakened breath sounds, dullness in percussion, and increased or decreased speech tremor.

CT Imaging Examination

In the early stage of pneumonia, chest images show multiple small patchy shadows and interstitial changes, remarkable in the

lung periphery. Severe cases can further develop to bilateral multiple ground-glass opacity, infiltrating shadows, and pulmonary consolidation, with infrequent pleural effusion. While chest CT Scan pulmonary lesions are shown more clearly by CT than x-ray examination, including ground-glass opacity and segmental consolidation in bilateral lungs, especially in the lung periphery.

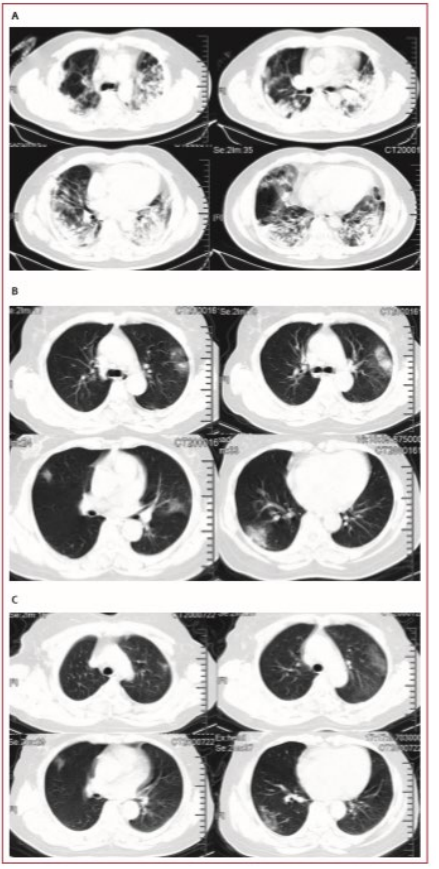

Fig 3: CT Scans of three different patients infected with SARS-CoV-2 in China (Huang et al 2020)

In a study of 41 patients, 40 (98%) had bilateral involvement. The typical findings of chest CT images of ICU patients on admission were bilateral multiple lobular and subsegmental areas of consolidation (A in figure 3). The representative chest CT findings of non-ICU patients showed bilateral ground glass opacity and mental areas of consolidation (B in fig. 3). Later chest CT images showed bilateral ground-glass opacity, whereas the consolidation had been resolved (see fig. 3). This data was extracted from Huang et. Al. (2020).

Laboratory Diagnosis

Current diagnostic strategies involve the exclusion of other known viral causes of pneumonia, such as influenza virus, parainfluenza virus, adenovirus, respiratory syncytial virus, rhinovirus, or SARS-CoV. Also, bacterial infections such as mycoplasma pneumonia, chlamydia pneumonia, and bacterial pneumonia should be tested for prior to conducting a COVID-19 test. A variety of specimens such as nasal swabs, nasopharynx or trachea extracts, sputum or lung tissue, blood, and feces are commonly used for testing.

Should those causes be ruled out, then samples can be collected from the upper respiratory tract (oropharyngeal and nasopharyngeal) or lower respiratory tract (endotracheal aspirate, expectorated sputum, or bronchoalveolar lavage). The standard diagnosis is the CDC 2019-nCov Real-Time RT-PCR Diagnostic Panel, a molecular in vitro diagnostic test, based on the widely used nucleic acid amplification technology.

Treatment and Prevention

At this time, there is no vaccine or antiviral treatment for human and animal coronavirus. The World Health Organization (WHO) has announced that a vaccine for SARS-CoV-2 should

be available in 18 months. The currently available clinical treatment options essentially focus on dealing with the symptoms arising from the infection. This ranges from bed rest, antiviral therapy, antibiotics application, immunomodulating therapy, organ function support, respiratory support, bronchoalveolar lavage (BAL), blood purification, and extracorporeal membrane oxygenation (ECMO). Prevention is mostly about self-isolation, social distancing, and minimizing the exposure to a potential infection while living a healthy lifestyle.

Development of Future Treatment Options

Kupferschmidt and Cohen (2020) review four of the most promising therapies that WHO has identified: an experimental antiviral compound called remdesivir, the mala-ria medication chloroquine and hydroxychloroquine, a combination treatment consisting of lopinavir and ritonavir. and lastly a combination of lopinavir, ritonavir and interferon-beta. All drugs are being tested, although the WHO opted not to conduct randomized double-blind studies in the interest of time.

Remdesivir: It was initially tested to treat Ebola with no confirmed efficacy. There is a reported case in the US with a positive health outcome after a treatment with the drug. More data is required. It is a drug that is being administered intravenously.

Chloroquine and Hydroxychloroquine: Studies in cell cultures have suggested that there is some effect on SARS-Cov-2 at very high doses close to toxic dose ranges. Multiple smaller studies in various countries such as China and France have been conducted that showed some encouraging results. However, there is overall insufficient evidence that warrants a broad usage as of today.

Lopinavir and Ritonavir: This is a drug that was approved in 2000 to treat HIV. It has shown efficacy to treat MERS virus infections. Rigorous data collection is required. The drug can cause severe liver damage.

Lopinavir, Ritonavir and Interferon-Beta: This combination is already in trials to treat MERS. It could be potentially helpful to treat a COVID-19 infection although experts point out that a late application of interferon-beta could actually lead to worse tissue damage.

In the meantime, there is a substantial global effort underway to develop a vaccine that could provide population-level protection against this novel virus. However, there is a major concern. It is a known risk that coronavirus vaccines potentially make the disease. The mechanism that causes that risk is not fully understood and is one of the stumbling blocks that has prevented the successful development of a coronavirus vaccine. Normally, researchers would take months to test for the possibility of vaccine enhancement in animals. Given the urgency to stem the spread of the new coronavirus, some drug makers are moving straight into small-scale human tests, without waiting for the completion of such animal tests (see Steenhuysen, 2020).

