The Role of Host Genetic Factors in Coronavirus Susceptibility: Review of Animal and Systematic Review of Human Literature

Marissa LoPresti\textsuperscript{1}, David B. Beck\textsuperscript{2}, Priya Duggal\textsuperscript{3}, Derek A. T. Cummings\textsuperscript{4,5}, Benjamin D. Solomon\textsuperscript{6}

\textsuperscript{1} University of Florida College of Veterinary Medicine, Gainesville, Florida, United States of America

\textsuperscript{2} Inflammatory Disease Section, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America

\textsuperscript{3} Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America

\textsuperscript{4} Department of Biology, University of Florida, Gainesville, Florida, United States of America

\textsuperscript{5} Emerging Pathogens Institute, University of Florida, Gainesville, Florida, United States of America

\textsuperscript{6} Office of the Clinical Director, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States America

Corresponding author:

Benjamin D. Solomon

National Human Genome Research Institute

Building 10 - CRC, Suite 3-2551

10 Center Drive

Bethesda, MD 20892

solomonb@mail.nih.gov
Abstract

Background: The recent SARS-CoV-2 pandemic raises many scientific and clinical questions. One set of questions involves host genetic factors that may affect disease susceptibility and pathogenesis. New work is emerging related to SARS-CoV-2; previous work on other coronaviruses in humans or other host species may be relevant.

Objectives: To review existing literature on host genetic factors and their association with infection and disease with coronaviruses in humans and in other host species.

Methods: We conducted a systematic review of literature on host genetic factors in humans associated with coronavirus outcomes. We also reviewed studies of host genetic factors associated with coronavirus outcomes in non-human species. We categorized articles, summarized themes related to animal studies, and extracted data from human studies for analyses.

Results: We identified 1,187 articles of potential relevance. Forty-five studies examined human host genetic factors related to coronavirus, of which 35 involved analysis of specific genes or loci; aside from one meta-analysis on respiratory infections, all were candidate-driven studies, typically investigating small numbers of research subjects and loci. Multiple significant loci were identified, including 16 related to susceptibility to coronavirus (of which 7 identified protective alleles), and 16 related to outcomes or clinical variables (of which 3 identified protective alleles). The types of cases and controls used varied considerably; four studies used traditional replication/validation cohorts. Of the other studies, 28 involved both human and non-human host genetic factors related to coronavirus, and 174 involved study of non-human (animal) host genetic factors related to coronavirus.

Key findings: We have outlined key genes and loci from animal and human host genetic studies that may bear investigation in the nascent host genetic factor studies of COVID-19. Previous human studies
have been limited by relatively low numbers of eligible participants and limited availability of advanced genomic methods. These limitations may be less important to studies of SARS-CoV-2.

**Key words:** Coronavirus; COVID-19; Host genetic factors; SARS-CoV-2
Introduction

The ongoing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic raises many scientific and clinical questions. One unknown is the extent to which individuals vary in susceptibility to infection and disease (COVID-19). Various hypotheses have been suggested to explain observed differences, including sex, age, comorbidities, and genetic factors. As with many complex diseases, the explanations likely involve a combination of genetic and non-genetic factors. In this context, genetic factors involve an interplay between virus and host genetics.

Large, international studies and collaborations have formed to investigate host genetic factors related to COVID-19, including disease severity and susceptibility. These investigations include analyses of existing public and private datasets, as well as the establishment of new cohorts (e.g., https://blog.23andme.com/23andme-research/genetics-and-covid-19-severity/).

While SARS-CoV-2 has seized recent attention, there are many other coronaviruses and a significant related body of literature exists about host genetic factors and their association with infection and outcomes in both humans and non-human host species. The Coronavirinae subfamily of the Coronaviridae family consists of four genera. The alphacoronaviruses include two major human coronaviruses, HCoV-229E (of which multiple HCoV-229E-like strains have been identified) and HCoV-NL63. Alphacoronaviruses that affect other species include mouse hepatitis virus (MHV), feline coronavirus (FCoV), which includes feline infectious peritonitis virus (FIPV) and feline enteric coronavirus (FECV), canine coronavirus (CCoV), and transmissible gastroenteritis coronavirus (TGEV) and porcine transmissible gastroenteritis coronavirus (TGEV) in pigs. The betacoronaviruses consist of four lineages: lineage A (HCoV-OC43 and HCoV-HKU1), lineage B (SARS-CoV-1 and SARS-CoV-2), lineage C (Middle East Respiratory Syndrome (MERS) and many bat coronaviruses), and lineage D (coronaviruses only identified in bats to date). HCoV-OC43, HCoV-229E, HCoV-HKU1, and HCoV-NL63 can result in a variety of
presentations, including “common cold” and severe but rarely fatal disease; they are also frequently detected as co-infections with other viruses. There are other rare coronaviruses observed in humans as well as in other species. Relative to other coronaviruses, SARS-CoV-2 has unique biological properties and related clinical impact, but data regarding other coronaviruses may be relevant.

In various species, much work has focused on the genes encoding the relevant coronavirus receptor, including effects of viral and host genetic changes. Among other cell surface determinants, these receptor genes include ACE2 for HCoV-NL63, SARS-CoV-1, and SARS-CoV-2, ANPEP for HCoV-229, FIPV, CCoV, and TGEV, DPP4 for MERS, and Ceacam1 for MHV. Host genetic studies have - to varying degrees and in different ways - analyzed these genes, as well as other genes identified through targeted and agnostic methods. Studies to date have been disparate in terms of the virus and species studied, as well as the aims of the particular study. This has resulted in a rich body of literature that is difficult to efficiently leverage for SARS-CoV-2-related work.

We aimed to perform a review of the literature to outline previous studies of host genetic factors related to coronaviruses, with the objective of performing a systematic review to encapsulate genes and loci interrogated through these efforts. We do not attempt to fully describe the findings nor recapitulate what is known about the underlying host biology related to coronavirus infection and disease. As the majority of studies are candidate-driven, we did not attempt to conduct a meta-analysis. However, one goal is that the data presented here can help populate lists of genes that - along with data from related work - may bear scrutiny in the developing and important large-scale host genetic studies related to SARS-CoV-2. We present an overview of themes and interrogated genes/loci from animal studies, and perform a systematic review on human studies.

Methods
We conducted an initial search of the PubMed database (last queried May 4, 2020) using each of the following phrases: “host genetics”; “genetic resistance”; “genetic susceptibility”; “genetic factors”; “genetics”; “GWAS” along with each of the following terms: “coronavirus”; “SARS”; “MERS”; “COVID-19”; “COVID19”. We also identified additional articles by searching for specific coronaviruses or coronavirus-associated conditions (e.g., “canine coronavirus”; “middle east respiratory syndrome”) along with the term “genetics”. Articles were included in the search regardless of publication date.

Articles included electronic, ahead-of-print publications available in the PubMed database. We also identified and categorized relevant articles from the references of initially selected articles. We did not include articles only available on non-peer reviewed preprint servers, though recognize that a substantial number of these manuscripts will be on PubMed soon.

