The role of the X chromosome in infectious diseases

Hay, Megan; Kumar, Vinod; Ricaño-Ponce, Isis

Published in:
Briefings in Functional Genomics

DOI:
10.1093/bfgp/elab039

Publication date:
2022

Briefings in Functional Genomics, 00(00), 2021,1–16

https://doi.org/10.1093/bfgp/elab039
Review Paper

The role of the X chromosome in infectious diseases

Megan Hay, Vinod Kumar and Isis Ricaño-Ponce

Corresponding author: Vinod Kumar, Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA, Nijmegen, The Netherlands. Tel.: +31-617646295; Email: v.kumar@radboudumc.nl

Abstract

Many infectious diseases in humans present with a sex bias. This bias arises from a combination of environmental factors, hormones and genetics. In this study, we review the contribution of the X chromosome to the genetic factor associated with infectious diseases. First, we give an overview of the X-linked genes that have been described in the context of infectious diseases and group them in four main pathways that seem to be dysregulated in infectious diseases: nuclear factor kappa-B, interleukin 2 and interferon γ cascade, toll-like receptors and programmed death ligand 1. Then, we review the infectious disease associations in existing genome-wide association studies (GWAS) from the GWAS Catalog and the Pan-UK Biobank, describing the main associations and their possible implications for the disease. Finally, we highlight the importance of including the X chromosome in GWAS analysis and the importance of sex-specific analysis.

Key words: X chromosome; infectious diseases; GWAS

Introduction

The role of the X chromosome in complex diseases

A long list of human infectious and autoimmune diseases are more prevalent in one sex than the other [1–3]. Various environmental factors have an effect on this sex bias, for instance smoking, travel, access to health care and adherence to treatment [3]. Concomitantly, sex chromosomes constitute the genetic basis for the difference between males and females. In humans, the Y chromosome contains only around 100 genes, whereas the X chromosome harbors over 1500 genes, with a high density of loci modulating the immune response [3–5]. Since females have two copies of the X chromosome and males only have one, it is vital for proper function to realign the difference in dosage of gene expression between sexes. This role is made evident in X-dosage imbalances caused by aneuploidies of the X chromosome such as Turner syndrome (45, X) and Klinefelter syndrome (47, XXY) [6]. Individuals with either condition have an increased risk to develop autoimmune diseases [6–9], and recent research begins to elucidate the shared regulatory network affected in both conditions [9].

The dosage discrepancy between sexes is ordinarily compensated via several mechanisms, most importantly the process of X chromosome inactivation in females during early embryonal development [4]. Normally, one random X chromosome is silenced and females become functional mosaics for expression of the X chromosome [4]. The X chromosome inactivation process is mediated by X-inactive specific transcript (*Xist*), a long intergenic non-coding RNAs (lincRNA) encoded in the X chromosome inactivation center. LincRNAs are non-coding transcripts of more than 200 nucleotides in length. *Xist* causes the silencing of a large proportion of the X chromosomes [10] through long-range chromatin interactions that bind two distal genomic regions together [11]. However, up to 15% of genes on the X chromosome escape silencing [6]. The dosage difference in such genes is the suspected genetic cause of the differences between males and females in response to complex diseases.

A striking X chromosome-specific characteristic is that it contains 10% of all genomic micro-RNAs (miRNA), a disproportionately high amount considering its size compared to the rest of the chromosomes [2, 3]. In humans, as well as in other

Megan Hay is currently pursuing a Master's degree in Medical Biology at Radboud University, Nijmegen NL, with particular emphasis on Human Biology and Science Management.
Vinod Kumar is a group leader and a PI at the department of Internal Medicine at RadboudUMC, and the department of Genetics at UMC Groningen. His group applies system genetic approaches to characterize the genetic basis for infectious disease susceptibility.
Isis Ricaño-Ponce is a postdoc fellow at the Department of Internal Medicine at RadboudUMC. Her work focuses on multiomics data integration on infectious diseases.

seven mammalian species, the density of miRNA on the X-chromosome is greater than two-fold of those from the autosomal chromosomes [12]. These miRNAs are small non-coding RNA molecules with a regulatory function, achieved by targeting complementary sequences in usually multiple mRNAs and altering splicing, instigating their degradation or blocking translation [2, 3, 6]. Some miRNAs are located in genes escaping X chromosome inactivation and, considering their effects on diverse signaling cascades, have the potential to be major regulators of sex differences [6]. In contrast, the Y chromosome does not appear to contain any miRNAs [12]. Most miRNAs on the X chromosome have yet to be described; however, their suppression of protein synthesis appears to be an important regulatory step in many cancer, fertility and immune-related pathways [6, 7, 13].

While mutations on the X chromosome contribute to almost 10% of Mendelian disorders [14], the number of known associations on the X chromosome with complex diseases is lower than associations on the autosomes. For example, a few years ago, a study found that if we compare the X chromosome to chromosome 22, which is three times smaller, chromosome 22 has four times more reported associations in the GWAS catalog [5]. Associations on the X chromosome are scarce as it is often excluded from GWAS due to its sex-specific characteristics and concomitant technical difficulties: Dosage differences of the sex chromosomes present difficulties to statistical analysis between males and females, as well as sex chromosomes compared to autosomes [3, 5]. X chromosome inactivation and sometimes skewed or tissue-specific inactivation patterns in females present a challenge for association variants with individual phenotypes, because standard sequencing technologies cannot differentiate between the silenced and the active X chromosome [5, 14]. Moreover, the density of markers on the X chromosome in the genotyping platforms was very small compared to other chromosomes, though this has since improved [5, 14]. Finally, there was simply a lack of power and available software to perform robust associations of the X chromosome [3, 5, 14]. Nevertheless, 437 associations to 268 diseases or traits have been reported in the GWAS catalog [15] by November 2020. Including some immunological and neurological diseases such as celiac disease, vitiligo, Type 1 diabetes, bipolar disorder, Schizophrenia and Alzheimer's disease, among others [16–20].

Beyond the genetic level, it has been proposed that sex hormones act as an additional compensatory mechanism to ameliorate the effect of gene dosage discrepancies [4]. Testosterone, present at higher concentrations in males, is anti-inflammatory, and the estrogens more produced in females are primarily pro-inflammatory [2, 3]. Interestingly, characteristic sex differences in the microbiome have been shown to develop after puberty, i.e. at the same time that sex hormone levels rise [2]. While the effect of the immune system on the microbiome is not a new discovery, the influences of the microbiome on the immune system are becoming increasingly recognized in complex diseases including cancer, autoimmune disorders and microbial infections, among others [21–24]. Further, the mitochondria have an important role in innate and adaptive immune cell regulation and, in particular, antiviral defense [2]. Though mitochondrial DNA is separate from genomic DNA, this is of note in sex-dependent immune responses since mitochondria are maternally inherited and as such display bias towards variations advantageous to females, even when deleterious to males [2].

This evidence combined purports multifaceted interactions between diverse environmental and genetic factors that contribute to the sex bias observed in many complex diseases. The influence of environmental factors and hormones is bigger than the influence of genetics on the majority of complex diseases; nevertheless, the underlying genetic differences cannot be ignored, especially those of the X chromosome. Thus, further studies are needed to elucidate the role of X-linked genes in disease pathogenesis.

Importance of the X chromosome in infectious diseases

Females are more susceptible to developing autoimmune disorders; however, they perform better in defense against infectious diseases [3, 4, 6, 7]. This makes sense, considering the evolutionary bias of the X chromosome towards benefitting females and its high density of immune genes. Though sex hormones appear to contribute to this phenomenon, they do not fully explain it [3]. Apart from the dosage difference of certain X-linked genes, the biallelic expression of X-linked genes makes females non-phenotypical for recessive mutations and is thought to increase the genetic diversity of immune cells [4]. This diversity may be advantageous for the detection of infectious pathogens.

In the following, we briefly review X-linked genes with a role in susceptibility to infectious diseases that have been described in published literature. This is followed by a detailed assessment of SNP associations on the X chromosome affecting traits related to infectious diseases. To this end, we evaluate the relevant associations published in the GWAS Catalog to date (November 2020), as well as the associations published by the Pan-UK Biobank [25]. The latter is a large genetic association database based on a cohort of 500 000 individuals, for which the UK Biobank collected various samples and health information.