In the US, the National Institute of Allergy and Infectious Diseases (NIAID) within the National Institutes of Health (NIH) is overseeing the funding of federal research and response to COVID-19. There are also some companies in the US that are conducting their own COVID-19 research. Internationally, the UK Medicines and Healthcare products Regulatory Agency (MHRA) and European Medicines Agency (EMA) are supporting efforts to develop therapies against COVID-19. In general, it is expected that it takes 12-18 months to develop a

vaccine. This tracker lists the major vaccine candidates currently in development.

Chapter 3: Leveraging NGS-Technology in the Fight Against COVID-19

Next-Gen Sequencing (NGS) of the novel virus can make a tremendous contribution in enhancing our understanding of the underlying pathways in which it impacts humans. In a short period of time we have made significant progress. On January 24, the SARS-CoV-2 genome was published in the New England Journal of Medicine (see Zhu et al., 2020).

Through the Global Initiative to Share All Influenza Data (GISAID) and GenBank, researchers are sharing their understanding of the origin of the new virus, the epidemiology and transmission routes, and facilitate development of diagnostic and treatment strategies (see https://www.gisaid.org/). This website provides a wealth of information as well as latest news on this subject.

Understanding the genome of SARS-CoV-2 early enables us to understand the viral spread and impacted response strategies. Here are a few examples in this context.

Zhou et al. (2020) discusses the whole genome sequences from COVID-19 that were obtained from five patients at an early stage of the outbreak. The sequences are almost identical and share 79.6% sequence identity to SARS-CoV. Furthermore, they were able to show that 2019-nCoV is 96% identical at the whole genome level to a bat coronavirus. Pairwise protein sequence analysis of seven conserved non-structural proteins domains show that this virus belongs to the species of SARS-CoV. In addition, the COVID-19 virus isolated from the bronchoalveolar lavage fluid of a critically ill patient could be neutralized by sera

from several patients. Notably, they confirmed that COVID-19 uses the same cell entry receptor-angiotensin converting enzyme II (ACE2) as SARS-CoV. This study shows how future research studies can be designed.

Tang et al. (2020) conducted a population genetic analyses of 103 SARS-CoV-2 genomes and classified out two prevalent evolvement types of SARS-CoV-2, L type (~ 70%) and S type (~ 30%). The strains in L type, derived from S type, are evolutionarily more aggressive and contagious. This is another example of a study design that could be useful in future work to understand on a deeper level how this virus functions.

NGS allows researchers to better understand the phylogeny of this virus (see fig. 4).

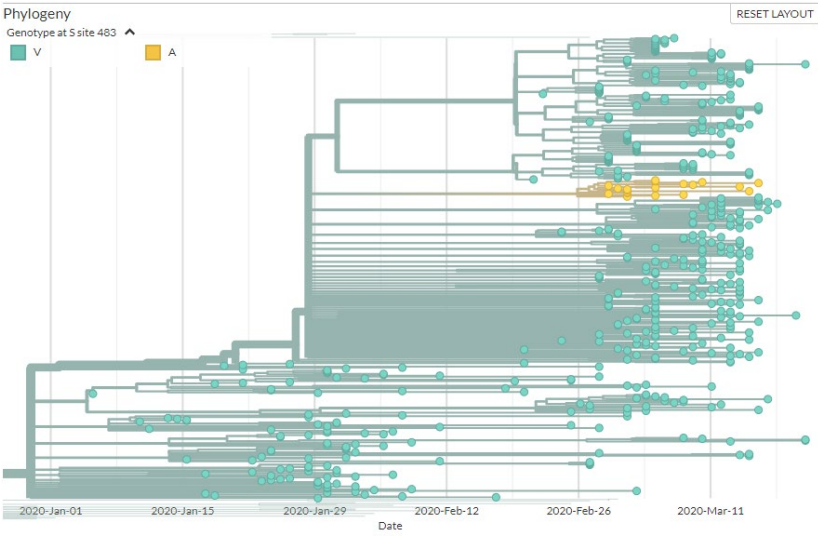

Fig 4: Phylogeny of COVID-19

This is useful to establish a clear picture of so-called transmission routes (see fig. 5).

Fig 5: Possible Transmission Routes of COVID-19

Additionally, there is a growing number of diagnostic use cases for NGS-based tests. Paragon Genomics has recently introduced an NGS Panel for the research and surveillance of COVID-19. Similarly, Fulgent offers an NGS test to detect viral presence. It is safe to assume that we need many more labs to follow those examples and build out testing capabilities leveraging the NGS paradigm.

We are still in the early stages of the global fight against COVID-19 (SARS-CoV-2). There are many unknowns. NGS is one of the key paradigms that can lead us to a deeper understanding of this novel virus.

Chapter 4: Capabilities of Golden Helix in the Infectious Disease Space

Next-Generation Sequencing (NGS) technology has decreased the price of nucleotide sequencing exponentially in the last 10 years. The clinical applications are broad, from diagnosis of rare diseases, to carrier screening and hereditary disease risk, and finally, for the personalized treatment of cancer with molecular profiling of tumors for therapeutic, diagnostic, and prognostic genomic biomarkers.

Discovery and Sharing of the Reference Genome for SARS-CoV-2

With sequencing machines broadly available to clinical and research labs, the identification and sequencing of the complete genome of novel pathogens can be done by small groups in a matter of days. Shortly after the outbreak of severe illness, Chinese scientists were able to identify a novel coronavirus in samples taken from the first patients in Wuhan province. Using standard NGS machines and a metagenomic RNA sequencing protocol, a complete genome of the novel virus was assembled and shared with the world on January 12, 2020, by the Chinese authorities. A second group corroborated this finding shortly thereafter. On January 29, 2020, five days after the French Ministry of Health confirmed the first cases of the Wuhan coronavirus, the Institute Pasteur sequenced and shared the whole genome of the coronavirus of two of the first three confirmed cases in France (2019-NCOV Press Release). In comparison, in 2004 it was considered a heroic feat of the international consortium of scientists to sequence the SARS

virus genome in 31 days after the investigation of the virus outbreak started (WIRED, 2004).

