Review Article

Kerstin U. Ludwig*, Axel Schmidt, Per Hoffmann and Markus M. Nöthen

Combating the SARS-CoV-2 pandemic: How can the field of Human Genetics contribute?

https://doi.org/10.1515/medgen-2020-2027
Received July 16, 2020; accepted August 20, 2020

Abstract: The ongoing SARS-CoV-2 pandemic has triggered several months of global turmoil, and governments across the world are now actively attempting to develop strategies to combat both the virus and its societal impact. Since SARS-CoV-2 is a novel pathogen, basic research is essential, and manifold international efforts are now underway in order to elucidate the molecular mechanisms of both the SARS-CoV-2 infection process and the resulting disease, COVID-19. In this article we discuss how the field of Human Genetics can contribute to this endeavor, and summarize available findings from human genetic COVID-19 research. Finally, we briefly outline how through the use of existing national research infrastructure, German scientists can facilitate progress in this novel and fast-moving field.

Keywords: SARS-CoV-2, COVID-19, host genetics, coronavirus, genetic variation, genotyping, sequencing

Do genetic factors play a role? The results of a UK twin study and previous findings for other viral infections suggest that this is the case

As in other infectious diseases, wide variability is observed for SARS-CoV-2/COVID-19 in terms of infection risk and clinical course. Reported risk factors for a severe COVID-19 disease course include advanced age, male sex, underlying pre-existing health conditions, non-European ethnicity, and a high deprivation index [1, 2]. However, available data suggest that these clinical and demographic factors are not definitive predictors of COVID-19 outcome. This indicates that additional components of disease etiology, such as genetic factors, are implicated in the determination of SARS-CoV-2 infection risk (i. e., susceptibility) and COVID-19 disease course (i. e., severity).

At writing, formal evidence for the role of host genetics in SARS-CoV-2/COVID-19 remains limited. In a recent twin study [3], 2,633 participants of the TwinsUK adult twin register were asked to conduct a daily self-assessment of the presence of COVID-19 associated symptoms, as recorded using a digital tracking app. The study was performed between March 25th and April 3rd, 2020 (i. e., several weeks after the commencement of the pandemic). Subsequently, heritability estimates were generated for individual symptoms. The highest heritability estimates were obtained for delirium (49 %, 95 % confidence interval [CI] = 24 %–75 %); anosmia (47 %, CI = 27 %–67 %); shortness of breath (43 %, CI = 8 %–77 %); and fever (41 %, CI = 12 %–70 %). Notably, even higher heritability was found for the phenotype of "predicted COVID-19" (point estimate 50 %, CI = 29 %–70 %), which was determined using a symptom-based algorithm that predicts a probable SARS-CoV-2 infection [4].

The hypothesis that genetic factors play an important role in SARS-CoV-2 infection [5] is further supported by previous findings for other viral infections. These include a reported role for common variants of the human leukocyte antigen (HLA) region in multiple viral infections [6, 7]; mutations in the IFN-induced transmembrane protein-3 gene in influenza [8, 9]; and a protective effect of a loss-of function variant in *CCR5* in HIV-1 infection [10]. The latter was exploited to cure HIV-1 infected patients via stem cell transplantation [11], and was the basis for anti-HIV-1 drug development [12], thus illustrating the potential of human genetic approaches in infectious disease research.

***Corresponding author: Kerstin U. Ludwig,** Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, building 76, D-53127 Bonn, Germany, URL: www.humangenetics.uni-bonn.de, e-mail: kerstin.ludwig@uni-bonn.de
Axel Schmidt, Per Hoffmann, Markus M. Nöthen, Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, D-53127 Bonn, Germany

What molecular genetic findings have already emerged? Genome-wide association studies and comprehensive sequencing in patients have proven to be successful approaches