X-linked genes associated with infectious diseases

Genes on the X chromosome involved in central host defense pathways

Many X-linked genes with a role in susceptibility to infectious diseases have been described, and even longer is the list of immune genes with potential impact on the response to infection; we give an overview of the most important genes (Figure 1, Supplementary Table 1). In the following, we highlight four integral pathways in the immune response to infection and the related genes on the X chromosome.

NF-κB

Nuclear factor kappa-B (*NF-κB*) is a transcription factor central to many cellular processes, including the inflammatory response to pathogens [26, 27]. The NF-κB essential modulator (NEMO) is encoded by the X-linked Inhibitor of NF-κB Kinase γ (*IKBKG*) gene and inhibits an NF-κB kinase, thereby upregulating the expression of *NF-κB* [4, 27]. Amorphic mutations are male lethal, while milder mutations in *IKBKG* cause primary immunodeficiencies in diseased individuals [4, 28]. Clinically these mutations manifest as a Mendelian susceptibility to mycobacterial infections, often accompanied by ectodermal dysplasia [4, 28]. Interestingly, the ectodysplasin gene (*EDA*) and its receptor *EDA2R*, both located on the X chromosome, are upstream of NEMO and NF-κB in this signaling pathway [27, 29]. Mutations in this gene and receptor can likewise result in ectodermal dysplasia, though there are no reports of immunodeficiencies, potentially due to compensatory signals in the cascade [6, 27]. *NF-κB* can also be upregulated by *miR-18*, an X-linked miRNA involved in cancer and immunity [6, 30]. In contrast, NF-κB Repressing Factor

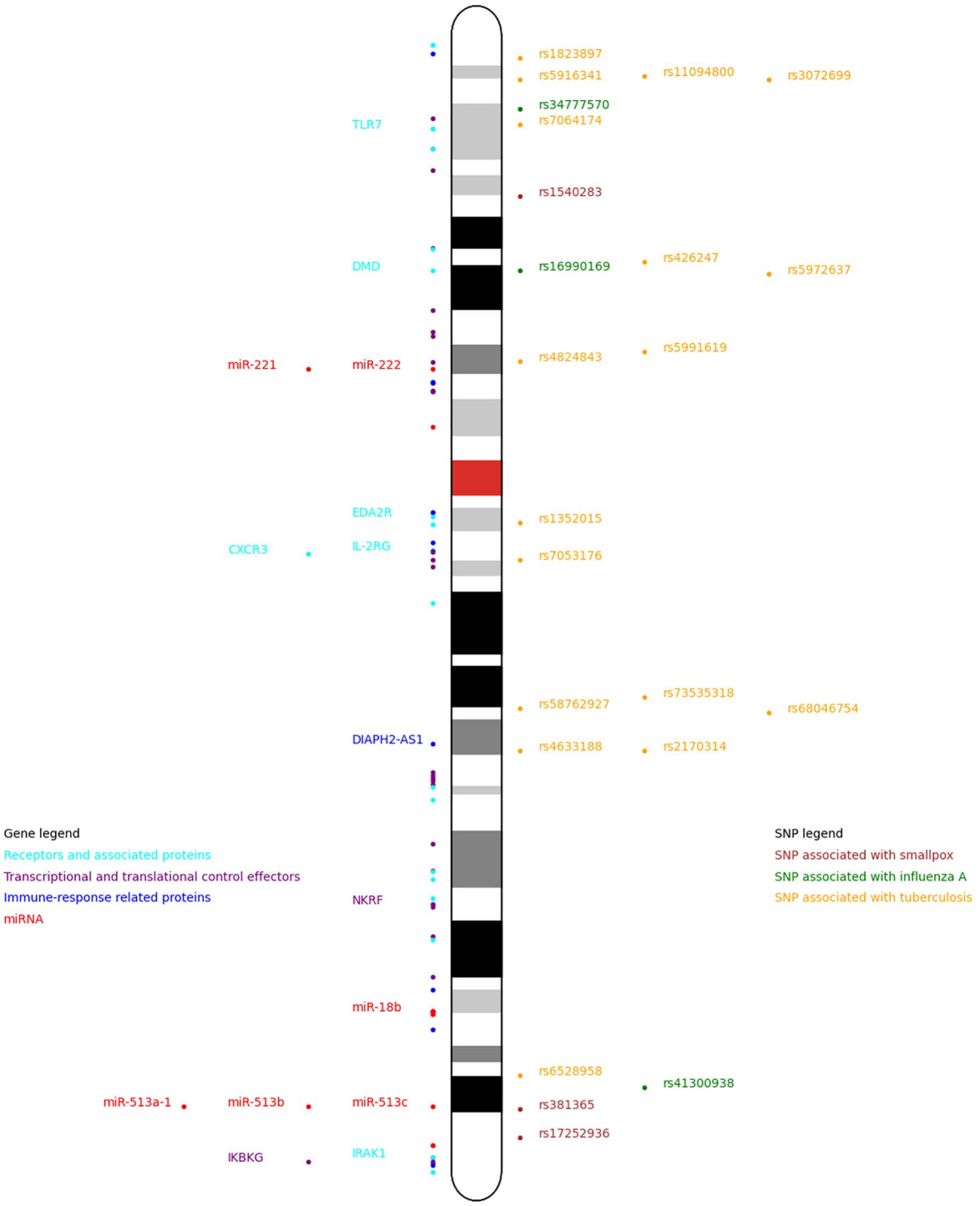

Figure 1. Ideogram of the X chromosome. Genes (left) and SNPs (right) associated with susceptibility to infectious diseases discussed in this review are highlighted and labeled. The asterisk marks two SNPs both located in the *DMD* gene, which is hypothesized to be involved in a common condition in both diseases: muscle wasting. Further genes involved in the immune response to infectious diseases are highlighted (for the detailed list refer to Supplementary Table 1). Categorization inspired by Fish et al. [1].

(*NKRF*) is an X-linked inhibitor of NF-κB and exults an anti-inflammatory effect [1, 31]. Specifically, excess levels of NKRF impede the immune response to *Mycobacterium tuberculosis* by inhibiting NF-κB-dependent induction of the cytokines Interferon γ-induced protein 10 (IP-10) and Interleukin-8 (IL-8) in monocytes [32].

Il-2/IFNγ

The cytokines IL-2 and interferon γ (IFNγ) are part of a signaling cascade that downregulates immune cells, especially $T_{helper}17$ cells [33]. This subset of T cells is involved in the host defense to extracellular microbes, and moreover, there are several studies linking mutations in IFNγ receptor gene to

increased susceptibility to the intracellular *Mycobacteria* [33–36]. Stimulation with *M. tuberculosis* antigens correlates with increased IFNγ expression, and this response is dependent on IL-2 [37]. Beyond the bacterial realm, IL-2 is a 'key component of effective immunity' for instance in T cell susceptibility to the Human Immunodeficiency Virus (HIV) [38]. Mutations in the IL-2 receptor subunit γ gene (*IL2RG*), located on the X chromosome, are associated with severe X-linked immuno-deficiencies [4, 26, 39].

Toll-like receptors

Toll-like receptor (TLR) pathways play an important role in the innate immune response by recognizing pathogen-specific molecular patterns with downstream activation of NF-κB and its pro-inflammatory network [40, 41]. TLR7 is such a pattern-recognition receptor, located on endosomes, and specifically detects the presence of viral single-stranded RNA and DNA [1, 41, 42]. For instance, TLR7 is a protein integral to recognition of Influenza virus A infection [43, 44]. Though TLR7 is normally acknowledged for its viral recognition, it can also efficiently recognize bacterial RNA in conventional dendritic cells, leading to the expression of pro-inflammatory cytokines [42]. Another X-linked gene, IL-1 receptor-associated kinase (*IRAK1*) performs a key regulatory function downstream of TLR7 in the TLR-dependent signaling pathway [3, 41]. Although no genome-wide significant associations have been associated with infectious diseases in TLR genes from GWAS, candidate studies support this association. For example, a meta-analysis of 32 candidate gene studies investigating the influence on TLR and tuberculosis susceptibility comprising 18 907 individuals reported 6 significant associations in TLR1, TLR2, TLR4 and TLR9 [45].

