doi:10.20944/preprints202005.0273.v2

The frequency of combined IFITM3 haplotype involving the reference alleles of both rs12252 and rs34481144 is in line with COVID-19 standardized mortality ratio of ethnic groups in England

Dimitris Nikoloudis[1], Dimitrios Kountouras[1] and Asimina Hiona[1]

[1] Center for Preventive Medicine & Longevity, Bioiatriki Healthcare Group, 11525 Athens, Greece

Abstract

Evidence was brought forward in England and the USA that Black, Asian, Latino and Minority Ethnic people exhibit higher mortality risk from COVID-19 than White people. While socioeconomic factors were suggested to contribute to this trend, they arguably do not explain the range of the differences observed, allowing for possible genetic implications. Almost concurrently, the analysis of a cohort in Chinese COVID-19 patients proposed an association between the severity of the disease and the presence of the minor allele of rs12252 of the Interferon-induced transmembrane protein 3 (IFITM3) gene. This SNP, together with rs34481144, are the two most studied polymorphisms of IFITM3 and have been associated in the past with increased severity in Influenza, Dengue, Ebola, and HIV viruses. IFITM3 is an immune effector protein that is pivotal for the restriction of viral replication, but also for the regulation of cytokine production. Following up to these two developments in the ongoing SARS-CoV-2 pandemic, the present study investigates a possible association between the differences in mortality of ethnic groups in England and the combined haplotypes of rs12252 and rs34481144. The respective allele frequencies were collected for 26 populations from 1000 Genomes Project and subgroups were pooled wherever possible to create correspondences with ethnic groups in England. A significant correlation (r=0.9687, p= 0.0003) was observed between the reported Standardized Mortality Ratios and the frequency of the combined haplotype of both reference alleles, suggesting that the combination of reference alleles of the specific SNPs may be implicated in more severe outcomes of COVID-19. This study calls for further focus on the role of IFITM3 variants in the mechanism of cellular invasion of SARS-CoV-2, their impact in COVID-19 severity and their possible implications in vaccination efficacy.

Introduction

Emerging scientific evidence from international **(**Kirby, 2020**)** and UK **(**Aldridge et al.**,** 2020**)** COVID-19 patient reports and death records, indicate a disproportionate effect of the novel coronavirus on ethnic minorities. According to CDC (CDC, 2020), Black, Asian and Minority Ethnic (BAME) people are at higher risk of death from COVID-19. Importantly, an Indirect Standardization of NHS mortality data in England **(**Aldridge et al.**,** 2020**)**, revealed that the adjusted for age and region Standardized Mortality Ratios (SMRs), were highest in Black African, Black Caribbean, Pakistani, Bangladeshi, and Indian minority ethnic groups. On the contrary, White Irish and White British ethnic groups exhibited a significantly lower risk of death. Similarly, in the USA (Garg et al., 2020), preliminary data compiled from hospitals in 14 US states, confirmed the UK study outcomes, showing that African Americans are also disproportionately affected by COVID-19. Specifically, African Americans represented 33% of COVID-19 hospitalizations, despite only making up 18% of the total population studied. In a subsequent analysis, among COVID-19 deaths in New York City, for which race and ethnicity data were available, death rates from COVID-19 among black or African Americans and Hispanic or Latinos were substantially higher than that of white or Asian people (Garg et al., 2020).

Several reasons have been proposed to explain these ethnic discrepancies in COVID-19 mortality risk arising from these preliminary studies. Chronic pre-existing conditions, such as Cardio Vascular Diseases (CVD), diabetes, hypertension, obesity, etc. are more common in minorities compared to Caucasian populations and have all been associated with adverse outcomes in COVID-19 (Centers for Disease Control and Prevention, 2020; Kirby, 2020**).** However, race disparities in those diseases are not large enough to fully explain the COVID-19 death disparity (Aldridge et al., 2020)**.** Factors such as housing and living conditions, use of public transportation, lack of regular access to primary health, and occupation-related differences that prohibit the work from home, or require more frequent and/or close social contact, may have all played an important role in producing disproportionate death rates among BAME groups (Aldridge et al., 2020; Kirby, 2020; Niedzwiedz et al., 2020; Khunti et al., 2020). Nevertheless, it is suggested that inequalities in socioeconomic status parameters do not seem to adequately explain the range of differences, and in some instances, the extreme variations observed among ethnic minorities in mortality rates from COVID-19 infection (Kirby, 2020)**.**

As the importance of genetic polymorphisms (SNPs) in the modulation of individual susceptibility to, and severity of, infectious diseases has been well established (Chapman et al., 2012; Zhao et al, 2018), we turned our focus into two very highly studied polymorphisms of the interferon-induced transmembrane protein 3 (IFITM3) gene: rs12252 and rs34481144. IFITM3 encodes an immune effector protein that is pivotal for restriction of viral replication (Brass et al., 2009) of many enveloped RNA viruses including HIV-1, influenza A virus (IAV), Ebola and Dengue virus **(**Brass et al., 2009; Feeley et al., 2011; Huang et al., 2011; Everitt et al., 2012; Compton et al., 2014)**.** IFITM3 has been demonstrated to affect severity of infection and improve the host cellular defenses against viruses **(**Brass et al., 2009; Everitt et al., 2012; Compton et al., 2014)**.** Interestingly, IFITM3 has also been shown to act as a regulator of antiviral immunity that controls cytokine production to restrict viral pathogenesis, in CMV (Stacey et al., 2017) and Sendai virus (Jiang et al., 2017). This finding is particularly important since cytokine storm in influenza can lead to a rapid progression of the infection in humans (Wang et al., 2014) and the same observation is also apparent in COVID-19 severe and deadly cases (Giamarellos-Bourboulis et al., 2020; Blanco-Melo et al., 2020). Moreover, IFITM3 was found to be explicitly upregulated in SARS-CoV-2 infected cells (Blanco-Melo et al., 2020; Hachim et al., 2020; He *et al.*, 2020).

The minor allele of rs12252 (C in minus, or G in plus strand orientation) has been associated with rapid progression of acute HIV infection (Zhang et al., 2015), with the severity of influenza (Zhang et al., 2013) and recently with COVID-19 severity (Zhang et al, 2020)**.** The minor allele of rs34481144 (A in minus, or T in plus strand orientation) was previously found to be correlated with increased severity of IAV infection (Allen et al., 2017). Moreover, the minor allele of rs34481144 is also associated with enhanced methylation on the IFITM3 promoter of CD8+ T cells, and general transcriptional repression of the broader locus surrounding IFITM3, which includes several genes known to be involved in host responses to viral infection (Wellington et al., 2019).

