Genetic Predisposition to Infectious Disease

Nikolai Klebanov [1]

1. Dermatology, Massachusetts General Hospital, Boston, USA

✉ **Corresponding author:** Nikolai Klebanov, nikolai.klebanov@tufts.edu
Disclosures can be found in Additional Information at the end of the article

Abstract

In contemporary medical practice, approaches to infectious disease management have been primarily rooted in a pathogen-centered model. However, host genetics also contribute significantly to infectious disease burden. The fast expansion of bioinformatics techniques and the popularization of the genome-wide association study (GWAS) in recent decades have allowed for rapid and affordable high-throughput genomic analyses. This review focuses on the host model of infectious disease with particular emphasis placed on the genetic variations underlying observed infectious disease predisposition. First, we introduce observational twin-twin concordance studies of diseases such as poliomyelitis, tuberculosis, and hepatitis which suggest the important role of host genetics. We review the well-established links between specific genetic alterations and predisposition to malaria (*P. falciparum* and *P. vivax*), Creutzfeldt-Jacob disease (CJD), human immunodeficiency virus (HIV), and Norwalk virus. Finally, we discuss the novel findings yielded by modern GWAS studies, which suggest the strong contribution of immunologic variation in the major histocompatibility complex (MHC) to host genetic infectious disease susceptibility. Future large-scale genomic studies hold promise in providing insights into immunology-pathogen links and may allow for the development of personalized genomic approaches to infectious disease prevention and treatment.

Introduction And Background

The pathogen or germ theory of infectious disease was not always as universally accepted as it is today. Hippocrates (c.460–370 BC) considered *phtisis*, a disease now known as tuberculosis, to be hereditary rather than infectious as he observed that it commonly clustered within families [1]. Nearly five hundred years later, Galen (129–210 CE) proposed that *phtisis* may have a contagious nature, and he warned against close contact with those afflicted with this disease. The pathogen-focused infectious disease theory truly took its strong hold in contemporary medicine in the late 19th century. Groundbreaking work by German physician and microbiologist Robert Koch led to the isolation of *Mycobacterium tuberculosis* as the causative tuberculosis pathogen, and thus medical microbiology became established as an independent discipline [2]. The germ infectious disease theory has led to unparalleled benefits for mankind in the development of highly effective antibiotics and antiviral and anti-parasitic medications.

Infectious diseases exert significant selective genetic pressure, and the genes that are involved in immune response are exquisitely diverse [3]. These observations suggest a strong role for host genetic variability in the susceptibility to exogenous pathogens. These genetic links have been observed grossly in twin-twin concordance studies. Specific genetic mutations and

How to cite this article
Klebanov N (August 27, 2018) Genetic Predisposition to Infectious Disease. Cureus 10(8): e3210. DOI 10.7759/cureus.3210

variations have been strongly implicated in the literature to confer susceptibility or resistance to diseases such as malaria (*P. falciparum* and *P. vivax*), Creutzfeldt-Jacob disease (CJD), human immunodeficiency virus (HIV), and Norwalk virus. Multiple novel genomic studies have demonstrated strong associations between polymorphisms in multiple histocompatibility complex (MHC) and human leukocyte antigen (HLA) genes with infectious disease susceptibility. Here, we provide an overview of these findings and discuss the future directions of this research and implications for clinical practice.

Review

Concordance studies

Several studies focusing on infectious disease among monozygotic (MZ) and dizygotic (DZ) twin pairs have suggested a genetic contribution to disease burden. A twin-family study of poliomyelitis revealed that 36% of MZ twin pairs presented with paralytic poliomyelitis, compared to only 6% among DZ twins [4]. Tuberculosis concordance was likewise shown to be significantly higher among MZ than DZ pairs [5]. An analysis of hepatitis B among a Chinese population showed higher rates of seroconcordance between MZ and DZ pairs compared to non-twin sibling controls [6]. A landmark adoptee study in a Danish population reported in 1988 by Sørensen et al. reveals an interesting finding regarding the likely heritability of infectious disease predisposition [7]. The study authors followed 960 families where children were placed early in life with adopted parents. The risk of the adopted child dying from certain causes was evaluated in the context of a biologic or adoptive parent dying from the same cause before the age of 50. If a biologic parent died from infectious disease before age 50, the adopted child had 5.81 greater risk of also dying from infection. Meanwhile, the relative risk (RR) for the child if the adoptive parent died from infection was close to unity. The high RR of 5.81 was greater than the RR associated with cardiac, cerebrovascular, oncologic, and natural causes for the biologic parent. The results of this groundbreaking study suggest a strong host genetic component in infection predisposition, possibly stronger than for those diseases commonly believed to be highly heritable such as malignancy or cardiovascular disease.

The "big six" of infectious disease genetics

Six important genetic relationships with infectious disease susceptibility have been discovered and repeatedly validated in the literature, and thus constitute very strong findings [8]. These associations are summarized in Table *1*.

Year	Disease	Gene	Effect	Notes
1954	*Plasmodium falciparum malaria*	Hemoglobin subunit beta (HBB)	Protective	Increased red blood cell (RBC) turnover and abnormal shape à increased uptake by macrophages; decreased parasite growth
1976	*Plasmodium vivax malaria*	Duffy antigen receptor (DARC)	Protective	Duffy glycoprotein is an RBC surface receptor for *P. vivax;* abnormal or missing receptor leads to impaired parasite penetration
1991	Creutzfeldt-Jacob disease (CJD)	PRiON protein (PRNP)	Susceptible	Homozygous PRNP mutations predispose to CJD. 51% of the general population is PRNP heterozygous
1995	*Plasmodium falciparum malaria*	Band 3 anion transport protein (SLC4A1)	Protective	Southeast Asian ovalocytosis (SAO) is similar (but distinct from) hereditary elliptocytosis. Entry of *P. falciparum* is impaired.
1996	Human immunodeficiency virus 1 (HIV-1)	Chemokine receptor 5 (CCR5)	Protective	Altered CCR5 co-receptor (such as seen in CCR5Δ32 mutation) leads to impaired viral T-cell entry and slower progression to acquired immunodeficiency syndrome (AIDS)
2003	Norwalk virus (norovirus)	Fucosyltransferase 2 (FUT2)	Protective	*FUT2* variants lead to non-secretor phenotype; gut cells do not express ligands necessary for binding of the GII.4 norovirus strain

TABLE 1: The "big six" genetic variants relevant to infectious disease

Selective pressure in populations residing in areas with high prevalence of malaria has favored the sickle cell trait [9], and there is evidence that this genetic selection still continues in the present day [10]. Variation in the HBB gene has also been found to be associated with malaria protection by GWAS [11], and interestingly, has been associated with a significant host to vector *P. falciparum* transmission [12]. Another protective variant against *P. falciparum* malaria is the genetic condition termed Melanesian ovalocytosis (also known as Southeast Asian ovalocytosis or SAO). In SAO, there is a defect in the band 3 protein, an RBC membrane protein, which causes the band 3 protein to ankyrin protein bond to be stronger than normal. Multiple consequences result from this genetic change including greater RBC robustness, reduced anion exchange, intracellular ATP partial depletion, and a decrease in antigen expression [13]. Due to a combination of these forces, the entry of *P. falciparum* malaria into the RBC is impaired. SAO is maintained as a balanced polymorphism among human populations [14], and to highlight the genetic selective pressure for this condition, there is a 35% incidence of SAO in the north coast of Madang Province in Papua New Guinea, a geographic location where malaria is endemic [15].

The Duffy blood group genotype is another important host genetic contribution to malaria susceptibility, specifically to the *Plasmodium vivax* species. In 1975, Miller et al. studied the blood types of 11 black and six white volunteers who had been exposed to bites of *P. vivax*-infected mosquitoes. The study authors demonstrated that individuals with Duffy-blood-group-negative erythrocytes were resistant to parasite invasion [16]. They coined the Duffy group negative genotype as FyFy. This study led to the discovery that Duffy group antigens that are present on the RBC surface are important recognition sites and entry points for the *P. vivax* malaria parasite. The Duffy genotype is another important example of genetic selection as *P.*

2018 Klebanov et al. Cureus 10(8): e3210. DOI 10.7759/cureus.3210

3 of 8

vivax, which is widespread throughout tropical and subtropical areas, is absent in West Africa where more than 95% of the population is Duffy negative [17].

Viral disease also presents with three important host genetic variants affecting disease susceptibility. Mutations in the PRNP (PRiON protein) gene, which are strikingly common at 51% among the normal population, are known to be a predisposing factor for sporadic Creutzfeld-Jakob disease [18]. The human immunodeficiency virus 1 (HIV-1) uses the CCR5 co-receptor on the surface of CD4+ T-lymphocytes to recognize and enter the T-cell. Thus, variants in the CCR5 gene such as the CCR5Δ32 mutation have been strongly linked to slower (acquired immunodeficiency syndrome) AIDS progression and protection against HIV infection. The CCR5 gene has been studied in vitro as a promising target for HIV treatment [19]. Finally, susceptibility to infections with norovirus has been shown to be affected by the FUT2 genotype. Cells lining the gastrointestinal tract in humans present A, B, or O blood group antigens on their surface. Fucosyltransferase 2, the protein product of FUT2, is an important mediator of this antigen presentation. Mutant FUT2 thus yields gut lining cells that do not present these antigens, and individuals with this phenotype are termed non-secretors. The non-secretor phenotype has been shown to be protective against norovirus infection specifically with the GII.4 strain [20]. The strain-specificity limits the clinical application of the knowledge of this genetic host factor, as in a case series of four patients who were affected by norovirus gastroenteritis, all four patients tested negative on a consumer genetic test. Three of these four individuals were infected with the GI.6 norovirus strain [21].

Genome-wide association study (GWAS) of infectious disease susceptibility

GWAS have been instrumental in the discovery of several important host genetic associations with infectious disease course. The methodology of GWAS allows for the simultaneous assay of over 500,000 single nucleotide polymorphisms (SNPs) across thousands of individuals. The statistical analysis of GWAS involves curating those SNPs that are significantly correlated with cases of disease with a p-value below 5×10^{-8}, which is termed "genome-wide significance", This stringent significance cutoff is important for controlling the number of false-positive hits given that hundreds of thousands or sometimes one to two million SNPs are interrogated at once (although alternative cutoffs to the standard 5×10^{-8} have been proposed in the literature [22-23]). The GWAS approach is particularly strong in its ability to provide for a systematic, global, and unbiased search for strong disease association, and GWAS has been instrumental in identifying many significantly associated variants of modest effect size [24].

Notable confirmatory and new findings yielded by GWAS include the significant association of common SNPs with leprosy, malaria, meningococcal disease, dengue, and chronic hepatitis B and C [24-25]. GWAS of leprosy [26-27] (caused by *Mycobacterium leprae*) revealed strong associations with HLA-DR and NOD2 which encode recognition sites for pathogen-associated motifs, and IL23R, RIP2K, and TNFSF15 which are involved in the downstream pathogen immune inflammatory response. With regards to malaria, GWAS have confirmed ABO blood-group association with *P. falciparum* infection (such as variation at the HBB gene locus), but no novel insights have emerged [11,28]. In meningococcal disease (referring to invasive *Neisseria meningitidis* brain or bloodstream infections), strong associations emerged between Factor H (CFH) and other CFH-related genes such as CFHR3 and CFHR1 [29]. The identification of this locus as significant confirms the known important interaction between *N. meningitidis* and the human complement cascade. Analysis of severe dengue infection manifesting as dengue shock with GWAS has identified strong associations at the MICB (located within the MHC complex) and PLCE1 loci [30]. These results suggest that infection with *Flaviviridae* family members may rely on the immune response arm constituted by MICB signaling via natural killer (NK) cells.

Regarding chronic hepatitis B virus (HBV), GWAS have revealed that HLA-DP variation strongly influences chronic HBV infection [31], supporting prior observations that antigen-presentation is an important interaction point between HBV infection and human host response.

The IL-28B association with hepatitis C warrants special mention. The IL-28B gene codes for the cytokine interferon lambda (IFN-lambda). Single nucleotide polymorphisms (SNPs) in the IL-28B locus have been strongly associated with response to IFN-alpha treatment in hepatitis C [32,33]. Race including African-American and Hispanic race has a significant effect-modulatory contribution to this association. Although this is a genetically-robust association which has been validated with GWAS study, in the context of contemporary direct-acting antiviral (DAA) therapy, the effects of differing IL-28B genotypes has shown very small impact. Given this small impact once appropriate DAA therapy is initiated, routine genotyping for IL-28B is not recommended in clinical practice.

Apart from these specific findings, GWAS has yielded mainly associations with many human leukocyte antigen (HLA) and multiple histocompatibility locus (MHC) genes to infectious disease [8,25]. Nearly half of loci identified by GWAS focused on infectious disease susceptibility are located within the broad MHC region of chromosome 6 [24]. These findings emphasize the undeniably important role of complex host immune response in the interaction with invading pathogens.

Summary of host genetic patterns contributing to infectious disease response

The main thematic groups of host genetic variations underlying disease susceptibility are summarized in Figure 1.
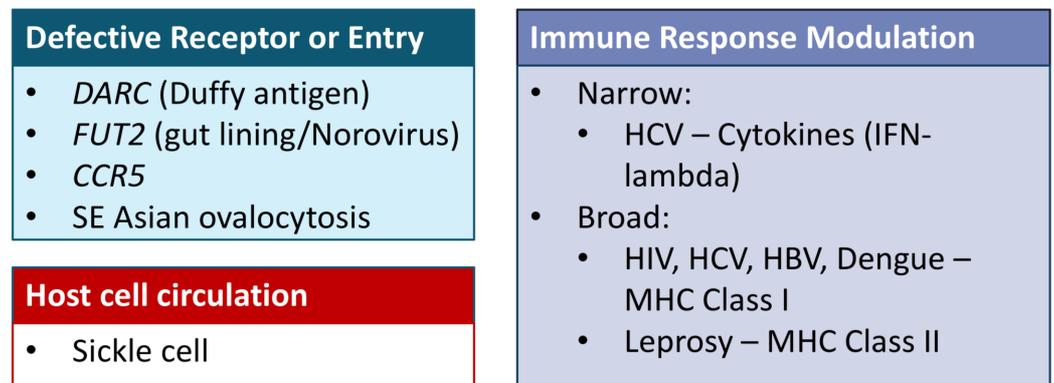

Defective Receptor or Entry
- *DARC* (Duffy antigen)
- *FUT2* (gut lining/Norovirus)
- *CCR5*
- SE Asian ovalocytosis

Host cell circulation
- Sickle cell

Immune Response Modulation
- Narrow:
  - HCV – Cytokines (IFN-lambda)
- Broad:
  - HIV, HCV, HBV, Dengue – MHC Class I
  - Leprosy – MHC Class II

FIGURE 1: Summary of host genetic patterns contributing to infectious disease response

The relationships of DARC to *P. vivax*, FUT2 to norovirus, CCR5 to HIV-1, and SAO to *P. falciparum* constitute defective receptor or entry points which thus are protective against infection. This group of genetic variants is characterized by specific cell-surface protein alterations. Findings of such specific entry-point mutations are promising as possible therapeutic targets. The sickle cell anemia to *P. falciparum* malaria relationship falls under the rare category of alterations in the human host cell circulation patterns (specifically RBC longevity). Given the specific targeting of malaria to the RBC, this relationship is a unique one, and it is unlikely that other not yet discovered genetic variants would fall under this category. Finally, the category of immune response modulation captures the majority of the findings

yielded by GWAS. There is the narrow immune system modulation such as the IL-28B or IFN-gamma relationship to chronic hepatitis C (HCV) infection. There is also a component of broad immune system variation such as the relation of multiple loci in the MHC region of chromosome 6 to many infectious disease entities.

Conclusions

In this review, we have discussed the host-focused approach to infectious disease study (rather than the pathogen-focused one), with particular emphasis placed on specific human host genetic variations that predispose to changes in the response to invading pathogens. Candidate-gene studies throughout the past decades have led to the discovery and validation of well-established links between specific genetic alterations and predisposition to malaria (*P. falciparum* and *P. vivax*), Creutzfeldt-Jacob disease (CJD), human immunodeficiency virus (HIV), and Norwalk virus. Modern studies using the GWAS approach have yielded several novel findings and many confirmatory ones. Most GWAS-identified loci are located within the broad MHC region of chromosome 6, suggesting that infectious disease susceptibility is highly dependent across many pathogen entities to complex variation within the human immune response. These are interesting findings, but with only nascent robust ways to specifically target and modulate the immune system such as immunotherapy for malignancies, these findings are not currently clinically actionable. Future studies with in vitro and in vivo validation of these GWAS findings, and the expected further development of more highly-specific immunotherapies, are likely to yield promising results in the use of immune modulation for infectious disease treatment.

Additional Information

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Herzog H: History of tuberculosis. Respiration. 1998, 65:5-15. 10.1159/000029220
2. Kaufmann SH: Robert Koch, the Nobel Prize, and the ongoing threat of tuberculosis . N Engl J Med. 2005, 353:2423-2426. 10.1056/NEJMp058131
3. Burgner D, Jamieson SE, Blackwell JM: Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better?. Lancet Infect Dis. 2006, 6:653-663. 10.1016/S1473-3099(06)70601-6
4. Herndon CN, Jennings RG: A twin-family study of susceptibility to poliomyelitis . Am J Hum Genet. 1951, 3:17-46.
5. Comstock GW: Tuberculosis in twins: a re-analysis of the Prophit survey . Am Rev Respir Dis. 1978, 117:621-624.
6. Lin TM, Chen CJ, Wu MM, et al.: Hepatitis B virus markers in Chinese twins . Anticancer Res. 1989, 9:737-741.
7. Sorensen TI, Nielsen GG, Andersen PK, Teasdale TW: Genetic and environmental influences on premature death in adult adoptees. N Engl J Med. 1988, 318:727-732. 10.1056/NEJM198803243181202
8. Hill AV: Evolution, revolution and heresy in the genetics of infectious disease susceptibility . Philos Trans R Soc Lond B Biol Sci. 2012, 367:840-849. 10.1098/rstb.2011.0275
9. Luzzatto L: Sickle cell anaemia and malaria. Mediterr J Hematol Infect Dis. 2012, 4:2012065.

10.4084/MJHID.2012.065

10. Elguero E, Delicat-Loembet LM, Rougeron V, et al.: Malaria continues to select for sickle cell trait in Central Africa. Proc Natl Acad Sci U S A. 2015, 112:7051-7054. 10.1073/pnas.1505665112

11. Jallow M, Teo YY, Small KS, et al.: Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat Genet. 2009. 41:657-665. 10.1038/ng.388

12. Gouagna LC, Bancone G, Yao F, et al.: Genetic variation in human HBB is associated with Plasmodium falciparum transmission. Nat Genet. 2010, 42:328-331. 10.1038/ng.554

13. Liu SC, Palek J, Yi SJ, et al.: Molecular basis of altered red blood cell membrane properties in Southeast Asian ovalocytosis: role of the mutant band 3 protein in band 3 oligomerization and retention by the membrane skeleton. Blood. 1995, 86:349-358.

14. Paquette AM, Harahap A, Laosombat V, et al.: The evolutionary origins of Southeast Asian Ovalocytosis. Infect Genet Evol. 2015, 34:153-159. 10.1016/j.meegid.2015.06.002

15. Mgone CS, Koki G, Paniu MM, et al.: Occurrence of the erythrocyte band 3 (AE1) gene deletion in relation to malaria endemicity in Papua New Guinea. Trans R Soc Trop Med Hyg. 1996, 90:228-231.

16. Miller LH, Mason SJ, Clyde DF, McGinniss MH: The resistance factor to Plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy. N Engl J Med. 1976, 295:302-304. 10.1056/NEJM197608052950602

17. Langhi DM, Jr., Bordin JO: Duffy blood group and malaria . Hematology. 2006, 11:389-398. 10.1080/10245330500469841

18. Palmer MS, Dryden AJ, Hughes JT, Collinge J: Homozygous prion protein genotype predisposes to sporadic Creutzfeldt-Jakob disease. Nature. 1991, 352:340-342. 10.1038/352340a0

19. Barmania F, Pepper MS: C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection. Appl Transl Genom. 2013, 2:3-16. 10.1016/j.atg.2013.05.004

20. Currier RL, Payne DC, Staat MA, et al.: Innate susceptibility to Norovirus infections influenced by FUT2 genotype in a United States pediatric population. Clin Infect Dis. 2015, 60:1631-1638. 10.1093/cid/civ165

21. Prystajecky N, Brinkman FS, Auk B, Isaac-Renton JL, Tang P: Personalized genetic testing and norovirus susceptibility. Can J Infect Dis Med Microbiol. 2014, 25:222-224.

22. Fadista J, Manning AK, Florez JC, Groop L: The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. Eur J Hum Genet. 2016, 24:1202-1205. 10.1038/ejhg.2015.269

23. Panagiotou OA, Ioannidis JP, Genome-Wide Significance Project: What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. Int J Epidemiol. 2012, 41:273-286. 10.1093/ije/dyr178

24. Khor CC, Hibberd ML: Host-pathogen interactions revealed by human genome-wide surveys . Trends Genet. 2012, 28:233-243. 10.1016/j.tig.2012.02.001

25. Chapman SJ, Hill AV: Human genetic susceptibility to infectious disease . Nat Rev Genet. 2012, 13:175-188. 10.1038/nrg3114

26. Zhang FR, Huang W, Chen SM, et al.: Genomewide association study of leprosy . N Engl J Med. 2009, 361:2609-2618. 10.1056/NEJMoa0903753

27. Zhang F, Liu H, Chen S, et al.: Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. Nat Genet. 2011, 43:1247-1251. 10.1038/ng.973

28. Fry AE, Griffiths MJ, Auburn S, et al.: Common variation in the ABO glycosyltransferase is associated with susceptibility to severe Plasmodium falciparum malaria. Hum Mol Genet. 2008, 17:567-576. 10.1093/hmg/ddm331

29. Davila S, Wright VJ, Khor CC, et al.: Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. Nat Genet. 2010, 42:772-776. 10.1038/ng.640

30. Khor CC, Chau TN, Pang J, et al.: Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat Genet. 2011, 43:1139-1141. 10.1038/ng.960

31. Kamatani Y, Wattanapokayakit S, Ochi H, et al.: A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat Genet. 2009, 41:591-595. 10.1038/ng.348

32. Ge D, Fellay J, Thompson AJ, et al.: Genetic variation in IL28B predicts hepatitis C treatment-

induced viral clearance. Nature. 2009, 461:399-401. 10.1038/nature08309

33. Rauch A, Kutalik Z, Descombes P, et al.: Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology. 2010, 138:1345-1331. 10.1053/j.gastro.2009.12.056