Interestingly, variants in TLR have been associated with infectious diseases in a sex-specific manner. For instance, a study involving 23 SNPs located in 5 TLR genes comprising 729 tuberculosis cases and 487 healthy controls from a South African population, identified 5 SNPs within TLR8 that have sex-specific effects. Three variants were associated only in females, while two were specific for males [46]. Similar sex-specific associations to TLR8 have been observed in different populations [47–49].

More recently, rare variants in TLR7 have been associated with critical Covid-19 in males below the age of 60 years disrupting the type I and II IFN (interferons) responses [50–53]. In a case control study of 156 young males (<60 years) from the Italian GEN-COVID cohort, 5 out of 79 male patients have rare missense variants predicted to impact protein function and none in 77 oligo-asymptomatic Covid-19 patients that did not require hospitalization [50]. Similar results were observed in another study, where enrichment of very rare or private nonsynonymous variants in TLR7 was observed in 1.8% of male patients below the age of 60 years with critical Covid-19 pneumonia [51]. These studies highlight the importance of sex-specific analysis in infectious diseases.

PD-L1

The programmed death ligand 1 (*PD-L1*) gene has a well-described role in tumor immune evasion, but also has a complex role in the antiviral immune response [54, 55]. PD-L1 effectively downregulates the immune response in persistent infections, which is achieved by lowering the stimulatory ability of dendritic cells, expressing the ligand, and T cells, expressing the receptor PD-1 [56–58]. *PD-L1* is partly induced by STAT3 [58, 59]. This pathway is subject to inhibition by three X-linked miRNAs: *miR-221* and *miR-222* interfere with STAT3, while *miR-513* targets the *PD-L1* mRNA directly [6, 30, 55].

There are two additional genes of interest in the context of this review. Firstly, the chemokine receptor gene *CXCR3* is expressed in effector T cells and escapes X chromosome inactivation specifically during infection [60–62]. And secondly, diaphanous-related formin 2, antisense RNA 1 (*DIAPH2-AS1*) plays a role in early lymphoid activation [1, 63]. Both CXCR3 and DIAPH2-AS1 are discussed further on with specific SNP associations.

Among this selection of X-linked genes with a role in host defense to infectious pathogens, a number have been observed to escape silencing of the X chromosome in females (*miR-221* during meiotic sex chromosome inactivation [64], *TLR7* [61, 65], *IRAK1* escape reports are inconsistent [4, 61], *IKBKG* [4] and *EDA2R* in some individuals [66]). Though escape or reactivation is not an immutable status, the genes that do escape inactivation might have dosage differences between sexes and are potential contributors in sex-biased responses to infectious agents.

Genetic associations with infectious diseases in the GWAS Catalog

Since 2008, the GWAS Catalog curates SNP-trait associations from published studies and makes them widely accessible [15]. We use these data as a primary source to assess the currently available information on SNP associations on the X chromosome with phenotypes pertaining to infectious diseases. After filtering for X chromosome associations, the phenotypes listed in the GWAS Catalog were manually curated for phenotypes linked with infectious diseases.

Our findings implicate 65 associations on the X chromosome with 20 infectious diseases with suggestive level of significance (Supplementary Table 2). Among these associations, we find 70 unique SNPs, which encompasses 12 SNP×SNP interactions. The majority of these are intronic (34 SNPs) or intergenic (27 SNPs), though we find four variants in regulatory regions, two variants in 3' UTRs, two missense variants and one synonymous coding variant. These results are in accordance with previous observations in GWAS studies, where it has been shown that the majority of GWAS SNPs colocalize with regulatory regions [67]. These 65 associations are a conglomerate of 17 studies, only 6 of which included non-European cohorts: Hispanic, African American, South African, Oceanian individuals and Individuals with African and Asian ancestry. Notably, just a single study has been replicated with a larger cohort.

For further analysis, the criteria were narrowed down to SNPs of genome-wide significance ($P \leq 5 \times 10^{-8}$), resulting in 18 unique associations encompassing 23 SNPs associated with three infectious diseases: three SNP associations with smallpox virus infection, three with influenza virus A subtype H1N1, and 12 SNP × SNP interactions associated with tuberculosis (Table 1). In the following, we discuss these 23 SNPs in detail.

Associations with the smallpox virus

The smallpox virus was endemic until widespread vaccination with live pox viruses was achieved in 1980, though the campaign was not without some adverse events [68]. Taking advantage of the new technological possibilities GWAS offers, Kennedy et al. investigated the cytokine response to vaccinia virus in individuals that have been vaccinated against smallpox to inform the safety of future inoculations. On the X chromosome, three

Table 1. Genome-wide significant SNP associations with infectious diseases on the X chromosome published in the GWAS catalog

Phenotype	Sample size	Mapped/nearest gene	SNPs	Risk allele	P-value	OR	Regulatory motifs altered [99]	Reference
Immune response to smallpox (secreted IFN-α)	Up to 512 European ancestry individuals, up to 199 African American individuals	Z97180.1 - CXorf51B	rs381365	A	2E-12	NA	Foxa, Foxj2, Foxk1, GATA, Pou2f2	[68]
		MAMLD1	rs17252936	G	6E-10	NA	NSRF, Pax-5	
		PHEX	rs1540283	G	9E-09	NA	Egr-1, RBP-Jκ	
Influenza A (H1N1) infection	49 European ancestry severe cases, 107 European ancestry mild cases, 549 European ancestry population controls	DMD	rs16990169	A	4E-11	11.05	Foxo, YY1	[79]
Mild influenza (H1N1) infection	107 European ancestry mild cases, 549 European ancestry population controls	DMD	rs16990169	A	5E-14	14.55	Foxo, YY1	
		WWC3 - CLCN4	rs34777570	Gindel	1E-09	NA	EBF, Ik-1, Pax-4, Zic	
Severe influenza A (H1N1) infection	49 European ancestry severe cases, 549 European ancestry population controls	SLITRK4	rs41300938	T	2E-16	NA	NF-κB	
Tuberculosis (SNP × SNP interaction)	242 South African admixed female cases, 168 South African admixed male cases, 223 South African admixed female controls, 182 South African admixed male controls	DMD × UBE2V1P9 - RNU6-555P	rs5972637 × rs73535318	?	2E-11	NA	DMRT2, DMRT4, E2F, Foxp1, Homez, Pou2f2, TATA	[102]
		EDA2R × PCDH11X	rs1352015 × rs68046754	?	2E-11	NA	FAC1 × PLZF	
		NLGN4X × PCDH11X	rs5916341 × rs68046754	?	3E-11	NA	Fox, Foxa, Foxf1, Foxi1, Foxj2, Foxk1, Foxl1, Foxp1, HNF1, Nkx3, PEBP, TATA, Zfp105 × PLZF	
		NLGN4X × PCDH11X	rs3072699 × rs68046754	?	3E-11	NA	× PLZF	
		AC112653.1 - LINC01546 × CXCR3 - BX276092.7	rs1823897 × rs7053176	?	3E-11	NA		
		AC112653.1 - LINC01546 × FRMPD4	rs1823897 × rs7064174	?	7E-14	NA	x GCNF, SF1, YY1, Zfp410	
		TAB3 - FTHL17 × PCDH11X	rs426247 × rs68046754	?	2E-12	NA	Maf, RFX5 × PLZF	
		AL023574.1 × UBE2V1P9 - RNU6-555P	rs5991619 × rs73535318	?	6E-12	NA	Pou2f2, Pou5f1	
		AC110995.1 × PCDH11X	rs11094800 × rs68046754	?	7E-12	NA	× PLZF	
		PCDH11X × AL031073.2	rs68046754 × rs6528958	?	8E-12	NA	PLZF × PU.1, STAT	
		FUNDC1 - CHTF8P1 × PCDH11X	rs4824843 × rs58762927	?	1E-11	NA	Alx4, GATA, Gfi1 × AIRE	

SNPs affecting IFNα secretion in response to challenge with the vaccinia virus were identified in a cohort of circa 200 individuals of African American ancestry [68].

The two genes Z97189.1, which appears to encode a non-protein coding transcript, and *CXorf51B*, a protein-coding gene expressed only in the testes, map closely to the intergenic variant rs381365, which is the strongest association in the locus [69, 70]. Although there are no studies suggesting a role of these genes in the host response to infectious diseases, the SNP contains altered regulatory motifs for four transcription factors (FoxA, FoxK, Pou2f2 and GATA) that have been implicated in host response or recovery from pathogenic challenge [71–75]. Further investigation is required to rule out the relevance of this SNP.