To date, only a limited number of studies have investigated the impact of specific host genetic factors on SARS-CoV-2 susceptibility and COVID-19 severity. The first published *genome-wide association study* (GWAS) was a collaborative effort involving hospitalized COVID-19 patients from Spain and Italy, and genetic and analysis expertise from Norway and Germany [13]. This study found no suggestive association for the HLA region. However, the authors identified genome-wide significant genetic risk variants on chromosomes 3p21.31 and 9q34.2. The associated region on 3p21.31 is relatively large (nearly 50 kb), and includes at least six genes, several of which represent strong candidates for a role in COVID-19 based on prior functional evidence (e. g., *SLC6A20* [encoding SIT1, which interacts with the cell surface receptor ACE2 that mediates SARS-CoV-2 cell entry [14]]) or code for members of the chemokine receptor family (*CCR9, CXCR6* [13]). Importantly, the association signal at 3p21.32 has since been confirmed by independent data from the international consortium on COVID-19 host genetics ([5], COVID-19 Host Genetics Initiative [HGI], www.covid19hg.org; data release v3, end of June 2020). Interestingly, recent research has shown that the risk haplotype entered the modern human population through Neanderthals. Consistent with this origin, the research also showed that the frequency of the risk haplotype varies widely across different populations (e. g., 30 % in South Asia; almost absent in the African population). This may contribute to cross-population differences in infection rates [15].

The associated variants at the 9q34.2 locus map to the region around the gene *ABO*, thus providing genetic support for prior observational evidence implicating blood type in COVID-19 [16]. Together, these studies suggest that patients with blood type A have an increased risk of a severe COVID-19 disease course, while blood type O appears to confer a protective effect. To date, scant support for an effect of the 9q34.2 locus on COVID-19 severity has been obtained from the independent cohorts in COVID-19 HGI. However, subtype analyses suggest that this locus may be associated predominantly with SARS-CoV-2 susceptibility (as opposed to COVID-19 severity). This hypothesis is further supported by data from 23andMe, which showed that individuals with blood type O were 9 %–18 % less likely to have tested positive for SARS-CoV-2 infection than individuals with other blood types [17].

For both loci, the biological underpinnings of the reported statistical associations remain unclear. Thus, a priority for the research community is to translate these findings into mechanistic insights. This work has already commenced, as summarized on the COVID-19 HGI website (www.covid19hg.org). However, these efforts will face the challenges encountered for most risk loci in common disorders over recent years, i. e., dissection of linkage disequilibrium, moderate effect sizes, identification of relevant cell types, and accessibility of appropriate cellular systems for functional follow-up studies. Community efforts, such as GTEx, the RoadMap Epigenetics Consortium, and the Human Cell Atlas, will play a central role in functional follow-up studies, and have already proven highly useful in terms of the dissection of the expression pattern of genes central to SARS-CoV-2 cellular entry, such as *ACE2* and *TMPRSS2* [18, 19]. In parallel, the identification of further common risk variation can be anticipated as cohort sizes increase within the context of COVID-19 HGI. This will also facilitate another objective of GWAS, namely, the generation of large data sets for the purpose of polygenic risk score calculation. In the future, these scores may be used to identify groups with an increased risk for SARS-CoV-2/COVID-19.

Another approach to the identification of causal genes in SARS-CoV-2 susceptibility and COVID-19 severity is the search for rare variants with stronger effect sizes using cohorts of patients with extreme phenotypes, an undertaking which would typically involve the *large-scale sequencing of individual genomes* (exome or whole genome sequencing). For the analysis of susceptibility, this could involve individuals who have not become infected, despite pronounced exposure to viral load. Given that a substantial proportion of infected individuals remain asymptomatic [20, 21], non-infection status should be confirmed via serological antibody testing. To identify risk factors for disease severity, patients who develop a severe disease course despite the absence of reported clinical or demographic risk factors for COVID-19 severity should be investigated. In rare situations, pedigrees with a monogenic, inherited susceptibility to severe COVID-19 may be observed. Such families would be characterized by multiple occurrences of severe COVID-19 and/or the tendency among family members to an extremely severe COVID-19 disease course. In the first comprehensive study of this aspect of COVID-19, two male sib-pairs were analyzed. All

four men were aged < 35 years, had no previous history of any COVID-19 relevant pre-existing condition, and required mechanical ventilation due to COVID-19 [22]. In all four men, rapid exome analysis identified rare, probable loss-of-function variants in the X-linked gene *TLR7*. Primary peripheral blood mononuclear cells of these individuals displayed markedly reduced interferon type I and type II responses to the TLR7-agonist Imiquimod. Previous research showed that TLR7 is a pattern recognition receptor for SARS-CoV-2 related viruses Middle East respiratory syndrome CoV (MERS-CoV) and severe acute respiratory syndrome CoV (SARS-CoV) [23, 24], thus supporting the hypothesis that TLR7 plays an important role in the immune response upon SARS-CoV-2 infection. Besides severe familial cases, severely affected children are another group in whom monogenic susceptibility might be enriched. Most children who test positive for SARS-CoV-2 remain asymptomatic or develop only a very mild disease course [25, 26]. However, a small proportion of pediatric cases develop severe COVID-19 and require intensive care unit management with prolonged ventilation [25].

An exome sequencing study of 35 severely affected patients from Italy identified an average of 2.5 pathogenic mutations per patient in genes implicated in viral infection and/or susceptibility [27]. These and other promising genes warrant further investigation in larger cohorts. The generated data should then be integrated with results from functional studies in COVID-19 patients, including findings generated from bulk or single cell expression analyses (e. g., [28, 29]). Together, this will facilitate the interpretation of rare variation in candidate genes, and allow both the pinpointing of targets for therapeutic interventions and the identification of individuals who are at risk for a severe COVID-19 disease course. The latter is a possible translational approach that requires competence in human genetic diagnostics.

What can the German Human Genetics community contribute?

In Germany, the SARS-CoV-2 pandemic arrived somewhat later than in southern European countries, such as Italy or Spain, and has – at least to date – impacted a lower proportion of the population. Together with the decentralized nature of the national health system, this has contributed to the difficulties experienced by German research teams in recruiting large COVID-19 patient cohorts for studies of host genetics. To accommodate the nature of the German health system and to complement ongoing efforts by clinicians to recruit COVID-19 patients for research projects, web-based recruitment campaigns have been initiated (e. g., the Bonn study of COVID genetics [BoSCO], www.bosco-studie.de).

At the international level, researchers with well-established national population-based cohorts and associated biobanks (e. g., FinnGen, UK Biobank) responded rapidly to the COVID-19 pandemic. Due to their well-established research structures, these research teams were able to launch genetic studies shortly after the initial disease outbreak. These studies made important contributions to the first COVID-19 HGI analyses. In Germany, a cohort of comparable size is only just being established (German National Cohort, www.nako.de), and no genetic data have yet been generated from the available biosamples. However, this resource will certainly play an important role in any future pandemic.

In Germany, a number of localized SARS-CoV-2 outbreaks have occurred (e. g., the outbreak in the municipality of Gangelt [21]), and these represent opportunities for key scientific investigations. In general, the individuals who are investigated during these outbreaks undergo a comprehensive phenotype assessment, and thus these cohorts will be useful in terms of the dissection of the genetic contribution to individual symptoms. This applies in particular to mild symptoms (such as persistent loss of smell and/or taste), since patients with mild disease are underrepresented in clinical samples. Furthermore, these studies typically include multiple individuals from individual households, and will thus be extremely valuable in terms of understanding the interaction between genetics, exposure, and environmental factors. For this reason, the collection of materials from patients who are suitable subjects for genetic studies is an important aspect of study design.