Each abstract was reviewed by a single reviewer. Full articles were reviewed when insufficient data were available in the abstract, or when no abstract was available. Publications were classified into the following categories: 1) Study of human host genetic factors related to coronavirus; 2) Study of non-human (animal) host genetic factors related to coronavirus; 3) Study of non-genetic (including non-DNA-based analyses - see further explanation below) host factors related to coronavirus, including involving immunopathogenesis; 4) Study of other pathogens (not coronavirus); 5) Other studies of coronavirus. Articles containing information in both categories 1 and 2 were identified as such; articles were otherwise categorized according to the lowest numerical category (e.g., an article involving both human host genetic factors to coronavirus as well as immunopathogenesis would be categorized into group 1. Articles that did not involve investigations of specific DNA-based genetic changes (e.g., transcriptomic or proteomic studies) were categorized into group 3, as were studies that only included analyses of sex without other genetic analyses. Other publications, including: 6) Untranslated studies in another language (not English); 7) Not relevant (unrelated to coronavirus or other pathogens); 8) No data available; were removed from further analysis after categorization into these latter four categories.
Data from category 1 publications were manually extracted for relevant information pertaining to:
coronavirus studied; general methods and questions analyzed; gene(s), variant(s), or loci analyzed; size
of cohorts studied; geographic or ancestral composition of cohorts; statistical results, including (where
available) odds ratios, confidence intervals, and p-values.

Results

Our search identified 1,187 articles of potential relevance (Figure 1, Supplementary Table 1). Of these,
45 involved study of human host genetic factors related to coronavirus (Table 1); 35 of the 45 human
studies involved analysis of specific genes or loci (only one was a non-candidate study), while 10
involved biological, computational, or case report studies of human host genetic factors. Twenty-eight
involved both human and non-human host genetic factors related to coronavirus (these largely
investigated inter-species differences in disease susceptibility and pathogenesis, such as related to
differences in ACE2); 174 involved study of non-human (animal) host genetic factors related to
coronavirus; 584 involved study of non-genetic host factors related to coronavirus, including involving
immunopathogenesis; 16 involved study of other pathogens (not coronavirus); 321 involved other
studies of coronavirus. 18 studies were assigned to the other categories and removed.

We organized our analysis and findings into the schema presented below.

Animal studies

Coronaviruses affect many species, from Beluga whales to spotted hyenas to turkeys, and sequelae of
disease can range from apparently asymptomatic infections to severe or lethal effects on different organ
systems, potentially manifesting as diarrheal, encephalitic, nephritic, respiratory, and other types of
disease.26,27 There are numerous non-observational animal studies of coronaviruses, such as involving
hamsters,28-30 guinea pigs,31 rats,10,32-35 and non-human primates.36-38 However, formal host genetic
studies have been described for some but not all species. Many studies have involved examination of differences in species susceptibility and pathogenesis to human and non-human coronaviruses.\textsuperscript{14,39-41}

Among the host genetic studies in animals, the objectives and methods used differ significantly depending on the species studied. For example, in chickens and other livestock, the types of published studies predictably differ from those conducted on experimental mice. That is, while MHV represents a problem for mouse colonies, the rationale of the livestock studies may focus more purely on economic repercussions versus attempts to use a model organism to understand immunopathogenesis of infectious disease.\textsuperscript{42} The degree to which results may be reported through the scientific literature (versus other routes) is also anticipated to differ depending on the species studied and the reason for the study. See Figure 2 for a summary of interrogated loci in animal studies.

We describe representative studies and key findings below, but the descriptions should not be considered as truly comprehensive; additionally, as noted above, many studies compared susceptibility across species, both through cell-based assays and experimental animals. Many investigations using other methods (e.g., transcriptomics or proteomics) have identified key molecules involved in coronavirus susceptibility and pathogenesis. Though beyond the scope of this article, these molecules should also be considered in future SARS-CoV-2 host genetic studies.

\textit{Model animal strains, experimental animals, and domesticated animals}

\textit{Chicken}

In chickens, the infectious bronchitis virus (IBV) coronavirus can cause disease affecting different organ systems and tissues, such as IBV-associated nephritis. As with other species, inbred status and specific chicken lines impact host susceptibility, immune response, and outcomes, and virus/host genetic interactions have been described.\textsuperscript{43-47} Breeding experiments have suggested different inheritance patterns related to susceptibility and outcomes, and have implicated both MHC and non-MHC loci.\textsuperscript{48,49}
Multiple GWAS investigating the immune response to IBV have identified significantly-associated polymorphisms in the breeds studied; the implicated or nearest genes include: AKT1, AvBD12, CEP170B, CRYL1, CWF19L2, DHR5X, FAM19A2, GABRB3, INT59, NMNAT3, PINX1, RAB39A, VRK1, YEATS2; and SETBP1.

Domestic cat

Felinas can be infected by feline coronavirus (FCoV), which include feline infectious peritonitis (FIPV) and feline enteric coronavirus (FECV). As with other species, cats demonstrate a range of potential effects. In addition to association with traits such as age, sex, and reproductive status, purebred status and loss of heterozygosity has been shown to be associated with the effects of disease. Susceptibility and outcomes also appear to vary between different breeds. A small study of feline leukocyte antigen (FLA)-DRB alleles did not show a statistically significant association between the number of FLA-DRB alleles and FCoV infection outcome. Polymorphisms in IFNG (investigated as FIP can result in decreased interferon-gamma levels) were shown to correlate with plasma interferon-gamma levels and outcomes. Polymorphisms in TNFA and CD209 were also shown to be associated with outcomes in one inbred line.

In addition to candidate studies, several GWAS have been performed in cats. One small study on outcomes in experimentally-induced infections in random-bred cats identified one associated genomic region (which did not harbor any obvious candidate genes). Another small study on an inbred breed identified multiple candidate genes (ELMO1, ERAP1, ERAP2, RRAGA, TNSF10) but none was fully concordant with the FIP disease phenotype. Recent studies on SARS-CoV-1 and SARS-CoV-2 have investigated the susceptibility of cats as well as other animals; see further details below (under Ferrets).

Dromedary camel
Camels are an important reservoir of coronaviruses that can infect humans; this became especially relevant in the context of MERS.\textsuperscript{66-68} Many studies have analyzed factors that contribute to spread,\textsuperscript{69} though the searches employed in this analysis identified relatively few host genetic studies separate from analyses of DPP4 receptor characteristics and tropism, including comparisons between camels, humans, and other species.\textsuperscript{70-74}

\textit{Ferret}

Several studies have investigated the susceptibility of various species to coronaviruses. One objective relates to identifying useful animal models of disease, in which non-human species show similar infection and disease outcomes to humans upon exposure to coronaviruses.\textsuperscript{65,75,76} For example, bat, camel, and humans can be infected by MERS, unlike mouse, ferret, hamster, and guinea pig. SARS-CoV-2 replicates better in ferrets and cats than in dogs, pigs, chickens, and ducks. One explanation involves genetic characteristics of the host receptor for the relevant virus.\textsuperscript{76,77} Additionally, within an infected animal, the site of viral replication appears to vary according to the species and coronavirus, and is additionally potentially related to tissue-specific receptor expression.\textsuperscript{78} This line of reasoning may also be relevant to age-specific differences observed with SARS-CoV-2 and human infections.\textsuperscript{79}