The final genome of sequenced SARS-CoV-2 consists of a single, positive-stranded RNA that is 29,811 nucleotides long, broken down as follows (GenBank Accession MG772933, RefSeq Accession NC_045512):

- 8,903 (29.86%) adenosines
- 5,482 (18.39%) cytosines
- 5,852 (19.63%) guanines
- 9,574 (32.12%) thymines

Golden Helix's Data Curation Team brings together genome assemblies and relevant annotations of many species beyond human. Over the course of many years providing support for research and clinical genome analysis applications, it has published curations of over 100 genomes and 1,700 unique annotations including gene models, functional evidence, variation frequencies and clinical interpretations. The SARS-CoV-2 genome and gene model was curated and published within 24 hours of the first customer request.

Other genomes in the infectious disease space have previously been curated by Golden Helix. In figure 6, you find a few examples.

COMMON NAME	GENOME SIZE
Malaria parasite	23Mbp
Staphylococcus aureus	2.9MBp
Tuberculosis causative agent	4.4Mbp
SARS-CoV-2	29.9Kbp
Ebola virus	18kbp

Fig 6: Examples of curated pathogen reference genomes

A Golden Helix user interested in doing analysis on a novel genome has a couple of options. They can contact our Support Team and expect a short turn-around for the curated and published genome to be available for global access. Alternatively, there are existing capabilities built into the software to curate new reference sequences with a guided user-friendly wizard.

Visualization of Reference and Gene Model

Golden Helix provides different analysis tools for research and clinical genomic analysis applications. For researchers, SNP & Variation Suite (SVS) provides general purpose data normalization and annotations, statistical analysis for a broad set of research questions and expansive visualizations for exploring analysis results.

On the clinical side, VarSeq and VSClinical support the analysis of individual patients' genomic variants for the entire spectrum of clinical genomic tests ranging from focused gene panels, exomes, to complete genomes for hereditary and cancer workflows. Both of these tools embed an integrated genomic visualization tool, GenomeBrowse that is also available as

standalone application, free for researchers. Figure 7, below, demonstrates the visualization capabilities of GenomeBrowse for the curated reference genome and genes for SARS-CoV-2 (see fig 7).

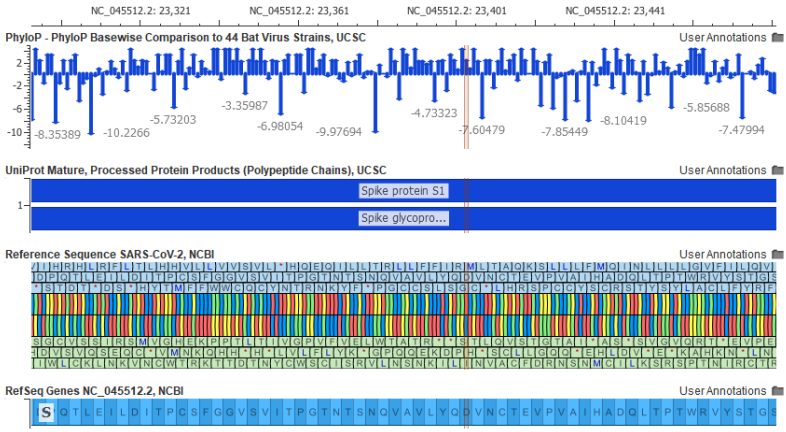

Fig 7: Reference Sequence and Gene Model of SARS-CoV-2 Viral Genome, zoomed to conserved region of Spike Protein S1 with GenomeBrowse

NGS Analysis Pipeline

Now that a reference sequence has been established, easier comparisons can be made between samples and standard analysis workflows and tools for processing NGS data can be employed. This processing, often called "secondary analysis," takes the millions of short (often 150 nucleotide) sequences produced from the machine and "align" or map them to their position on the reference genome. When multiple reads overlap, the confidence in the true sequence of the sample is raised at that position. Sometimes, individual letters in a sample will be

different than the reference, causing a "variation." A tool specializing in modeling the chance that these variations are real and not technical artifacts of one form or another is called a variant caller, and it produces a Variant Call File (VCF) that can be processed in conjunction with the variants of other samples in a downstream analysis tool such as SVS or VarSeq.

Golden Helix provides industry leading secondary analysis tools for doing the following industry standard workflow for SARS-CoV-2 NGS samples:

- Align reads to the reference genome using the BMA-MEM short read alignment algorithm (Li, 2013)
- Remove unwanted identical (duplicate) reads caused by PCR amplification
- Call variants using an accelerated GATK algorithm implemented by Sentieon (Poplin et al. 2017)

Fig 8: The Wuhan patient's meta-transcriptomic NGS reads aligned to the SARS-CoV-2 Genome created from this sample

Here is what the alignment of the raw NGS data taken from the first Wuhan patient's meta-transcriptome looks like aligned to the SARS-CoV-2 genome following this workflow (see fig 8).

Research and Clinical Analysis Workflows

The data sharing precedent of the Chinese and French scientists has been carried forward as COVID-19 reached global pandemic status. Sequences from patient samples are posted to public repositories such as GISAID on a daily basis, and countries and networks are systematizing the collection of samples for further research and analysis.

Protocols for more efficiently sequencing the virus are being developed and shared to provide complete genome coverage with much fewer NGS reads, allowing more samples to be sequenced per run of a sequencing machine and thus more economically. The following analysis examples are based on samples shared from the Utah Public Health Laboratory (BioProject PRJNA614995), prepared using the ARTIC Protocol (SARS-CoV-2 Sequencing Resources) for targeted RNA amplification of SARS-CoV-2 and run through the above secondary analysis pipeline. On average 98% of the NGS reads sequenced aligned to the curated reference, with an average depth of coverage of 2,800X.