SARS-CoV-2 uses primarily the ACE2 receptor as main point of entry and the serine protease TMPRSS2 for spike (S) protein priming (Hoffmann et al., 2020). Interestingly, it was recently shown that TMPRSS2 is specifically allowing evasion of IFITM3 restriction for bat SARS-Like WIV1 coronavirus (Zheng et al., 2020), opening the possibility for a similar mechanism in the case of SARS-CoV-2. Further potential involvement of IFITM3 in COVID-19 outcome was revealed in the context of syncytial pneumocytes in severe cases with extensive lung damage, where it was suggested that the cellular location of IFITMs 1-3 could be playing a role in syncytia formation (Buchrieser et al., 2020). Indeed, the accumulation of many direct and indirect layers of evidence linking IFITM3 with COVID-19 outcome, has also led to explicit calls for further investigation of the role of this highly relevant first-line of cellular defense protein (Zhao, 2020). Following up to the analysis of COVID-19 NHS mortality data in BAME groups **(**Aldridge et al.**,** 2020**)**, the purpose of the present study was to investigate a possible association between the stand-alone and combined frequencies of the alleles of the IFITM3 gene variants rs12252 and rs34481144, with COVID-19 standardized mortality ratio of ethnic groups in England.

Methods

The rs12252 (A>G) and rs34481144 (C>T) allele and haplotype frequencies were collected for all available 1000 Genomes Project ancestral populations, from LDlink, specifically 5 major groups, *i.e.* African (AFR), Ad Mixed American (AMR), European (EUR), East Asian (EAS) and South Asian (SAS), comprising 26 subgroups in total (Machiela et al., 2015) (Table 1). The plus orientation for the reference and minor alleles was retained throughout this analysis, for better data handling and in compliance with dbSNP. Combined rs12252_rs34481144 haplotypes were defined as A_C (H1), G_C (H2), A_T (H3), while G_T haplotype was not represented. Rankings were examined by sorting all populations by individual reference allele (rs12252:A, rs34481144:C) and by combined haplotype frequency ratios (rs12252_rs34481144: h1_ratio=A_C/(A_T+G_C), h2_ratio=G_C/(A_C+A_T), h3_ratio=A_T/(G_C+A_C)) (Figure 1), and subsequently compared visually to the reported Standardized Mortality Ratios (SMR) of ethnic groups in England (Figure 2). Additionally, an attempt was made to correlate directly the two rankings, *i.e.* SMR and IFITM3 haplotype frequencies by specific reported ethnic subgroup. UK demographics sources were therefore consulted (Office for National Statistics, UK, 2011; Chanda & Ghosh, 2012) in order to pool, wherever possible, the ancestral subgroups to the reported ethnic groups in England. With all reservations tied to the inevitable discrepancies of this type of simplified socio-genetic correspondences, the following pools were formed: [AFR-YRI, AFR-LWK, AFR-GWD, AFR-MSL, AFR-ESN]>"African", [SAS-STU, SAS-GIH, SAS-

PJL]>"Indian", [EAS-CDX, EAS-CHS, EAS-CHB]>"Chinese", [EUR-CEU, EUR-IBS, EUR-TSI, EUR-FIN]>"White Other" (see Table 1 for full subgroup descriptions, Table 2 for pooled subgroups and correspondences to ethnic minorities). A pool for the reported Pakistani group failed to form from ancestral populations, as the Punjabi (SAS-PJL), being the only related subgroup, account roughly for just 45% of Pakistan's demographics, while in London the community includes comparable numbers of Punjabis, Pathans and Kashmiris, with small communities of Sindhis and Balochis (Department for Communities and Local Government, UK, 2009). Moreover, the Punjabi form also a considerable part of Indians' pool (at least 40% of Delhi's total population), therefore a single-ended direct correspondence between Punjabi and British Pakistani was not warranted in this case. Indian Telugu (SAS-ITU) were not included in Indians' pool, as no demographic report was suggestive of comparable numbers to the other 3 included subgroups, for people of Indian origin in England. The same rationale applied for the non-inclusion of AFR-ASW (Americans of African Ancestry in SW USA) in the African pool. The haplotype frequencies were simply averaged within pooled groups, and both the ratios and SMR were normalized to the White British result (represented uniquely by EUR-GBR subgroup) (Table 2).

Results and Discussion

The ranking that visually appeared in line with the reported SMR (Figures 2a, 2b) was produced by the A_C/(A_T+G_C) ratio (h1), specifically (a) African groups, followed by (b) South Asian, followed by (c) White Non-British, (d) Chinese, and finally (e) White British. It is also noteworthy, that EUR-IBS (Iberian Population in Spain) and EUR-TSI (Toscani in Italia), representative of two countries that suffered higher death rates than other European countries, share the highest h1 ratio between all European subgroups. The fact that these two rankings coincided is significant and can be viewed as an event of permutation (5 items) with a 1/120 chance to occur randomly, or a probability of 0.008. Subsequently, in order to assess the potential strength of the theorized correlation, the rankings of pooled h1 ratios and SMR, per group, were traced one versus the other (Figure 3) with Pearson correlation $r=0.9687$, $p=3*10^{-4}$ (Figure 4).

This calculated level of correlation appears to be remarkable, considering the possible discrepancies in the pooling of the available ancestral groups, but also the expected multi-parametric causes of the observed COVID-19 SMR in England's ethnic groups (as previously described, potentially involving prior health status, income level, household density, behavioral biases, questionable attribution of death to COVID-19, etc.). On one hand, the alignment of ethnic subgroup rankings between SMR and unpooled h1 ratio, was less than 1% probable to occur randomly, or to simulate by chance the combined effects of the theorized socioeconomic factors. On the other hand, the level of correlation (>3.5σ) between SMR and pooled h1 ratio, confirms the previous alignment and appears strong enough to suggest a possible causal link, albeit in this case, the pooling process may have introduced further discrepancies. The potential introduction of pooling discrepancies is expected and is impossible to quantify given the available data. However, the purpose of the pooled analysis was mainly to reinforce the primary association between IFITM3 and COVID-19 severity rather than to (indirectly) infer causality. Considering the parametric uncertainty of the pooled analysis it is probably inadequate to suggest a causal link, based on the strength of the observed correlation alone. Regardless, this set of results taken together constitutes a clear and valid starting point for designing further investigations and appears as one more piece of evidence that associates IFITM3 with COVID-19 outcomes; the main point of this analysis being that the functional effects of the two examined SNPs should preferably henceforth be studied under a combined haplotype and not separately, as performed so far for SARS-CoV-2 and for a great variety of other viruses.

Before proceeding to the discussion of the implications of the above conclusion, it is interesting to also view the present observations in the context of data from the USA. A preliminary analysis of death rates from COVID-19 in New York City shows 92.3 deaths per 100,000 population among black or African American people, followed by Hispanic or Latino people (74.3), then by white (45.2) or Asian (34.5) people (Kirby, 2020). The same trend was clearly displayed in the initial ranking by h1 ratio (Figure 2a), with American populations (AMR) occupying the middle of the chart, between African and European / Asian populations. Moreover, the lower death rate of Asian people compared to white people in the above New York data, an inversion of the respective numbers in England, could be in line with smaller proportion of White British in the reported "white phenotype" in the USA (1.5M in USA & Canada) and bigger proportion of Japanese people (>1.5M) in the reported "Asian phenotype" in the USA. Both White British (EUR-GBR) and Japanese (EAS-JPT) ethnic subgroups have among the lowest h1 ratio between all subgroups.