\textit{Hamster}

As noted above, hamsters have been used as model organisms to study coronaviruses, including studies of host receptors. This includes studies using standard hamster cell lines as well as other approaches involving hamster models.\textsuperscript{80-85} For example, hamster models have been used to study species susceptibility to MHV (related to \textit{Ceacam 1}),\textsuperscript{86} how alterations of specific Dpp4 amino acids in hamster affect susceptibility to MERS,\textsuperscript{71,87} and the roles of ACE2 and CD209L in SARS-CoV-1 susceptibility.\textsuperscript{82}

\textit{Mouse}
MHV has represented a challenge for the health of mouse colonies, though relatively recent improvements in animal care practices have been beneficial. Differences in the susceptibility of different mouse strains to MHV has been noted for seven decades. Studies have examined a number of different MHV strains. These strains demonstrate different tissue tropism and have different effects on various mouse lines. One distinct example is the JHM strain of MHV, which causes encephalitis in susceptible animals. In the discussion below, though susceptibility and outcome findings will be summarized, it is important to note that studies generally focus on the interactions between certain MHV strains and mouse lines, and it is not always clear how well these findings extrapolate to other strains and lines.

Many studies have investigated biological explanations for differences in MHV susceptibility and pathogenesis. Studies examining different laboratory mouse strains have suggested that multiple loci are involved. Early studies suggested various models, including potential monogenic/Mendelian explanations as well as more complex explanations involving interacting loci.

Among many studies aiming to understand the underlying pathophysiology, mouse studies originally focused on strains believed to be involved in host susceptibility and reaction to infection. Importantly, these studies have identified interactions of host genetic factors with other factors, such as the cellular environment, cell and tissue-specific effects related to viral as well as host genetics, and host age. Unsurprisingly, some aspects of the disease process appear to be independent of observed strain differences. These studies also showed that host genetic factors influence different parts of the disease process, from initial virus-receptor binding to cellular viral spreading and multiple aspects of the immune response. These studies enabled the cloning of Ceacam1, the MHV receptor gene, as well as related work regarding how genetic changes affecting this receptor confers MHV resistance in SJL mouse lines via inhibition of viral integration into host cells.
In addition to the above studies, MHV-based mouse studies have used transgenic models to directly test the role of implicated pathways (summarized in Table 2). Not surprisingly, the majority of work in mouse models have focused on pathways already implicated in viral infection susceptibility including adaptive immune responses including both humoral and cellular, specific cytokine and immune receptor pathways, viral receptors, complement pathway, apoptosis, autophagy, and tissue repair. These studies have prominently implicated Type I (\(\alpha_2\beta\)) and II (\(\gamma\)) interferon responses in host response and predominantly protection against MHV infection. However, not all pro-inflammatory pathways are protective. For example, complement activation promotes tissue damage caused by MHV infection, highlighting the complex interplay between the host and virus. In addition to targeted gene disruptions described above, a GWAS using a recombinant inbred mouse panel implicated \(\text{Trim55}\), which is involved in vascular cuffing and inflammation in response to SARS-CoV-1.\(^{137}\)

Additional transgenic studies have investigated multiple biologic effects as well as returning to questions regarding susceptibility of different strains.\(^{138}\) Other mouse models (including knockouts, specific knock-in mutations, humanized mice, and other models involving genetic manipulation) have been used to study human pathogens such as SARS-CoV and MERS; revealing similar properties for viral receptors, \(Dpp4\) for MERS, \(Ace2\) for SARS-CoV, cytokine and immune receptor pathways, and complement pathway as with mouse models of MHV. Intriguingly, there are differences between the importance of interferon pathways in host response to SARS-CoV1, where these pathways are dispensable as compared to MHV, where they are protective. Together, these different pathogen models have shown overlapping and unique pathways of host response between coronaviruses and highlight the potential relevance for SARS-CoV-2. See also the Additional papers on humans and other species section regarding further examples of studies involving mice and humans, as well as other species.

\(Pigs\)
Pigs can be infected by transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV), as well as the more recently-identified porcine deltacoronavirus (PDCoV). Like coronavirus disease in chickens, these diseases can have economic effects on the food industry, and analyses aim to address ways to ameliorate disease, such as the development of vaccines. Importantly, variants (both natural and experimentally-induced) may have different effects on different coronaviruses. For example, aminopeptidase N, encoded by ANPEP (also called APN) was reported as a functional receptor for TGEV and PEDV (as well as HCoV-229E), but multiple models, including CRISPR/Cas9-generated knock-outs, show differences in cellular susceptibility to TGEV and PEDV. In another study, infection by PEDV and TGEV correlated positively with ANPEP expression, but PEDV and TGEV could infect ANPEP-positive and negative enterocytes, with differences observed between viral strains. Overall, the results suggest the presence of an additional receptor. Building on this type of work, site-specific editing of ANPEP has been suggested as a potential means to breed resistant animals. Studies focusing on PEDV have shown that knock-out of CMAH (hypothesized to affect cellular binding) does not result in immunity, but may improve outcomes.

Rats

Rats can be affected by rat coronaviruses, and can be hosts to a number of different coronaviruses that affect other species. Rats have been used as model systems to investigate MHV, including through cellular-based assays. Several studies have examined rat susceptibility to various coronaviruses. As with many other studies, these have implicated key interactions between viral and host genetics that affect species and tissue tropism. In addition to computational approaches examining receptor characteristics, such as involving ACE2 in the context of SARS-CoV-1, experimental studies suggest that rats are not susceptible to MERS based on Dpp4 characteristics.

Non-domesticated animals
As described, many species can be infected by coronaviruses. These species include wild as well as domesticated animals. The below section provides select examples of genetic studies on wild animals. Others studies been conducted on coronaviruses (as well as other pathogens), especially related to host ranges or reservoirs and involving host/pathogen co-evolution. Related to host genetic studies that are particularly relevant to the current SARS-CoV-2 pandemic (e.g., pangolin), our searches did not identify relevant articles.

*Cheetah*

Among wild animals, severe population bottlenecks (resulting in reduced genetic diversity) in cheetahs has been used to explain their increased susceptibility to infection by FIPV as well as other infectious diseases. Several such bottlenecks appear to have occurred in cheetah, due to a combination of factors. Among possible explanations for this susceptibility, genetic uniformity of the major histocompatibility complex (MHC) has been suspected to be involved.

*Civet*

Studies have focused on palm civets (as well as other species) related to zoonotic implications as this species has been implicated as the reservoir associated with introduction of SARS-CoV-1 into humans. Specifically, questions about host receptor characteristics (ACE2) have been described in the context of SARS-CoV-1. As with other coronaviruses and species, the interactions of viral and host genetics have been shown to be important.

