Review Article

Genetic Predisposition and the Variable Course of Infectious Diseases

Axel Schmidt, Ana M. Groh, Julia S. Frick, Maria J. G. T. Vehreschild, Kerstin U. Ludwig

Institute of Human Genetics, Medical Faculty of the University of Bonn & Bonn University Hospital: Dr. med. Axel Schmidt, Dr. rer. nat. Kerstin U. Ludwig

Medical Department II, Infectiology, University Hospital-Frankfurt, Goethe University Frankfurt: Ana M. Groh, Prof. Dr. med. Maria J. G. T. Vehreschild

Interfaculty Institute for Microbiology and Infection Medicine, University Hospital and Faculty of Medicine Tübingen: Prof. Dr. med. Julia S. Frick

MVZ Laboratory Ludwigsburg GbR: Prof. Dr. med. Julia S. Frick

Summary

Background: Contact with a pathogen is followed by variable courses of infectious disease, which are only partly explicable by classical risk factors. The susceptibility to infection is variable, as is the course of disease after infection. In this review, we discuss the extent to which this variation is due to genetic factors of the affected individual (the host).

Methods: Selective review of the literature on host genetics in infectious disease, with special attention to the pathogens SARS-CoV-2, influenza viruses, Mycobacterium tuberculosis, and human immunodeficiency virus (HIV).

Results: Genetic variants of the host contribute to the pathogenesis of infectious diseases. For example, in HIV infection, a relatively common variant leading to a loss of function of the HIV co-receptor CCR5 affects the course of the disease, as do variants in genes of the major histocompatibility complex (MHC) region. Rare monogenic variants of the interferon immune response system contribute to severe disease courses in COVID-19 and influenza (type I interferon in these two cases) and in tuberculosis (type II interferon). An estimated 1.8% of life-threatening courses of COVID-19 in men under age 60 are caused by a deficiency of toll-like receptor 7. The scientific understanding of host genetic factors has already been beneficial to the development of effective drugs. In a small number of cases, genetic information has also been used for individual therapeutic decision-making and for the identification of persons at elevated risk.

Conclusion: A comprehensive understanding of host genetics can improve the care of patients with infectious diseases. Until the present, the clinical utility of host genetics has been limited to rare cases; in the future, polygenic risk scores summarizing the relevant genetic variants in each patient will enable a wider benefit. To make this possible, multicenter studies are needed that will systematically integrate clinical and genetic data.

Cite this as:
Schmidt A, Groh AM, Frick JS, Vehreschild MJGT, Ludwig KU: Genetic predisposition and the variable course of infectious diseases. Dtsch Arztebl Int 2022; 119: 117–23. DOI: 10.3238/arztebl.m2022.0105

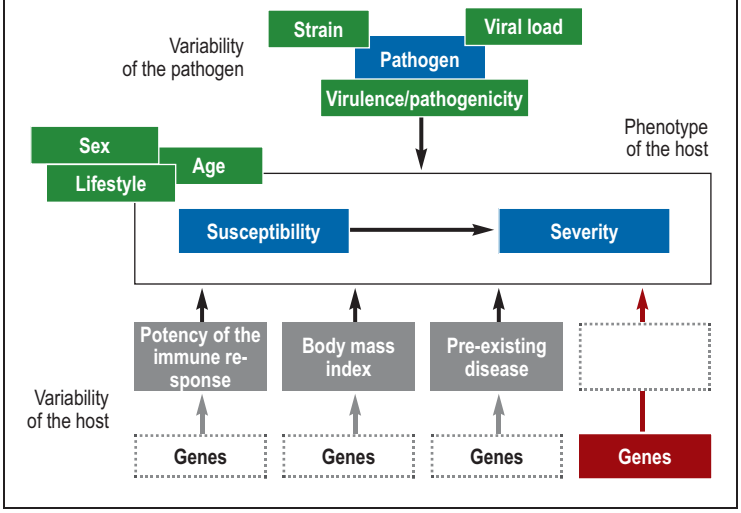
Due to the SARS-CoV-2 pandemic, infectious diseases have come into public focus. A frequently discussed aspect in this context is the heterogeneity of disease courses, given that the majority of people infected with SARS-CoV-2 develop either no or only mild symptoms, while others become extremely ill. Susceptibility upon contact with SARS-CoV-2 also varies; for example, some individuals rapidly become infected, others not at all (1).

This interindividual variability following contact with a pathogen can also be seen in a similar form in other infectious diseases. Severe disease courses can be explained, at least in part, by acquired factors, such as advanced age, unhealthy lifestyle (for example, smoking, overweight, alcohol consumption), or certain pre-existing diseases. However, there are disease courses that are more severe or milder than acquired factors would lead one to expect.

As early as the mid-20th century, genetic factors of the host were already being postulated as modulators of disease severity (2). A groundbreaking study conducted in 1988 on the contribution of genetics in infectious diseases showed that adoptees have an approximately five-fold higher risk of dying of an infectious disease if one biological parent died of an infectious disease at an early stage in life (3). These broad findings have also been supported by entity-specific studies. For example, approximately 30% of the clinical variability in COVID-19 is attributable to genetic factors of the host (4).

