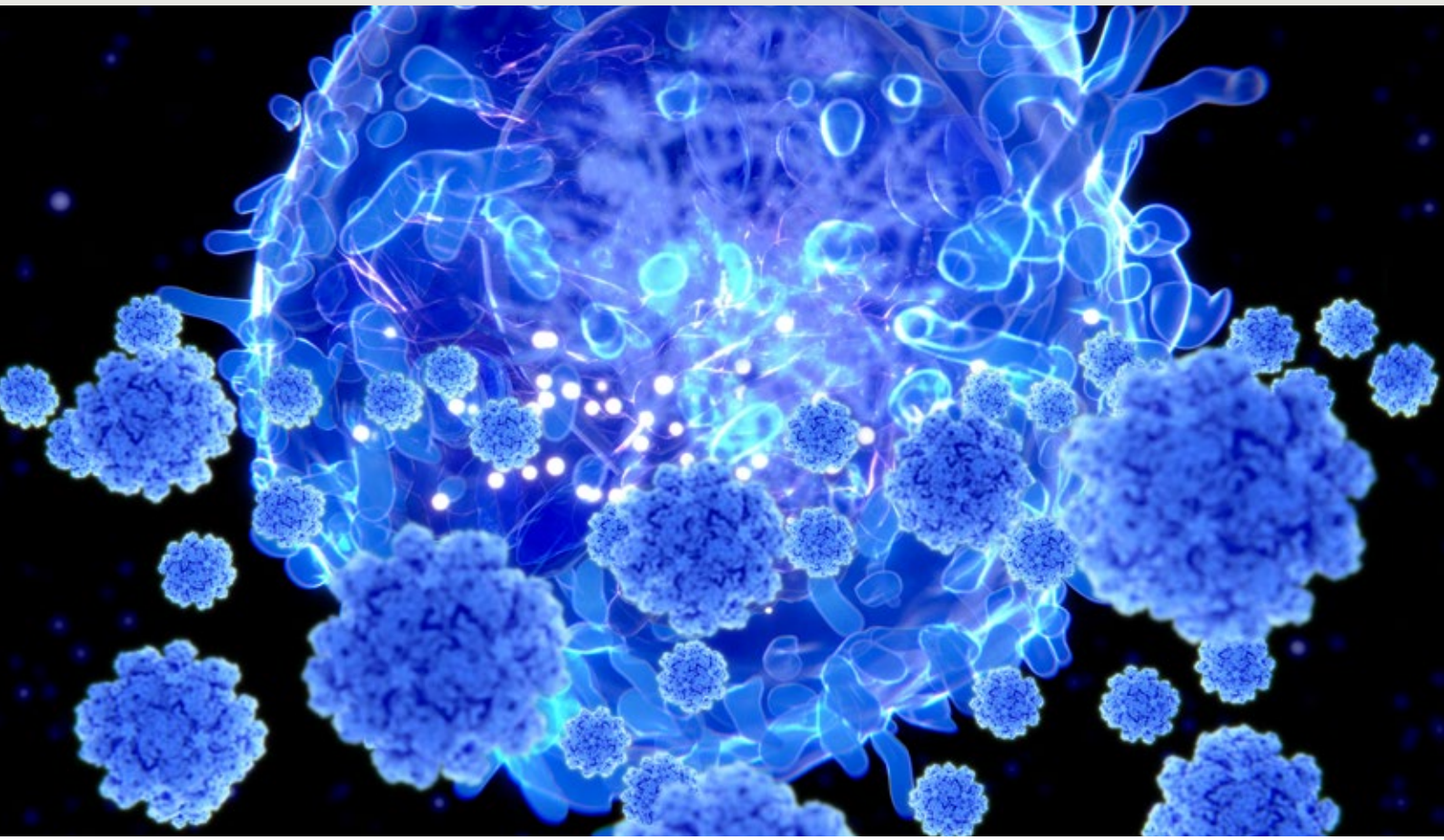
COVID-19: UNRAVELLING THE HOST IMMUNE RESPONSE

Using next-generation sequencing tools, scientists are exploring how SARS-CoV-2 interacts with the immune system to better understand the disease, identify those at higher risk, and minimize its impact.