*Bat*

As a natural reservoir for many coronaviruses, bats have been studied more extensively than other species outside of laboratory-based animals and livestock. Studies have included co-evolutionary studies between coronaviruses and the genomes of bat hosts (e.g., by correlating phylogenetic analyses of bat coronaviruses with CYTB in multiple bat species) as well as genetic/biologic studies related to
host genetic factors. These have involved well-studied genes such as ACE2 with SARS-CoV-1\textsuperscript{162,163} and DPP4 with MERS.\textsuperscript{164,165} In addition to allowing analyses of host susceptibility, these and similar studies help provide estimates for the time-frame of coronavirus circulation in species and populations, and explore cross-species transmission.\textsuperscript{150}

**Human**

Details of the human studies are presented in Table 1, Figure 3, and Supplementary Table 2. Forty-five studies were initially identified by the methods described. Of these, 35 involved association or other studies related to human host genetic factors (see summary in the next paragraph). Ten others involved biological, computational, or other non-genetic association studies. Many other studies were identified that used a combination of human and animal models, but were categorized separately; additionally, many studies that might be considered genetic studies - if the definition were applied less stringently - were grouped in category 3. For example, studies have examined how specific genes are involved in aspects of viral disease but did not strictly study how DNA-based host genetic variants affect this process. In summary, these ten included mapping of a susceptibility locus to HCoV-229E to chromosome 15.\textsuperscript{166} Multiple studies examined the biological effects of mutant genes. Studying the effects of mutant ACE2 on SARS-CoV-1 entry provided evidence that the cytoplasmic tail of ACE2 is not required for SARS-CoV-1 penetration.\textsuperscript{167} Studies of mutant TRIM56 on antiviral activity against HCoV-OC43 and other viruses showed that anti-HCoV-OC43 activity relies solely upon TRIM56 E3 ligase activity; this appears different from the mechanisms related to other viral pathogens.\textsuperscript{168} Knockout culture cells and nonsynonymous variant PPIA models result in limitation of HCoV-229E replication.\textsuperscript{169} (Please note that we did not separately or exhaustively investigate human genetic experiments involving cell culture systems.) Specific variants in IFITM genes (IFITM1 and IFITM3 were studied) modulate the entry of multiple human coronaviruses (HCoV-229E; HCoV-NL63; HCoV OC43; MERS-CoV; and SARS-CoV-1 were studied).\textsuperscript{170} Computational models suggest that, while most ACE2 variants have similar binding
affinity for SARS-CoV-2 spike protein, certain variants (rs73635825 and rs143936283) demonstrate different intermolecular interactions with the spike protein.\textsuperscript{171} An in silico analysis of viral peptide-MHC class I binding affinity related to HLA genotypes for SARS-CoV-2 peptides, as well as potential cross-protective immunity related to four common human coronaviruses, provides evidence that HLA-B*46:01 may be associated COVID-19 vulnerability, while HLA-B*15:03 may enable cross-protective T-cell based immunity.\textsuperscript{172} A recent study on viral cell entry showed that SARS-CoV-2 uses ACE2 for cell entry and TMPRSS2 for S protein priming; potential interventions based on these results include TMPRSS2 inhibition and convalescent sera.\textsuperscript{173} In addition to these examples, there are undoubtedly other biological, computational, and other studies examining how changes in and affecting key proteins may modulate disease.

Of the 35 human studies meeting the host genetic study criteria described above, 32 (91\%) involved SARS-CoV-1, while 3 (9\%) involved SARS-CoV-2. Two of the three SARS-CoV-2 studies were case reports (one on a single family, the other on two patients with a rare immunodeficiency) without specific studies related to host factors; it is anticipated that many more studies on SARS-CoV-2 will be published soon. All of the association studies except one were candidate-gene analyses based on genes hypothesized to be important in disease susceptibility or clinical variables/outcome. The exception was a meta-analysis of 386 studies on susceptibility to tuberculosis, influenza, respiratory syncytial virus, SARS-CoV-1, and pneumonia.\textsuperscript{174}

Candidate studies ranged from studies of single variants to studies of over 50 genes selected due to biological plausibility; seven of these studies focused on HLA alleles. Sixteen significant loci related to susceptibility to coronavirus were reported (of which 7 identified protective alleles). Sixteen significant loci related to outcomes or clinical variables were reported (of which 3 identified protective alleles). The types of cases and controls used varied. Only four studies used separate cohorts for replication/validation. However, the studies used many different types of cases and controls, including
within the same study. For example, some studies compared healthcare workers with SARS-CoV-1 infection with healthcare workers who tested negative. Others compared data from individuals with documented infection with data from control samples taken from blood donors. Four studies conducted laboratory-based biological studies in addition to association analyses. These studies are summarized in Table 1 and Figure 3; more details are available in Supplementary Table 2.

Additional papers on humans and other species

As described, human and animal studies have examined various host factors related to coronavirus infection. For example, human\textsuperscript{175} and animal\textsuperscript{33,52,125} studies have implicated age as having significant associations with outcomes; age appears to be strongly correlated with COVID-19 outcomes.\textsuperscript{176} The overall explanations remain unclear, but could at least partially involve age-related gene expression. Sex also appears to have a role. Human studies of SARS-CoV-1 and SARS-CoV2 suggest a correlation between sex and certain clinical parameters, perhaps rooted in sex-based or related immunologic differences.\textsuperscript{175,177,178} However, separating biological differences from sex-related cultural practices (e.g., different rates of social distancing) may be difficult.

Animal studies also suggest sex effects in multiple species, such as related to disease severity.\textsuperscript{57,179}

Multiple studies examined different genes/proteins to determine disease susceptibility, transmissibility, and pathogenesis in various species. In addition to humanized genes, such as used in mouse models, studies have involved a combination of computational and biological approaches, and have investigated the viral entry receptors \textit{ACE2} in SARS-CoV-1\textsuperscript{10,35,157,158,180-183} and SARS-CoV-2\textsuperscript{184} (for which there already exists a large body of unpublished and preprint work) and \textit{DPP4} in MERS.\textsuperscript{71,148,185-188} Among other findings, these studies examined specific protein residues that are critical in viral-host interactions \textsuperscript{[18448527].157} Other studies examined manipulations of various genes/proteins to study the functional biological effects, including of \textit{ANPEP},\textsuperscript{189} \textit{GLTSCR2},\textsuperscript{190} \textit{IFITM1}, \textit{IFITM2}, and \textit{IFITM3},\textsuperscript{191} and \textit{MAVS}.\textsuperscript{192}
Discussion: human studies

Traditional genome-wide methods have been applied to human viral infections generally, but results have not been specific to coronaviruses, and it is unclear to what extent the observations are relevant to the current pandemic. Several dozen studies have investigated human genetic factors related to coronavirus infection. However, these studies have been limited by several potential factors. For endemic human coronaviruses, the mildness of disease may have deprioritized these studies; similar observations may explain the relative dearth of serologic knowledge related to these pathogens. For coronaviruses associated with more severe human disease, such as MERS and SARS-CoV-1, the fact that these epidemics were limited more than the current pandemic crisis may have fortunately led to a lack of cases with which one might conduct traditional association studies (unlike some other respiratory infections leading to more widespread disease). Additionally, these two severe conditions primarily affected human populations prior to the technological developments that led to wide availability of much cheaper and faster genomic sequencing.