Using four infectious diseases as examples, we describe in this review article how host genetics can influence their course and how genetic insights can be used for treatment or prevention.

FIGURE 1



The multifactorial etiology of infectious diseases is driven not only by a variety of pathogen-specific factors but also by host-specific factors. Pathogen-specific factors, as well as some host-specific factors, are not—or only to as small extent—affected by genetic factors (marked in green). Acquired, classic risk factors often have a polygenic component (examples marked in gray), which is picked up via the clinical risk factor. However, there are also other host genetic factors whose effects are independent of these risk factors and which therefore represent an additional level in the clinical assessment.

Basic genetic principles

Infectious diseases are multifactorial entities, the etiology of which is contributed to not only by the pathogen but also by genetic as well as non-genetic factors *(Figure 1)*. In this context, the individual genetic contribution is often based on a combination of low-penetrance variants and, more rarely, on isolated high-penetrance variants (mutations). Accordingly, the quantitative contribution of individual variants, the effect size, is usually negatively correlated with the frequency of the variant in the population. Whether predominantly rare or indeed more frequent variants contribute to an infectious disease depends to a crucial extent on three factors:

- The biological complexity of the disease
- The (mostly unknown) extent of selection pressure exerted by the pathogens
- The interaction of the pathogen with the host.

A classic example of this would be mutations that cause hemoglobinopathy: since these simultaneously confer resistance to malaria, carrierships predisposing to hemoglobinopathies are relatively common in regions with high malaria prevalence compared to non-malarial areas (5).

Single common variants generally contribute in only a minor way to a person's individual risk. Overall, however, common variants can, in combination with non-genetic factors, significantly influence individual risk (6). To identify associated variants, DNA array-based genotyping of large cohorts is used *(Box)*. In this process, the allelic expression of up to one million variable positions in each genome are recorded and the allele frequencies then compared between affected and unaffected individuals. This approach, which is referred to as a genome-wide association study (GWAS), is extremely cost-effective and has already been successfully used for over 15 years to decipher the genetics of multifactorial diseases (7). The associated variants are mostly found in the non-coding regions of the genome and, as such, have no immediate effect on protein function. Although an effect on gene regulation is assumed, the causal relationships are not directly evident from the GWAS. Only comparatively few GWAS in infectious diseases have been conducted to date (with the exception of COVID-19). In the most extensive GWAS to date, which was carried out in 2017, numerous infectious diseases were investigated and 59 risk factors identified (8). The effect sizes of the individual risk variants (effect size 1.05–1.78) are, as expected, low and therefore are of only limited clinical informative value. It is only by considering individual variants in a combinatorial manner that significantly stronger effects can be predicted (compare polygenic risk scores). Although this approach offers great potential, it has not yet found its way into clinical evaluation.

In the case of individuals whose phenotype is significantly more strongly pronounced than one would expect based on non-genetic risk factors, a substantial genetic influence is assumed. In extreme cases, even monogenic forms of the disease may be present, that is to say, these patients carry a single genetic variant that, after pathogen contact, is in effect solely responsible for the course of disease. These are mostly coding variants that cause major functional changes to the protein and are rare in the population (usually < 1 %). In the case of infectious diseases, these monogenic predispositions to certain infections are attributed to congenital immunodeficiencies and can be inherited in, for example, an autosomal-dominant or autosomal-recessive manner *(Table)*. These predispositions cover a range of infectious phenotypes: on the one hand, there are high-penetrance forms that tend towards earlier, more severe infection and a uniform phenotype (for example, autosomal-recessive *IFNGR1* deficiency) and, on the other, low-penetrance forms (for example, autosomal-dominant *IFNGR2* deficiency) (9).

In clinical routine, an immune defect of this kind should be considered if patients exhibit unusual disease courses, infections with atypical pathogens, or a remarkable frequency of infections. Guidelines on this are available (10, 11). In order to elucidate the genetic cause of immunodeficiencies, next-generation sequencing (NGS), that is to say, comprehensive sequencing of either the entire protein-coding sequence (exome sequencing) or even the entire genome (whole-genome sequencing), is increasingly used, not least in diagnostics. In contrast to DNA array-based genotyping, sequencing is also able to identify rare variants, some of which have not yet been previously described.

Cell biological mechanisms in pathogen contact

The basis of defense against pathogens lies in the host-specific immune response, which comprises elements of the innate and the adaptive immune response. The former responds rapidly and non-specifically to pathogens, whereas the latter is mounted in a delayed manner but establishes a long-lasting immunological memory (12).

Mechanisms relevant to the host-specific response of innate immunity include the barrier function of the (mucosal) skin, a variety of cells in the immune system (for example, granulocytes, macrophages, and natural killer cells), as well as humoral components, to which the complement system and cytokines belong. The latter encompass a group of immunomodulatory proteins that also include the interferons *(Table)*.

To differentiate between foreign and endogenous components, the innate immune system uses receptors that recognize molecular patterns of pathogens (pattern recognition receptors), such as toll-like receptors (TLR) and receptors for proteins coded in the major histocompatibility complex (MHC) region. Once the pathogen has been recognized and classified, a proinflammatory cascade that protects the host during the initial days of infection is activated. In addition, cytokines and co-stimulatory molecules are produced in the further course for the activation of the adaptive immune system(13).