The fact that the proposed risk haplotype (A_C) involves the reference alleles of both studied SNPs, appears as counterintuitive. Especially so, since it has been suggested after analysis of a Chinese cohort, that it is the minor allele rs12252:G that is linked to increased COVID-19 severity (Zhang et al., 2020). In fact the minor allele rs12252:G was linked to worse outcome in almost every related study, such as increased influenza severity (Zhang et al., 2013), or more rapid HIV progression (Zhang et al., 2015), albeit always observed on Chinese patients and not in European or American cohorts. This is noteworthy, as minor allele rs12252:G is found frequently in Chinese populations (roughly 50%), but is on the contrary more rare in European populations (1-8%), or infrequent in South Asian (10-18%), or African groups (21%-33%). Interestingly, an inversed trend is observed in the other half of the discussed A_C haplotype, with namely rs34481144:T being rare in Chinese populations (1-2%), rare or infrequent in African groups (2-14%), but fairly frequent in European groups (38-56%). Rs34481144:T was found to correlate strongly with increased influenza severity in 3 independent cohorts (Allen et al., 2017). These 3 independent cohorts, however, did not confirm the link between rs12252:G and increased influenza severity, as was suggested in Chinese cohorts. To add to the controversy of the possible antiviral effects linked to rs12252, a detailed study on 293T cells of the putative truncated variant Δ1-21 that is theorized to result from the rs12252:G mutant, showed increased potential to restrict HIV replication and therefore an advantage compared to the complete IFITM3 protein carrying the reference allele (Compton et al., 2016). However, this truncated version was not observed later in the blood of IAV or HIV patients (Randolph et al., 2017; Makvandi-Nejad et al., 2018), while rs12252:G was, inversely, found to enhance HIV-1 infection in Chinese patients (Zhang et al., 2015). The reports for the functional role and consequences of minor rs12252:G allele are therefore conflicting. Although it is shown that Δ1-21 variant redistributes the protein to the plasma membrane, by prohibiting the phosphorylation of residue Y20 that produces a signal for endocytosis (Jia et al., 2012), a functional link between Δ1-21 and rs12252:G has yet to be established. In the other examined SNP, the minor rs34481144:T allele is currently believed to favor the binding of transcriptional repressor CTCF, also known as CCCTC-binding factor, at IFITM3 promoter, seemingly leading to an inactive IFITM3 profile (Allen et al., 2017). However, the exact functional effect of rs34481144:T is still not well understood.

As part of the IFITM family of proteins, one of the evolutionary ancient first lines of antiviral cellular defenses, the localization in endosomal or lysosomal membrane, or at the surface, *e.g.* of CD4$^+$ T cells, and the exact antiviral mechanism of IFITM3, is regulated by many different post translational modifications, mainly palmitoylation, ubiquitination and phosphorylation. It is shown that genotypic variants of IFITM3 play a role in diversifying a host's potential antiviral repertoire, in conjunction with selective post translation modifications, and therefore should not be considered *de facto* as risk factors but rather as trade-offs in antiviral specificity (Compton et al., 2016). Similarly, in the case of SARS-CoV-2, the observed strong correlation of reference haplotype H1 (A_C) with increased morbidity in ethnic groups in England, could be pointing at a specific antiviral advantage conferred by the presence of each minor allele. However, since both minor alleles are not observed together (haplotype G_T is not represented), it is harder to conceive an independently equivalent beneficial effect by each distinct minor allele in the mixed reference/minor haplotypes H2 (G_C, here minor allele> rs12252:G) and H3 (A_T, here minor allele> rs34481144:T). Instead, it is more plausible to consider an effective hijacking of IFITM3 by SARS-CoV-2 in order to infect the cell, or to replicate, or to spread, or involving more than one of these phases. Indeed, there are known examples of similar hijacking, for example by the coronavirus that causes the common cold, HCoV-OC43 (Zhao et al., 2014), or by human cytomegalovirus (HCMV) (Xie et al., 2015). More specifically for HCoV-OC43, it was shown that all three types of interferons, IFN-α, IFN-γ, and IFN-λ, actually enhance HCoV-OC43 infection, while IFITM3 possibly promotes the low-pH–activated membrane fusion between the viral envelope and endosomal membranes. In contrast, human cytomegalovirus hijacks BST-2/tetherin to promote its entry into host cells and co-opts viperin to facilitate its replication, with IFITM3 facilitating the formation of the virion assembly compartment, but the virus is otherwise less sensitive to IFNs. In the case of SARS-CoV-2, it is therefore not inconceivable that if there is in fact a pro-infection role of IFITM3, that the virus could have evolved to exploit the most abundant haplotype A_C (59% abundance across all populations). The different effect between the H1 haplotype and H2/H3 haplotypes most probably involves the cellular distribution of IFITM3, which is mainly controlled by post-translational modifications, which in turn may be influenced by key polymorphisms such as the two examined here. Of great relevance, in this context, is the recent finding that plasma membrane localization of IFITM3 enhances SARS-CoV-2 infection, while endocytosis of IFITM3 effectively restricts the virus (Shi et al., 2020). The same study confirms an even greater enhancement of SARS-CoV-2 in bypassing IFITM3 defense via TMPRSS2 activation of plasma membrane fusion, and reports compatibility with HCoV-OC43 mode of enhanced infection.

It remains, that the recent suggestion that rs12252:G is the risk allele in a n=80 COVID-19 cohort with Chinese patients (Zhang et al., 2020), is still challenging our conclusions, claiming the inverse effect. Considering that the cohort took place at Beijing You'an Hospital, if it is safe to assume that patients belonged to EAS-CHB group (Han Chinese in Beijing), the subgroup with the highest frequency in rs12252:G (54%), then an alternative reading of the result may be possible. With 28/80 patients hospitalized with pneumonia being homozygotes rs12252:GG, 37/80 being heterozygotes rs12252:AG and 15/80 being homozygous rs12252:AA, this results to 58% (93/160) abundance for the G allele and 42% (67/160) abundance for the C allele. As the prior probability for the G allele was as high as 54%, the previous result seems inconclusive (*i.e.* 58% observed vs 54% expected for rs12252:G) and therefore the suggestion that rs12252:G alone is a COVID-19 severity risk allele seems unfounded in this case. The same conclusion is reached, with whichever possible mix of the 3 available Chinese subgroups from 1000 Genomes Projects (EAS-CHB, EAS-CHS, EAS-CDX), as they all show high rs12252:G frequencies (0.47-0.54), surpassed only by the Japanese subgroup (EAS -JPT, 0.64).