As shown (Supplementary Table 2), the small sample sizes of previous studies may have led to the preponderance of candidate gene studies. The sample sizes may also have precluded significant findings due to limitations of statistical power and the ability to replicate or validate findings. As previous human studies occurred in areas of the world affected by the coronavirus studied, it is possible that results from these studies would not extrapolate to other populations. Finally, different genes and loci are involved than those previously hypothesized. That is, hypothesis-free approaches may identify significant loci that were not identified by candidate approaches.

Based on announcements about multiple large-scale projects on host genetic factors and SARS-CoV-2, as well as the existence of larger genomic datasets that can be mined quickly and new methods that can be
used to address biological questions, it is anticipated that considerable efforts - and an unfortunately large pool of research subjects - and may yield significant new results quickly.

Limitations

There are multiple limitations to our summaries and analyses. First, it is likely that relevant articles were missed by our search process, and that key findings - including the study of certain genes - were therefore omitted. Along these lines, important findings within identified articles may also have been missed. Second, this analysis focused on DNA-based variants. These DNA-based genetic changes include those studied and identified through association studies as well as genes that were manipulated in experimental approaches, such as via knockout models to understand disease pathogenesis. Related ‘omic approaches, such as targeted or broad transcriptomic or proteomic studies, are frequently used to understand important aspects of disease. These approaches can lead to knowledge regarding specific genetic changes. For example, observed transcriptomic changes may enable the identification of important DNA-based variants that explain disease by correlating transcriptomic data with results of DNA sequencing. However, we categorized non-DNA based ‘omic approaches separately from DNA-based studies, and did not attempt to comprehensively recapitulate what is known about host reaction to disease. Finally, as the studies varied in many aspects, such as how cases and controls were defined, and which loci were interrogated, we were careful about comparing or combining data between different studies.