Population Analysis Using SNP & Variation Suite (SVS)

SVS has been used for over a decade by researchers doing genomic research on populations of many species, resulting in over 1,500 publication citations in research journals. The

capabilities span from RNA to DNA analysis, extracting relationships between phenotypes and genotypes using complex statistical models that can account for multiple covariate phenotypes and cryptic relatedness between individuals.

For COVID-19 samples, the statistical framework can be applied to the sequenced variants of analyzed cohorts that can grow to hundreds of thousands or millions of samples. Common statistical methods used in these large-N workflows include:

- **Mutation Analysis and Tabulation**: The frequency and segregation of samples based on which variants they contain can be analyzed in various forms.
- **Principle Component Analysis**: By reducing high-dimensional data down to the eigenvectors that best represent the per-sample differences, clustering of the samples by these empirical factors can clearly track the lineage of mutational strains of the virus.
- **Statistical Association**: With sample phenotypes and clinical attributes such as incubation period and disease outcomes incorporated, statistically significant associations with specific variants may lead to key insights into behavior and pathogenesis of various strains of the virus.

Cohort and Sample Analysis Using VarSeq

VarSeq specializes in variant analysis workflows. While some features are focused on the clinical workflows associated with gene testing for rare disease diagnosis, carrier screening and cancer gene testing, many are generalized across analysis types and species. These include variant merging and normalization,

annotation against curated genomic resources, aggregation of allele and genotype counts and frequencies of sample cohorts, and the integrated visualization and reporting of filtered and selected variants in the genomic context.

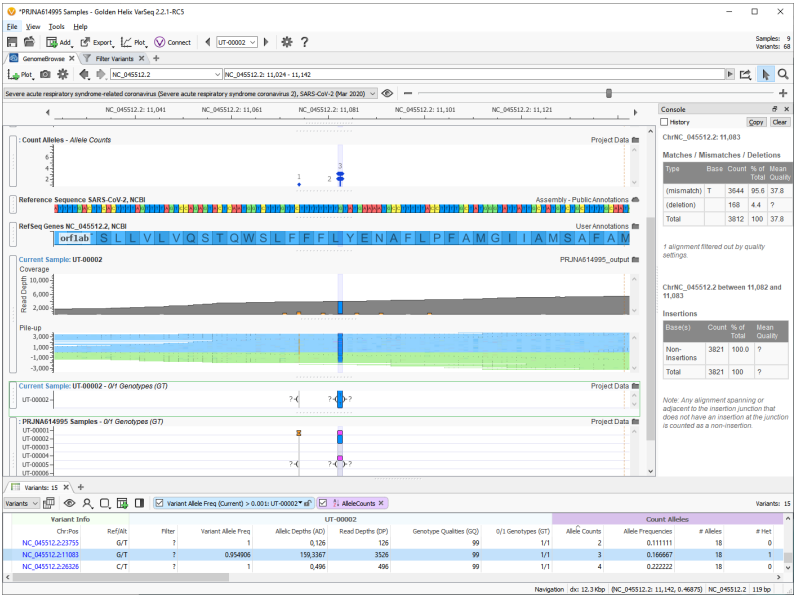

Fig 9: A Variant Analysis and Visualization in VarSeq of COVID-19 sample

VarSeq analysis of COVID-19 samples can include (see fig 9):

- Merging all samples into a single cohort matrix
- Annotating variants against the gene model for the virus, determining their impact as silent or missense (amino acid changing) to the protein sequence
- Counting the samples and frequency of mutations at each position

- Plotting these counts, along with the variant and NGS supporting read information
- Filtering variants for inclusion in a clinical report that identifies the strain and sub-type of the virus with supporting evidence of the sample quality and coverage statistics

Sample Collection and Aggregation with VSWarehouse

As the global effort to capture samples for complete genome sequencing of the virus continues, the need to aggregate many samples into a collaborative environment for ongoing research and clinical application development will warrant a genomic data warehouse.

VSWarehouse was designed to support the incremental aggregation of samples into a collaborative and sharable centralized resource. The solution scales to very large datasets that are enriched with genomic annotations and supports efficient queries and exports for various research workflows.

The basic COVID-19 workflow may include (see fig 10):

- Enriching sample collections with clinical and patient phenotypes
- Uploading samples through the VarSeq integration with VSWarehouse to a collaborative project
- Setting up the project to integrate the latest genomic annotations, providing critical variant functional impact annotations and updated sample and frequency counts of all variants

- Sharing the project across institutions, allowing direct annotation of any COVID-19 samples with the frequency and counts of variants from within VarSeq
- Supporting research workflows to export out subsets or cohorts for statistical analysis in tools such as SVS

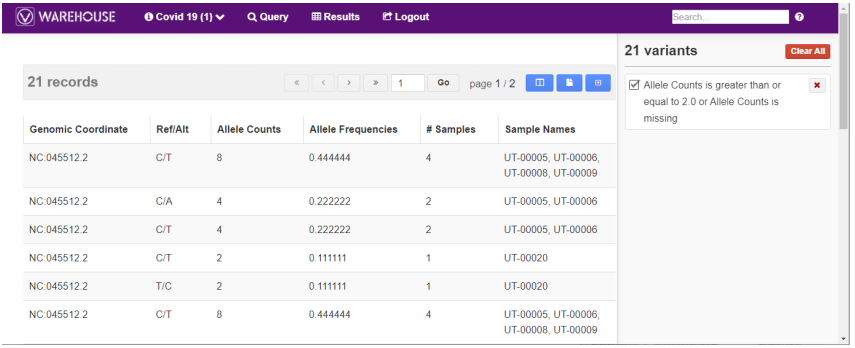

Fig 10: VSWarehouse table of variants from a cohort of COVID-19 samples

Clinical Testing with NGS Sequencing

In concert with the rapid-diagnosis capabilities of RT-PCR tests currently in use, the capabilities of NGS machines may be employed to capture the complete genomes of the virus as it spreads and evolves as well as confirm the virus presence. Longer term, these collected virus samples will provide critical data to understand the evolution of the virus, its biological properties and to aid in the development of therapeutic drugs and even vaccines.

