The COVID-19 pandemic provides a backdrop for the scientific community to come together and work toward a common goal. Successfully emerging from this challenging time will absolutely require interdisciplinary collaboration from many different scientific fields. The constantly evolving situation, growing collection of data, and need for quick action makes the role of computational biologists in this global collaboration even more important.

Computational methods can play multiple vital roles in the response to this situation. For example, sequencing data can be used to track how the SARS-CoV-2 virus spreads across the globe and how quickly it mutates. Likewise, algorithmic methods can aid in identifying candidate drugs for the disease. Computational approaches can even play a key role in how to best distribute limited resources (e.g., personal protective equipment or ventilators). As this situation unfolds, other opportunities will emerge for computation to play a leading role in combatting this virus.

Finally, the important role that algorithms and computational biology play in the response to the COVID-19 pandemic may help to inspire the next generation of scientists. This is a real-world scenario, where the impact of interdisciplinary science is on full display. I hope that one outcome from this pandemic will be that it opens minds and eyes to the importance and essential nature of science—and especially computational science.

One of the main challenges in curtailing and defeating the COVID-19 pandemic is identifying geographical clusters where the virus is quietly spreading—before symptoms appear. Our lab developed a computational tool that tracks evolving symptoms according to location, while maintaining patient privacy, for initial deployment in Israel. Our Predict Corona tool relies on individuals filling out a one-minute online survey once a day and has been used over 1,000,000 times. By modeling the data, we allowed Israel’s health-care authorities to identify clusters of contagion in the making. With that information and other layers of data available to them, they were able to direct more testing in some places and quickly implement more stringent social distancing and quarantining in several cities.

Our approach has since been adopted by over ten other countries. In partnership with researchers in some of these other countries, the project has evolved into the Coronavirus Census Collective, an open-access platform whereby voluntarily reported data can be used to predict the location of future outbreaks. CCC’s data sources include self-reported health status through surveys, diagnostic lab test results, and other static and geospatial data. We hope this effort will help predict hotspots of disease outbreak, identify factors underlying the infection rate, inform policy decisions, evaluate the effectiveness of public health measures, and provide insight into etiology.

Computational biologists suffer much less than their experimental colleagues from restrictions imposed on their activities by this pandemic. They only need a computer and an internet connection to advance their projects and can be 100% active, even when most research institutes are closed. Nevertheless, computational biology lives in continuous exchange with experimental biology, the source of its challenges and data. Without fuel, the car will sooner or later slow down and stop.

As computational biologists, we can and should react to this pandemic as a unique organism, joining our minds and hardware in a cohesive and well-organized collaborative effort, minimizing wastes of time and energy. Hackathons are a model for reacting to this challenge effectively, with many researchers working together on defined topics, exchanging ideas, developing code, and comparing results: not single bees, but a hive with an organized hierarchy of experience, capacity, and manpower. One of the first goals should be capturing known COVID-19 biology to identify important targets and epitopes. The COVID challenge presses for better algorithms for de novo genome assembly, structural annotation, variant analysis, and more. Computational biology projects often address specific problems, but these yield algorithms and tools that can be reused without much further developmental effort.

I would like to close with a note of hope: baby booms have occurred nine months after electricity blackouts. Can we expect a burst in computational biology performance and productivity after this pandemic?
Small molecule therapies are currently at the forefront of the fight against the coronavirus. Ivermectin, a microbial natural product discovered in 1975, has shown efficacy against the coronavirus. This brings us to the question: can we mine nature for more potent therapeutics against the coronavirus?

The accidental discovery of penicillin by Alexander Fleming has revolutionized the world of medicine. Surprisingly, small molecule discovery has not changed much since a century ago. The number of novel drugs approved by the FDA has been decreasing ever since the 1980s. This deficiency is partly due to the fact that the abundant bioactive molecules from cultivable microbes have already been discovered.

Mining nature for novel small molecules from uncultivable microbes requires more sensitive technologies, such as metagenomics and metabolomics. Metagenomics provides information about the microbial genes and their role in small molecule biosynthesis, while metabolomics provides information about the chemistry of their molecular products. While high-throughput sequencers and mass spectrometers have been accessible since a decade ago, searching for small molecules in the massive data produced by these technologies is analogous to searching for a needle in a haystack. Various areas within computational biology including genome assembly, genome annotation, and computational mass spectrometry, along with algorithmics and machine learning, will play a crucial role in streamlining drug discovery.

In the current global scenario of fast-emerging and fast-spreading viral outbreaks, more than ever, we need to break the traditional silos that prevent interdisciplinary cooperation. Experimentalists, data scientists, and mathematical modelers need not be compartmentalized. Data-driven approaches can be coupled with mechanistic models to generate new hypotheses that can be experimentally tested, creating loops where experiments, modeling, and data analysis synergistically augment each other. For instance, today we can create multi-scale hybrid models of the immune system that combine mechanistic descriptions of gene regulatory and cell-to-cell communication events, with AI-driven approaches to model the recognition of antigen. Furthermore, robotic experimental platforms driven by AI-algorithms can substantially speed up the testing and validation phase.

Models need to be extremely adaptive to rapidly integrate new data as it becomes available. This is especially important when designing vaccines against mutating viruses, where the vaccine developed today might not be effective against future strains. Finally, breaking silos also applies to different research areas. For instance, the frameworks developed to test and design anti-cancer compounds, e.g., https://ibm.biz/paccmann-aas and https://covid19-mol.mybluemix.net/, can be readily repurposed to design anti-viral compounds. Global initiatives such as the COVID-19 High Performance Computing Consortium are essential to combine forces and enable a fast and coordinated response to the current and future viral outbreaks. If we succeed, the impact will be unprecedented with far-reaching benefits beyond the current pandemic.