Table 1. Summary of human studies (related to specific genes or loci) on host genetic factors related to coronaviruses. More details are available in Supplementary Table 2.
<table>
<thead>
<tr>
<th>Coronavirus studied (other coronaviruses or pathogens)</th>
<th>SARS-CoV-1</th>
<th>Analysis of association of HLA gene polymorphisms with susceptibility to SARS-CoV-1 infection or clinical parameters</th>
<th>Association of HLA-B*4601 with severity of SARS-CoV-1 infection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SARS-CoV-1</td>
<td>Analysis of association of HLA gene polymorphisms with susceptibility to SARS-CoV-1 infection</td>
<td>HLA-B<em>0703, HLA-DRB1</em>0301 and co-inheritance of HLA-B*0703 and HLA-B60 were associated with susceptibility to SARS-CoV-1 infection</td>
</tr>
<tr>
<td></td>
<td>SARS-CoV-1</td>
<td>Analysis of association of ACE2 polymorphisms with SARS-CoV-1 clinical parameters</td>
<td>No association of ACE2 polymorphisms with SARS-CoV-1 outcomes</td>
</tr>
<tr>
<td></td>
<td>SARS-CoV-1</td>
<td>Analysis of association of ACE insertion/deletion (I/D) polymorphism with susceptibility to SARS-CoV-1 infection</td>
<td>ACE D allele was associated with hypoxemia in SARS-CoV-1 infections</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>OAS1, PKR, MX1</strong> polymorphisms with susceptibility to SARS-CoV-1 or clinical parameters</td>
<td><strong>OAS1</strong> rs3741981 and rs2660 were associated with SARS-CoV-1 susceptibility; <strong>MX1</strong> rs2071430 was associated in hypoxemia in SARS-CoV-1 infections</td>
<td>15766558&lt;sup&gt;203&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>ACE</strong> insertion/deletion (I/D) polymorphism with susceptibility to SARS-CoV-1 or clinical parameters</td>
<td>No association was found with <strong>ACE</strong> insertion/deletion (I/D) polymorphism and susceptibility to SARS-CoV-1 or clinical parameters</td>
<td>15819995&lt;sup&gt;175&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>MBL</strong> polymorphisms susceptibility to SARS-CoV-1 or clinical parameters and biological study of MBL</td>
<td>Serum MBL was lower in patients with SARS-CoV-1 infections than controls, and haplotypes associated with lower serum MBL were more frequent in patients with SARS-CoV-1 infections than in control subjects, but there was not association with mortality</td>
<td>15838797&lt;sup&gt;204&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>ACE2</strong> polymorphisms and</td>
<td>No association was found with <strong>ACE2</strong> polymorphisms and</td>
<td>15937940&lt;sup&gt;205&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>MBL</strong> polymorphisms and susceptibility to SARS-CoV-1 infection</td>
<td>MBL rs1800450 was associated with susceptibility to SARS-CoV-1 infection</td>
<td>16170752&lt;sup&gt;206&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>FCGR2A</strong> and <strong>MBL</strong> polymorphisms and susceptibility to SARS-CoV-1 infection or clinical parameters</td>
<td>Homozygosity for <strong>FCGR2A</strong> rs1801274, as well as a linear trend of <strong>FCGR2A</strong> genotypes, was associated with severe SARS-CoV-1 infection</td>
<td>16185324&lt;sup&gt;207&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>CLEC4M</strong> VNTR polymorphism with susceptibility to SARS-CoV-1 and biological studies of cells with these polymorphisms</td>
<td>Homozygosity for the <strong>CLEC4M</strong> VNTR polymorphism was associated with susceptibility to SARS-CoV-1, and homozygous cells had higher binding capacity for SARS-CoV-1, higher proteasome-dependent viral degradation, and lower capacity for trans infection.</td>
<td>16369534&lt;sup&gt;208&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1 Analysis of association of <strong>HLA</strong> polymorphisms with SARS-CoV-1 susceptibility</td>
<td>HLA-Cw*0801 was associated with susceptibility to SARS-CoV-1 infection</td>
<td>16455884(^209)</td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-1 Analysis of association of polymorphisms in 65 genes with SARS-CoV-1 viral shedding</td>
<td>SARS-CoV-1 shedding was associated with alleles of <strong>IL18</strong>, <strong>IL1A</strong>, <strong>RELB</strong>, and <strong>FLG2</strong></td>
<td>16652313(^210)</td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-1 Analysis of association of <strong>OAS1</strong> and <strong>MX1</strong> polymorphisms with susceptibility to SARS-CoV-1</td>
<td><strong>OAS1</strong> 3'-UTR rs2660 and <strong>MX1</strong> promoter rs2071430 were associated with susceptibility to SARS-CoV-1</td>
<td>16824203(^211)</td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-1 Analysis of association of <strong>CLEC4M</strong> VNTR polymorphism with susceptibility to SARS-CoV-1 infection</td>
<td>No association was found with homozygosity for the <strong>CLEC4M</strong> VNTR polymorphism and susceptibility to SARS-CoV-1</td>
<td>17534354(^212)</td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-1 Analysis of association of <strong>CLEC4M</strong> VNTR polymorphism with susceptibility to SARS-CoV-1 infection</td>
<td>No association was found with homozygosity for the <strong>CLEC4M</strong> VNTR polymorphism and susceptibility to SARS-CoV-1</td>
<td>17534355(^213)</td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>CCL5</strong>, <strong>CXCL9</strong>, and <strong>CXCL10</strong> polymorphisms with susceptibility to SARS-CoV-1 infection or clinical parameters</td>
<td><strong>CCL5</strong> rs2107538 was associated with susceptibility to SARS-CoV-1 in one cohort and severe outcomes of SARS-CoV-1 infection in another cohort</td>
<td>17540042&lt;sup&gt;214&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>FCER2</strong> and <strong>ICAM3</strong> polymorphisms with susceptibility to SARS-CoV-1 or clinical parameters</td>
<td>Homozygosity for <strong>ICAM rs2304237</strong> was associated with higher LDH levels and lower total WBC counts</td>
<td>17570115&lt;sup&gt;215&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>CD14</strong>, <strong>TLR2</strong>, and <strong>TLR4</strong> polymorphisms with susceptibility to SARS-CoV-1 or clinical parameters</td>
<td><strong>CD14</strong> rs2569190 was associated with severe SARS-CoV-1 infection (this data was also combined with previous data, suggesting that this and <strong>FCGR2A</strong>-RR131 are risk genotypes for severe SARS-CoV-1 infection)</td>
<td>17913858&lt;sup&gt;216&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <strong>TNF</strong> polymorphisms with susceptibility to SARS-CoV-1 infection or clinical parameters</td>
<td><strong>TNF</strong> polymorphisms were associated with susceptibility to SARS-CoV-1</td>
<td>18312678&lt;sup&gt;217&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of polymorphisms in IL12RB1 with susceptibility to SARS-CoV-1 or clinical outcomes</td>
<td>IL12RB1 (+1664) polymorphism was associated with susceptibility to SARS-CoV-1 infection</td>
<td>18478121&lt;sup&gt;218&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of polymorphisms in 4 C-type lectin genes with susceptibility to SARS-CoV-1 infection</td>
<td>No association of polymorphisms in C-type lectin genes with SARS-CoV-1 susceptibility</td>
<td>18697825&lt;sup&gt;219&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of polymorphisms in 9 inflammatory response genes with susceptibility to SARS-CoV-1 or clinical outcomes</td>
<td>No association of polymorphisms in inflammatory response genes with SARS-CoV-1 susceptibility or clinical outcomes</td>
<td>18708672&lt;sup&gt;220&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of polymorphisms in MASP2 with susceptibility to SARS-CoV-1 infection</td>
<td>No association of MASP2 polymorphisms with SARS-CoV-1 susceptibility</td>
<td>19405982&lt;sup&gt;221&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of HLA polymorphisms with SARS-CoV-1 susceptibility</td>
<td>HLA-DRB1<em>12 was more frequent in SARS-CoV-1 patients versus controls; HLA-DRB1</em>1202 showed the strongest association with SARS-CoV-1 infection in a dominant model</td>
<td>19445991&lt;sup&gt;222&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of polymorphisms in 64 genes with susceptibility to SARS-CoV-1 infection</td>
<td>CXCL10(-938AA) is protective (but appears jointly with other variants); FGL2(+158T/*) is associated with higher susceptibility unless combined with CXCL10/(-938AA), when jointly is associated with lower susceptibility</td>
<td>19590927&lt;sup&gt;223&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of CD209 rs4804803 with SARS-CoV-1 outcomes</td>
<td>CD209 polymorphism rs4804803 is associated with lower LDH levels (and therefore, worse prognosis)</td>
<td>20359516&lt;sup&gt;224&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Biological study and analysis of MX1 promoter polymorphisms with suppressed interferon beta induction and association of MX1 promoter polymorphisms with susceptibility to SARS-CoV-1 infection</td>
<td>Differences were observed in binding affinity to nuclear proteins related to IFN-beta stimulation; MX1 rs2071430 was associated with lower risk of SARS-CoV-1 infection</td>
<td>20462354225</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of HLA gene polymorphisms with SARS-CoV-1 susceptibility</td>
<td>No significant associations (after correction) HLA gene polymorphisms with SARS-CoV-1 susceptibility were identified</td>
<td>20864745226</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Biological study of in vitro functional effects of rs4804803 and analysis of association of CD209 rs4804803 with SARS-CoV-1 outcomes</td>
<td>CD209 polymorphism rs4804803 was associated with lower risk of high admission LDH levels, and may contribute to a reduced immune response/reduced lung injury during disease progression</td>
<td>20864747227</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of AHSG and CYP4F3A polymorphisms with SARS-CoV-1 susceptibility</td>
<td>AHSG polymorphism rs2248690 was associated with SARS-CoV-1 susceptibility (as well as higher</td>
<td>21904596228</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <em>HLA</em> polymorphisms with SARS-CoV-1 susceptibility</td>
<td>HLA-Cw*1502 conferred resistance against SARS infection is associated with resistance to SARS-CoV-1 infection</td>
<td>21958371&lt;sup&gt;229&lt;/sup&gt;</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <em>HLA</em> polymorphisms with SARS-CoV-1 susceptibility and outcome</td>
<td>No association of <em>HLA</em> polymorphisms with SARS-CoV-1 susceptibility and outcome were identified</td>
<td>24643938&lt;sup&gt;230&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1</td>
<td>Analysis of association of <em>CCL2</em> and <em>MBL</em> polymorphisms with susceptibility to SARS-CoV-1 infection</td>
<td>Variants in <em>MBL</em> (rs1800450) and <em>CCL2</em> (rs1024611) (CCL2) were cumulatively associated with SARS-CoV-1 susceptibility</td>
<td>25818534&lt;sup&gt;231&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-1 (and other respiratory pathogens)</td>
<td>Meta-analysis of 386 studies on susceptibility to tuberculosis, influenza, respiratory syncytial virus, SARS-CoV-1, and pneumonia</td>
<td>In a pooled model, variants in <em>IL4</em> were positively associated with susceptibility after multiple testing correction</td>
<td>26524966&lt;sup&gt;174&lt;/sup&gt;</td>
</tr>
<tr>
<td>SARS-CoV-2</td>
<td>Case report of death due to</td>
<td>Suggestion of genetic</td>
<td>32277694&lt;sup&gt;232&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>COVID-19 in three previously healthy adult brothers</td>
<td>predisposition due to apparent familial clustering</td>
<td>SARS-CoV-2</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Case reports</strong></td>
<td>Case reports of two patients with X-linked agammaglobulinemia (and documented pathogenic variants in BTK)</td>
<td>Patients recovered, suggesting that B cell response might not be required to overcome the SARS-CoV-2 infection</td>
<td>32319118^{233}</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td>Analysis of association of IFITM3 rs12252 with clinical outcomes of SARS-CoV-2 infection</td>
<td>Significant association of homozygosity IFITM3 rs12252 with disease severity</td>
<td>32348495^{234}</td>
</tr>
</tbody>
</table>

**Abbreviations:** CCoV: canine coronavirus; FCoV: feline coronavirus; human coronavirus 229E: HCoV-229E; human coronavirus NL63: HCoV NL63; human coronavirus OC43: HCoV OC43; LDH: lactate-dehydrogenase; MBL: Mannose-binding lectin; MERS-CoV: middle east respiratory syndrome coronavirus; SARS-CoV-1: severe acute respiratory syndrome coronavirus 1; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; SL-CoV: SARS-Cov-1-like coronaviruses; TGEV: porcine transmissible gastroenteritis coronavirus; WBC: white blood cell; WT: wild-type