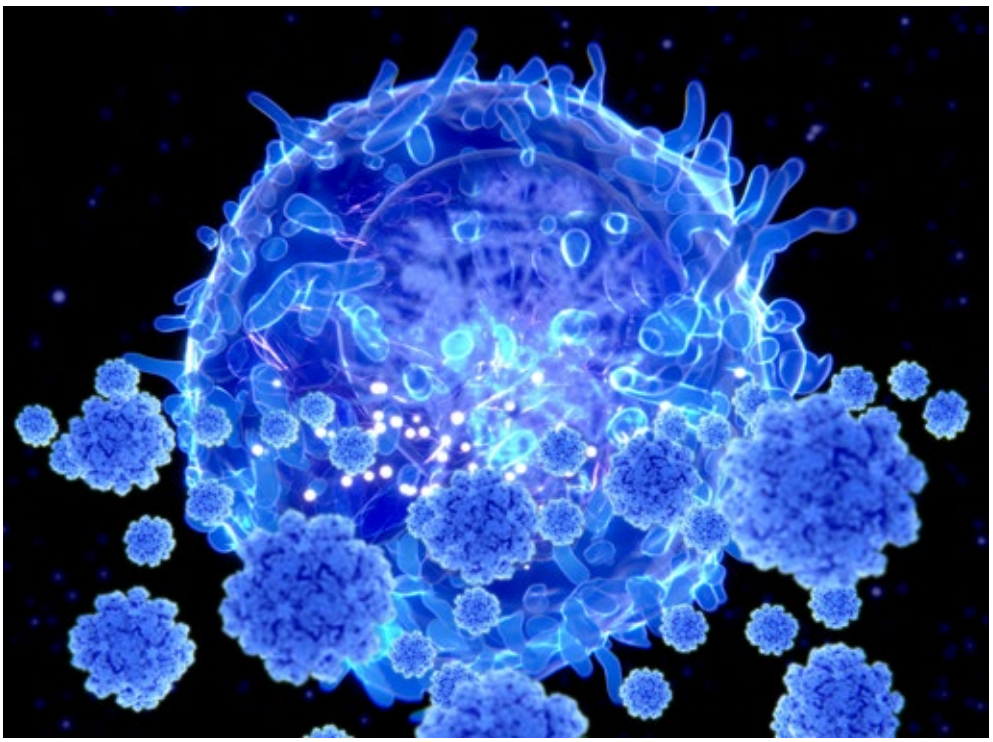


For illumina®

by nature research CUSTOM MEDIA

COVID-19: UNRAVELLING THE HOST IMMUNE RESPONSE

Using NEXT-GENERATION SEQUENCING TOOLS, scientists are exploring how SARS-CoV-2 interacts with the immune system to better understand the disease, identify those at higher risk, and minimize its impact.



SARS-CoV-2 first emerged in China in late 2019 and has since spread rapidly around the world. As the world scrambles to learn more about COVID-19, the disease it causes, the novel virus continues to infect and kill.

"This is a devastating pandemic," says Kenneth Baillie, Academic Consultant in Critical Care Medicine at the University of Edinburgh, UK. "Our only way out is either to find a treatment or a way of preventing the disease, such as a vaccine."

Researchers around the world are rising to the challenge, digging deep to fill in the knowledge gaps around this novel coronavirus. One of the biggest puzzles is why the response to infection varies so much from person to person.

"We know that human populations have a very diverse range of responses to infections," explains Baillie. "Some people will be remarkably unaffected, and some people will become very sick, which is exactly what we are seeing with SARS-CoV-2."

Scientists are using powerful genomics technologies to explore this and other major questions around the host immune response. Their discoveries will help inform the selection of therapies for clinical trials, vaccine development, and defining high-risk groups.

Huge variation in the host response

Most people who develop COVID-19 symptoms will have only mild to moderate disease. However, a significant minority will develop serious complications[1]. Around 15% of patients will progress to severe pneumonia and about 5% eventually develop acute respiratory distress syndrome, septic shock and/or multiple organ failure[2]. But anywhere between 40% and 80% of infected individuals will have no symptoms at all[3,4]. Emerging evidence suggests that these asymptomatic patients mount a weaker immune response to the virus[5].