The adaptive immune response itself is triggered by the presentation of foreign antigens by MHC complexes (coded by human leukocyte antigen [HLA] genes). It has a humoral component through the formation of specific antibodies in plasma cells, as well as a cellular component through the specific activation of cytotoxic T cells and T helper cells.

Host genetics in selected infectious diseases

For the period 2009–2013 in Europe, influenza, tuberculosis, and AIDS were the infectious diseases that caused the greatest number of lost healthy life years (e1) *(Figure 2)*. COVID-19 disease caused by SARS-CoV-2 has also already been held responsible for a considerable disease burden (e2).

Influenza

In addition to viral subtype, monogenic defects that reduce the effect or production of type I/type III interferons have been identified as causal for severe influenza. The affected genes influence this defense pathway (14), for example, by impairing detection of viral RNA (*TLR3*), restricting the induction of interferon I/III expression (*IRF7*), or affecting the action of interferon I/III (*IRF9*). Inexplicably severe courses of influenza have also been observed in individuals with mutations in *GATA2* (e3) and *DBR1* (e4); in both mutations, the pathomechanism is independent of the interferon defense system *(Table)*. Estimates on how many cases of severe influenza are caused by monogenic defects are currently not available. However, it is assumed that they account for only a small portion of the variance. Despite the comparatively high incidence of infections with influenza viruses, only a handful of GWAS have been published to date, with inconclusive (equivocal) results. The strongest evidence has been identified for the non-coding variants near the *CD55* and *IFITM3* genes; however, this has not been independently confirmed as yet, nor has the causal role of these genes been demonstrated (15).

BOX

Basic principles of host genetics

Central to the investigation of host genetics is a large cohort in which subjects are known to have had prior contact with a pathogen. Data from individuals that have not had pathogen contact are of only limited informative value in the identification of genetic risk factors. Some individuals become infected on contact with a pathogen, while others do not (susceptibility phenotype). The genetic investigation of non-infected individuals can yield valuable information regarding possible resistance factors. Infected individuals develop disease with varying degrees of severity (severity phenotype), with this interindividual variability partially, but often not completely, explained by classic risk factors. One can assume that host genetic factors, as part of a multifactorial etiology, also play a role in this variability. In isolated cases, one even sees disease courses that are completely unexpected from a clinical perspective; a strong genetic—possibly even monogenic—contribution can be assumed in these individuals

HIV infections

As early on as in 1996, a homozygous loss-of-function variant in the *CCR5* gene was identified as a protective cause (16). The CCR5 protein, together with the surface antigen CD4, serves as a crucial receptor for HIV adsorption—therefore, the absence of CCR5 prevents the virus from entering the host cell. These insights formed the basis for the development of the CCR5 antagonist maraviroc as an HIV drug. Furthermore, in patients with hemato-oncologic disease receiving stem cell transplantation, it was possible to treat HIV infection by selecting donors with a homozygous CCR5 deletion (17, e5). Using GWAS, frequent low-penetrance variants associated with the progression of HIV infection were also identified (18). For example, there are other independent associations at the *CCR5* locus that control the expression of *CCR5* in CD4+ T cells. Thus, a smaller amount of CCR5 is associated with a lower potential for in vitro HIV infection (19).

Consistent associations with alleles in the MHC Class I region have also been described (20). Here, the biological mechanisms at the affected genes (*HLA-A*, *HLA-B*, *HLA-C*) differ: at the *HLA-A* and *HLA-B* loci, the mechanism is characterized by the expression of two and three amino acids, respectively, whereby various alleles here can influence the course of HIV infection both positively and negatively (20).

TABLE

Gene systems of relevance for host genetics (selection)[*1]

Gene system	Function	Clinical picture	Causal genes[*2]
Pattern recognition receptors	Recognition of pathogens based on specific molecular patterns, important examples include membrane-bound Toll-like receptors (TLRs) and cytosolic Rig-I-like receptors (RLR).	Severe viral infections (influenza viruses, SARS-CoV-2, herpes simplex viruses)	AD: *TLR3, TICAM1* XL: *TLR7*
		Bacterial infections (e.g. *Streptococcus pneumoniae, Staphylococcus aureus*)	AR: *MYD88, IRAK4, TIRAP* XL: *IRAK1*
Cytokines: interferons	As components of the innate and adaptive immune system, interferons (IFN) contribute to defense against viruses and endogenous bacteria. A distinction is made between type-I, type-II, and type-III interferons.	Severe viral infections (influenza viruses, SARS-CoV-2, herpes simplex viruses)	AR: *STAT1, STAT2, IRF9, IRF7, IFNAR1, IFNAR2* AD: *TRAF3, IRF3, TBK1*
		Mycobacterial infections (in particular type-II IFN)	AR: *IL12RB1, IL12B, IL12RB2, IL23R, IFNGR1, IFNGR2, ISG15* AD: *IFNGR1, STAT1*
Cytokines: interleukin-17 family	Interleukin-17 cytokines are proinflammatory and serve in, among other things, the defense against extracellular bacteria and fungi.	Chronic mucocutaneous candidiasis	AR: *IL17RA, IL17RC, TRAF3IP2* AD: *IL17F, STAT1*
Complement system	The complement system is a system of plasma proteins that attacks the surface of microorganisms.	Increased number of *Neisseria* infections (particularly meningitis)	AR: *C5, C6, C7, C8A, C8B, C9, CFD* XL: *CFP*
Antigen presentation	Antigen presentation via MHC molecules is part of the adaptive immune response. Both intracellular (MHC Class I) and extracellular antigens (MHC Class II) are presented and subsequently recognized by CD8+ and CD4+T cells, respectively.	Immune defect syndrome with frequent, non-specific infections	AR: *CIITA, RFX5, RFXAP, RFXANK*
Others	Linearization of certain circular RNA molecules (RNA lariat metabolism)	Brainstem infections with herpes simplex viruses, influenza viruses, or noroviruses	AR: *DBR1*