The role of the examined IFITM3 variants in the severity of COVID-19 should be elucidated, by in vitro investigation of the effect of H1 (A_C), vs H2 (G_C), vs H3 (A_T) rs12252_rs34481144 haplotypes in SARS-CoV-2 infectivity and viral spread, including the infected cells' tendency for syncytia formation. Related in silico investigations should include the examined combined haplotype in their analysis and consider non-additive interactions. It is notable how various ongoing GWAS studies (*e.g.* The COVID-19 Host Genetics Initiative, 2020; 23andMe, 2020). did not confirm in their meta-analysis some other independently established genetic variant effects, which are otherwise broadly accepted by the scientific community as influencing COVID-19 severity, such as the ABO blood group (Li et al., 2020; Wu et al., 2020), or APOE e4 genotype (Kuo et al., 2020). If functional differences between the examined IFITM3 haplotypes are shown to produce distinct profiles of COVID-19 progression in severe patients, then an improved understanding of the underlying mechanisms may allow more adequate or

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2020 doi:10.20944/preprints202005.0273.v2

personalized treatment protocols. Last but not least important, is the need for investigation of the examined variant implications in raising an effective immune response after future vaccination against SARS-CoV-2, as it was recently demonstrated that homozygote rs12252:GG, specifically, is reducing the level of antibody response after influenza vaccination (Lei et al., 2020). It could be important to verify whether this also stands for the upcoming SARS-CoV-2 vaccines, or whether the effect could be the opposite; namely with the reference alleles to be associated with reduced antibody response, as was suggested throughout this analysis regarding COVID-19 severity. In conclusion, this study presented one more piece of evidence associating IFTM3 variants with the severity of COVID-19, suggested that the two most highly studied IFITM3 polymorphisms should be considered as a combined haplotype, and is calling for further research focus on this important first line of cellular antiviral defense protein, in the context of the ongoing SRAS-CoV-2 pandemic.

Major group	Subgroup	rs12252:A	rs12252:G	rs34481144:C	rs34481144:T	A_T	A_C	G_C
EUR	CEU (Utah residents from north and west Europe)	0.955	0.045	0.505	0.495	0.49	0.46	0.05
EUR	TSI (Toscani in Italia)	0.967	0.033	0.617	0.383	0.38	0.58	0.03
EUR	FIN (Finnish in Finland)	0.919	0.081	0.505	0.495	0.49	0.42	0.08
EUR	GBR (British in England & Scotland)	0.989	0.011	0.440	0.560	0.56	0.43	0.01
EUR	IBS (Iberian Population in Spain)	0.967	0.033	0.603	0.397	0.40	0.57	0.03
EAS	CHB (Han Chinese in Beijing)	0.461	0.539	0.990	0.010	0.00	0.45	0.54
EAS	JPT (Japanese in Tokyo)	0.361	0.639	1.000	0.000	0.00	0.36	0.64
EAS	CHS (Southern Han Chinese)	0.495	0.505	1.000	0.000	0.00	0.50	0.50
EAS	CDX (Chinese Dai in Xishuangbanna	0.521	0.478	0.995	0.005	0.00	0.52	0.48
EAS	KHV (Kinh in Ho Tsi Minh ciry, Vietnam)	0.530	0.470	0.985	0.015	0.02	0.52	0.47
AMR	MXL (Mexican ancestry from Los Angeles)	0.781	0.219	0.820	0.180	0.18	0.60	0.22
AMR	PUR (Puerto Ricans from Puerto Rico)	0.889	0.111	0.702	0.298	0.30	0.59	0.11
AMR	CLM (Colombians from Medelin, Colombia)	0.925	0.074	0.681	0.319	0.32	0.61	0.07
AMR	PEL (Peruvians from Lima Peru)	0.659	0.341	0.900	0.100	0.10	0.56	0.34
SAS	GIH (Gujarati Indian from Houston Texas)	0.835	0.165	0.806	0.194	0.19	0.64	0.17
SAS	PJL (Punjabi from Lahore, Pakistan)	0.823	0.177	0.812	0.188	0.19	0.64	0.18
SAS	BEB (Bengali from Bangladesh)	0.837	0.163	0.831	0.169	0.17	0.67	0.16
SAS	STU (Sri Lankan Tamil from the UK)	0.868	0.132	0.735	0.265	0.26	0.60	0.13
SAS	ITU (Indian Telugu from the UK)	0.897	0.103	0.789	0.211	0.21	0.69	0.10
AFR	ASW (Americans of African Ancestry in SW USA)	0.697	0.303	0.861	0.139	0.14	0.56	0.30
AFR	ACB (African Carribeans in Barbados)	0.776	0.224	0.906	0.094	0.09	0.68	0.22
AFR	ESN (Esan in Nigera)	0.788	0.212	0.990	0.010	0.01	0.78	0.21
AFR	MSL (Mende in Sierra Leone)	0.753	0.247	0.982	0.018	0.02	0.74	0.25
AFR	GWD (Gambian in Western Gambia)	0.783	0.217	0.969	0.031	0.03	0.75	0.22
AFR	LWK (Luhya in Webuye Kenya)	0.702	0.298	0.985	0.015	0.02	0.69	0.30
AFR	YRI (Yoruba in Ibadan, Nigera)	0.667	0.333	0.968	0.032	0.03	0.63	0.33

Table 1. Detailed allele and haplotype frequencies per ethnic subgroup derived from 1000 Genomes Project, for rs12252 and rs34481144.

Ethnic group	SMR -White British Normalized	A_C / (A_T + G_C) - White British Normalized	1000 Genomes Populations	rs12252:A	rs12252:G	rs34481144:C	rs34481144:T	A_T	A_C	G_C	rs12252_rs34481144 : A_C / (A_T + G_C)
African	3.68	3.38	*AFR-YRI, AFR-LWK, AFR-GWD, AFR-MSL, AFR-ESN*	0.74	0.26	0.98	0.02	0.02	0.72	0.26	2.54
Bangladeshi	2.74	2.69	*SAS-BEB*	0.84	0.16	0.83	0.17	0.17	0.67	0.16	2.02
Caribbean	2.51	2.86	*AFR-ACB*	0.78	0.22	0.91	0.09	0.09	0.68	0.22	2.15
Indian	1.93	2.23	*SAS-STU, SAS-GIH, SAS-PJL*	0.84	0.16	0.78	0.22	0.22	0.63	0.16	1.68
White other	1.53	1.38	*EUR-CEU, EUR-IBS, EUR-TSI, EUR-FIN*	0.95	0.05	0.56	0.44	0.44	0.51	0.05	1.04
Chinese	1.30	1.28	*EAS-CDX, EAS-CHS, EAS-CHB*	0.49	0.51	1.00	0.01	0.00	0.49	0.51	0.96
White British	1.00	1.00	*EUR-GBR*	0.99	0.01	0.44	0.56	0.56	0.43	0.01	0.75
Average	2.10	2.12		0.80	0.20	0.78	0.22	0.21	0.59	0.20	1.59

Table 2. Pools of ethnic subgroups were formed to emulate the ethnic populations that are reported in England. Both the h1 ratio (A_C/A_T+G_C) and the Standardized Mortality Ratios (SMR) were normalized by the corresponding numbers of White British, to allow a direct comparison.