These differences in response to infection do not appear to be related to variation in viral genetics. "The virus that is circulating right now appears to be relatively stable," says Alessandra Renieri, director of the Medical Genetics Unit at the University of Siena, Italy. "So there must be reasons on the host side that can help to explain this huge variability in disease susceptibility and outcomes."

Studies have so far identified older age and deprivation, male sex, Black or South Asian ethnicities and pre-existing medical conditions - such as obesity, severe asthma or diabetes - as risk factors for death from COVID-19[6,7]. But while social and clinical factors will account for some of the variations in response, genetic variants are also likely to play a role. Preliminary results from a twin study, which is yet to be peer reviewed, indicate that host genetic factors could explain 50% of such differences[8].

"Identifying key genetic variants involved in the development of COVID-19 will help us in several different ways," says Renieri. "It will help to improve our understanding of the biology of the disease, selecting drugs for repurposing — and knowing who is most at risk."

Searching for genetic variants

Examining the genomes of patients who have a severe response to COVID-19 will help identify crucial immune-related variants. However, asymptomatic people, the ideal control group for comparison, are extremely difficult to find. Until testing improves, researchers are instead using genetic datasets that are already available from the general population.

The COVID-19 Host Genetics Initiative (COVID-19 HG) is a global consortium of scientists studying the relationship between the host genome and SARS-CoV-2 infection[9]. The initiative currently comprises more than 100 studies, many of which mix biobank data with active connections to health systems.

"There's no doubt that we'll get answers from these retrospective studies," says Baillie. "But the trouble is we just have to wait for people in those registries to get the disease."

Due to the urgency of the situation, many teams have initiated prospective collections that directly take in patients with severe SARS-CoV-2 infection. The GenOMICC study was already recruiting people with a critical illness caused by emerging infections. It now aims to enrol up to 20,000 patients with COVID-19.

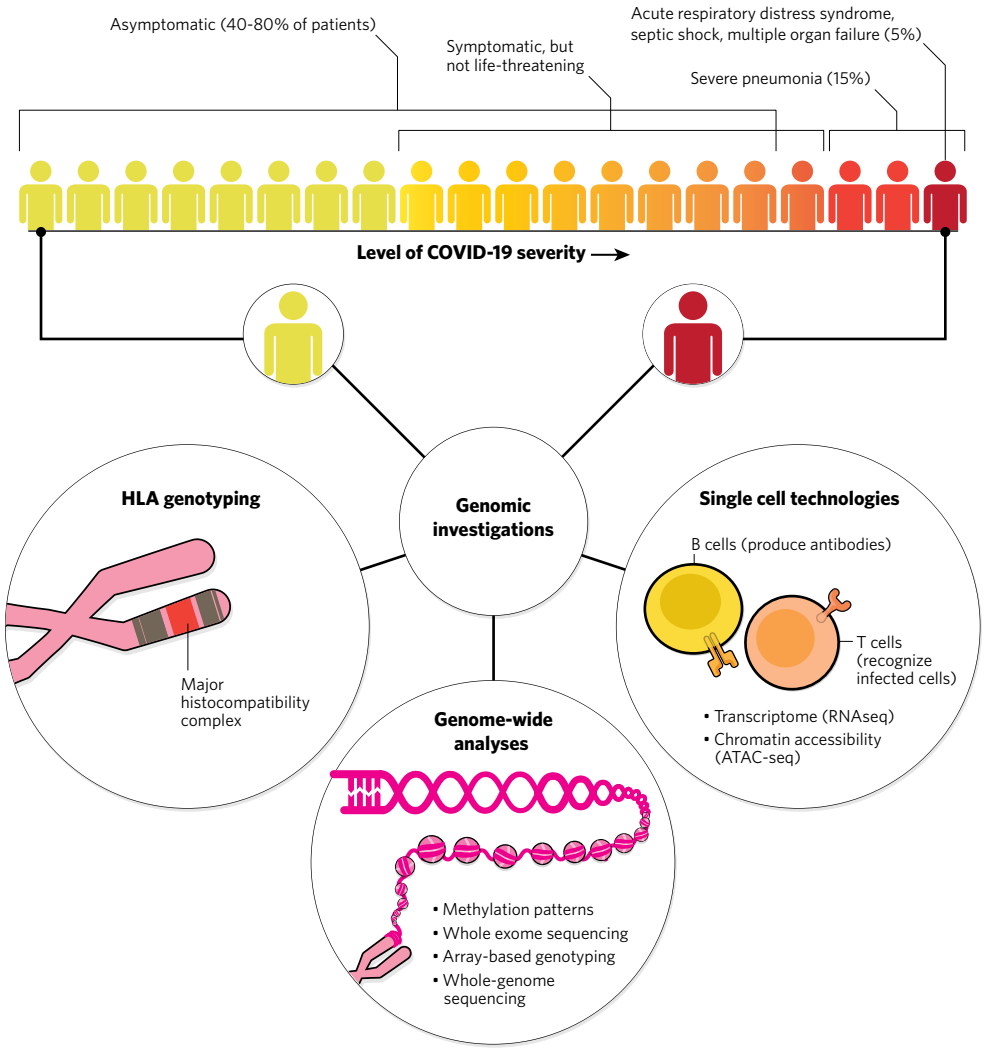
"We originally set the study up to enable clinical researchers anywhere in the world to do high-quality genetic research in intensive care," explains Baillie. "The reason for our approach is that, by definition, the people who end up on life-support machines are extremely susceptible to the disease, and that's where the strongest genetic signals will lie."