[*1] Gene systems in which mutations can cause susceptibility to infection in the absence of other clinical or laboratory abnormalities have been preferentially selected (modified from [39, 40]).
[*2]Selected genes, grouped according to pattern of inheritance; in some genes, mutations can be inherited in both a dominant and a recessive manner.
AR, autosomal-recessive; AD, autosomal-dominant; MHC, major histocompatibility complex; XL, X-linked

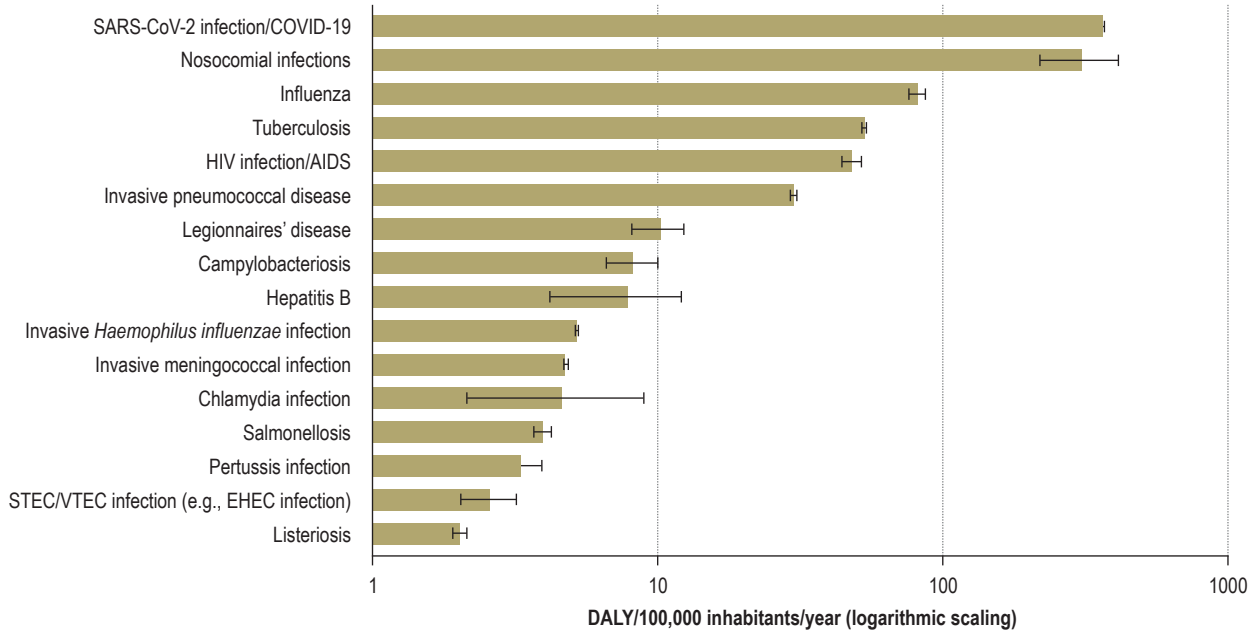
In contrast, the effect at the *HLA-C* locus is mediated by the amount of expression, that is to say, in a regulatory manner. High HLA-C expression is associated with milder disease (21). Milder disease has also been reported in individuals with different HLA alleles (the so-called heterozygote advantage) (22). Genetic findings on HIV are already being applied in treatment: for example, HIV-infected patients are only treated with abacavir if they do not carry the HLA-B*57:01 allele, since approximately 50% of HLA-B*57:01-allele carriers develop severe adverse effects (23). From a mechanistic perspective, abacavir binds to HLA-B*57:01 and alters the endogenous peptide repertoire presented on the surface. As a result, T cells react to the newly presented peptides and trigger a hypersensitivity reaction (24).

Tuberculosis

Two independent GWAS (8, e6) have been conducted to date for infections with the *Mycobacterium tuberculosis* bacterium. These studies identified associations in genes in MHC class II, notably two amino acid exchanges in *HLA-DR* and *HLA-DQ*, as well as one non-coding variant. Associations in the *HLA-D*R region have also been demonstrated in infections with the related strain, *Mycobacterium leprae* (e7). It is becoming increasingly evident that in the presence of associations in the MHC region, these primarily involve class I molecules in viral infections and class II molecules in bacterial infections. This is supported by the biological mechanisms of antigen presentation, which differ between bacteria and viruses (25).