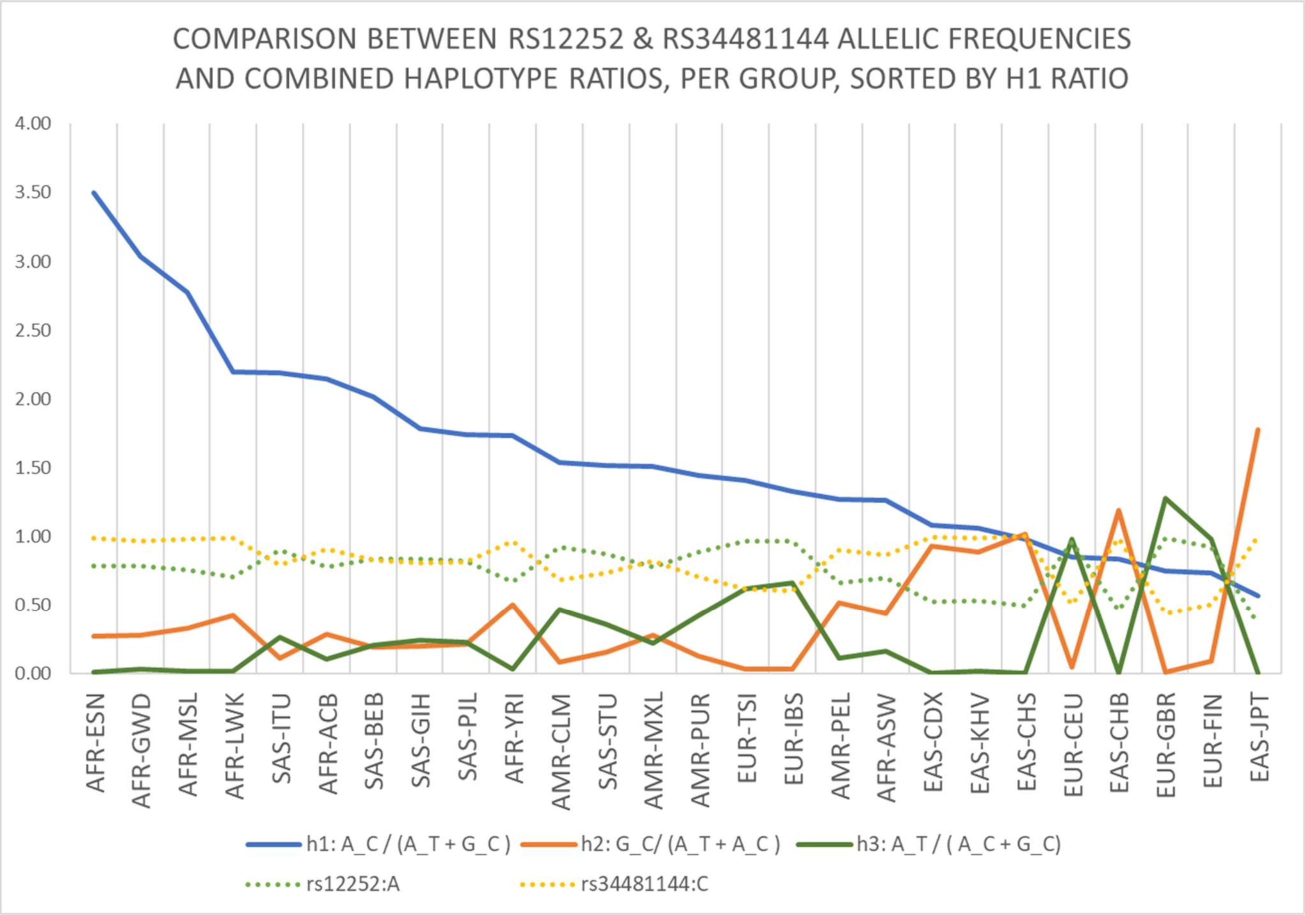

Figure 1. Comparison of h1, h2 and h3 haplotype ratios, for all available ethnic subgroups in 1000 Genomes Project. The fluctuation of frequencies of major alleles rs12252:A and rs34481144:C can be seen following a trend similar to h3 and h2 ratio, respectively. Inversely, the fluctuation of frequencies of minor alleles rs12252:G and rs34481144:T follows a trend similar to h2 and h3 ratio, respectively (not shown on chart).