Given the challenges of working in this demanding environment, the study makes recruitment as easy as possible, requiring consent, a single blood DNA sample, and a short online form to capture key clinical information. The team has already collected around 3,600 samples and started array-based genotyping.

"This approach offers us a fast, cost-effective way to screen for common genetic variants that might be important for disease pathogenesis," says Baillie. "But the limitation is that it doesn't include very rare variants that might also be important for an individual's response."

INVESTIGATIONS INTO HOST RESPONSE TO SARS-COV-2

Clues as to why the disease varies so widely between people might lie in their DNA (e.g. variations in immune-related genes). Gene expression patterns within their immune cells may also determine how they react to the virus. Here are some of the technologies that are helping to reveal new insights.



In order to look more deeply, the team has set up a new partnership for whole-genome sequencing. They also plan to collaborate to carry out genome-wide DNA methylation analyses to look for epigenetic variations that might play a role.

A similar prospective approach is taken by a team in Italy, which aims to recruit 2,000 patients from across the clinical spectrum of COVID-19.

"This includes very severe cases who require intubation, as well as those who require milder respiratory assistance — or none at all," explains Renieri, who is leading the study. "We have also started to recruit non-hospitalized patients who have only very mild infections."

As well as whole-exome sequencing, the Italian team will perform a genome-wide association study (GWAS) in collaboration with a group in Finland.

"Our initial results have already identified some good

candidates, which we will need to validate with more data," says Renieri. "But what we are seeing is that it is likely there is a composite model with both common and rare genetic variants that contribute to disease susceptibility."

Such studies are already suggesting some interesting correlations. The first peer-reviewed COVID-19 GWAS paper found variants in the *ABO* blood-group system and in a cluster of genes on chromosome 3 that are more common among patients with respiratory failure than in the local population[10].

Immunogenomics

While many researchers are using genome-wide approaches to search for susceptibility variants, others are looking specifically at the human leukocyte antigen (HLA) system. This group of molecules is encoded by the major histocompatibility complex (MHC) and plays a vital part in shaping the immune response.

"Variation in this genomic region has already been associated with several other autoimmune conditions and infectious diseases," explains Fumihiko Matsuda, director of the Center for Genomic Medicine at the University of Kyoto, Japan. "So it's possible that certain HLA alleles could contribute to an individual's susceptibility to COVID-19."

The MHC region is one of the most highly variable in the human genome, with more than 10,000 known alleles. As its worldwide distribution patterns correspond to specific groups of people with shared ancestry, Matsuda's team is hoping to forge international collaborations to access DNA samples from diverse geographical populations.

The team has previously established a high-throughput workflow that enables the accurate genotyping of HLA alleles, which involves multiplex long-range PCR amplification followed by next-generation sequencing using the Illumina MiSeq platform[11]. Their previous experience gives them a good idea about how many samples they will need.