In addition, over 20 monogenic defects that reduce the effect or production of type II interferon are known; these result in increased susceptibility to mycobacterial infections (for example, mutations in *IFNGR1*, *IL12RB2*, *IL23R*, *IRF8*, *STAT1*) *(Table)* (26). In particular, defects that cause reduced interferon production (for example, in the *IL12B* gene) can be treated adjunctively with recombinant interferons (9). The penetrance of the individual immune defects differs—it can be complete, but it can also be significantly reduced. A prominent example would be autosomal recessive *IFNGR1* deficiency, which is associated with complete penetrance and an overall poor prognosis. Although hematopoietic stem cell transplantation can be performed as a curative treatment, complications are common (9).

FIGURE 2



Most important infectious diseases in Europe

The figure lists infectious diseases expressed as disability-adjusted life years (DALY, one DALY = a loss equivalent to one healthy year of life).
Error bars depict a 95% confidence interval (no data available for SARS-CoV-2 infection), represented logarithmically. The figure is based on the following publications: SARS-CoV-2 (e2), nosocomial infections (e10), and other infections that caused the highest number of healthy years of life lost in the period 2009–2013 (e1). Nosocomial infections include, among others, healthcare-associated *Clostridioides difficile* infections, pneumonia, urinary tract infections, and wound infections (e11).
AIDS, acquired Immunodeficiency syndrome; EHEC, enterohemorrhagic *Escherichia coli*; HIV, human immunodeficiency virus;
STEC, Shiga-toxigenic *Escherichia coli*; VTEC, verotoxigenic *Escherichia coli*

COVID-19

Since mid-2020, GWAS have been used to investigate common risk variants for COVID-19. One of the largest GWAS was published in July 2021 by an international consortium (COVID-19 Host Genetics Initiative): by investigating just under 50,000 affected individuals, it was possible to identify 13 loci, some of which are associated with susceptibility to, and others with severity of, COVID-19 (27) *(eTable)*. Some variants could only be identified by including non-European populations, since those particular variants are too rare in the European population to reach statistical significance. What is also interesting is the absence (to date) of genetic findings in the MHC region. Not surprisingly, the associations in COVID-19 also lie primarily in the non-coding region of the genome and are still not understood in terms of their function. However, the regions cover a multitude of genes whose involvement in the pathogenesis of COVID-19 seems plausible on the basis of biological insights (for example, *OAS1* and *IFNAR2*). The strongest association to date has been reported for variants on chromosome 3 (1, e8, 28). A correlation between the risk allele and extensive clinical findings showed a particularly strong association with respiratory failure in under 60-year-olds (effect size 2.7), which is in line with the effect size of established clinical risk factors (29).

An increased susceptibility to severe COVID-19 due to monogenic defects that reduce the effect or production of type I interferon has also been described: exome sequencing identified hemizygous mutations in the *TLR7* gene as causal for extremely severe disease in two pairs of brothers (aged 20–35 years) (30). A number of studies have confirmed this finding (31, 32, e9). It is estimated that approximately 1–2% of life-threatening COVID-19 infections in males aged under 60 years can be attributed to *TLR7* deficiency (point estimate, 1.8%) (31). Awareness of *TLR7* mutation status has already been used for prevention by fast-tracking male mutation carriers for vaccination (32). Involvement of rarer variants has also been demonstrated in other genes in the interferon-1 defense system, including a homozygous *TBK1* mutation in a child with lethal COVID-19 (33). Synthetically produced type I interferon has already been used to treat isolated cases of individuals with defects in interferon-1 production (34).

Translation of genetic findings in infectiology

Although maraviroc is not currently used as the first-line drug in HIV treatment, this example illustrates the clinical benefit of genetic findings. At present, less than 10% of approved drugs are based on genetic insights. However, active substances selected for clinical trials on the basis of genetic evidence are approximately two- to four-fold more likely to be approved as drugs compared to active substances that were selected without considering genetic data. Thus, the success rate for drug development is significantly higher if the selection of drug targets is supported by genetic evidence (35).

In addition to the abovementioned diseases, there are a number of other translational examples: for example, homozygosity for certain variants on the *IL28B* locus is the strongest predictor for the efficacy of combination therapy with interferon and ribavirin in hepatitis C infection (36), and successful treatment with ruxolitinib has been reported in patients with chronic mucocutaneous candidiasis caused by gain-of-function mutations in *STAT1* (37). To date, only a handful of genetic variants have been included in pharmacogenetic assessments of this kind. However, diagnostics in the future will incorporate a greater number of different variants together with non-genetic factors, thereby enabling more precise prediction.

Summary

Infectious diseases will continue to pose a challenge for health care systems in the future, due, among other reasons, to the appearance of new pathogens, increasing resistance to treatment strategies for known pathogens, and the worldwide disparity in access to vaccines and drugs (38). In addition to the significance of the infection itself, etiological links between an infection and other, mostly multifactorial diseases are increasingly coming to light (for example, Epstein-Barr virus infection was recently identified as the likely main cause of multiple sclerosis [e10]). Greater understanding of host genetics, both at the individual and at the population level, will substantially contribute to personalized medicine and, as such, is of central clinical importance. In order to achieve this goal, inter-site studies that combine clinical, infectious disease, and genetic information are needed.