The second association corresponds to rs17252936, an SNP that locates to an intron of the mastermind-like domain-containing protein 1 (*MAMLD1*) gene, which appears to be involved hypospadias (a malformation of the urethra) in males and has not been described in the context of infection [76].

The third SNP identified by Kennedy et al., rs1540283, is linked with the phosphate-regulating gene with homologies to endopeptidase on the X chromosome (*PHEX*) gene. This gene is known for its role in phosphate metabolism and rickets and has not been implicated in the host defense [77]. However, PHEX was observed to be downregulated in response to *Mycobacterium leprae*, making it a putative candidate for the neuropathy and bone deformities experienced chronically as a consequence of leprosy [78].

Existing literature does not yet validate an effect of these three SNPs on host susceptibility or immune response to infectious agents, thus further functional studies are needed to clarify their roles in immune challenge with the smallpox virus.

Associations with influenza A virus subtype H1N1

H1N1 is the strain of influenza A virus that was pandemic in 2009, which distinguished itself from other influenza viruses especially in its severity in children and certain adult populations [79]. While no genetic risk factor specific for disease severity could be identified by comparing 49 severely affected patients to 107 patients with mild disease progression, Garcia-Etxebarria et al. identified three SNPs on the X chromosome that were more common in combined influenza cases (49 severe and 107 mild) compared to a control group of 549 individuals of a general European population.

rs16990169 was significantly associated with both mild and all influenza cases compared to the general population. This SNP represents a missense variation (arginine to cysteine) in the Duchenne Muscular Dystrophy (*DMD*) gene, which encodes an essential structural protein in striated muscle cells called dystrophin [79]. Mutations in this gene can result in severe muscle weakness and can be causative for the equally named DMD disorder [80]. Interestingly, patients suffering from this disease experience high morbidity from respiratory infections [81]. A zebrafish model supports the hypothesis that the immune system could exacerbate muscle wasting, suggesting a synergistic effect of influenza A infection and *DMD* deficiency [82]. Therefore, existing literature hints at an involvement of dystrophin in H1N1 infection and merits experimental investigation.

rs34777570 is an intergenic variant associated with mild influenza cases compared to the general population sample. The SNP is located upstream of the WW and C2 domain containing 3 (*WWC3*) gene and downstream of chloride channel 4 (*CLCN4*) [79]. *WWC3* regulates cell proliferation and particularly low expression of this gene is a common marker of lung cancers [83]. *CLCN4* encodes a chloride channel, where mutations can be causal for intellectual disability and epilepsy [84]. While both genes have been researched extensively, no implication in infectious diseases has been reported.

The synonymous variant rs41300938 is associated with severe influenza compared to the general population. It maps to the SLIT and NTRK-like family, member 4 (*SLITRK4*) gene, which encodes a transmembrane protein highly expressed in the central nervous system [85]. This SNP contains an altered regulatory motif for NF-κB, which is a common target of viruses to 'hijack' cellular pathways, making this locus attractive for further investigation [86].

Though there have been numerous studies investigating the sex-specific response to the influenza virus, results are inconsistent and may differ between populations [87]. A consensus exists that young boys are more affected than girls and that pregnancy poses a substantial risk in the context of viral infection, and several studies have found women of reproductive age to be more strongly affected by influenza than males of the same age group in some, but not all populations [87, 88]. On the individual level, women mount a higher inflammatory response to influenza infection than men. This is advantageous for females unless the response becomes excessive and causes tissue damage (immunopathogenic), whereas males are more at risk of an insufficient immune response [89]. It has been proposed that this immunosuppression in males may be regulated in part by steroid hormone receptors on immune cells and their (proinflammatory) downstream effector protein NF-κB, as well as by steroid hormone regulation of miRNAs, for example by estradiol [89]. Additional genetic studies in diverse populations, as well as sex- or age-stratified analyses might lead to the discovery of new associations.

Associations with tuberculosis

Tuberculosis (TB) is a deadly bacterial infection with strong sex bias towards men. Therefore, Schurz et al. set a particular focus on the overlooked X chromosome in their sex-stratified TB susceptibility GWAS in an admixed South African population sample of circa 400 cases and controls each. There was no genome-wide significant association of a single SNP on the X chromosome, potentially due to the complexity of the disease, yet the group identified 12 SNP $\times$ SNP interactions below this threshold. As the authors note, there is no established significance threshold for these epistatic SNPs, and despite additional correction, these interactions are not directly comparable with individual SNPs. Potential involvement of these SNPs in infection reported in published literature is summarized in Table 2.

Conclusions on the associations reported in the GWAS Catalog

We have discussed all 6 genome-wide significant SNP associations and 12 SNP $\times$ SNP interactions on the X chromosome that are associated with three infectious-disease phenotypes in the GWAS Catalog (as of November 2020). Not surprisingly, several undescribed genes, pseudogenes and non-coding genes were nearby the SNPs identified in the corresponding studies. However, it has been shown that non-coding RNAs can play an important role in the immune response to pathogens [90, 91].

In the context of X-specific high density of miRNAs, we probed for overlap of the previously discussed SNPs with miRNA using SNP Nexus, although we did not find any matches with annotated miRNAs [92–96]. It is worth mentioning that both pseudogenes and lncRNA can act as decoys for miRNAs [97, 98]. Crosstalk between lncRNAs and miRNAs has gained interest in recent years and expanded our view of complex regulatory

Table 2. Literature reports on the SNPs involved in significant SNP × SNP interactions in tuberculosis as identified by Schurz et al. [102]

SNP	Variant	SNP	Variant
rs68046754	Intronic variant in the Protocadherin 11 X-linked (*PCDH11X*) gene [102]. PCDH11X is a transmembrane protein pertaining to the cell adherin superfamily that has been implicated in cell–cell recognition, speech, cerebral asymmetry and connectivity [110–112], and is highly expressed in the female brain [113]. PCDH11X has previously been identified as a suggestive association with the host response to HIV [114]	rs1352015	Intronic variant in the ectodysplasin-A2 receptor (*EDA2R*). EDA2R is involved in hair follicle development, hair loss and its ligand was found to induce apoptosis in osteosarcoma [115–117]. EDA2R plays a role upstream of NF-κB, though consequences for infectious disease susceptibility have not been reported [27].
		rs6528958	Intronic variant in the scarcely described, unprocessed pseudogene *AL031073.1* [118]. *AL031073.1* and the close-by, protein-coding gene Melanoma-associated Antigen C2 (*MAGEC2*) appear to be expressed regularly in the testes and overexpressed in various cancers, where the latter promotes tumor metastasis [118, 119]. So far, no connection to infectious diseases has been published for this SNP.
		rs11094800	Maps most closely to an intron of the long non-coding RNA *AC110995.1*, that has not been characterized in detail [120]. The nearest coding gene is Neuroligin-4 (*NLGN4*), further discussed in context with the two following epistatic associations with rs68046754 [15].
		rs5916341	Two intronic variants in the synaptic cell adhesion protein encoded by *NLGN4*, which has been established as causal in X-linked cases of autism and mental retardation in humans and replicated in a murine model of monogenetic autism [121, 122]. Autism is sex-biased towards males, which is common in (partially) X-linked diseases [123]. There are several parallels between *NLGN4* and *PCDH11X*. In terms of protein function the proteins perform a similar task, and both have a close homolog on the Y chromosome. However, the homologs cannot compensate for a deficiency in the X-linked gene [123]. Nevertheless, a mechanism involving the proteins in the response to infection is unclear. Still, a controversial body of evidence exists implicating certain viral infections in causing autistic disorder, analogous to the more accepted concept that infections can trigger immune disorders [124–126]. The Genome-Wide Repository of Associations Between SNPs and Phenotypes (GRASP) reports one eQTL (expression quantitative trait locus) hit for rs5916341 with C19orf6, a transmembrane protein in lymphoblastoid cell lines ($P = 9.8 \times 10^{-8}$) [127].
		rs3072699	
		rs426247	Localizes to an intergenic area closest to the TAK1-Binding Protein gene (*TAB3*) upstream and the Ferritin Heavy Chain-Like 17 gene (*FTHL17*) downstream of the SNP [15]. The TAB3 protein forms part of the TAK1 complex, essential for NF-κB activation and subsequent cytokine expression [128]. Lei et al. show interference in this signaling cascade by enterovirus 71 by cleavage of the TAB proteins, resulting in suppression of the cytokine response [128]. FTHL17, on the other side, is an iron transport protein. Iron withholding is a key factor in defense against bacterial infection, making it a reasonable candidate gene that has previously been identified in an RNAi screen investigating the *in vitro* response of human epithelial cells to the bacterial pathogen *Salmonella enterica* [129]. Thus, both *TAB3* and *FTHL17* are good candidate genes for involvement in infectious disease susceptibility.