The key to scaling up any NGS pipeline for clinical diagnosis and rapid turnaround is the automation of as many parts of the bioinformatic process as possible. This not only reduces time-to-

result but also removes potential for human error in otherwise manually performed steps. Golden Helix provides VSPipeline as the high-throughput automation tool that enables repeatable workflows of the NGS analysis capabilities previously discussed.

With molecular barcoding, the throughput of NGS machines allows for many samples to be multiplexed on a single run. The automated analysis workflow after sequencing run may look like:

- De-multiplex raw reads into per-sample sequence files
- Align sequence reads to the SARS-CoV-2
- Remove duplicate reads
- Call variants
- Run customized workflow with VSPipeline to:
    - Call Positive/Negative for SARS-CoV-2 based on coverage analysis
    - Analysis of strain based on variants
    - Output PDF clinical report

While these workflows presented here are certainly not comprehensive, they provide examples of common analysis strategies for different use cases of COVID-19 NGS sequencing. The space of NGS analysis is constantly changing, and Golden Helix's tools have been built with maximum flexibility to tackle the unknown problems of the future.

Chapter 5: Summary

COVID-19 challenges the healthcare system, governments and the global economy in an unprecedented way. The novel coronavirus SARS-CoV-2 has just recently been discovered. Vast resources are being directed towards understanding this new disease and the virus that caused it.

This eBook outlines, to the best of my knowledge, the current state of our understanding of COVID-19 and SARS-CoV-2. Chapter 1 gives a brief introduction into this topic. Chapter 2 summarizes key facts about COVID-19; it reviews epidemiology, reservoir hosts, transmission routes, and clinical manifestation. Chapter 3 answers the question of how Next-Generation Sequencing can be utilized in this area. Lastly, Chapter 4 reviews Golden Helix's capabilities that are relevant in this context.

By the time you read this sentence, there will be new findings, studies, and research papers. Also, any of the population level statistics that have been cited will be dwarfed by the numbers that are reported by the time you read these words. But one thing is certain: Next-Gen Sequencing technologies will allow us to gain a deeper understanding of this virus and to develop advanced diagnostic capabilities to help patients and provide us the ability to conduct research at the same time. Golden Helix software can be utilized in the process. On our end, we plan to contribute to the global effort defeating this virus, by providing capabilities that allow clinicians and researchers to do their job.

Learn More About Golden Helix's Solutions for COVID-19

Bibliography

Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, Nicholas Davies, Amy Gimma, Kevin van Zandvoort, Hamish Gibbs, Joel Hellewell, Christopher I Jarvis, Sam Clifford, Billy J Quilty, Nikos I Bosse, Sam Abbott, Petra Klepac, Stefan Flasche. (2020) Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The Lancet Infectious Diseases.

Alessia Lai, Annalisa Bergna, Carla Acciarri, Massimo Galli, Gianguglielmo Zehender. (2020) Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2. Journal of Medical Virology 2020.

Anita Patel, Daniel B. Jernigan, , Fatuma Abdirizak, Glen Abedi, Sharad Aggarwal, Denise Albina, Elizabeth Allen, Lauren Andersen, Jade Anderson, Megan Anderson, Tara Anderson, Kayla Anderson, Ana Cecilia Bardossy, Vaughn Barry, Karlyn Beer, Michael Bell, Sherri Berger, Joseph Bertulfo, Holly Biggs, Jennifer Bornemann, Josh Bornstein, Willie Bower, Joseph Bresee, Clive Brown, Alicia Budd, Jennifer Buigut, Stephen Burke, Rachel Burke, Erin Burns, Jay Butler, Russell Cantrell, Cristina Cardemil, Jordan Cates, Marty Cetron, Kevin Chatham-Stephens, Kevin Chatham-Stevens, Nora Chea, Bryan Christensen, Victoria Chu, Kevin Clarke, Angela Cleveland, Nicole Cohen, Max Cohen, Amanda Cohn, Jennifer Collins, Erin Conners, Aaron Curns, Rebecca Dahl, Walter Daley, Vishal Dasari, Elizabeth Davlantes, Patrick Dawson, Lisa Delaney, Matthew Donahue, Chad Dowell, Jonathan Dyal, William Edens, Rachel Eidex, Lauren Epstein, Mary Evans, Ryan Fagan, Kevin Farris, Leora Feldstein, LeAnne Fox, Mark Frank, Brandi Freeman, Alicia Fry, James Fuller, Romeo Galang, Sue Gerber, Runa Gokhale, Sue Goldstein, Sue Gorman, William Gregg, William Greim, Steven Grube, Aron Hall, Amber Haynes, Sherrasa Hill, Jennifer Hornsby-Myers, Jennifer Hunter, Christopher Ionta, Cheryl Isenhour, Max Jacobs, Kara Jacobs Slifka, Daniel Jernigan, Michael Jhung, Jamie Jones-Wormley, Anita Kambhampati, Shifaq Kamili, Pamela Kennedy, Charlotte Kent, Marie Killerby, Lindsay Kim, Hannah Kirking, Lisa Koonin, Ram Koppaka, Christine Kosmos, David Kuhar, Wendi Kuhnert-Tallman, Stephanie Kujawski, Archana Kumar, Alexander Landon, Leslie Lee, Jessica Leung, Stephen Lindstrom, Ruth Link-Gelles, Joana Lively, Xiaoyan Lu, Brian Lynch, Lakshmi Malapati, Samantha Mandel, Brian Manns, Nina Marano, Mariel Marlow, Barbara Marston, Nancy McClung, Liz McClure, Emily McDonald, Oliva McGovern, Nancy Messonnier, Claire Midgley, Danielle Moulia, Janna Murray, Kate Noelte, Michelle Noonan-Smith, Kristen Nordlund, Emily Norton, Sara Oliver, Mark Pallansch, Umesh Parashar, Anita Patel, Manisha Patel, Kristen Pettrone, Taran Pierce, Harald Pietz, Satish Pillai, Lewis Radonovich, Sarah Reagan-Steiner, Amy Reel, Heather Reese, Brian Rha, Philip Ricks, Melissa Rolfes, Shahrokh Roohi, Lauren Roper, Lisa Rotz, Janell Routh, Senthil Kumar Sakthivel, Luisa Sarmiento, Jessica Schindelar, Eileen Schneider, Anne Schuchat, Sarah Scott, Varun Shetty, Caitlin Shockey, Jill Shugart, Mark Stenger, Matthew Stuckey, Brittany Sunshine, Tamara Sykes, Timothy Uyeki, Grace Vahey, Amy Valderrama, Julie Villanueva, Tunicia Walker, Megan Wallace, Lijuan Wang, John Watson, Angie Weber, Cindy Weinbaum, William Weldon, Caroline Westnedge, Brett Whitaker, Michael Whitaker, Alcia Williams, Ian Willams, Karen Wong, Amy Xie, Anna Yousef. (2020) Initial public health response and interim clinical guidance for the 2019 novel coronavirus outbreak — United States, December 31, 2019–February 4, 2020. American Journal of Transplantation 20:3, 889-895.