Funding

Dr. Schmidt received a Gerok grant from the BONFOR program of the Medical Faculty of the University of Bonn (account number O-149.0134). Dr. Ludwig is supported by the Emmy Noether program of the German Research Foundation (DFG) (LU 1944/3–1). Both authors are members of the German COVID-19 Multi-Omics Initiative (DeCOI-Host Genetics). Dr. Ludwig also receives funding from the German Federal Ministry of Education and Research (BMBF) for the COVIMMUNE project.

Acknowledgments

We would like to thank Professor Markus Nöthen for discussing the manuscript with us, Dr. Anna Vyvers for her critical reading of the manuscript, as well as Mrs. Christine Fischer for her support in designing the figures.

Conflict of interest statement

Dr. Schmidt holds shares in Twist Bioscience.

Prof. Frick received study support (third-party funding) from the DFG.

Dr. Ludwig holds shares in LAMPseq Diagnostics GmbH. She has received lecture fees from the Dr. Hans Riegel-Stiftung

The remaining authors declare that no conflict of interests exists.

Manuscript received on 31 August 2021, revised version accepted on 11 January 2022.

Translated from the original German by Christine Rye.

References

1. Shelton JF, Shastri AJ, Ye C, et al.: Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. Nat Genet 2021; 53: 801–8.
2. Barua D, Paguio AS: ABO blood groups and cholera. Ann Hum Biol 1977; 4: 489–92.
3. Sørensen TI, Nielsen GG, Andersen PK, Teasdale TW: Genetic and environmental influences on premature death in adult adoptees. N Engl J Med 1988; 318: 727–32.
4. Williams FMK, Freidin MB, Mangino M, et al.: Self-reported symptoms of COVID-19, including symptoms most predictive of SARS-CoV-2 infection, are heritable. Twin Res Hum Genet 2020; 23: 316–21.
5. Hedrick PW: Population genetics of malaria resistance in humans. Heredity (Edinb) 2011; 107: 283–304.
6. Khera AV, Chaffin M, Aragam KG, et al.: Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018; 50: 1219–24.
7. Loos RJF: 15 years of genome-wide association studies and no signs of slowing down. Nat Commun 2020; 11: 5900.
8. 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.
9. Bustamante J, Boisson-Dupuis S, Abel L, Casanova JL: Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. Semin Immunol 2014; 26: 454–70.
10. Farmand S, Baumann U, von Bernuth H, et al.: [Interdisciplinary AWMF guideline for the diagnostics of primary immunodeficiency]. Klin Padiatr 2011; 223: 378–85.
11. Hanitsch L, Baumann U, Boztug K, et al.: Treatment and management of primary antibody deficiency: German interdisciplinary evidence-based consensus guideline. Eur J Immunol 2020; 50: 1432–46.
12. Netea MG, Schlitzer A, Placek K, Joosten LAB, Schultze JL: Innate and adaptive immune memory: an evolutionary continuum in the host's response to pathogens. Cell Host Microbe 2019; 25: 13–26.
13. Takeda K, Kaisho T, Akira S: Toll-like receptors. Annu Rev Immunol 2003; 21: 335–76.
14. Lim HK, Huang SXL, Chen J, et al.: Severe influenza pneumonitis in children with inherited TLR3 deficiency. J Exp Med 2019; 216: 2038–56.
15. Garcia-Etxebarria K, Bracho MA, Galán JC, et al.: No major host genetic risk factor contributed to A(H1N1)2009 influenza severity. PloS One 2015; 10: e0135983.
16. Liu R, Paxton WA, Choe S, et al.: Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 1996; 86: 367–77.
17. Hütter 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.
18. McLaren PJ, Pulit SL, Gurdasani D, et al.: Evaluating the impact of functional genetic variation on HIV-1 control. J Infect Dis 2017; 216: 1063–9.
19. Kulkarni S, Lied A, Kulkarni V, et al.: CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. Nat Immunol 2019; 20: 824–34.
20. McLaren PJ, Coulonges C, Bartha I, et al.: Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. Proc Natl Acad Sci U S A 2015; 112: 14658–63.
21. Apps R, Qi Y, Carlson JM, et al.: Influence of HLA-C expression level on HIV control. Science 2013; 340: 87–91.
22. Carrington M, Nelson GW, Martin MP, et al.: HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science 1999; 283: 1748–52.
23. Mallal S, Nolan D, Witt C, et al.: Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. Lancet 2002; 359: 727–32.
24. Illing PT, Vivian JP, Dudek NL, et al.: Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature 2012; 486: 554–8.
25. Blum JS, Wearsch PA, Cresswell P: Pathways of antigen processing. Annu Rev Immunol 2013; 31: 443–73.