Finally, in their attempts to obtain further knowledge of SARS-CoV-2 and COVID-19, German experts from the field of Human Genetics can exploit their excellent clinical, scientific, and technological infrastructure. For example, extensive sequencing capacity is available through the Next Generation Sequencing Competence Network (NGS-CN, www.ngs-kn.de), which is financed by the German Research Foundation. In addition, the recently established German Human Genome Archive (GHGA, https://ghga.dkfz.de/) will serve as a central infrastructure. Here, human access-controlled research data from Germany can be archived, and this will facilitate analyses at both the national and the international level. One specific collaborative project between the NGS-CN and the GHGA is the German COVID-19 OMICS Initiative (DeCOI, https://decoi.

eu/). This national network was founded in March, 2020 in order to facilitate the use of NGS-based Omics data in COVID-19 research, and involves the input of several German Human Genetics centers.

Outlook

The use of human genetic approaches in research into the basic science of SARS-CoV-2 infection has already yielded some initial success. By enlarging available samples and generating large-scale molecular data, many more findings will be obtained in the future. Here, the expertise of human geneticists will be of central importance in terms of the generation – and even more importantly, the interpretation – of these results. Besides the identification of novel genetic risk factors, the field of Human Genetics can contribute to diverse types of follow-up analyses, such as the identification of genotype–phenotype correlations, functional follow-up of individual risk loci, and the systematic investigation of whether the mechanisms underlying SARS-CoV-2 infection and COVID-19 disease course are shared with, or distinct from, those of other rare or common phenotypes. In addition, over recent years, human geneticists have gathered additional types of expertise that can be applied within collaborative approaches, e. g., the use of NGS-based technologies for the development of mass SARS-CoV-2 infection testing [30]. Given that a large number of research departments across the country offer multidisciplinary expertise, the German Human Genetics community can be expected to make a substantial contribution to the international fight against the SARS-CoV-2 pandemic. To maximize the impact of German COVID-19 research, cross-institution collaborations should be encouraged.

Funding: Kerstin U. Ludwig, Per Hoffmann, and Markus M. Nöthen are members of the West German Genome Center (WGGC), which is one of four NGS competence centers funded by the German Research Foundation (DFG). Kerstin U. Ludwig is further supported by DFG-grant LU-1944-3/1. In addition, Kerstin U. Ludwig and Markus M. Nöthen are members of the German COVID-19 OMICS Initiative (DeCOI).

Conflict of interest: The authors have no competing interests to declare.

References

[1] Williamson EJ, Walker AJ, Bhaskaran K et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. Nature. 2020. https://doi.org/10.1038/s41586-020-2521-4.
[2] Zhou F, Yu T, Du R et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395:1054–62.
[3] Williams FMK, Freydin M, Mangino M et al. Self-reported symptoms of Covid-19 including symptoms most predictive of SARS-CoV-2 infection, are heritable. Preprint at medRxiv. 2020. https://doi.org/10.1101/2020.04.22.20072124.
[4] Menni C, Valdes AM, Freidin MB et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med. 2020. https://doi.org/10.1038/s41591-020-0916-2.
[5] COVID-HGI. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet. 2020;28:715–8.
[6] International HIV Consortium. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science. 2010;330:1551–7.
[7] Tian C, Hromatka BS, Kiefer AK et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. Nat Commun. 2017;8:599.
[8] Everitt AR, Clare S, Pertel T et al. IFITM3 restricts the morbidity and mortality associated with influenza. Nature. 2012;484:519–23.
[9] Wang Z, Zhang A, Wan Y et al. Early hypercytokinemia is associated with interferon-induced transmembrane protein-3 dysfunction and predictive of fatal H7N9 infection. Proc Natl Acad Sci USA. 2014;111:769–74.
[10] Huang Y, Paxton WA, Wolinsky SM et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med. 1996;2:1240–3.
[11] Hutter G, Nowak D, Mossner M et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med. 2009;360:692–8.
[12] Dorr P, Westby M, Dobbs S et al. Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. Antimicrob Agents Chemother. 2005;49:4721–32.
[13] Ellinghaus D, Degenhardt F, Bujanda L et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020. https://doi.org/10.1056/NEJMoa2020283.
[14] Kuba K, Imai Y, Ohto-Nakanishi T et al. Trilogy of ACE2: a peptidase in the renin-angiotensin system, a SARS receptor, and a partner for amino acid transporters. Pharmacol Ther. 2010;128:119–28.
[15] Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neandertals. Preprint at bioRxiv. 2020. https://doi.org/10.1101/2020.07.03.186296.
[16] Zietz M, Tatonetti NP. Testing the association between blood type and COVID-19 infection,