**Table 2.** Summary of relevant mouse studies related to coronavirus. Note that the different studies have disparate objectives, many of which more directly involve aspects of immunopathogenesis versus standard host genetic questions regarding why specific genetic variants may affect disease susceptibility and outcomes.
<table>
<thead>
<tr>
<th>Mouse (Human gene)</th>
<th>Method(s) or approach(es)</th>
<th>Pathway: Key findings</th>
<th>PMID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ace2 (ACE2)</strong></td>
<td>Humanized mice, SARS-CoV1</td>
<td>Viral receptor: humanized Ace2 mice, increased infection, permissive gene</td>
<td>18495771&lt;sup&gt;235&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Atg5 (ATG5)</strong></td>
<td>KO, MHV infection</td>
<td>Autophagy: required for MHV replication, permissive gene</td>
<td>14699140&lt;sup&gt;236&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Atp1a1 (ATP1A1)</strong></td>
<td>knockdown and chemical inhibition across many coronaviruses</td>
<td>Ion channel: chemical inhibition or gene silencing, results in blocking viral entry, permissive gene</td>
<td>25653449&lt;sup&gt;237&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>B2m (B2M)</strong></td>
<td>KO, MHV infection</td>
<td>Adaptive immunity: MHC Class I/CD8 T-cells required for host immune response, protective gene</td>
<td>8799201&lt;sup&gt;238&lt;/sup&gt;; 10023135&lt;sup&gt;239&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Bnip3 (BNIP3)</strong></td>
<td>Cull culture model, MHV infection</td>
<td>Apoptosis: pro-apoptotic gene is suppressed upon viral entry, likely protective</td>
<td>14599795&lt;sup&gt;240&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>C3 (C3)</strong></td>
<td>KO, SARS-CoV1</td>
<td>Complement pathway: decreased complement activation leads to less</td>
<td>30301856&lt;sup&gt;241&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gene (Symbol)</td>
<td>KO, MHV infection</td>
<td>Description</td>
<td>Gene ID</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>C5ar1 (CSAR1)</td>
<td>KO, MHV infection</td>
<td>Complement pathway: Complement pathway exacerbates hepatitis, KO decreases manifestations, decreased susceptibility, permissive gene</td>
<td>24604562</td>
</tr>
<tr>
<td>Ccr1 (CCR1)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Loss of Ccr1 increased mortality, protective gene</td>
<td>18158733</td>
</tr>
<tr>
<td>Ccr2 (CCR2)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Ccr2 required for clearance of the virus from CNS, KO increased susceptibility, protective gene</td>
<td>15518805</td>
</tr>
<tr>
<td>Ccr5 (CCR5)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: KO decreased severity of demyelination disease, permissive gene</td>
<td>11543653</td>
</tr>
<tr>
<td>Cd200r1 (CD200R1)</td>
<td>KO, MHV infection</td>
<td>Immune receptor: Cd200 KO increases clearance of MHV, decreases susceptibility, permissive gene</td>
<td>22615569</td>
</tr>
<tr>
<td>Ceacam1</td>
<td>Isoform specific</td>
<td>Viral receptor: KO are fully resistant to</td>
<td>11483763</td>
</tr>
<tr>
<td>Gene</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T2), KO leads to increased mortality, protective gene</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>CEACAM1</td>
<td>transgenic and KO, MHV infection</td>
<td>infection, liver, and CNS manifestations, permissive gene</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>KO, MHV infection</th>
<th>Cytokine pathways: Interferon related (T2), KO leads to increased mortality, protective gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cxcl10 (CXCL10)</td>
<td>KO, MHV infection</td>
<td>17142734&lt;sup&gt;248&lt;/sup&gt;, 17617609&lt;sup&gt;249&lt;/sup&gt;;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>KO, MHV infection</th>
<th>Cytokine pathways: Interferon related (T2), KO had increased MHV associated mortality, protective gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cxcl9 (CXCL9)</td>
<td>KO, MHV infection</td>
<td>18973912&lt;sup&gt;250&lt;/sup&gt;;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Various transgenic and humanized models, MERS infection</th>
<th>Viral receptor: humanized Dpp4 or mutations, deletions in mouse Dpp4 leads to MERS induced ARDS, permissive gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dpp4 (DPP4)</td>
<td>KO, MHV infection</td>
<td>24574399&lt;sup&gt;251&lt;/sup&gt;, 25653445&lt;sup&gt;252&lt;/sup&gt;;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>KO, MHV infection</th>
<th>Cytokine pathways: Interferon related (T2), KO leads to increased mortality, protective gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ebi3 (EBI3)</td>
<td>KO, MHV infection</td>
<td>23102608&lt;sup&gt;256&lt;/sup&gt;;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>KO, MHV infection</th>
<th>Adaptive immunity: Athymic mice lacking T-cells unable to clear infection cause severe disseminated disease, protective gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foxn1 (FOXN1)</td>
<td>KO, MHV infection</td>
<td>8799201&lt;sup&gt;238&lt;/sup&gt;, 15070459&lt;sup&gt;257&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Gene</td>
<td>KO, Virus</td>
<td>Pathway/Effect</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>H2-Ab1 (H2AB1)</td>
<td>KO, MHV infection</td>
<td>Adaptive immunity: MHC Class I/CD4 T-cells required for host immune response, protective gene</td>
</tr>
<tr>
<td>Ifih1 (IFIH1)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T1), KO more severe, disseminated MHV infection, decreased survival, protective gene</td>
</tr>
<tr>
<td>Ifnar (IFNAR)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T1), KO leads to increased mortality and higher viral titers, protective gene</td>
</tr>
<tr>
<td>Ifnar1 (IFNAR1)</td>
<td>KO, SARS-CoV</td>
<td>Interferon pathway: Type 1, II and III interferon does not alter infection for SARS-CoV-1, in contrast to MHV</td>
</tr>
<tr>
<td>Ifng (IFNG)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T2), KO has increased mortality, decreased viral clearance, protective gene</td>
</tr>
<tr>
<td>Ifngr1 (IFNGR1)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T2), KO has increased mortality, decreased viral clearance, protective</td>
</tr>
<tr>
<td>Gene</td>
<td>KO, Infection</td>
<td>Pathway</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Ifngr1 (IFNGR1)</td>
<td>KO, SARS-CoV1</td>
<td>Interferon pathway: Type 1, II and III interferon does not alter infection for SARS-CoV-1, in contrast to MHV</td>
</tr>
<tr>
<td>Ighm (IGHM)</td>
<td>KO, MHV infection</td>
<td>Adaptive immunity: B-cell deficient develop subclinical infection and transmit virus for increased time span, protective gene</td>
</tr>
<tr>
<td>Il1r1 (IL1R1)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: KO shows reduced viral replication, mortality, and disease progression, permissive gene</td>
</tr>
<tr>
<td>Mavs (MAVS)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related (T1), viral sensor, studied in the presence of attenuated virus, protective gene</td>
</tr>
<tr>
<td>Myd88 (MYD88)</td>
<td>KO, rMA15 infection</td>
<td>Cytokine pathways: downstream of multiple pathways, KO increased susceptibility to MHV infection and mortality, protective gene</td>
</tr>
</tbody>
</table>

For use under a CC0 license.