B. Coutard, C. Valle, X. de Lamballerie, B. Canard, N.G. Seidah, E. Decroly. (2020) The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Research 176, 104742.

Catrin Sohrabi, Zaid Alsafi, Niamh O'Neill, Mehdi Khan, Ahmed Kerwan, Ahmed Al-Jabir, Christos Iosifidis, Riaz Agha. (2020) World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). International Journal of Surgery 76, 71-76.

Cui J, Li F, Shi ZL 2019 Origin and evolution of pathogenic coronaviruses, Nat Rev Microbiol 2019 Mar;17(3): 181-192.

Dawei Wang, Bo Hu, Chang Hu, Fangfang Zhu, Xing Liu, Jing Zhang, Binbin Wang, Hui Xiang, Zhenshun Cheng, Yong Xiong, Yan Zhao, Yirong Li, Xinghuan Wang, Zhiyong Peng. (2020) Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. JAMA 323:11, 1061.

Di Wu, Tiantian Wu, Qun Liu, Zhicong Yang. (2020) The SARS-CoV-2 outbreak: what we know. International Journal of Infectious Diseases.

Fei Zhou, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, Lulu Guan, Yuan Wei, Hui Li, Xudong Wu, Jiuyang Xu, Shengjin Tu, Yi Zhang, Hua Chen, Bin Cao. (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet.

Fuk-Woo CJ, Shuofeng Y, Kin-Hang K et. Al (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person: as study of a family cluster. The Lancet. 2020; 515-523.

Guan WJ, Ni ZY, Hu Y et al (2020) Clinical Characteristics of Coronavirus Disease in 2019 in China, N Engl J Med. 2020.

Harald Brüssow. (2020) The Novel Coronavirus – A Snapshot of Current Knowledge. Microbial Biotechnology.

Harding, A. Lanese, N. (2020) The 12 deadliest viruses on Earth, Live Sciences, March 2020.

Hengbo Zhu, Li Wei, Ping Niu. (2020) The novel coronavirus outbreak in Wuhan, China. Global Health Research and Policy 5:1.

Hiroshi Nishiura, Natalie M. Linton, Andrei R. Akhmetzhanov. (2020) Serial interval of novel coronavirus (COVID-19) infections. International Journal of Infectious Diseases 93, 284-286.

Huan Liang, Ganesh Acharya. (2020) Novel corona virus disease (COVID-19) in pregnancy: What clinical recommendations to follow?. Acta Obstetricia et Gynecologica Scandinavica 99:4, 439-442.

Huang et. al. (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, The Lancelet, Vol 395 February 2020.

Huwen Wang, Zezhou Wang, Yinqiao Dong, Ruijie Chang, Chen Xu, Xiaoyue Yu, Shuxian Zhang, Lhakpa Tsamlag, Meili Shang, Jinyan Huang, Ying Wang, Gang Xu, Tian Shen, Xinxin Zhang, Yong Cai. (2020) Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. Cell Discovery 6:1.

Jennifer Harcourt, Azaibi Tamin, Xiaoyan Lu, Shifaq Kamili, Senthil K. Sakthivel, Janna Murray, Krista Queen, Ying Tao, Clinton R. Paden, Jing Zhang, Yan Li, Anna Uehara, Haibin Wang, Cynthia Goldsmith, Hannah A. Bullock, Lijuan Wang, Brett Whitaker, Brian Lynch, Rashi Gautam, Craig Schindewolf, Kumari G. Lokugamage, Dionna Scharton, Jessica A. Plante, Divya Mirchandani, Steven G. Widen, Krishna Narayanan, Shinji Makino, Thomas G. Ksiazek, Kenneth S. Plante, Scott C. Weaver, Stephen Lindstrom, Suxiang Tong, Vineet D. Menachery, Natalie J. Thornburg. (2020) Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with 2019 Novel Coronavirus Disease, United States. Emerging Infectious Diseases 26:6.

Jiang, F., Deng, L., Zhang, L. et al. Review of the Clinical Characteristics of Coronavirus Disease 2019 (COVID-19). J GEN INTERN MED (2020).

Joseph T. Wu, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M. de Salazar, Benjamin J. Cowling, Marc Lipsitch, Gabriel M. Leung. (2020) Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nature Medicine 162.

Jun She, Jinjun Jiang, Ling Ye, Lijuan Hu, Chunxue Bai, Yuanlin Song. (2020) 2019 novel coronavirus of pneumonia in Wuhan, China: emerging attack and management strategies. Clinical and Translational Medicine 9:1.

Kaiyuan Sun, Jenny Chen, Cécile Viboud. (2020) Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. The Lancet Digital Health 2:4, e201-e208.