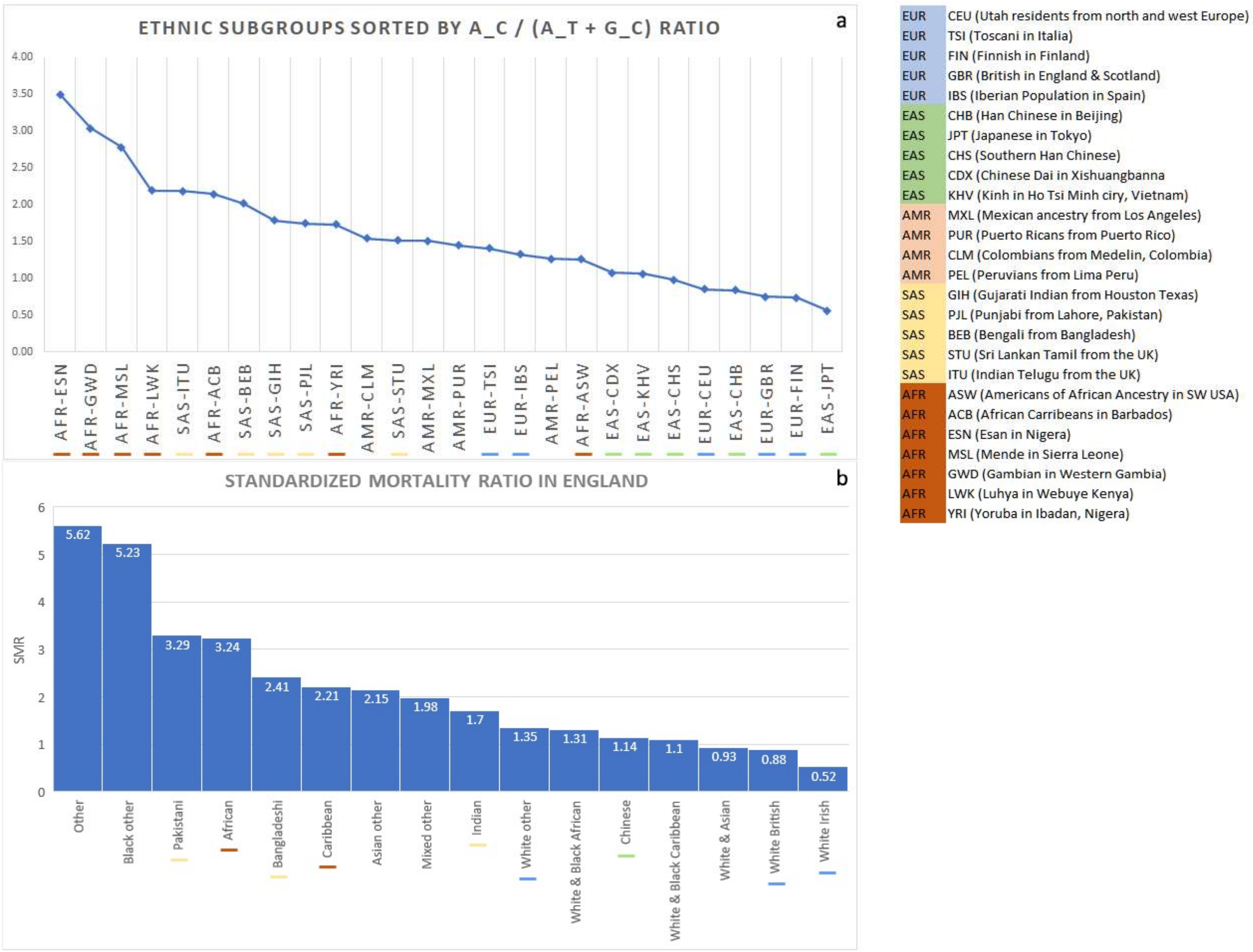

Figure 2. Visual comparison of ethnic subgroup rankings. a) Ethnic subgroups from 1000 Genomes Project sorted by h1 ratio Colored tags are added beneath most subgroups to aid visual identification of major groups and follow the color code on the right. **b)** Standardized Mortality Ratio (SMR) of ethnic groups in England, adjusted for Age and NHS Region.

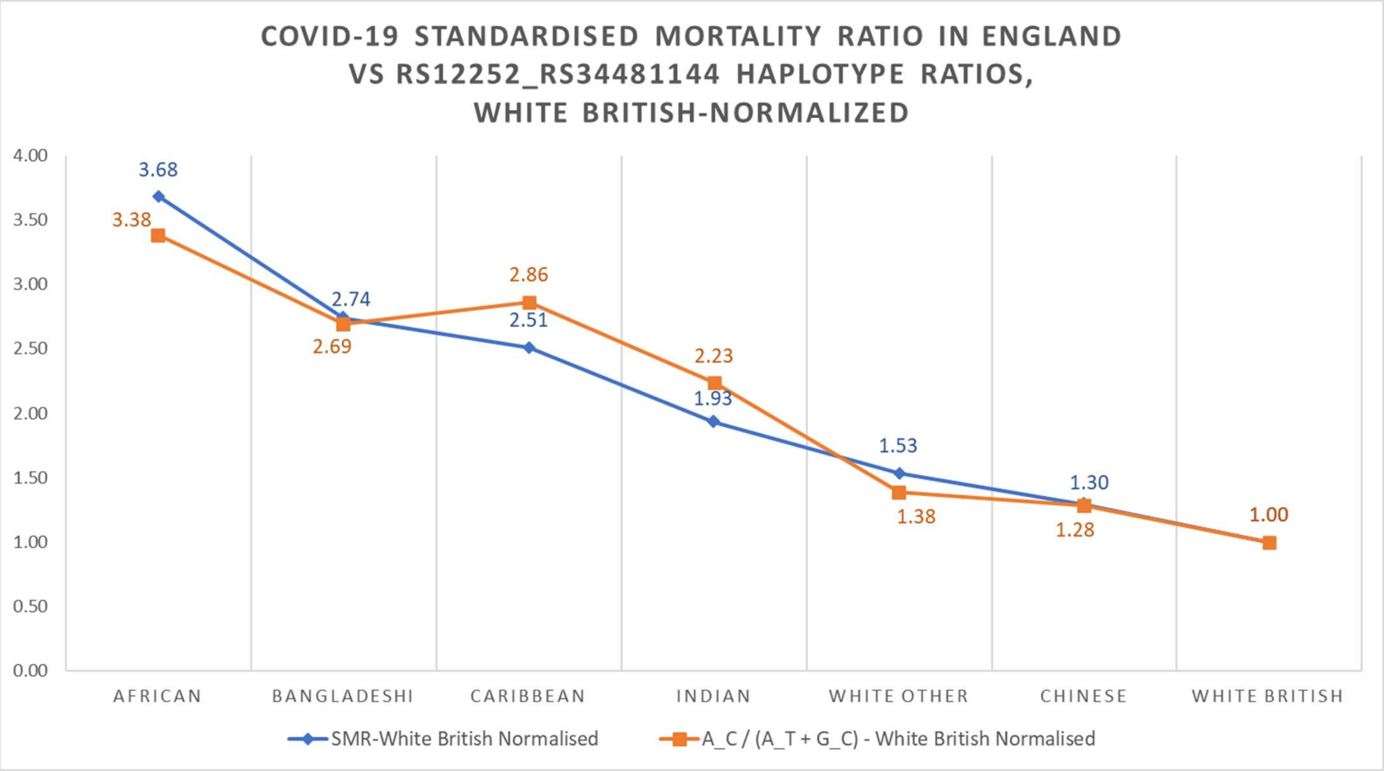

Figure 3. Comparison between Standardized Mortality Ratio (SMR) of ethnic groups in England and h1 haplotype ratio. H1 haplotype ratios are derived from the best corresponding pools of ethnic subgroups from 1000 Genomes Project. Both ratios are normalized by the corresponding numbers of White British, to allow a direct comparison.