This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.
<table>
<thead>
<tr>
<th>Gene</th>
<th>KO, Infection</th>
<th>Phenotype</th>
<th>Gene ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prkdc (PRKDC)</td>
<td>KO, MHV infection</td>
<td>Adaptive immunity: Loss of T- and B-cells cause severe disseminated infection, protective gene</td>
<td>8799201(^{238})</td>
</tr>
<tr>
<td>Rag1 (RAG1)</td>
<td>KO, MHV infection</td>
<td>Adaptive immunity: Loss of mature T- and B-cells leads to failure to clear infection, protective gene</td>
<td>17142734(^{248}), 18973912(^{250}), 25428866(^{271}), 27604627(^{272})</td>
</tr>
<tr>
<td>Serpine1 (SERPINE1)</td>
<td>KO, SARS-CoV1 infection</td>
<td>Tissue remodeling: KO mice are more susceptible to infection and inflammation, protective gene</td>
<td>23919993(^{273})</td>
</tr>
<tr>
<td>Stat1 (STAT1)</td>
<td>KO/KI, HCoV-229E infection</td>
<td>Cytokine pathways: Interferon related (T1), KO increased susceptibility HCoV in transgenic APN model, protective</td>
<td>15919828(^{274})</td>
</tr>
<tr>
<td>Stat1 (STAT1)</td>
<td>KO, SARS-CoV-1</td>
<td>Cytokine pathways, KO worsens disease, increases susceptibility, protective gene</td>
<td>20386712(^{262}), 23142821(^{275})</td>
</tr>
<tr>
<td>Stat6 (STAT6)</td>
<td>Conditional KO, LysM and FoxJ1, Stat1/Stat6 -/- double knockout</td>
<td>Cytokine pathways: conditional KO of Stat1 in macrophages but not ciliated epithelial cells showed pulmonary disease, double knockout of Stat1 and</td>
<td>23015710(^{276})</td>
</tr>
<tr>
<td>Immune Receptor</td>
<td>Gene</td>
<td>KO, Virus</td>
<td>Pheno-</td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>-----------</td>
<td>------</td>
</tr>
<tr>
<td>Stat6</td>
<td>Ticam2 (TICAM2)</td>
<td>KO, SARS-CoV1</td>
<td>Relieves pulmonary disease, implicates alternatively activated macrophages, permissive gene</td>
</tr>
<tr>
<td>Tlr2 (TLR2)</td>
<td>KO, MHV infection</td>
<td>Immune receptor: KO decreases inflammatory response, protective gene</td>
<td>19740307(^{278})</td>
</tr>
<tr>
<td>Tlr3 (TLR3)</td>
<td>KO, SARS-CoV1</td>
<td>Immune receptor: TLR mediated, KO more susceptible for SARS-CoV-1 infection, although no increased mortality, protective gene</td>
<td>26015500(^{279})</td>
</tr>
<tr>
<td>Tlr4 (TLR4)</td>
<td>KO, SARS-CoV1</td>
<td>Immune receptor: TLR mediated, KO more susceptible for SARS-CoV-1 infection, although no increased mortality, protective gene</td>
<td>26015500(^{279})</td>
</tr>
<tr>
<td>Tlr7 (TLR7)</td>
<td>KO, MHV infection</td>
<td>Immune receptor: viral sensor, KO prolonged infection, protective gene</td>
<td>29717007(^{269})</td>
</tr>
<tr>
<td>Gene (Abbreviation)</td>
<td>Phenotype</td>
<td>Pathway</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td><em>Tram1</em> (<em>TRAM1</em>)</td>
<td>KO, SARS-CoV-1</td>
<td>Immune receptor: TLR mediated</td>
<td>KO more susceptible for SARS-CoV-1 infection, although no increased mortality, protective gene</td>
</tr>
<tr>
<td><em>Trif</em> (<em>TRIF</em>)</td>
<td>KO, SARS-CoV-1</td>
<td>Immune receptor: TLR mediated</td>
<td>KO more susceptible to SARS-CoV-1 infection, more severe infection with increased interferon signaling, protective gene</td>
</tr>
<tr>
<td><em>Trim55</em> (<em>TRIM55</em>)</td>
<td>KO, SARS-CoV-1</td>
<td>Uncharacterized pathway: contributed to lung pathology</td>
<td>KO decreased severity, permissive gene</td>
</tr>
<tr>
<td><em>Usp18</em> (<em>USP18</em>)</td>
<td>KO, MHV infection</td>
<td>Cytokine pathways: Interferon related</td>
<td>(T1), KO leads to increased survival, decreased pathology and viral titer, gene is permissive</td>
</tr>
</tbody>
</table>

**Abbreviations:**
- ARDS: acute respiratory distress syndrome
- CNS: central nervous system
- KI: knock-in
- KO: knockout
- MERS: middle east respiratory syndrome
- MHC: major histocompatibility complex
- MHV: mouse hepatitis virus
- SARS-CoV-1: severe acute respiratory syndrome coronavirus 1
- T1: type 1
- T2: type 2
- TLR: Toll-like receptor

**Figure Legends**
**Figure 1.** Description of articles identified through PubMed searches described in Methods.

**Figure 2.** Genes investigated in animal studies related to coronavirus disease. Human genes are shown only for those studies in multiple species analyses; other human gene details are presented elsewhere.

**Figure 3.** 3A: Significant genetic associations with human susceptibility to coronavirus disease. Both protective and permissive genes are shown. Only studies reporting odds ratios (OR) and confidence intervals are shown. 3B: Significant genetic associations with human clinical variables and outcomes related to coronavirus disease. Both protective and permissive genes are shown. Only studies reporting odds ratios (OR) and confidence intervals (CI) are shown (PMID 32348495 did not include CI).

**Acknowledgments**

This research was supported [in part] by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health. PD was supported by the US National Institutes of Health award number 2R01AI148049-21A1. DATC was supported by the US National Institutes of Health award number R01-AI114703-01.

**References**


244. Held KS, Chen BP, Kuziel WA, Rollins BJ, Lane TE. Differential roles of CCL2 and CCR2 in host defense to coronavirus infection. Virology 2004;329:251-60.

245. Glass WG, Liu MT, Kuziel WA, Lane TE. Reduced macrophage infiltration and demyelination in mice lacking the chemokine receptor CCR5 following infection with a neurotropic coronavirus. Virology 2001;288:8-17.


250. Muse M, Kane JA, Carr DJ, Farber JM, Lane TE. Insertion of the CXC chemokine ligand 9 (CXCL9) into the mouse hepatitis virus genome results in protection from viral-induced encephalitis and hepatitis. Virology 2008;382:132-44.


265. de Wit MC, Horzinek MC, Haagmans BL, Schijns V. Host-dependent type 1 cytokine responses driven by inactivated viruses may fail to default in the absence of IL-12 or IFN-alpha/beta. J Gen Virol 2004;85:795-803.