Kaiyue Diao, Peilun Han, Tong Pang, Yuan Li, Zhigang Yang. (2020) HRCT imaging features in representative imported cases of 2019 novel coronavirus pneumonia. Precision Clinical Medicine 3:1, 9-13.

Kupferschmidt K, Cohen J (2020) WHO launches globar megatrial of the four most promising coronavirus treatments, Science Magazine, Mar 22, 2020.

Lei Tang, Xiaoyong Zhang, Yvquan Wang, Xianchun Zeng. (2020) Severe COVID-19 Pneumonia: Assessing Inflammation Burden with Volume-rendered Chest CT. Radiology: Cardiothoracic Imaging 2:2, e200044.

Li Heng (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, 2013 arXiv:1303.3997

Lin Li, Lixin Qin, Zeguo Xu, Youbing Yin, Xin Wang, Bin Kong, Junjie Bai, Yi Lu, Zhenghan Fang, Qi Song, Kunlin Cao, Daliang Liu, Guisheng Wang, Qizhong Xu, Xisheng Fang, Shiqin Zhang, Juan Xia, Jun Xia. (2020) Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. Radiology, 200905.

Mary E Wilson, Lin H Chen. (2020) Travellers give wings to novel coronavirus (2019-nCoV). Journal of Travel Medicine 27:2.

Moritz U. G. Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, , Louis du Plessis, Nuno R. Faria, Ruoran Li, William P. Hanage, John S. Brownstein, Maylis Layan, Alessandro Vespignani, Huaiyu Tian, Christopher Dye, Oliver G. Pybus, Samuel V. Scarpino. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. Science, eabb4218.

Nathan Kapata, Chikwe Ihekweazu, Francine Ntoumi, Tajudeen Raji, Pascalina Chanda-Kapata, Peter Mwaba, Victor Mukonka, Matthew Bates, John Tembo, Victor Corman, Sayoki Mfinanga, Danny Asogun, Linzy Elton, Lia Bárbara Arruda, Margaret J. Thomason, Leonard Mboera, Alexei Yavlinsky, Najmul Haider, David Simons, Lara Hollmann, Swaib A. Lule, Francisco Veas, Muzamil Mahdi Abdel Hamid, Osman Dar, Sarah Edwards, Francesco Vairo, Timothy D. McHugh, Christian Drosten, Richard Kock, Giuseppe Ippolito, Alimuddin Zumla. (2020) Is Africa prepared for tackling the COVID-19 (SARS-CoV-2) epidemic. Lessons from past outbreaks, ongoing pan-African public health efforts, and implications for the future. International Journal of Infectious Diseases 93, 233-236.

Nian Shao, Min Zhong, Yue Yan, HanShuang Pan, Jin Cheng, Wenbin Chen. (2020) Dynamic models for Coronavirus Disease 2019 and data analysis. Mathematical Methods in the Applied Sciences 50.

Parham Habibzadeh, Emily K. Stoneman. (2020) The Novel Coronavirus: A Bird's Eye View. The International Journal of Occupational and Environmental Medicine 11:2, 65-71.

Phelan Al, Kat Rebecca, Gostin L (2020) The Novel Coronavirus Origninating in Wuhan China: Challenges for Global Health Governance, Jama 2020.

Poplin et al (2017) Scaling accurate genetic variant discovery to tens of thousands of samples, 2017 bioRxiv

Qianying Lin, Shi Zhao, Daozhou Gao, Yijun Lou, Shu Yang, Salihu S. Musa, Maggie H. Wang, Yongli Cai, Weiming Wang, Lin Yang, Daihai He. (2020) A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. International Journal of Infectious Diseases 93, 211-216.

Qingmei Han, Qingqing Lin, Shenhe Jin, Liangshun You. (2020) Coronavirus 2019-nCoV: A brief perspective from the front line. Journal of Infection 80:4, 373-377.

Qun Li, M.Med., Xuhua Guan, Ph.D., Peng Wu, Ph.D., Xiaoye Wang, M.P.H., Lei Zhou, M.Med., Yeqing Tong, Ph.D., Ruiqi Ren, M.Med., Kathy S.M. Leung, Ph.D., Eric H.Y. Lau, Ph.D., Jessica Y. Wong, Ph.D., Xuesen Xing, Ph.D., Nijuan Xiang, M.Med., Yang Wu, M.Sc., Chao Li, M.P.H., Qi Chen, M.Sc., Dan Li, M.P.H., Tian Liu, B.Med., Jing Zhao, M.Sc., Man Liu, M.Sc., Wenxiao Tu, M.Med., Chuding Chen, M.Sc., Lianmei Jin, M.Med., Rui Yang, M.Med., Qi Wang, M.P.H., Suhua Zhou, M.Med., Rui Wang, M.D., Hui Liu, M.Med., Yinbo Luo, M.Sc., Yuan Liu, M.Med., Ge Shao, B.Med., Huan Li, M.P.H., Zhongfa Tao, M.P.H., Yang Yang, M.Med., Zhiqiang Deng, M.Med., Boxi Liu, M.P.H., Zhitao Ma, M.Med., Yanping Zhang, M.Med., Guoqing Shi, M.P.H., Tommy T.Y. Lam, Ph.D., Joseph T. Wu, Ph.D., George F. Gao, D.Phil., Benjamin J. Cowling, Ph.D., Bo Yang, M.Sc., Gabriel M. Leung, M.D., and Zijian Feng, M.Med, New England Journal of Medicine, March 2020. Vol. 382, No. 13

Roy M Anderson, Hans Heesterbeek, Don Klinkenberg, T Déirdre Hollingsworth. (2020) How will country-based mitigation measures influence the course of the COVID-19 epidemic?. The Lancet 395:10228, 931-934.

Rui Li, Songlin Qiao, Gaiping Zhang. (2020) Analysis of angiotensin-converting enzyme 2 (ACE2) from different species sheds some light on cross-species receptor usage of a novel coronavirus 2019-nCoV. Journal of Infection 80:4, 469-496.