Continued

Table 2. Continued

SNP	Variant	SNP	Variant
rs1823897	It maps to a region between *AC112653.1*, an unexplored gene, and *LINC01546*, which is a long non-coding RNA (lncRNA) induced by the Akt/STAT3 pathway and associated with poor prognosis in lung adenocarcinoma [130]. RNA sequencing data show its regular expression pattern specifically in brain and skin [131]. Although the closest coding gene to this SNP is the arylsulfatase F (*ARSF*) gene, Schurz et al. also report *ARSE* as a potential causal gene in the GWAS Catalog [102, 132]. The genes are close homologs that putatively contribute to *Chondrodysplasia punctate*, where the latter gene is considered causative of the bone deformities [133]. Among the possible gene associations, no involvement in infection has been described in published literature for these genes.	rs7064174	Intronic variant in the FERM And PDZ Domain Containing 4 gene (*FRMPD4*) [15]. It encodes a neural scaffolding protein critically involved in glutamatergic signaling and mutations in this gene can cause for X-linked intellectual disability [134].
		rs7053176	Located in a regulatory region [15]. Its most proximate gene is *CXCR3*, which encodes a chemokine receptor expressed on various T-cell subsets and affects their trafficking and function [60, 135]. Because of its role in inflammation, dysfunction of this gene can manifest in autoimmune disorders and more recently was shown to impede the host response to *Salmonella enterica* in murine models [136].
rs73535318	Interacts with two further SNPs. Information about this region is scarce, as the SNP is located between two inferred and still undescribed pseudogenes, *UBE2V1P9* and *RNU6-555P* [137, 138].	rs5991619	An intronic variant rs5991619 in a long non-coding RNA (lncRNA), *AL023574.1*, which appears to be a novel gene [139]. However, fine-mapping efforts link this SNP with the downstream monoamine oxidase A gene (*MAOA*), responsible for lowering serotonin, norepinephrine and dopamine levels in the brain [140]. MAOA deficiency is linked to aggressive and antisocial behavior in mouse models as well as human cases [140, 141]. Expression may be downregulated by infection with Epstein–Barr virus, which enhances metastasis in nasopharyngeal carcinoma, though the tumor-suppressive or promoting effect of MAOA varies in different cancers [142].
		rs5972637	A variant within the *DMD* gene, whose potential involvement in infection has been discussed previously in context with H1N1. Notably, tuberculosis infection is often associated with severe muscle wasting and a study investigating this shared symptom between muscular dystrophies and cancer cachexia suggests that the causal mechanisms of muscle wasting may be shared across several chronic diseases, e.g. HIV, multiple sclerosis, sepsis and tuberculosis [143, 144]

Continued

Table 2. Continued

SNP	Variant	SNP	Variant
rs58762927	Intronic variant in *PCDH11X* (discussed in detail with the SNP rs68046754)	rs4824843	An intergenic SNP rs4824843 located between FUN14 Domain Containing 1 (*FUNDC1*) and a largely undescribed pseudogene *CHTF8P1* [145]. FUNDC1 is a mitochondrial membrane protein crucially involved in mitophagy, a process which removes damaged mitochondria [146, 147]. Disturbances in mitophagy are associated with neurodegenerative and metabolic disorders, and *FUNDC1* overexpression was observed in certain cancers [147–149]. *FUNDC1* expression was increased in tumorigenic cells after bovine papillomavirus infection, and in addition human parainfluenza virus 3 has also been shown to 'hack' the process of mitophagy, though the latter has not been shown to act directly on FUNDC1 [148, 150].
rs4633188		rs2170314	Two epistatic variants both in the intergenic region between *DIAPH2-AS1* (Diaphanous-related formin 2—antisense RNA 1) and the pseudogene *AL354685.1*. In a previous study, a human hepatocyte cell line (PH5CH8) was challenged with IFNβ to investigate the proinflammatory immune response elicited by the cytokine, and *AL354685.1* was among the genes found to be upregulated 12 hours after IFNβ signaling, indicating its potential involvement in the process of inflammation. This increased expression is more pronounced in *IRF1*-deficient cells in the same study, and it is suggested that IRF1 is involved in early antiviral mechanisms and induction of chemokines to recruit immune cells [151]. On the other hand, *DIAPH2-AS1* encodes four lncRNAs, which are thought to often regulate the expression of their gene complement [152, 153]. *DIAPH2* was recently found to be upregulated in latent tuberculosis infection in a separate study [154]. *DIAPH2-AS1* itself is connected with early lymphoid cell activation [63]. Whether there is a direct connection between *DIAPH2-AS1* and tuberculosis remains to be elucidated.

networks [98]. Only one SNP yielded an eQTL hit using HaploReg [99]. Analysis of published literature surrounding the close-by genes revealed a number of good candidate genes for further investigation. However, none of the studies reporting these associations studied cohorts larger than a few hundred individuals, and none have been replicated in populations of the same or different ancestry.

Genetic associations with infectious diseases in the pan-UK biobank

We hypothesized that the low number of associations on the X chromosome with infectious diseases might be in part due to a lack of power from previous studies, as most of the GWAS conducted on infectious diseases had relatively small sample sizes. Thus, we analyzed a large UK-based study: the Pan-UK Biobank. The UK Biobank contains genotypes obtained from roughly half a million UK residents between 40 and 69 years of age, including >20 000 individuals of non-European ancestry [100]. Non-European populations are often omitted from GWAS, thereby putting these populations at a disadvantage in disease-risk analyses based on results from European population studies and overlooking important insights [101]. Using the UK Biobank data on all available populations, the Pan-UK Biobank team conducted GWAS on over 7000 distinct phenotypes and made the data publicly available [25].

We identified 84 phenotypes associated with infectious diseases among over 7000 phenotypes in the Pan-UK Biobank and extracted only the SNPs on X chromosome from each per-phenotype file. Filtering the Pan-UK Biobank data for genome-wide significance ($P \leq 5 \times 10^{-8}$), against expectations, yielded no associations matching the criteria. Due to the sheer size of the study and higher power of analysis among diverse populations, we had anticipated to substantiate some results published in the GWAS Catalog or identify new associations.

For a closer look, we identified the phenotypes comparable to smallpox virus, influenza virus and tuberculosis. Associations with the smallpox virus were not explored in the Pan-UK Biobank, so the following analysis is limited to influenza virus and tuberculosis.

Influenza

Although we found 68 SNPs with suggestive associations ($P \leq 5 \times 10^{-5}$) with influenza viruses in the Pan-UK Biobank (Supplementary Table 3), strains of the influenza virus were not specified, thus the results cannot be compared to the results from the GWAS on influenza A strain H1N1. Additionally, in the study by Garcia-Etxebarria et al., only one SNP () was associated to differences between cases and controls, the other two SNP only became significant after comparing with in mild or severe cases.

Even though the GWAS in the Pan-UK Biobank was performed on a population of European ancestry, same as in the study by Garcia-Etxebarria et al. and the Pan-UK Biobank report the associations for ~900 000 SNP tested, the three SNPs associated to influenza A, subtype H1N1, from the GWAS Catalog were not included in the published results. It could be that the SNPs were excluded during quality control because the three variants have a very low frequency in the general population with minor allele frequency below 1% and they were mainly present in the cases.

Tuberculosis

Among 31 SNPs with suggestive associations ($P \leq 5 \times 10^{-5}$) with tuberculosis identified from the Pan-UK Biobank, none of them overlap with the associations by Schurz et al. (Supplementary Table 3). Although 16 of the 17 associations were tested in the Pan-UK Biobank, only one was significant (rs4633188, $P = 0.02$) in Europeans, two SNPs (rs7064174 and rs426247, $P = 0.01$) were significant in a population of Central/South Asian ancestry.