26. Rosain J, Kong XF, Martinez-Barricarte R, et al.: Mendelian susceptibility to mycobacterial disease: 2014–2018 update. Immunol Cell Biol 2019; 97: 360–7.
27. COVID-19 Host Genetics Initiative: Mapping the human genetic architecture of COVID-19. Nature 2021; 600: 472–7.
28. Severe Covid GG, Ellinghaus D, Degenhardt F, et al.: Genomewide association study of severe Covid-19 with respiratory failure. N Engl J Med 2020; 383: 1522–34.
29. Nakanishi T, Pigazzini S, Degenhardt F, et al.: Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. J Clin Invest 2021; 131.
30. van der Made CI, Simons A, Schuurs-Hoeijmakers J, et al.: Presence of genetic variants among young men with severe COVID-19. JAMA 2020; 324: 663–73.
31. Asano T, Boisson B, Onodi F, et al.: X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. Sci Immunol 2021; 6.
32. Solanich X, Vargas-Parra G, van der Made CI, et al.: Genetic screening for TLR7 variants in young and previously healthy men with severe COVID-19. Front Immunol 2021; 12: 719115.
33. Schmidt A, Peters S, Knaus A, et al.: TBK1 and TNFRSF13B mutations and an autoinflammatory disease in a child with lethal COVID-19. NPJ Genom Med 2021; 6: 55.
34. Lévy R, Bastard P, Lanternier F, Lecuit M, Zhang SY, Casanova JL: IFN-α2a therapy in two patients with inborn errors of TLR3 and IRF3 infected with SARS-CoV-2. J Clin Immunol 2021; 41: 26–7.
35. Nelson MR, Tipney H, Painter JL, et al.: The support of human genetic evidence for approved drug indications. Nat Genet 2015; 47: 856–60.
36. Tanaka Y, Nishida N, Sugiyama M, et al.: Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009; 41: 1105–9.
37. Bloomfield M, Kanderova V, Parackova Z, et al.: Utility of Ruxolitinib in a child with chronic mucocutaneous Ccandidiasis caused by a novel STAT1 gain-of-function mutation. J Clin Immunol 2018; 38: 589–601.
38. Becker K, Hu Y, Biller-Andorno N: Infectious diseases—a global challenge. Int J Med Microbiol 2006; 296: 179–85.
39. Casanova JL, Abel L: Lethal infectious diseases as inborn errors of immunity: toward a synthesis of the germ and genetic theories. Annu Rev Pathol 2021; 16: 23–50.
40. Kwok AJ, Mentzer A, Knight JC: Host genetics and infectious disease: new tools, insights and translational opportunities. Nat Rev Genet 2021; 22: 137–53.

Corresponding author
Dr. rer. nat. Kerstin U. Ludwig
Institut für Humangenetik, Department of Genomics
Universitätsklinikum Bonn
Venusberg-Campus 1, Gebäude 76
53127 Bonn, Germany
kerstin.ludwig@uni-bonn.de

Cite this as
Schmidt A, Groh AM, Frick JS, Vehreschild MJGT, Ludwig KU:
Genetic predisposition and the variable course of infectious diseases.
Dtsch Arztebl Int 2022; 119: 117–23. DOI: 10.3238/arztebl.m2022.0105

► **Supplementary material**

eReferences eTable:
www.aerzteblatt-international.de/m2022.0105

Questions on the article in issue 8/2022:

Genetic Predisposition and the Variable Course of Infectious Diseases

cme plus

The submission deadline is 24.2.2023. Only one answer is possible per question.
Please select the answer that is most appropriate.

Question 1

What does the abbreviation GWAS stand for?

a) Genome-wide analysis using sequencing
b) Genome-wide association study
c) Genome-wide analysis of sensitivity
d) Genome-wide analysis of selection factors
e) Genome-wide analysis of susceptibility

Question 2

Which of the following receptors in the innate immune system is involved in the recognition of molecular patterns of pathogens?

a) Canabinoid receptor type 2
b) Glutamate receptor
c) P2Y12 receptor
d) Capsaicin receptor
e) Toll-like receptor 3

Question 3

Which mutation has been identified as a protective factor in HIV-resistant individuals?

a) Homozygous loss-of-function variant in the CCR5 receptor gene
b) Heterozygous gain-of-function variant in the CCR4 receptor gene
c) Homozygous gain-of-function variant in the CD4 receptor gene
d) Homozygous gain-of-function variant in the CCR4 receptor gene
e) Homozygous loss-of-function variant in the IL-17 receptor gene

Question 4

What occurs more frequently in certain defects of the complement system?

a) Herpes simplex infections
b) Influenza infections
c) Candidiasis
d) Scabies
e) *Neisseria* infections

Question 5

According to the results of genetic association studies, which MHC class has been primarily associated with defense against bacterial infections?

a) Class I; b) Class II
c) Class III; d) Class IV
e) Class V

Question 6

Next-generation sequencing has identified genes with mutations that can predispose to particularly severe courses of COVID-19. An estimated 1–2% of life-threatening courses of COVID-19 in under 60-year-old males can be attributed to which genetic cause?

a) TLR5 gain-of-function mutation
b) TLR11 deficiency
c) TLR7 deficiency
d) TLR6 deficiency
e) TLR13 gain-of-function mutation

Question 7

Susceptibility to which infections is increased especially in the case of defects in the effect or production of type-II interferon?

a) Norovirus infections
b) *Pseudomonas* infections
c) Streptococcal infections
d) Mycobacterial infections
e) HPV infections

Question 8

What percentage of clinical variability in COVID-19 can be attributed to host genetic factors?

a) Approximately 0.5%
b) Approximately 5%
c) Approximately 10%
d) Approximately 30%
e) Approximately 50%