"Analysis of around 2,000 cases and controls is enough to identify differences in HLA allele frequencies," says Matsuda. "If we can't find anything with that number, then we should conclude that HLA is not associated with the disease."

"THIS IS THE NEXT GENERATION OF RESEARCH INTO COVID-19."

The results from these HLA association studies will complement in *silico* analyses that predict whether people with certain HLA alleles may be particularly vulnerable to COVID-19[12].

Single-cell genomics

An alternative to the large-scale population studies are single-cell technologies, to help address other important questions about the host immune response.

"We still don't know how the immune response is mounted, if it's lasting and, even more importantly, if it can be reactivated if you get the virus again," says Fabio Luciani of the School of Medical Sciences at the University of South Wales (UNSW) and the Kirby Institute in Sydney, Australia. "This is the next generation of research into COVID-19 as it underpins the work towards developing a protective vaccine."

As part of a larger study called COSiN, Luciani's team aims to isolate immune cells that recognize SARS-CoV-2 from the blood of patients who have recovered from COVID-19. They are specifically looking at T cells, that identify and eliminate infected cells, and B cells, which produce antibodies against viral particles.

"We need to know more about viral-specific T and B cells — how they respond to the virus, how diverse they are, whether every person has a different response — and whether these differences may affect vaccine efficacy," says Luciani.

Their experimental approach involves selecting viral antigens that are most likely to invoke an immune response and using these as 'bait' to lure the rare antigen-specific T cells.

"We're trying to fish a needle out of a haystack," says Luciani. "A person who has been exposed to the virus will have a lot of T and B cells in their blood, but only a very small number will be specific to SARS-CoV-2."

The researchers will then apply single-cell sequencing technologies, such as RNAseq, to analyse the transcriptome, and ATAC-seq to determine chromatin accessibility across the genome (see graphic).

"We can take a dataset from individual SARS-CoV-2-specific immune cells and use bioinformatics tools to translate that into biologically meaningful information," explains Luciani. "We hope to isolate and characterize memory T and B cells and, ultimately, determine whether these provide long-term protection against infection."

Pushing the boundaries of discovery

Scientists across the globe are redeploying their skills and resources towards research into COVID-19. The world is now looking towards them to find ideas to prevent or treat the infection.

"One step towards doing that is understanding the biology of the disease — and that's what genomics can help us to do," says Baillie. "We now have a tremendous responsibility to deliver analyses that make an impact."

With the help of today's advanced sequencing and array technologies, researchers are generating vast quantities of data and uncovering an unprecedented amount of knowledge about this novel virus over a relatively short period.

"COVID-19 has already generated as much data as we have for other viruses, such as HIV, in the last ten years," says Luciani. "The science is super exciting, but it's also challenging at the same time." ■

REFERENCES

1. Zhang, X., *et al. Nature* **583**: 437–440 (2020).
2. Cao, X. *Nat Rev Immunol* **20**;269–270 (2020).
3. Ing, A.J., Cocks, C. & Green, *J.P. Thorax* **75**:693-694 (2020)
4. Yang, R., *et al. JAMA Netw Open* **3**(6):e2014310 (2020)
5. Long, Q., *et al. Nat Med* https://doi.org/10.1038/s41591-020-0965-6 (2020)
6. Docherty A. B. *et al., BMJ* **369**:m1985 (2020).
7. Williamson, E.J. *et al. Nature* (2020) https://doi.org/10.1038/s41586-020-2521-4
8. Williams, F.M.K, *et al. MedRxiv* [preprint] doi: https://doi.org/10.1101/2020.04.22.20072124
9. The COVID-19 Host Genetics Initiative. *Eur J Hum Genet* **28**;715–718 (2020).
10. Ellinghaus, D. *et al. New Eng J Med* DOI: 10.1056/NEJMoa2020283 (2020)
11. Yin, Y., *et al. PloS One*;**11**(10):e0165810 (2016).
12. Nguyen, A., *et al. J Virol* **94**: e00510-20. https://doi .org/10.1128/JVI.00510-20 (2020)

illumina®