Since the Pan-UK Biobank did not analyze SNP × SNP interactions, the results are not directly comparable, and it is expected from the study by Schurz et al. that the SNPs individually do not exult a strong effect on tuberculosis pathogenesis. Furthermore, the SNP interactions shown in the GWAS Catalog were identified in an admixed South African population [102]. The Pan-UK Biobank team performed GWAS to tuberculosis on both European and Central/South Asian populations, which may well diverge and hence do not necessary replicate results in other population of other ancestries [101].

Conclusions

Limitations to replicating outcomes of GWAS Catalog using the pan-UK biobank

We have examined the significant genetic associations on the X chromosome with smallpox immune response, severity of influenza virus A infection subtype H1N1, and susceptibility to tuberculosis in detail. Although we look for replication of these results in a larger study, it was not possible due to differences in study design concerning phenotype, population ancestry and methodology.

The size of the study cohort affects the power of the analysis, where smaller studies are more prone to delivering false positives. In this case, we cannot conclude that any associations from the GWAS Catalog were false positives, as the studies differ in several factors impeding comparability. In the study by Garcia-Etxebarria et al., it was intriguing that such a low-frequency variants with such a small cohort were found (49 severe and 107 mild influenza cases against a general population of 549 individuals). Without replication, however, the results should be taken with caution due to the small sample size. As discussed previously, results do not always apply across populations and replication studies should ideally be performed in populations of similar ancestry [101]. Even varying methodology to ascribe individuals to a specific ancestry could diminish comparability between studies. The same concern is valid regarding methodology in quality control, for example accounting for age and sex. In the context of this review, especially the method of analyzing the X chromosome may warp results. Further factors complicating analysis are X chromosome inactivation and possible skewed inactivation.

The enormous complexity of infectious diseases also contributes to the difficulty in replicating GWAS to these diseases. This is supported by the observation that GWAS has been applied less for infectious diseases than 'non-communicable' diseases [103]. Many factors play a role in such complex diseases, including environmental circumstances, genetic predisposition, microbiota, as well as sex differences on the genetic, hormonal and environmental level [2, 3, 21, 22, 103]. Specifically in infectious diseases, variation in the pathogen altering the interaction with the host at the DNA, RNA or protein level must also be taken into account [13, 103, 104]. Among these factors, non-Mendelian heritable traits are the needle in the haystack. Despite the challenges, GWAS are the best approach currently available and should be further pursued.

The role of the X chromosome in infectious diseases warrants further investigation

On the basis of individual associations, summary statistics and background literature, we conclude that the X chromosome remains persistently underexplored in GWAS, despite the availability of the necessary technology. Numerous observations implicating the X chromosome in the host response to infectious diseases mandate its inclusion in large-scale studies, and importantly, the necessity of replication studies including diverse populations. While there is no uniform method of analysis for the X chromosome, a variety of tools is available. As statistically robust analysis becomes possible, we make the case that including the X chromosome in GWAS analysis can be expected to merit the additional effort. To this end, researchers conducting GWAS should be encouraged to include the X chromosome and an analysis pipeline should be developed to ensure ease, clarity and comparability of the studies.

Moreover, most of the current studies that include the X chromosome analyze males and females together, including sex as a covariate to correct for the co-funding factor. Most However, one of the limitations of this approach is that it does not allow the identification of sex-specific effects. Out of the 17 studies with reported associations in the GWAS catalog to infectious diseases on the X chromosome, only one performed a sex-stratified analysis. Schurz *et al.* performed sex-stratified analyses of the X chromosome in three different ways using the XGWAS software [3]. In the first analysis, the authors compared the frequency of alleles between cases and controls to determine if a specific allele co-occurs with a phenotype. In the second, they compared the effect size of a variant between the sexes to determine if a variant has a different effect on the risk between the sexes. In the third, they included X chromosome inactivation states as covariates. The authors showed that the first two approaches were able to identify sex-specific suggestive associations, while the analysis including X-inactivation states resulted in inflated p values and increased the chance of type 1 errors. Interestingly, many of these sex-specific effects showing different directions of the effects.

While the combined approach has been preferred mainly because with a bigger sample size comes bigger power to detect the associations, it is a tradeoff for the identification of sex-specific associations. If the sex-specific association is very strong in one sex, we might be able to capture it with the combined analysis, but this is usually not the case. While we acknowledge that the sample included in infectious diseases analysis is rather small compared to other complex diseases, we propose that additionally to the combined analysis, the authors perform sex-stratified analysis and then combine the results, as this approach has proven to be useful for other complex diseases.

Future perspectives

Different methodologies have been suggested for interpretation of GWAS results, one of which is the integration of the SNPs with quantitative trait locus (QTL) mapping. This type of analysis correlates the SNP genotype with for instance gene expression (eQTL), protein levels (pQTLs), cytokine levels (cQTLs) and methylation levels (meQTLs), allowing us to gain more insight into the disease biology. Unfortunately, most of the QTL studies published in the past years did not include the X chromosome into the analysis, since the underlying GWAS studies often excluded the X chromosome due to its sex-specific characteristics and concomitant technical difficulties. However, QTL studies can be designed to take into account more than just genotype–phenotype correlations [105]. Genotype-by-sex analysis as well as epistatic interactions between QTLs (here in comparison with tuberculosis SNPs) would be particularly well suited for analysis of sex differences caused by the X chromosome.

In a recent analysis of sepsis GWAS, Le et al. demonstrate the added value of QTL mapping and functional data [106]. From 39 independent loci with a suggestive association in the GWAS, they integrated eQTL data from blood samples to identify genes affected by the 39 SNPs. Several of these eQTL genes were then examined for differential expression in the transcriptome of peripheral blood mononuclear cells stimulated by one of three pathogens. As a result, 55 candidate genes were identified from eQTL mapping and correlation with RNA sequencing data. Though cytokines are thought to play an important role in sepsis, cQTL mapping showed that only nine of the 39 loci were shown to affect inflammatory cytokine levels. Genes nearby the 39 identified loci were enriched in the adherence junction pathway, which is an important factor in sepsis for maintaining the integrity of vasculature, and around half of these nearby genes overlapped with the 55 sepsis-associated genes [106]. This neatly illustrates the additional insight into biological effects of GWAS SNPs that can be achieved by incorporating functional study data.

Functional analysis including the X chromosome is gaining traction. Kukurba et al. exemplify eQTL mapping for the X chromosome, where they identified a depletion of *cis*-eQTLs on the X chromosome compared to autosomes and an enrichment of sex-interacting eQTLs on the X chromosome, especially in regulatory functions. Further, they show that the effect size of the same genotype can vary between sexes, prominently observed in autoimmune diseases [107]. This study underlines the importance of sex-stratified analysis and functional mapping to further our understanding of complex disease. Still, studies attempting functional genomics approaches on the X chromosome are scarce and we found none doing so in the context of infectious diseases [107]. Thus, we expect an important role of regulatory elements on the X chromosome. This is significant to QTL analysis since it has been shown that regulatory variants are often context-specific [108].

For functional integration of the GWAS SNPs discussed in this study, we could perform eQTL and cQTL mapping on the X chromosome using the 500FG cohort data by the Human Functional Genomics Project [109]. This cohort is suitable for investigation of disease-susceptibility QTLs since it incorporates a detailed set of data including pathogenic stimulation of immune cells in three distinct populations. The results of this analysis would be very interesting with respect to the possible shared molecular response ensuing in muscle wasting between influenza A and tuberculosis (marked $*$ in Figure 1). In conclusion, research should not only assess the genetic important of X chromosomes in GWAS studies, but also to make sure that it is included in all genomic studies and functional studies, which might lead to a better interpretation of the association results.

Key Points

- Known and tentative involvements of X-linked genes in infectious diseases are reviewed.
- The exclusion of the X chromosome from GWAS limits our understanding of its role in infectious diseases.
- Integrative tools are required to facilitate the inclusion of the X chromosome in large-scale analyses.

Supplementary data

Supplementary data is available at *BRIFUN* Journal online.

Acknowledgements

We thank the UK Biobank for the collection and public availability of the data, as well as the Pan-UK Biobank team for the analysis and again public availability of the results.

Funding

This work was funded by a research grant (2017) from the European Society of Clinical Microbiology and Infectious Diseases, a Radboud Universitair Medical Centrum Hypatia Grant (2018), and a grant from the research programme ZonMw COVID-19 call by the Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie (ZonMw) to V. K.

References

1. Fish EN. The X-files in immunity: sex-based differences predispose immune responses. *Nat Rev Immunol* 2008;**8**:737–44.
2. Kloc M, Ghobrial RM, Kubiak JZ. The role of genetic sex and mitochondria in response to COVID-19 infection. *Int Arch Allergy Immunol* 2020;**181**:629–34.
3. Schurz H, Salie M, Tromp G, *et al*. The X chromosome and sex-specific effects in infectious disease susceptibility. *Hum Genomics* 2019;**13**:2.
4. Libert C, Dejager L, Pinheiro I. The X chromosome in immune functions: when a chromosome makes the difference. *Nat Rev Immunol* 2010;**10**:594–604.
5. Wise AL, Gyi L, Manolio TA. eXclusion: toward integrating the X chromosome in genome-wide association analyses. *Am J Hum Genet* 2013;**92**:643–7.
6. Di Palo A, Siniscalchi C, Salerno M, *et al*. What microRNAs could tell us about the human X chromosome. *Cell Mol Life Sci* 2020;**77**:4069–80.
7. Bianchi I, Lleo A, Gershwin ME, Invernizzi P. The X chromosome and immune associated genes. *J Autoimmun* 2012;**38**:J187–92.
8. Mortensen KH, Cleemann L, Hjerrild BE, *et al*. Increased prevalence of autoimmunity in Turner syndrome – influence of age. *Clin Exp Immunol* 2009;**156**:205–10.
9. Zhang X, Hong D, Ma S, *et al*. Integrated functional genomic analyses of Klinefelter and Turner syndromes reveal global network effects of altered X chromosome dosage. *Proc Natl Acad Sci* 2020;**117**:4864–73.
10. Gendrel A-V, Heard E. Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu Rev Cell Dev Biol* 2014;**30**:561–80.
11. Engreitz JM, Pandya-Jones A, McDonel P, *et al*. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science (New York, NY)* 2013;**341**:1237973.
12. Guo X, Su B, Zhou Z, Sha J. Rapid evolution of mammalian X-linked testis microRNAs. *BMC Genomics* 2009;**10**:97.
13. Pinheiro I, Dejager L, Libert C. X-chromosome-located microRNAs in immunity: might they explain male/female differences? The X chromosome-genomic context may affect X-located miRNAs and downstream signaling, thereby contributing to the enhanced immune response of females. *BioEssays* 2011;**33**:791–802.
14. Accounting for sex in the genome. *Nat Med* 2017;**23**:1243–3.
15. Buniello A, MacArthur J, Cerezo M, *et al*. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;**47**:D1005–12.
16. Trynka G, Hunt KA, Bockett NA, *et al*. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* 2011;**43**:1193–201.
17. Jin Y, Andersen G, Yorgov D, *et al*. Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants. *Nat Genet* 2016;**48**:1418–24.
18. Barrett JC, Clayton DG, Concannon P, *et al*. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;**41**:703–7.
19. Wang K-S, Liu X-F, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res* 2010;**124**:192–9.
20. Nazarian A, Yashin AI, Kulminski AM. Genome-wide analysis of genetic predisposition to Alzheimer's disease and related sex disparities. *Alzheimers Res Ther* 2019;**11**:5.
21. Cebra JJ. Influences of microbiota on intestinal immune system development. *Am J Clin Nutr* 1999;**69**:1046s–51s.
22. Pfeiffer JK, Virgin HW. Transkingdom control of viral infection and immunity in the mammalian intestine. *Science (New York, NY)* 2016;**351**:aad5872.
23. Underhill DM, Pearlman E. Immune interactions with pathogenic and commensal fungi: a two-way street. *Immunity* 2015;**43**:845–58.
24. Thaiss CA, Zmora N, Levy M, Elinav E. The microbiome and innate immunity. *Nature* 2016;**535**:65–74.
25. Pan-UKB team. *Pan-UK Biobank*. https://pan.ukbb.broadinstitute.org/ (19 December 2020, last date accessed).
26. Lim CK, Abolhassani H, Appelberg SK, *et al*. IL2RG hypomorphic mutation: identification of a novel pathogenic mutation in exon 8 and a review of the literature. *Allergy Asthma Clin Immunol* 2019;**15**:2.
27. KEGG PATHWAY: *NF-kappa B signaling pathway*. https://www.genome.jp/kegg-bin/show_pathway?ko04064 (13 January 2021, last date accessed).
28. Braue J, Murugesan V, Holland S, *et al*. NF-κB essential modulator deficiency leading to disseminated cutaneous atypical mycobacteria. *Mediterranean J Hematol and Infectious Dis* 2015;**7**:e2015010.
29. Zhang Y, Tomann P, Andl T, *et al*. Reciprocal requirements for EDA/EDAR/NF-kappaB and Wnt/beta-catenin signaling pathways in hair follicle induction. *Dev Cell* 2009;**17**:49–61.
30. Carè A, Bellenghi M, Matarrese P, *et al*. Sex disparity in cancer: roles of microRNAs and related functional players. *Cell Death Differ* 2018;**25**:477–85.
31. Hazra B, Chakraborty S, Bhaskar M, *et al*. miR-301a regulates inflammatory response to Japanese encephalitis virus infection via suppression of NKRF activity. *J Immunol (Baltimore, Md: 1950)* 2019;**203**:2222–38.
32. Huang K-H, Wang C-H, Lin C-H, Kuo H-P. NF-κB repressing factor downregulates basal expression and mycobacterium tuberculosis induced IP-10 and IL-8 synthesis via interference with NF-κB in monocytes. *J Biomed Sci* 2014;**21**:71.
33. KEGG PATHWAY: *ko04659*. https://www.genome.jp/dbget-bin/www_bget?ko04659 (13 January 2021, last date accessed).
34. Newport MJ, Huxley CM, Huston S *et al*. *A Mutation in the Interferon-γ –Receptor Gene and Susceptibility to*