Question 9

Exome sequencing is restricted to identifying which type of genetic variants?

a) Variants in protein-coding sequences
b) Variants in gene regulatory sequences
c) Heterozygous variants
d) Homozygous variants
e) Variants encoded on the sex chromosomes

Question 10

What does the metric DALY describe?

a) Months of life spent in hospital due to a disease
b) Average disease duration (in days)
c) The number of healthy life years lost due to a disease
d) The reduction in life expectancy (in years) due to a disease
e) The curtailment of disease (in days) achieved through treatment

Supplementary material to:

Genetic Predisposition and the Variable Course of Infectious Diseases

by Axel Schmidt, Ana M. Groh, Julia S. Frick, Maria J. G. T. Vehreschild, Kerstin U. Ludwig

Dtsch Arztebl Int 2022; 119: 117–23. DOI: 10.3238/arztebl.m2022.0105

eReferences

e1. Cassini A, Colzani E, Pini A, et al.: Impact of infectious diseases on population health using incidence-based disability-adjusted life years (DALYs): results from the burden of communicable diseases in Europe study, European Union and European economic area countries, 2009 to 2013. Euro Surveill 2018; 23.

e2. Rommel A, Lippe EV, Plass D, et al.: The COVID-19 disease burden in Germany in 2020-years of life lost to death and disease over the course of the pandemic. Dtsch Arztebl Int 2021; 118.

e3. Sologuren I, Martínez-Saavedra MT, Solé-Violan J, et al.: Lethal influenza in two related adults with inherited GATA2 deficiency. J Clin Immunol 2018; 38: 513–26.

e4. Zhang SY, Clark NE, Freije CA, et al.: Inborn errors of RNA lariat metabolism in humans with brainstem viral infection. Cell 2018; 172: 952–65 e18.

e5. Gupta RK, Abdul-Jawad S, McCoy LE, et al.: HIV-1 remission following CCR5Delta32/Delta32 haematopoietic stem-cell transplantation. Nature 2019; 568: 244–8.

e6. Sveinbjornsson G, Gudbjartsson DF, Halldorsson BV, et al.: HLA class II sequence variants influence tuberculosis risk in populations of European ancestry. Nat Genet 2016; 48: 318–22.

e7. Wang Z, Sun Y, Fu X, et al.: A large-scale genome-wide association and meta-analysis identified four novel susceptibility loci for leprosy. Nat Commun 2016; 7: 13760.

e8. Pairo-Castineira E, Clohisey S, Klaric L, et al.: Genetic mechanisms of critical illness in COVID-19. Nature 2021; 591: 92–8.

e9. Fallerini C, Daga S, Mantovani S, et al.: Association of toll-like receptor 7 variants with life-threatening COVID-19 disease in males: findings from a nested case-control study. Elife 2021; 10.

e10. Bjornevik K, Cortese M, Healy BC, et al.: Longitudinal analysis reveals high prevalence of Epstein-Barr virus. Science 2022; 375: 296–301.

e11. Zacher B, Haller S, Willrich N, et al.: Application of a new methodology and R package reveals a high burden of healthcare-associated infections (HAI) in Germany compared to the average in the European Union/European Economic Area, 2011 to 2012. Euro Surveill 2019; 24.

eTABLE

A selection of common risk variants for SARS-CoV-2/COVID-19

Gene locus (chromosomal bands)	Mean global frequency of the effect allele	Candidate genes at the gene locus	Effect size
a) Susceptibility			
3p21.31	0.12	*SLC6A20, CCR3*	1.15
3q12.3	0.35	*ZBTB1, RPL24, CEP97, NXPE3*	0.94
9q34.2	0.65	***ABO***	0.90
19q13.33	0.18	*HSD17B14, PLEKHA4, PPP1R15A, TULP2, NUCB1*	0.95
b) Severity (hospitalization, severe disease)			
3p21.31	0.08	*LZTFL1, CXCR6*	1.88
6p21.1	0.04	***FOXP4***	1.26
8q24.13	0.01	***TMEM65***	1.37
12q24.13	0.65	*OAS1, OAS2, OAS3*	1.20
17q21.31	0.19	*ARHGAP27, PLEKHM1, LINC02210-CRHR1, CRHR1, SPPL2C, MAPT, STH, KANSL1, LRRC37A, ARL17B, LRRC37A2, ARL17A, NSF, WNT3*	0.88
17q21.33	0.03	*KAT7, TAC4, DLX3, FLJ45513*	1.45
19p13.3	0.32	***DPP9***	1.26
19p13.2	0.05	*ICAM1, ICAM3, ICAM4, ICAM5, ZGLP1, FDX2, RAVER1, TYK2*	1.43
22q22.11	0.66	***IFNAR2***	0.82

Through genome-wide association studies (GWAS), it was possible to identify variants at several gene loci whose allelic expression is associated with an increased risk (effect size > 1) or a protective effect (effect size < 1) on susceptibility to infection with SARS-CoV-2 (susceptibility) or the clinical course of COVID-19 (severity). The genes localized at the gene loci are referred to as candidate genes since the pathomechanism mediated by the variants is not yet known. For the genes marked in **bold**, there is independent biological evidence for their involvement in the pathogenesis of infectious diseases. Data based on (34).