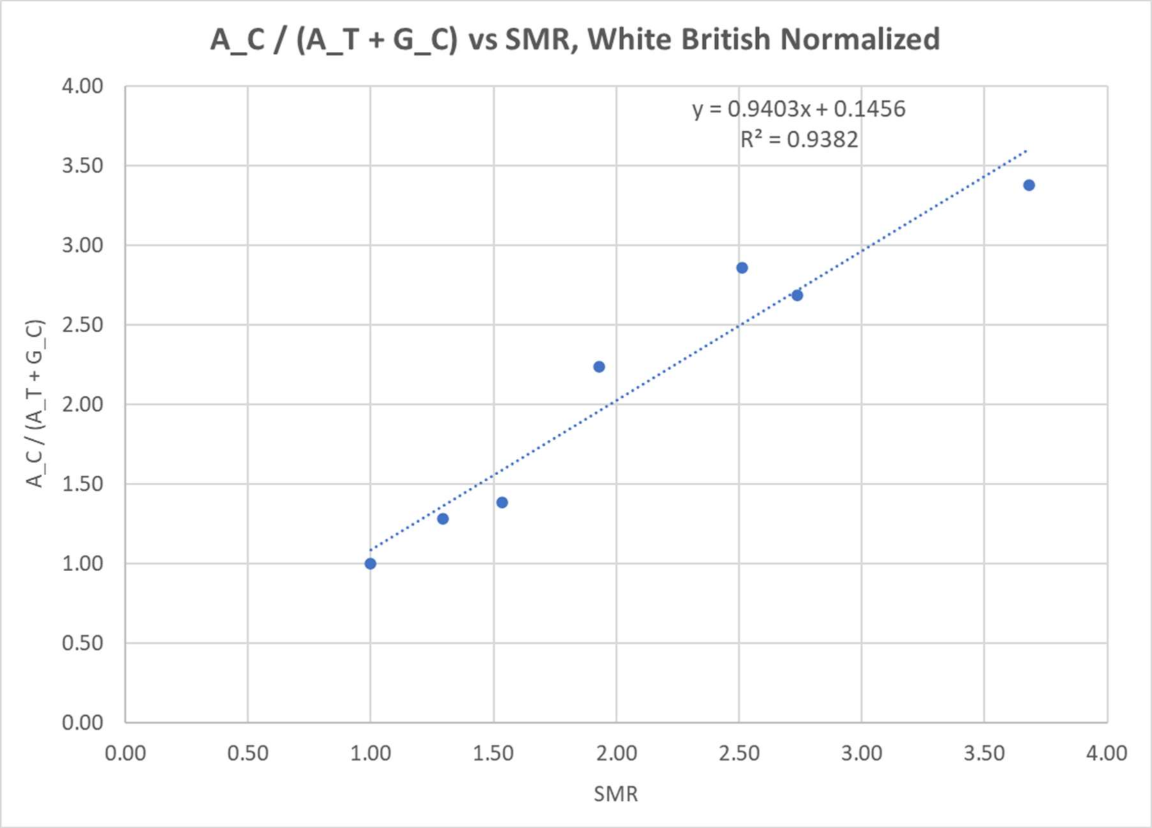

Figure 4. Correlation between h1 haplotype ratios and Standardized Mortality Ratios (SMR), with Pearson r=0.9687, p=3*10^{-4}.

References

23andMe.23andMe Engages Its Millions of Research Participants to Study COVID-19. 2020 April 6 [cited 27 May 2020]. In: 23andMe Blog [Internet]. Available at: https://blog.23andme.com/23andme-research/covid-19-genetic-study.

Aldridge RW, Lewer D, Katikireddi SV, et al. Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data [version 2; peer review: 2 approved, 1 approved with reservations]. Wellcome Open Res. 2020; 5:88. doi:10.12688/wellcomeopenres.15922.2.

Allen EK, Randolph AG, Bhangale T, Dogra P, Ohlson M, Oshansky CM, et al. SNP-mediated disruption of CTCF binding at the IFITM3 promoter is associated with risk of severe influenza in humans. Nature Medicine. 2017; 23: 975e83. doi:10.1038/nm.4370.

Blanco-Melo D, Nilsson-Payant BE, Liu WC, et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell Press. 2020. doi:10.1016/j.cell.2020.04.026.

Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, Ryan BJ, Weyer JL, van der Weyden L, Fikrig E, Adams DJ, Xavier RJ, Farzan M, Elledge SJ. The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. Cell. 2019; 139(7): 1243-54. doi:10.1016/j.cell.2009.12.017.

Buchrieser J, Dufloo J, Hubert M, Monel B, Planas D, Rajah MM, Planchais C, Porrot F, Guivel-Benhassine F, Van der Werf,S, Casartelli N, Mouquet H, Bruel T, Schwartz O. Syncytia formation by SARS-CoV-2 infected cells. BioRxiv. 2020. doi:10.1101/2020.07.14.202028.

CDC. COVID-19 in Racial and Ethnic Minority Groups. Accessed April 26, 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html.

Chanda R, Ghosh S. The Punjabi Diaspora in the UK: An overview of characteristics and contributions to India. Migration Policy Centre, CARIM-India Research Report, Country Reports, 2012; 23. Available from Cadmus, European University Institute Research Repository: http://hdl.handle.net/1814/24845.

Chapman SJ, Hill AV. Human genetic susceptibility to infectious disease. Nature Reviews Genetics. 2012; 13(3): 175-88. doi:10.1038/nrg3114.

Compton AA, Bruel T, Porrot F, Mallet A, Sachse M, Euvrard M, et al. IFITM proteins incorporated into HIV-1 virions impair viral fusion and spread. Cell Host Microbe. 2014; 16: 736-747. doi:10.1016/j.chom.2014.11.001.

Compton AA, Roy N, Porrot F, et al. Natural mutations in IFITM3 modulate post-translational regulation and toggle antiviral specificity. EMBO Reports. 2016; 17(11): 1657-1671. doi:10.15252/embr.201642771.

Department for Communities and Local Government, UK. The Pakistani Muslim Community in England. 2009; 38-39. Available from: https://webarchive.nationalarchives.gov.uk/20120920001118/http://www.communities.gov.uk/documents/communities/pdf/1170952.pdf

Everitt AR, Clare S, Pertel T, John SP, Wash RS, Smith SE, et al. IFITM3 restricts the morbidity and mortality associated with influenza. Nature. 2012; 484: 519-523. doi:10.1038/nature10921.

Feeley EM, Sims JS, John SP, Chin CR, Pertel T, Chen LM, et al. IFITM3 inhibits influenza A virus infection by preventing cytosolic entry. PLoS Pathogens. 2011; 7: e1002337. doi:10.1371/journal.ppat.1002337.

Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1–30, 2020. Morbidity and Mortality Weekly Report. 2020; 69: 458–464. doi:10.15585/mmwr.mm6915e3.

Giamarellos-Bourboulis EJ, Netea MG, Rovina N, et al. Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. Cell Host & Microbe. 2020; In Press. doi:10.1016/j.chom.2020.04.009.

Hachim MY, Al Heialy S, Hachim IY, Halwani R, Senok AC, Maghazachi AA, Hamid Q. Interferon-Induced Transmembrane Protein (IFITM3) Is Upregulated Explicitly in SARS-CoV-2 Infected Lung Epithelial Cells. Frontiers in Immunology. 2020; 11: 1372. doi: 10.3389/fimmu.2020.01372.

He J, Cai S, Feng H et al. Single-cell analysis reveals bronchoalveolar epithelial dysfunction in COVID-19 patients. Protein Cell. 2020; doi:10.1007/s13238-020-00752-4.

Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 2020; 181 (2):271-280.e8. doi:10.1016/j.cell.2020.02.052.

Huang IC, Bailey CC, Weyer JL, Radoshitzky SR, Becker MM, Chiang JJ, et al. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. PLoS Pathogens. 2011; 7: e1001258. doi:10.1371/journal.ppat.1001258.

Jia R, Pan Q, Ding S, Rong L, Liu SL, Geng Y, et al. The N-Terminal Region of IFITM3 Modulates Its Antiviral Activity by Regulating IFITM3 Cellular Localization. Journal of Virology. 2012; 86(24): 13697-707. doi: 10.1128/JVI.01828-12.

Jiang LQ, Xia T, Hu YH, Sun MS, Yan S, Lei CQ, et al. IFITM3 inhibits virus-triggered induction of type I interferon by mediating autophagosome-dependent degradation of IRF3. Cellular & Molecular Immunology. 2017; 15: 858e67. doi:10.1038/cmi.2017.15.

Khunti K, Singh AK, Pareek M, et al. Is ethnicity linked to incidence or outcomes of covid-19? BMJ. 2020; 369: m1548. doi:10.1136/bmj.m1548.

Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. The Lancet Respiratory Medicine, News, Published online May 8, 2020. doi:10.1016/S2213-2600(20)30228-9.

Kuo CL, Pilling LC, Atkins JL, Masoli JAH, Delgado J, Kuchel GA, Melzer D. APOE e4 Genotype Predicts Severe COVID-19 in the UK Biobank Community Cohort. The Journals of Gerontology:Series A. 2020. doi:10.1093/gerona/glaa131.

Lei N., Li Y, Sun Q, Lu J, Zhou J, Li Z, Liu L, Guo J, Qin K, Wang H, Zhao J, Li C, Sun L, Wang D, Zhao Z. & Shu Y. IFITM3 affects the level of antibody response after influenza vaccination. Emerging Microbes & Infections. 2020; 9(1): 976-987. doi: 10.1080/22221751.2020.1756696.

Li J, Wang X, Chen J. Cai Y, Deng A, Yang M. (2020), Association between ABO blood groups and risk of SARS-CoV-2 pneumonia. British Journal of Haematology. 2020; 190: 24-27. doi:10.1111/bjh.16797.

Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015; 31(21): 3555-3557. doi:10.1093/bioinformatics/btv402.

Makvandi-Nejad S, Laurenson-Schafer H, Wang L, Wellington D, Zhao Y, Jin B, Qin L, Kite K, Moghadam HK, Song C, Clark K, Hublitz P, Townsend AR, Wu H, McMichael AJ, Zhang Y, Dong T. Lack of Truncated IFITM3

Transcripts in Cells Homozygous for the rs12252-C Variant That is Associated With Severe Influenza Infection, The Journal of Infectious Diseases. 2018; 217(2): 257–262, doi:10.1093/infdis/jix512.

Niedzwiedz CL, O'Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. MedRxiv [Preprint]. 2020 [cited 2020 March 15]. Available from: https://www.medrxiv.org/content/10.1101/2020.04.22.20075663v2 doi:10.1101/2020.04.22.20075663.

Office for National Statistics, UK. Ethnicity and National Identity in England and Wales: 2011. Retrieved from Office for National Statistics, at: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/ethnicityandnationalidentityinenglandandwales/2012-12-11.

Randolph AG, Yip WK, Allen EK, et al. Evaluation of IFITM3 rs12252 Association With Severe Pediatric Influenza Infection. The Journal of Infectious Diseases. 2017; 216(1): 14-21. doi:10.1093/infdis/jix242.

Shi G., Kenney AD, Kudryashova E, Zhang L, Hall-Stoodley L, Robinson R, Kudryashov D, Compton AA, Yount JS. Opposing activities of IFITM proteins in SARS-CoV-2 infection. bioRxiv. 2020. doi:10.1101/2020.08.11.246678.

Stacey MA, Clare S, Clement M, Marsden M, Abdul-Karim J, Kane L, et al. The antiviral restriction factor IFN-induced transmembrane protein 3 prevents cytokine-driven CMV pathogenesis. Journal of Clinical Investigation. 2017; 127: 1463e74. doi: 10.1172/JCI84889.

The COVID-19 Host Genetics Initiative. 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. European Journal Human Genetics. 2020. doi: 10.1038/s41431-020-0636-6.

Wang Z, Zhang A, Wan Y, Liu X, Qiu C, Xi X, et al. Early hypercytokinemia is associated with interferon-induced transmembrane protein-3 dysfunction and predictive of fatal H7N9 infection. Proceedings of the National Academy of Sciences of the USA. 2014; 111: 769e74. doi:10.1073/pnas.1321748111.

Wellington D, Laurenson-Schafer H, Abdel-Haq A, Dong T. IFITM3: How genetics influence influenza infection demographically. Biomed Journal. 2019; 42(1): 19-26. doi:10.1016/j.bj.2019.01.004.

Wu Y, Feng Z, Li P, Yu Q. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. Clinica Chimica Acta. 2020; 509: 220-223. doi:10.1016/j.cca.2020.06.026.

Xie M, Xuan B, Shan J, Pan D, Sun Y, Shan Z, et al. Human cytomegalovirus exploits interferon-induced transmembrane proteins to facilitate morphogenesis of the virion assembly compartment. Journal of Virology. 2015; 89: 3049–3061. doi:10.1128/JVI.03416-14.

Zhang YH, Zhao Y, Li N, Peng YC, Giannoulatou E, Jin RH, Yan HP, Wu H, Liu JH, Liu N, Wang DY, Shu YL, Ho LP, Kellam P, McMichael A, Dong T. Interferon induced transmembrane protein 3 genetic variant rs12252 C is associated with severe influenza in Chinese individuals. Nature Communications. 2013; 4: 1418. doi:10.1038/ncomms2433.

Zhang Y, Makvandi-Nejad S, Qin L, et al. Interferon-induced transmembrane protein-3 rs12252-C is associated with rapid progression of acute HIV-1 infection in Chinese MSM cohort. AIDS. 2015; 29(8): 889-894. doi:10.1097/QAD.0000000000000632.

Zhang Y, Qin L, Zhao Y, Zhang P, Xu B, Li K, Liang L, Zhang C, Dai Y, Feng Y, Sun J, Hu Z, Xiang H, Knight CJ, Dong T, Jin R. Interferon-induced transmembrane protein-3 genetic variant rs12252-C is associated with disease severity in COVID-19, The Journal of Infectious Diseases. 2020; 224. doi:10.1093/infdis/jiaa224.

Zhao T. IFITM3 related to COVID-19, The Journal of Infectious Diseases. 2020. doi:10.1093/infdis/jiaa454.

Zhao X, Guo F, Liu F, Cuconati A, Chang J, Block TM, et al. Interferon induction of IFITM proteins promotes infection by human coronavirus OC43. Proceedings of the National Academy of Sciences of the U.S.A. 2014; 111: 6756–6761. doi:10.1073/pnas.1320856111.

Zhao X, Sehgal M, Hou Z, Cheng J, Shu S, Wu S, Guo F, Le Marchand SJ, Lin H, Chang J, Guo JT. Identification of Residues Controlling Restriction versus Enhancing Activities of IFITM Proteins on Entry of Human Coronaviruses. Journal of Virology. 2018; 92(6): e01535-17. doi:10.1128/JVI.01535-17.

Zheng M, Zhao X, Zheng S, Chen D, Du P, Li X, Jiang D, Guo JT, Zeng H & Lin H. Bat SARS-Like WIV1 coronavirus uses the ACE2 of multiple animal species as receptor and evades IFITM3 restriction via TMPRSS2 activation of membrane fusion. Emerging Microbes & Infections. 2020; 9:1, 1567-1579. doi:10.1080/22221751.2020.1787797.