intubation, and death. Preprint at medRxiv. 2020. https://doi.org/10.1101/2020.04.08.20058073.
[17] 23andMe. 23andMe finds evidence that blood type plays a role in COVID-19. 2020. https://blog.23andme.com/23andme-research/23andme-finds-evidence-that-blood-type-plays-a-role-in-covid-19/.
[18] He Y, Arvanitis M, Parsana P, Omdahl A et al. Tissue specific expression and genetic regulation of SARS-CoV-2 receptors ACE2 and TMPRSS2. Preprint at Medium.com. 2020. https://medium.com/@yuanhe716/tissue-specific-expression-and-genetic-regulation-of-sars-cov-2-receptors-ace2-and-tmprss2-d75d76bf45fb.
[19] Lukassen S, Chua RL, Trefzer T et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. EMBO J. 2020;39:e105114.
[20] Mizumoto K, Kagaya K, Zarebski A et al. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill. 2020;25:2000180.
[21] Streeck H, Schulte B, Kuemmerer B et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. Preprint at medRxiv. 2020. https://doi.org/10.1101/2020.05.04.20090076.
[22] Van Der Made CI Simons A, Schuurs-Hoeijmakers J et al. Presence of Genetic Variants Among Young Men With Severe COVID-19. JAMA. 2020. https://doi.org/10.1001/jama.2020.13719.
[23] Channappanavar R, Fehr AR, Zheng J et al. IFN-I response timing relative to virus replication determines MERS coronavirus infection outcomes. J Clin Invest. 2019;129:3625–39.
[24] Li Y, Chen M, Cao H et al. Extraordinary GU-rich single-strand RNA identified from SARS coronavirus contributes an excessive innate immune response. Microbes Infect. 2013;15:88–95.
[25] Götzinger F, Santiago-Garcia B, Noguera-Julian A et al. COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort study. Lancet Child Adolesc Health. 2020. https://doi.org/10.1016/S2352-4642(20)30177-2.
[26] Lu X, Zhang L, Du H et al. SARS-CoV-2, Infection in Children. N Engl J Med. 2020;382:1663–5.
[27] Benetti E, Giliberti A, Emiliozzi A et al. Clinical and molecular characterization of COVID-19 hospitalized patients. Preprint at medRxiv. 2020. https://doi.org/10.1101/2020.05.22.20108845.
[28] Chua RL, Lukassen S, Trump S et al. COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020. https://doi.org/10.1038/s41587-020-0602-4.
[29] Schulte-Schrepping J, Reusch N, Paclik D et al. Severe COVID-19 Is Markedby a Dysregulated Myeloid Cell Compartment. Cell. 2020. https://doi.org/10.1016/j.cell.2020.08.001.
[30] Schmid-Burgk JL, Schmithausen RM, Li D et al. LAMP-Seq: Population-Scale COVID-19 Diagnostics Using Combinatorial Barcoding. Preprint at bioRxiv. 2020. https://doi.org/10.1101/2020.04.06.025635.

Emmy-Noether group leader Kerstin U. Ludwig
Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, building 76, D-53127 Bonn, Germany
kerstin.ludwig@uni-bonn.de

Axel Schmidt
Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, building 13, D-53127 Bonn, Germany

Per Hoffmann
Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, building 76, D-53127 Bonn, Germany

Professor and Director Markus M. Nöthen
Institute of Human Genetics, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, building 13, D-53127 Bonn, Germany