Rui Liu, Huan Han, Fang Liu, Zhihua Lv, Kailang Wu, Yingle Liu, Yong Feng, Chengliang Zhu. (2020) Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. Clinica Chimica Acta 505, 172-175.

Sasmita Poudel Adhikari, Sha Meng, Yu-Ju Wu, Yu-Ping Mao, Rui-Xue Ye, Qing-Zhi Wang, Chang Sun, Sean Sylvia, Scott Rozelle, Hein Raat, Huan Zhou. (2020) Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. Infectious Diseases of Poverty 9:1.

Sheng Zhang, MengYuan Diao, Wenbo Yu, Lei Pei, Zhaofen Lin, Dechang Chen. (2020) Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. International Journal of Infectious Diseases 93, 201-204.

Shen K, Yang Y, Wang T (2020) Diagnosis, treatment and prevention of 2019 novel coronavirus infection in children: experts' consensus statement. World J Pediatr. 2020

Sijia Tian, Nan Hu, Jing Lou, Kun Chen, Xuqin Kang, Zhenjun Xiang, Hui Chen, Dali Wang, Ning Liu, Dong Liu, Gang Chen, Yongliang Zhang, Dou Li, Jianren Li, Huixin Lian, Shengmei Niu, Luxi Zhang, Jinjun Zhang. (2020) Characteristics of COVID-19 infection in Beijing. Journal of Infection 80:4, 401-406.

Simiao Chen, Juntao Yang, Weizhong Yang, Chen Wang, Till Bärnighausen. (2020) COVID-19 control in China during mass population movements at New Year. The Lancet395:10226, 764-766.

Steenhuysen, J (2020) AS pressure for coronavirus Vaccine mounts, scientists debate risks of accelerating testing, Reuters.com, Health News, March 11, 2020.

Tang X, Wu C, Lix X, Song Y, Yao X, Wu X et. Al (2020) On the origin and continuing evolution of SARS-CoV-2, Natl Science Rev, 2020.

Tanu Singhal. (2020) A Review of Coronavirus Disease-2019 (COVID-19). The Indian Journal of Pediatrics 87:4, 281-286.

Tian-Mu Chen, Jia Rui, Qiu-Peng Wang, Ze-Yu Zhao, Jing-An Cui, Ling Yin. (2020) A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infectious Diseases of Poverty 9:1.

Weilong Shang, Yi Yang, Yifan Rao, Xiancai Rao. (2020) The outbreak of SARS-CoV-2 pneumonia calls for viral vaccines. npj Vaccines 5:1.

Wen-Chien Ko, Jean-Marc Rolain, Nan-Yao Lee, Po-Lin Chen, Ching-Tai Huang, Ping-Ing Lee, Po-Ren Hsueh. (2020) Arguments in favour of remdesivir for treating SARS-CoV-2 infections. International Journal of Antimicrobial Agents, 105933.

Xingguang Li, Junjie Zai, Qiang Zhao, Qing Nie, Yi Li, Brian T. Foley, Antoine Chaillon. (2020) Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. Journal of Medical Virology 6.

Xingguang Li, Wei Wang, Xiaofang Zhao, Junjie Zai, Qiang Zhao, Yi Li, Antoine Chaillon. (2020) Transmission dynamics and evolutionary history of 2019-nCoV. Journal of Medical Virology 92:5, 501-511.

Yan-Rong Guo, Qing-Dong Cao, Zhong-Si Hong, Yuan-Yang Tan, Shou-Deng Chen, Hong-Jun Jin, Kai-Sen Tan, De-Yun Wang, Yan Yan. (2020) The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. Military Medical Research 7:1.

Yi Yang, Weilong Shang, Xiancai Rao. (2020) Facing the COVID-19 outbreak: What should we know and what could we do?. Journal of Medical Virology.

Yicheng Fang, Huangqi Zhang, Yunyu Xu, Jicheng Xie, Peipei Pang, Wenbin Ji. (2020) CT Manifestations of Two Cases of 2019 Novel Coronavirus (2019-nCoV) Pneumonia. Radiology 295:1, 208-209.

Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklöv. (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of Travel Medicine 27:2.

Zachary T. Bloomgarden. (2020) Diabetes and COVID-19. Journal of Diabetes 12:4, 347-348.

Zhou et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 2020 Mar; 579(7798): 2070:273.

Zhu H, Wang L, Fang C. etc al (2020) Clinical analysis of 10 neonates born to mothers with 2019-nCoV pneumonia, Translational Pediatrics, 2020; 9: 51-60

Zhu et al (2020) A Novel Coronavirus form Patients with Pneumonia in China, 2019. NEJM, Jan 24, 2020.

About the Author

Dr. Andreas Scherer is CEO of Golden Helix. The company has been delivering industry leading bioinformatics solutions for the advancement of life science research and translational medicine for over two decades. Its innovative technologies and analytic services empower clinicians and scientists at all levels to derive meaning from the rapidly increasing volumes of genomic data produced from next-generation sequencing and micro-arrays. With its solutions, hundreds of the world's hospitals, testing labs, academic research organizations and governments are able to harness the full potential of genomics to identify the cause of disease, develop genomic diagnostics, and advance the quest for personalized medicine. Golden Helix products and services have been cited in thousands of peer-reviewed publications.

He is also Managing Partner of Salto Partners, a management consulting firm headquartered in the DC metro area. He has extensive experience successfully managing growth as well as orchestrating complex turnaround situations.

Dr. Scherer holds a PhD in computer science from the University of Hagen, Germany, and a Master of Computer Science from the University of Dortmund, Germany. He is author and co- author of over 20 international publications and has written books on project management, the Internet, and artificial intelligence. His latest book, "Be Fast Or Be Gone", is a prizewinner in the 2012 Eric Hoffer Book Awards competition, and has been named a finalist in the 2012 Next Generation Indie Book Awards!

Connect with Dr. Scherer: