Host genetic effects in pneumonia

Hung-Hsin Chen,[1] Douglas M. Shaw,[1] Lauren E. Petty,[1] Misa Graff,[2] Ryan J. Bohlender,[3] Hannah G. Polikowsky,[1] Xue Zhong,[1] Daeeun Kim,[2] Victoria L. Buchanan,[2] Michael H. Preuss,[4] Megan M. Shuey,[1] Ruth J.F. Loos,[4] Chad D. Huff,[3] Nancy J. Cox,[1] Julie A. Bastarache,[5] Lisa Bastarache,[6] Kari E. North,[2] and Jennifer E. Below[1,*]

Summary

Given the coronavirus disease 2019 (COVID-19) pandemic, investigations into host susceptibility to infectious diseases and downstream sequelae have never been more relevant. Pneumonia is a lung disease that can cause respiratory failure and hypoxia and is a common complication of infectious diseases, including COVID-19. Few genome-wide association studies (GWASs) of host susceptibility and severity of pneumonia have been conducted. We performed GWASs of pneumonia susceptibility and severity in the Vanderbilt University biobank (BioVU) with linked electronic health records (EHRs), including Illumina Expanded Multi-Ethnic Global Array ($MEGA^{EX}$)-genotyped European ancestry (EA, n= 69,819) and African ancestry (AA, n = 15,603) individuals. Two regions of large effect were identified: the *CFTR* locus in EA (rs113827944; OR = 1.84, p value = 1.2×10^{-36}) and *HBB* in AA (rs334 [p.Glu7Val]; OR = 1.63, p value = 3.5×10^{-13}). Mutations in these genes cause cystic fibrosis (CF) and sickle cell disease (SCD), respectively. After removing individuals diagnosed with CF and SCD, we assessed heterozygosity effects at our lead variants. Further GWASs after removing individuals with CF uncovered an additional association in *R3HCC1L* (rs10786398; OR = 1.22, p value = 3.5×10^{-8}), which was replicated in two independent datasets: UK Biobank (n = 459,741) and 7,985 non-overlapping BioVU subjects, who are genotyped on arrays other than $MEGA^{EX}$. This variant was also validated in GWASs of COVID-19 hospitalization and lung function. Our results highlight the importance of the host genome in infectious disease susceptibility and severity and offer crucial insight into genetic effects that could potentially influence severity of COVID-19 sequelae.

The COVID-19 pandemic is a serious threat to public health; over 58 million confirmed cases from 191 countries have been reported (see "Coronavirus Disease 2019" and "COVID-19 Map" in Web resources). Pneumonia is a common complication of COVID-19 and may lead to acute respiratory distress syndrome (ARDS) and death.[1] In the context of the COVID-19 pandemic, identifying factors that influence host susceptibility to and severity of pneumonia has never been more important. Several clinical factors and underlying conditions that influence susceptibility to and severity of community-acquired pneumonia in adults have been identified, including age, chronic bronchitis or chronic obstructive pulmonary disease (COPD), asthma, obesity, diabetes, and others.[2] These risk factors have also been observed in COVID-19-associated pneumonia.[3–6] However, little is known about the role of the host genome in the susceptibility to and severity of pneumonia.

There is a paucity of studies interrogating host genetic susceptibility to and severity of pneumonia. While previous studies identified suggestive associations with childhood pneumonia, survival from sepsis due to pneumonia, and severe pneumonia following influenza A/H1N1 infection,[7–9] three genome-wide significant associations with pneumonia, one in the *HLA* class I region, have been mapped,[10] and two additional genome-wide significant independent hits on chromosome 15 were identified in a meta-analysis of the UK Biobank and FinnGen.[11] To identify additional genetic loci impacting pneumonia in existing data resources, we aimed to identify genetic variants associated with susceptibility to and severity of pneumonia by leveraging electronic health records (EHRs) from a large-scale biobank.

The Vanderbilt University Medical Center biobank (BioVU) comprises over 110,000 participants with linked EHRs genotyped on the Illumina Expanded Multi-Ethnic Genotyping Array ($MEGA^{EX}$) or another genome-wide array. We defined pneumonia cases by using 81 pneumonia-related diagnosis codes (Supplemental box) from the International Classification of Diseases Ninth Revision, Clinical Modification (ICD-9 CM) and used the subjects without any diagnosis of pneumonia as population-based controls. We used hospitalization status as a proxy for severity, determined via relevant Current Procedural Terminology codes recorded within 5 days of pneumonia diagnosis (Supplemental box). In 69,819 $MEGA^{EX}$-genotyped European ancestry (EA) individuals, 8,889 individuals with pneumonia were identified, including 5,774 with pneumonia-associated hospitalization (inpatients). In 15,603 $MEGA^{EX}$-genotyped African ancestry (AA) individuals, we

[1]Vanderbilt Genetics Institute and Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA; [2]Department of Epidemiology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27516, USA; [3]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA; [4]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [5]Allergy, Pulmonary and Critical Care Medicine, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; [6]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN 37232, USA
*Correspondence: jennifer.e.below@vanderbilt.edu
https://doi.org/10.1016/j.ajhg.2020.12.010.

Table 1. Significant findings from discovery pneumonia GWAS, replication, and COVID-19 validation

					BioVU MEGA^{EX}		BioVU non-MEGA^{EX}		UK Biobank[a]		COVID-19 HGI[b]	
Gene	**SNP**	**Chr**	**Pos**	**Pop**	**OR**	**p value**	**OR**	**p value**	**OR**	**p value**	**OR**	**p value**
Susceptibility to pneumonia/COVID-19												
CFTR	rs113827944	7	117,299,434	EA1	1.84	1.20E−36	1.92	2.42E−06	1.04	4.90E−01	1.14	1.46E-01
HBB	rs334	11	5,248,232	AA1	1.63	3.50E−13	1.15	5.37E−01	0.89	7.37E−01	N/A	N/A
UQCRFS1	rs148218440	19	29,589,778	AA2	2.21	3.60E−08	0.70	4.27E−01	2.23	1.43E−01	N/A	N/A
Severity of pneumonia/COVID-19												
CFTR	rs113827944	7	117,299,434	EA2	1.70	4.01E−09	1.92	2.70E−02	1.04	4.90E−01	1.66	7.59E-02
R3HCC1L	rs10786398	10	99,926,570	EA3	1.22	3.50E−08	1.01	9.50E−01	1.02	2.48E−01	1.19	3.72E-02
R3HCC1L	rs884811	10	99,923,763	EA3	1.16	4.80E−06	1.10	3.64E−01	1.05	3.49E−03	N/A	N/A
R3HCC1L	rs7086391	10	100,013,563	AA3	1.42	1.10E−03	0.85	5.96E−01	1.10	6.73E−01	N/A	N/A
HBB	rs334	11	5,248,232	AA4	2.57	4.50E−12	1.73	2.96E−01	0.89	7.37E−01	N/A	N/A

EA1, European individuals with pneumonia versus control individuals; EA2, European pneumonia inpatients versus outpatients; EA3, European pneumonia inpatients versus outpatients, individuals with CF removed; AA1, African American individuals with pneumonia versus control individuals; AA2, African American pneumonia inpatients versus matched control individuals, individuals with CF and SCD removed; AA3, African American pneumonia inpatients versus outpatients, individuals with CF and SCD removed; AA4, African American pneumonia inpatients versus outpatients; Chr, chromosome; Pos, position; Pop, population.
[a]In UK Biobank, only the comparison between individuals with pneumonia versus the general population is available.
[b]In COVID-19 HGI, we used "C2" phenotypes, which compared individuals with COVID-19 to population controls, as susceptibility and "B1" results, which compared hospitalized individuals with COVID-19 to non-hospitalized control individuals with COVID-19, as severity.

identified 1,710 individuals with pneumonia, of which 1,043 were inpatients. In both EA and AA subjects, we observed a significantly higher prevalence of obesity, COPD, diabetes, asthma, and liver and renal diseases in subjects with a pneumonia diagnosis (Tables S1 and S2). We sought replication of our top findings in two independent datasets: UK Biobank (n = 451,305) and 7,985 non-overlapping BioVU subjects, who were genotyped on arrays other than MEGA^{EX} (non-MEGA^{EX}) (Tables S3–S5).

Genetic imputation in MEGA^{EX}-genotyped subjects was conducted with minimac4 on the Michigan Imputation Server[12] with a reference panel of Haplotype Reference Consortium r1.1. A total of 39,635,008 SNPs was imputed. In EA, only 7,167,360 SNPs with an imputation info score greater than 0.4 and minor allele frequency (MAF) greater than 1% were used for further GWAS and GReX imputation. In AA, 13,633,982 SNPs passed quality control filter. A more stringent MAF cutoff was applied in the comparison of AA pneumonia inpatient versus outpatient because of the smaller sample size (n = 1,710; 7,594,451 variants with MAF > 5% and imputation info score > 0.4). Because of the existence of genetic relatedness in BioVU, we utilized a generalized estimating equations (GEE) model to perform GWASs with SUGEN.[13] Since SUGEN requires known family networks representing close relatedness within a dataset, we used PRIMUS to reconstruct non-directional family networks, including all first- and second-degree,[14] and we used ERSA to verify the families with more than five members.[15] Among MEGA^{EX}-genotyped subjects, 5,019 families (size 2 or greater) were identified in BioVU EAs and 1,699 families in BioVU AAs. We included age, sex, and three principal components to capture genetic ancestry as covariates in the association tests.

Separate GWASs for susceptibility and severity of pneumonia were performed in EA and AA, each of which identified a major locus. In EA, rs113827944 (MAF = 2.1%) was significantly associated with both pneumonia susceptibility and severity (affected individuals versus control individuals, odds ratio [OR] = 1.84, p value = 1.2×10^{-36}; inpatients versus outpatients, OR = 1.70, p value = 4.0×10^{-9}). Replication of this lead SNP was observed in BioVU non-MEGA^{EX}-genotyped EA (affected individuals versus control individuals, OR = 1.92, p value = 2.4×10^{-6}). Another SNP, rs334 (p.Glu7Val, MAF = 5.8%) was significantly associated with both susceptibility and severity in AA (affected individuals versus control individuals, OR = 1.63, p value = 3.5×10^{-13}; inpatients versus outpatients, OR = 2.56, p value = 4.5×10^{-12}) (Table 1, Figures 1 and 2, and Figures S1–S3).

The lead SNP in EA, rs11382794, is located in the intron of cystic fibrosis transmembrane conductance regulator (*CFTR* [MIM: 602421]), the causal gene for cystic fibrosis (CF [MIM: 219700]). CF causes abnormally thick mucus that blocks airways, leading to chronic infections, persistent inflammation, airway remodeling, and progressive respiratory failure.[16,17] Both acute and chronic lung infections are major contributors to morbidity and mortality in individuals with CF.[18,19] In a study of 19,802 CF carriers (individuals with one defective copy of *CFTR*) and 99,010 control individuals, CF carriers were more likely than non-carriers to have pneumonia (OR = 1.16), a personal history of recurrent pneumonia (OR = 2.76), and other respiratory infections.[20]

The top finding in AA, rs334, is a nonsynonymous variant in hemoglobin subunit beta (*HBB* [MIM:141900]) and the causal mutation for SCD, including sickle cell

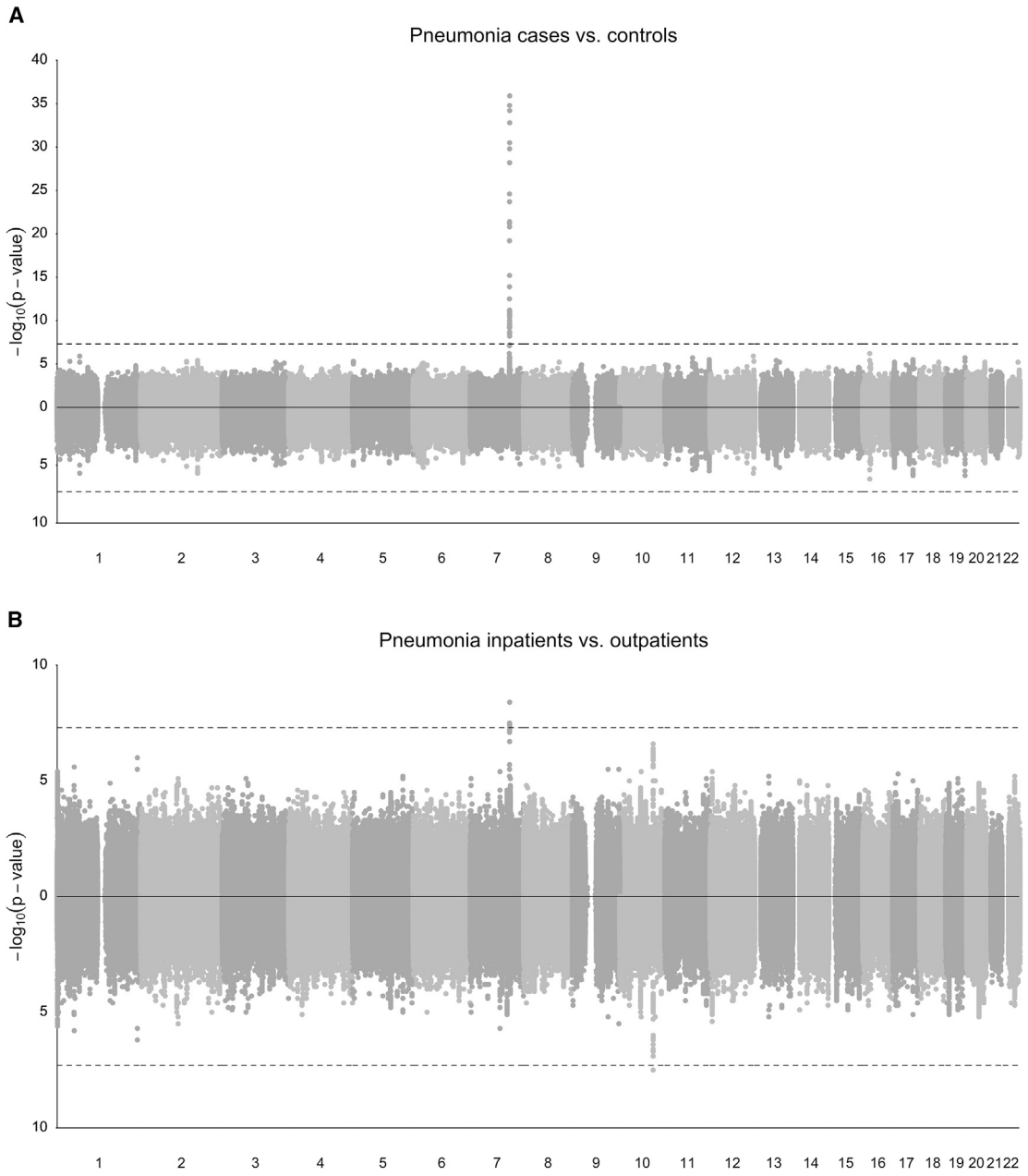

Figure 1. SNP associations with pneumonia in BioVU MEGA^{EX}-genotyped EA
(A) Miami plot of pneumonia susceptibility (individual with pneumonia versus control individuals) with (top) and without (bottom) individuals diagnosed with CF.
(B) Miami plot of pneumonia severity (pneumonia inpatients versus outpatients) with (top) and without (bottom) individuals diagnosed with CF.

anemia (SCA [MIM: 603903]). The relationship between SCA and pneumonia risk has been previously described in epidemiological studies.[21] Children with SCA are more likely to have pneumonia and influenza (OR = 7.38) and acute respiratory infections (OR = 1.29).[22] Acute chest syndrome (ACS) is a common complication of SCD and can be triggered by pneumonia and vaso-occlusive crises and is the leading cause of death in individuals with SCA.[23] Clinically differentiating between ACS and pneumonia can be difficult, and they often overlap in the EHR.

Our two initial findings provide genetic support of previously observed epidemiological predictors of severe pneumonia risk, namely, the two autosomal recessive disorders CF and SCD. Previous studies have also reported associations between carrier status for *CFTR* and CF-associated symptoms.[20] In contrast, SCA carriers had slightly lower risk of pneumonia and influenza (OR = 0.93) compared to subjects with normal hemoglobin,[22] although pneumonia severity was not studied. We applied Fisher's exact test to investigate the effect of the risk allele in

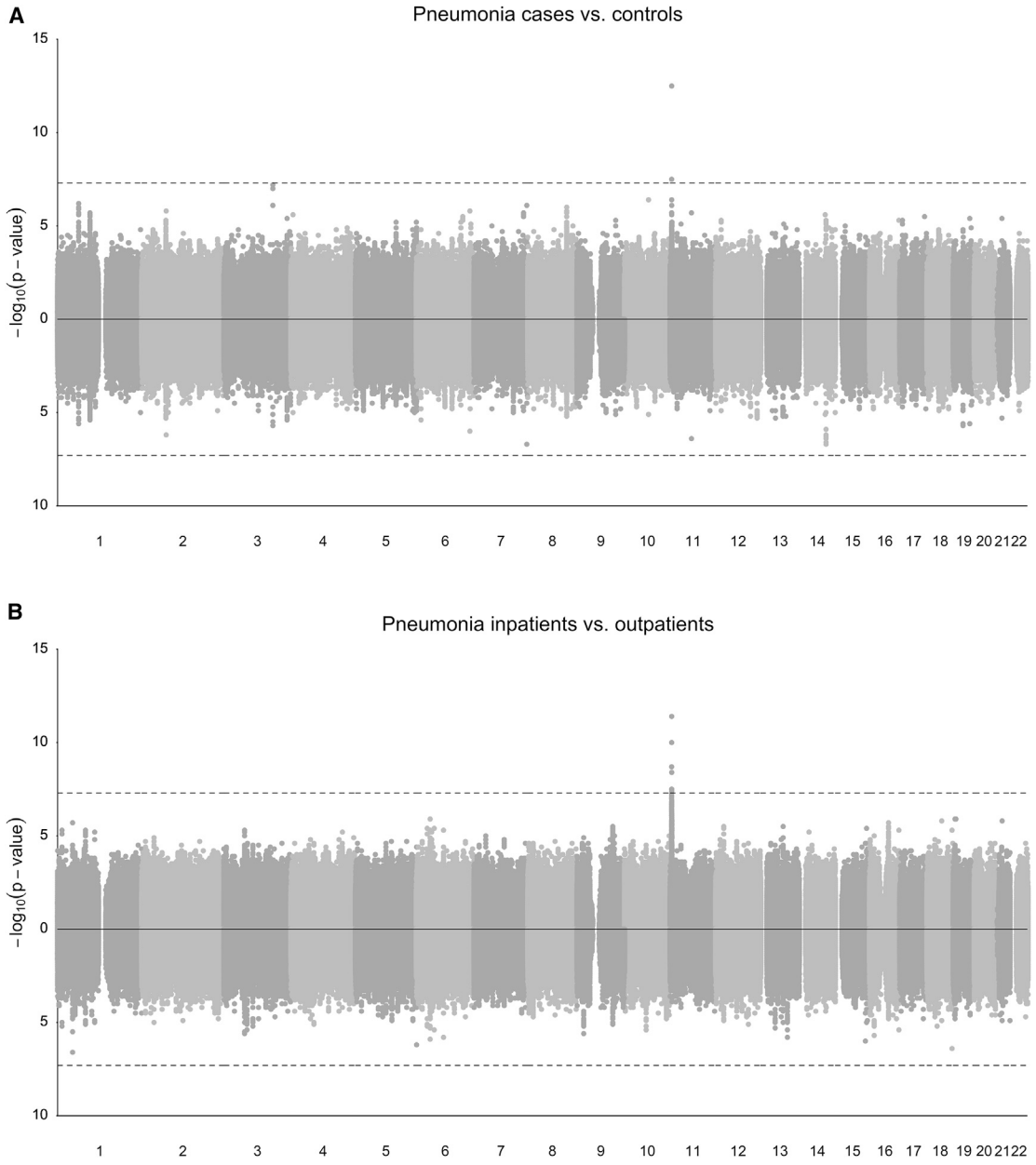

Figure 2. SNP associations with pneumonia in BioVU MEGA^{EX}-genotyped AA

(A) Miami plot of pneumonia susceptibility (individual with pneumonia versus control individuals) with (top) and without (bottom) individuals diagnosed with CF and SCD.

(B) Miami plot of pneumonia severity (pneumonia inpatients versus outpatients) with (top) and without (bottom) individuals diagnosed with CF and SCD.

heterozygous carriers of the risk allele compared to those homozygous for the reference allele. We found heterozygous carriers of rs113827944 (*CFTR*) were at greater risk of developing pneumonia (OR = 1.38, p value = 1.9×10^{-7}), including in sensitivity analyses excluding all individuals with diagnosed CF (OR = 1.17, p value = 0.019; Table 2). Heterozygous carriers of rs334 (*HBB*) were at greater risk of pneumonia-associated hospitalization (OR = 1.55, p value = 0.023), and sensitivity analyses excluding 138 individuals diagnosed with CF and SCD slightly attenuated this effect (OR = 1.49, p value = 0.109), most likely because of a reduction in power. Heterozygote effects observed herein are consistent with an intermediate phenotype or an undetected second variant contributing to compound heterozygosity.[20,24,25]

To identify additional effects masked by the strong effects of *CFTR* and *HBB*, we repeated GWASs after removing individuals diagnosed with CF and SCD. In EA, after removal of individuals with CF, we identified a genome-wide significant signal in *R3HCC1L* (rs10786398, MAF = 30.7%, OR = 1.22, p value = 3.5×10^{-8}, Table 1, and Figures 1 and 3). Replication of effects at this gene were

Table 2. Homozygous and heterozygous effects of CFTR and HBB lead SNPs

Gene	SNP	A1	A2	Pop	Case/control	n	A1A1	A1A2	A2A2	Homozygous[a] OR	Homozygous[a] p value	Heterozygous[b] OR	Heterozygous[b] p value
CFTR	rs113827944	A	G	EA	pneumonia	8,685	50	339	8,296	23.07	3.7E−32	1.38	1.9E−07
					population controls	59,781	15	1,713	58,053				
					pneumonia, individuals with CF removed	8,504	1	281	8,222	2.34	0.413	1.17	0.019
					population controls, individuals with CF removed	59,695	3	1,694	57,998				
					pneumonia inpatient	5,641	46	236	5,359	6.25	1.7E−05	1.26	0.063
					pneumonia outpatient	3,044	4	103	2,937				
					pneumonia inpatient, individuals with CF removed	5,494	1	187	5,306	N/A	N/A	1.09	0.526
					pneumonia outpatient, individuals with CF removed	3,010	0	94	2,916				
HBB	rs334	A	T	AA	pneumonia	1,617	53	139	1,425	7.43	3.9E−22	1.16	0.124
					population controls	13,214	60	1,021	12,133				
					pneumonia, individuals with CF and SCD removed	1,479	1	84	1,394	2.17	0.421	0.93	0.565
					population controls, individuals with CF and SCD removed	12,822	4	782	12,036				
					pneumonia inpatient	984	48	96	840	6.44	2.1E−06	1.55	0.023
					pneumonia outpatient	633	5	43	585				
					pneumonia inpatient, individuals with CF and SCD removed	875	1	57	817	N/A	N/A	1.49	0.109
					pneumonia outpatient, individuals with CF and SCD removed	604	0	27	577				

[a]Homozygous: A1A1 versus others.
[b]Heterozygous: A1A2 versus A2A2.

observed in a distinct haplotype in BioVU AA (rs7086391, MAF = 16.9%, OR = 1.42, p value = 1.14×10^{-3}, multiple-testing-corrected p value = 0.0490, Figure S4) and at a distinct sentinel variant in UK Biobank EA (rs884811, MAF = 44.2%, OR = 1.05, p value = 3.49×10^{-3}, multiple-testing-corrected p value = 0.0198, Table 1). We note that this region of the genome exhibits low linkage disequilibrium: the lead variants from our discovery analyses and from the UK Biobank validation had an R^2 of 0.55 in the BioVU data despite their being less than 3 kb apart (Figure S5). We also observed effects of rs10786398 in the latest release of the COVID-19 Host Genetics Initiative GWAS comparing hospitalized individuals with COVID-19 to non-hospitalized affected individuals (OR = 1.19, MAF = 31.4%, p value = 0.037),[26] indicating relevance of this variant for susceptibility to and severity of COVID-19. A previously published meta-analysis of lung function in UK Biobank and the SpiroMeta consortium provided additional validation of lead variants in *R3HCC1L* (Table S6).[27] *R3HCC1L* encodes a coiled-coil domain-containing protein. Variants in its promoter region have been reported to be associated with inflammatory skin disease[28] and body mass index.[29,30] However, the molecular function of *R3HCC1L* is still unclear and more studies are needed to clarify its role in the pathogenesis of pneumonia.

In AA, after removal of individuals with CF and SCD, an additional SNP near *UQCRFS1* (MIM: 191327) and non-coding RNA LOC105372352 (rs148218440, MAF = 2.4%; OR = 2.21, p value = 3.6×10^{-8}; Figures S6 and S7) was identified in the comparison of AA pneumonia inpatients and age-, sex-, and ancestry-matched control individuals. *UQCRFS1* encodes a Rieske iron-sulfur protein, which is part of the mitochondrial respiratory chain.[31] Rare mutations in *UQCRFS1* have been previously reported as the cause of a mitochondrial disorder (MIM: 618775).[32] Further, *Uqcrfs1* deletion in mice abolishes mitochondrial reactive oxygen species that are required for antigen specific T cell responses, a hallmark of the adaptive immune response.[33] We did not observe replication of rs148218440 in BioVU non-MEGA^{EX} AA (Table 1), however we note that this analysis was underpowered to detect effect (n = 292 affected individuals and 1,145 control individuals, Table S4); UK Biobank did not have enough

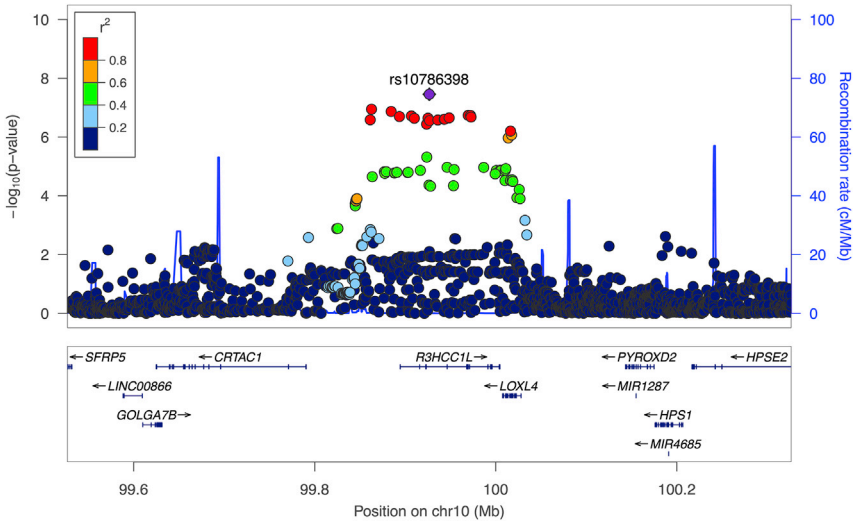

Figure 3. LocusZoom plot of *R3HCC1L* with the genome-wide association study results of pneumonia inpatients versus outpatients from BioVU MEGA^{EX}-genotyped EA excluding CF-affected individuals
Linkage disequilibrium and recombination rate estimates come from 1000 Genomes Project phase 3 European superpopulation reference data.

AA individuals with pneumonia to perform a replication analysis (n = 76 affected individuals and 9,272 control individuals, Table S5).

To further investigate the functional role of associated variation at our lead genes, we utilized PrediXcan to impute the genetically regulated expression (GReX) in MEGA^{EX}-genotyped EA and AA with the tissue-specific models trained by GTEx V8.[34,35] The association tests were conducted with SUGEN and adjusted for three principal components, sex, and age (Table S7 and Figures S8 and S9). The GReX of *R3HCC1L* in lungs is significantly associated with severity of pneumonia in EA ($\beta = 1.140$, p value $= 2.0 \times 10^{-7}$). In addition, the association of the GReX of *HBB* with pneumonia susceptibility in EA offers additional evidence of its role ($\beta = 30.670$, p value $= 7.3 \times 10^{-4}$). GReX of *CFTR* is not well captured in lung tissue,[34] however GReX of this gene in other tissues are significantly associated with pneumonia susceptibility in EA (heart left ventricle, $\beta = 0.464$, p value $= 3.7 \times 10^{10}$).

Moreover, we estimated the heritability of pneumonia susceptibility and severity in MEGA^{EX}-genotyped EA. We used genome-wide complex trait analysis (GCTA)[36] to construct a genetic relatedness matrix and calculated the proportion of phenotypic variance explained by the matrix. The SNP-based heritability (h^2) for susceptibility to pneumonia is estimated to be 0.029 and 0.026 after removing individuals diagnosed with CF (Table S8). As expected due to greater similarity of exposure in affected individuals and control individuals, genetic variation explains a larger portion of pneumonia severity compared to susceptibility ($h^2 = 0.121$; after removing individuals with CF, $h^2 = 0.150$). Our results are similar to previous estimates in UK Biobank (0.075 for self-reported pneumonia, 0.15 for pneumonia inpatient).[37]

The data used for the present discoveries exhibit several key limitations. Extracting accurate phenotypes from EHR data is a well-known challenge[38] and most likely results in some misclassification of cases. Such misclassification most likely reduces power to detect effects but is unlikely to introduce systemic bias that would invalidate results. Additionally, distinguishing viral and bacterial pneumonia via clinical criteria is challenging, and data on positive cultures of blood or pleural fluid for bacterial or positive nasopharyngeal sample or positive serology for viral were not widely available in BioVU. Despite these challenges, we were able to detect four loci linked to pneumonia, including two regions. Further research will be required to differentiate the genetic architecture of specific pneumonia etiologies and to confirm the *UQCRFS1* association, and functional studies will be needed to determine the biological mechanism underlying both signals.

In summary, we leveraged a large-scale biobank to identify genetic variants associated with the susceptibility and severity of pneumonia. Two clinically relevant Mendelian disease genes, *CFTR* and *HBB*, were implicated. These important genetic results indicate that individuals with CF and SCA are at heightened risk for development of and severe outcomes from pneumonia, an effect which may translate to COVID-19 outcomes. Heterozygous carriers of the CF risk allele demonstrated elevated risk of pneumonia susceptibility, and carriers of the SCD risk allele demonstrated elevated risk for pneumonia severity. These findings may have important implications for genetically informed patient care in infectious lung disease. They are also critically important in the context of the COVID-19 pandemic, and future studies will be needed to establish whether these carriers exhibit a silent, heightened risk for poor outcomes from COVID-19 as well. We also identified two additional pneumonia-related genes: *R3HCC1L* and *UQCRFS1*. Although we were most likely underpowered to replicate effects of *UQCRFS1* variation in AA in our data resources, we successfully replicated *R3HCC1L* effects in two independent datasets and validated effects of *R3HCC1L* in independent GWASs of COVID-19[26] and lung function.[27] Characterizing host genome effects on infectious disease susceptibility and severity can offer important insight into the molecular etiology of risk; our findings may help elucidate pathophysiological processes for pneumonia, an important COVID-19 sequelae.

Data and code availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2020.12.010.

Acknowledgments

The authors would like to thank Austin McPhilamy for his help identifying ICD codes relevant to pneumonia phenotyping. This project was supported by CTSA award UL1 TR002243 from the National Center for Advancing Translational Sciences. H.-H.C., D.M.S., L.E.P., R.J.B., C.D.H., and J.E.B. were supported by funding from the National Institute of General Medical Sciences, National Institutes of Health, R01GM133169. D.M.S. was supported by funding from the National Institutes of Health, 5T32GM080178.

Declaration of interests

J.A.B. receives support from Omniox for an ARDS and pneumonia project. The other authors declare no competing interests.

Received: August 11, 2020
Accepted: December 2, 2020
Published: December 13, 2020

Web resources

Coronavirus Disease 2019, https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/index.html
COVID-19 Map, https://coronavirus.jhu.edu/map.html
Online Mendelian Inheritance in Man, https://omim.org

References

1. Guo, Y.R., Cao, Q.D., Hong, Z.S., Tan, Y.Y., Chen, S.D., Jin, H.J., Tan, K.S., Wang, D.Y., and Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. Mil. Med. Res. *7*, 11.
2. Torres, A., Blasi, F., Dartois, N., and Akova, M. (2015). Which individuals are at increased risk of pneumococcal disease and why? Impact of COPD, asthma, smoking, diabetes, and/or chronic heart disease on community-acquired pneumonia and invasive pneumococcal disease. Thorax *70*, 984–989.
3. Hussain, A., Bhowmik, B., and do Vale Moreira, N.C. (2020). COVID-19 and diabetes: Knowledge in progress. Diabetes Res. Clin. Pract. *162*, 108142.
4. Chhiba, K.D., Patel, G.B., Vu, T.H.T., Chen, M.M., Guo, A., Kudlaty, E., Mai, Q., Yeh, C., Muhammad, L.N., Harris, K.E., et al. (2020). Prevalence and characterization of asthma in hospitalized and non-hospitalized patients with COVID-19. J. Allergy Clin. Immunol. *146*, 307–314.e4.
5. Zhao, Q., Meng, M., Kumar, R., Wu, Y., Huang, J., Lian, N., Deng, Y., and Lin, S. (2020). The impact of COPD and smoking history on the severity of COVID-19: A systemic review and meta-analysis. J. Med. Virol. *92*, 1915–1921.
6. Dietz, W., and Santos-Burgoa, C. (2020). Obesity and its Implications for COVID-19 Mortality. Obesity (Silver Spring) *28*, 1005.
7. Rautanen, A., Mills, T.C., Gordon, A.C., Hutton, P., Steffens, M., Nuamah, R., Chiche, J.D., Parks, T., Chapman, S.J., Davenport, E.E., et al.; ESICM/ECCRN GenOSept Investigators (2015). Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. Lancet Respir. Med. *3*, 53–60.
8. Zúñiga, J., Buendía-Roldán, I., Zhao, Y., Jiménez, L., Torres, D., Romo, J., Ramírez, G., Cruz, A., Vargas-Alarcon, G., Sheu, C.C., et al. (2012). Genetic variants associated with severe pneumonia in A/H1N1 influenza infection. Eur. Respir. J. *39*, 604–610.
9. Hayden, L.P., Cho, M.H., McDonald, M.N., Crapo, J.D., Beaty, T.H., Silverman, E.K., Hersh, C.P.; and COPDGene Investigators * (2017). Susceptibility to Childhood Pneumonia: A Genome-Wide Analysis. Am. J. Respir. Cell Mol. Biol. *56*, 20–28.
10. Tian, C., Hromatka, B.S., Kiefer, A.K., Eriksson, N., Noble, S.M., Tung, J.Y., and Hinds, D.A. (2017). Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. Nat. Commun. *8*, 599.
11. Campos, A., Kho, P.F., Vazquez-Prada, K.X., Garcia-Marin, L.M., Martin, N.G., Cuellar-Partida, G., and Renteria, M.E. (2020). Genetic susceptibility to pneumonia: A GWAS meta-analysis between UK Biobank and FinnGen. medRxiv. https://doi.org/10.1101/2020.06.22.20103556.
12. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat. Genet. *48*, 1284–1287.
13. Lin, D.Y., Tao, R., Kalsbeek, W.D., Zeng, D., Gonzalez, F., 2nd, Fernández-Rhodes, L., Graff, M., Koch, G.G., North, K.E., and Heiss, G. (2014). Genetic association analysis under complex survey sampling: the Hispanic Community Health Study/Study of Latinos. Am. J. Hum. Genet. *95*, 675–688.
14. Staples, J., Qiao, D., Cho, M.H., Silverman, E.K., Nickerson, D.A., Below, J.E.; and University of Washington Center for Mendelian Genomics (2014). PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. Am. J. Hum. Genet. *95*, 553–564.
15. Li, H., Glusman, G., Hu, H., Shankaracharya, Caballero, J., Hubley, R., Witherspoon, D., Guthery, S.L., Mauldin, D.E., Jorde, L.B., et al. (2014). Relationship estimation from whole-genome sequence data. PLoS Genet. *10*, e1004144.
16. Ratjen, F., and Döring, G. (2003). Cystic fibrosis. Lancet *361*, 681–689.
17. Rowe, S.M., Miller, S., and Sorscher, E.J. (2005). Cystic fibrosis. N. Engl. J. Med. *352*, 1992–2001.
18. Davis, P.B. (2006). Cystic fibrosis since 1938. Am. J. Respir. Crit. Care Med. *173*, 475–482.
19. Sanders, D.B., and Fink, A.K. (2016). Background and Epidemiology. Pediatr. Clin. North Am. *63*, 567–584.
20. Miller, A.C., Comellas, A.P., Hornick, D.B., Stoltz, D.A., Cavanaugh, J.E., Gerke, A.K., Welsh, M.J., Zabner, J., and Polgreen, P.M. (2020). Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. Proc. Natl. Acad. Sci. USA *117*, 1621–1627.
21. Poehling, K.A., Light, L.S., Rhodes, M., Snively, B.M., Halasa, N.B., Mitchel, E., Schaffner, W., Craig, A.S., and Griffin, M.R.

(2010). Sickle cell trait, hemoglobin C trait, and invasive pneumococcal disease. Epidemiology *21*, 340–346.
22. Reeves, S.L., Jary, H.K., Gondhi, J.P., Kleyn, M., and Dombkowski, K.J. (2019). Health outcomes and services in children with sickle cell trait, sickle cell anemia, and normal hemoglobin. Blood Adv. *3*, 1574–1580.
23. Jain, S., Bakshi, N., and Krishnamurti, L. (2017). Acute Chest Syndrome in Children with Sickle Cell Disease. Pediatr. Allergy Immunol. Pulmonol. *30*, 191–201.
24. Agarwal, S., Tamhankar, P.M., Kumar, R., and Dalal, A. (2010). Clinical and haematological features in a compound heterozygote (HBB:c.92 + 5G > C/HBB:c.93-2A > C) case of thalassaemia major. Int. J. Lab. Hematol. *32*, 369–372.
25. Raju, S.V., Tate, J.H., Peacock, S.K., Fang, P., Oster, R.A., Dransfield, M.T., and Rowe, S.M. (2014). Impact of heterozygote CFTR mutations in COPD patients with chronic bronchitis. Respir. Res. *15*, 18.
26. Covid-19 Host Genetics Initiative (2020). The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur. J. Hum. Genet. *28*, 715–718.
27. Shrine, N., Guyatt, A.L., Erzurumluoglu, A.M., Jackson, V.E., Hobbs, B.D., Melbourne, C.A., Batini, C., Fawcett, K.A., Song, K., Sakornsakolpat, P., et al.; Understanding Society Scientific Group (2019). New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. Nat. Genet. *51*, 481–493.
28. Baurecht, H., Hotze, M., Brand, S., Büning, C., Cormican, P., Corvin, A., Ellinghaus, D., Ellinghaus, E., Esparza-Gordillo, J., Fölster-Holst, R., et al.; Psoriasis Association Genetics Extension (2015). Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. Am. J. Hum. Genet. *96*, 104–120.
29. Hoffmann, T.J., Choquet, H., Yin, J., Banda, Y., Kvale, M.N., Glymour, M., Schaefer, C., Risch, N., and Jorgenson, E. (2018). A Large Multiethnic Genome-Wide Association Study of Adult Body Mass Index Identifies Novel Loci. Genetics *210*, 499–515.
30. Zhu, Z., Guo, Y., Shi, H., Liu, C.L., Panganiban, R.A., Chung, W., O'Connor, L.J., Himes, B.E., Gazal, S., Hasegawa, K., et al. (2020). Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. J. Allergy Clin. Immunol. *145*, 537–549.
31. Sánchez, E., Lobo, T., Fox, J.L., Zeviani, M., Winge, D.R., and Fernández-Vizarra, E. (2013). LYRM7/MZM1L is a UQCRFS1 chaperone involved in the last steps of mitochondrial Complex III assembly in human cells. Biochim. Biophys. Acta *1827*, 285–293.
32. Gusic, M., Schottmann, G., Feichtinger, R.G., Du, C., Scholz, C., Wagner, M., Mayr, J.A., Lee, C.Y., Yépez, V.A., Lorenz, N., et al. (2020). Bi-Allelic UQCRFS1 Variants Are Associated with Mitochondrial Complex III Deficiency, Cardiomyopathy, and Alopecia Totalis. Am. J. Hum. Genet. *106*, 102–111.
33. Sena, L.A., Li, S., Jairaman, A., Prakriya, M., Ezponda, T., Hildeman, D.A., Wang, C.R., Schumacker, P.T., Licht, J.D., Perlman, H., et al. (2013). Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling. Immunity *38*, 225–236.
34. Barbeira, A.N., Bonazzola, R., Gamazon, E.R., Liang, Y., Park, Y., Kim-Hellmuth, S., Wang, G., Jiang, Z., Zhou, D., Hormozdiari, F., et al. (2019). Widespread dose-dependent effects of RNA expression and splicing on complex diseases and traits. bioRxiv. https://doi.org/10.1101/814350.
35. Aguet, F., Barbeira, A.N., Bonazzola, R., Brown, A., Castel, S.E., Jo, B., Kasela, S., Kim-Hellmuth, S., Liang, Y., Oliva, M., et al. (2019). The GTEx Consortium atlas of genetic regulatory effects across human tissues. bioRxiv. https://doi.org/10.1101/787903.
36. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. *88*, 76–82.
37. Ge, T., Chen, C.Y., Neale, B.M., Sabuncu, M.R., and Smoller, J.W. (2018). Correction: Phenome-wide heritability analysis of the UK Biobank. PLoS Genet. *14*, e1007228.
38. DeLisle, S., Kim, B., Deepak, J., Siddiqui, T., Gundlapalli, A., Samore, M., and D'Avolio, L. (2013). Using the electronic medical record to identify community-acquired pneumonia: toward a replicable automated strategy. PLoS ONE *8*, e70944.