Mycobacterial Infection. 1996. https://www.nejm.org/doi/10.1056/NEJM199612263352602 (12 January 2021, last date accessed).
35. Dorman SE, Holland SM. Mutation in the signal-transducing chain of the interferon-gamma receptor and susceptibility to mycobacterial infection. *J Clin Investig* 1998;**101**:2364–9.
36. Altare F, Durandy A, Lammas D, *et al*. Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. *Science (New York, NY)* 1998;**280**:1432–5.
37. Jiang J, Cao Z, Shan W, *et al*. 4-1BB expression on MAIT cells is associated with enhanced IFN-γ production and depends on IL-2. *Cell Immunol* 2018;**328**:58–69.
38. Day CL, Walker BD. Progress in defining CD4 helper cell responses in chronic viral infections. *J Exp Med* 2003;**198**:1773–7.
39. Tuovinen EA, Grönholm J, Öhman T, *et al*. Novel Hemizygous IL2RG p.(Pro58Ser) mutation impairs IL-2 receptor complex expression on lymphocytes causing X-linked combined immunodeficiency. *J Clin Immunol* 2020;**40**:503–14.
40. Kawasaki T, Kawai T. Toll-like receptor Signaling pathways. *Front Immunol* 2014;**5**:461.
41. KEGG PATHWAY: *Toll-like receptor signaling pathway*. https://www.genome.jp/kegg-bin/show_pathway?ko04620 (13 January 2021, last date accessed).
42. Mancuso G, Gambuzza M, Midiri A, *et al*. Bacterial recognition by TLR7 in the lysosomes of conventional dendritic cells. *Nat Immunol* 2009;**10**:587–94.
43. Pang IK, Pillai PS, Iwasaki A. Efficient influenza a virus replication in the respiratory tract requires signals from TLR7 and RIG-I. *Proc Natl Acad Sci* 2013;**110**:13910–5.
44. Stegemann-Koniszewski S, Behrens S, Boehme JD, *et al*. Respiratory influenza a virus infection triggers local and systemic natural killer cell activation via toll-like receptor 7. *Front Immunol* 2018;**9**:245.
45. Schurz H, Daya M, Möller M, *et al*. TLR1, 2, 4, 6 and 9 variants associated with tuberculosis susceptibility: a systematic review and meta-analysis. *PLoS One* 2015;**10**:e0139711.
46. Salie M, Daya M, Lucas LA, *et al*. Association of toll-like receptors with susceptibility to tuberculosis suggests sex-specific effects of TLR8 polymorphisms. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 2015;**34**:221–9.
47. Davila S, Hibberd ML, Hari Dass R, *et al*. Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. *PLoS Genet* 2008;**4**:e1000218.
48. Dalgic N, Tekin D, Kayaalti Z, *et al*. Relationship between toll-like receptor 8 gene polymorphisms and pediatric pulmonary tuberculosis. *Dis Markers* 2011;**31**:33–8.
49. Lai Y-F, Lin T-M, Wang C-H, *et al*. Functional polymorphisms of the TLR7 and TLR8 genes contribute to mycobacterium tuberculosis infection. *Tuberculosis (Edinb)* 2016;**98**:125–31.
50. Fallerini C, Daga S, Mantovani S, *et al*. Association of toll-like receptor 7 variants with life-threatening COVID-19 disease in males: findings from a nested case-control study. *elife* 2021;**10**:e67569.
51. Asano T, Boisson B, Onodi F, *et al*. X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. *Science Immunol* 2021;**6**:eabl4348.
52. van der Made CI, Simons A, Schuurs-Hoeijmakers J, *et al*. Presence of genetic variants among young men with severe COVID-19. *JAMA* 2020;**324**:663–73.
53. Solanich X, Vargas-Parra G, van der Made CI, *et al*. Genetic screening for TLR7 variants in young and previously healthy men with severe COVID-19. *Front Immunol* 2021;**12**:2965.
54. Schönrich G, Raftery MJ. The PD-1/PD-L1 Axis and virus infections: a delicate balance. *Front Cell Infect Microbiol* 2019;**9**:207.
55. Chen J, Jiang CC, Jin L, Zhang XD. Regulation of PD-L1: a novel role of pro-survival signalling in cancer. *Ann Oncol* 2016;**27**:409–16.
56. Bazhin AV, von Ahn K, Fritz J, *et al*. Interferon-α up-regulates the expression of PD-L1 molecules on immune cells through STAT3 and p38 Signaling. *Front Immunol* 2018;**9**:2129.
57. Brooks DG, Ha S-J, Elsaesser H, *et al*. IL-10 and PD-L1 operate through distinct pathways to suppress T-cell activity during persistent viral infection. *Proc Natl Acad Sci* 2008;**105**:20428–33.
58. Copland A, Sparrow A, Hart P, *et al*. Bacillus Calmette-Guérin induces PD-L1 expression on antigen-presenting cells via autocrine and paracrine interleukin-STAT3 circuits. *Sci Rep* 2019;**9**:3655.
59. Fang W, Zhang J, Hong S, *et al*. EBV-driven LMP1 and IFN-γ up-regulate PD-L1 in nasopharyngeal carcinoma: implications for oncotargeted therapy. *Oncotarget* 2014;**5**:12189–202.
60. Groom JR, Luster AD. CXCR3 in T cell function. *Exp Cell Res* 2011;**317**:620–31.
61. Syrett CM, Sindhava V, Sierra I, *et al*. Diversity of epigenetic features of the inactive X-chromosome in NK cells, dendritic cells, and macrophages. *Front Immunol* 2019;**9**:3087.
62. Oghumu S, Varikuti S, Stock JC, *et al*. CXCR3 escapes X chromosome inactivation in T cells during infection: potential implications for sex differences in immune responses. *Journal of Immunology (Baltimore, Md: 1950)* 2019;**203**:789–94.
63. Bennett JS, Tredway TL, Dizikes GJ, Nawrocki JF. Properties of a novel gene isolated from a Hodgkin's disease cell line that is expressed early during lymphoid cell activation. *J Immunol (Baltimore, Md: 1950)* 1994;**152**:2229–40.
64. Berletch JB, Yang F, Disteche CM. Escape from X inactivation in mice and humans. *Genome Biol* 2010;**11**:213.
65. Hagen SH, Henseling F, Hennesen J, *et al*. Heterogeneous escape from X chromosome inactivation results in sex differences in type I IFN responses at the single human pDC level. *Cell Rep* 2020;**33**:108485.
66. Zhang Y, Castillo-Morales A, Jiang M, *et al*. Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving. *Mol Biol Evol* 2013;**30**:2588–601.
67. Cano-Gamez E, Trynka G. From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. *Front Genet* 2020;**11**:424.
68. Kennedy RB, Ovsyannikova IG, Pankratz VS, *et al*. Genome-wide analysis of polymorphisms associated with cytokine responses in smallpox vaccine recipients. *Hum Genet* 2012;**131**:1403–21.
69. ANKRD11P2 (ENSG00000231686) - Summary - Homo_sapiens. *Ensembl genome browser 102*. http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000231686;r=X:146619597-146620584;t=ENST00000445173 (8 December 2020, date last accessed).
70. CXorf51B [Homo sapiens]. *NCBI Genes*. https://www.ncbi.nlm.nih.gov/gene/100133053#gene-expression (8 December 2020, date last accessed).

71. Panda D, Gold B, Tartell MA, *et al*. The transcription factor FoxK participates with Nup98 to regulate antiviral gene expression. *MBio* 2015;**6**:e02509–14.
72. McFadden VC, Shalaby RE, Iram S, *et al*. Hepatic deficiency of the pioneer transcription factor FoxA restricts hepatitis B virus biosynthesis by the developmental regulation of viral DNA methylation. *PLoS Pathog* 2017;**13**:e1006239.
73. Choi W, Choe S, Lau GW. Inactivation of FOXA2 by respiratory bacterial pathogens and dysregulation of pulmonary mucus homeostasis. *Front Immunol* 2020;**11**:515.
74. Hodson DJ, Shaffer AL, Xiao W, *et al*. Regulation of normal B-cell differentiation and malignant B-cell survival by OCT2. *Proc Natl Acad Sci U S A* 2016;**113**:E2039–46.
75. Head BP, Olaitan AO, Aballay A. Role of GATA transcription factor ELT-2 and p38 MAPK PMK-1 in recovery from acute P. aeruginosa infection in C. elegans. *Virulence* 2017;**8**:261–74.
76. Fukami M, Wada Y, Miyabayashi K, *et al*. CXorf6 is a causative gene for hypospadias. *Nat Genet* 2006;**38**:1369–71.
77. Francis F, Hennig S, Korn B, *et al*. A gene (PEX) with homologies to endopeptidases is mutated in patients with X–linked hypophosphatemic rickets. *Nat Genet* 1995;**11**:130–6.
78. Silva SRB, Tempone AJ, Silva TP, *et al*. Mycobacterium leprae downregulates the expression of PHEX in Schwann cells and osteoblasts. *Mem Inst Oswaldo Cruz* 2010;**105**:627–32.
79. Garcia-Etxebarria K, Bracho MA, Galán JC, *et al*. No major host genetic risk factor contributed to a(H1N1)2009 influenza severity. *PLoS One* 2015;**10**:e0135983.
80. *The Dystrophin Complex: structure, function and implications for therapy*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4767260/ (8 December 2020, date last accessed).
81. McDonald CM, Meier T, Voit T, *et al*. Idebenone reduces respiratory complications in patients with Duchenne muscular dystrophy. *Neuromuscul Disord* 2016;**26**:473–80.
82. Goody M, Jurczyszak D, Kim C, Henry C. Influenza a virus infection damages zebrafish skeletal muscle and exacerbates disease in zebrafish Modeling Duchenne muscular dystrophy. *PLoS Currents* 2017;**9**: ecurrents.md.8a7e35c50fa2b48156799d3c39788175.
83. Han Q, Lin X, Zhang X, *et al*. WWC3 regulates the Wnt and hippo pathways via dishevelled proteins and large tumour suppressor 1, to suppress lung cancer invasion and metastasis. *J Pathol* 2017;**242**:435–47.
84. Palmer EE, Stuhlmann T, Weinert S, *et al*. De novo and inherited mutations in the X-linked gene CLCN4 are associated with syndromic intellectual disability and behavior and seizure disorders in males and females. *Mol Psychiatry* 2018;**23**:222–30.
85. Proenca CC, Gao KP, Shmelkov SV, *et al*. Slitrks as emerging candidate genes involved in neuropsychiatric disorders. *Trends Neurosci* 2011;**34**:143–53.
86. Santoro MG, Rossi A, Amici C. NF-kappaB and virus infection: who controls whom. *EMBO J* 2003;**22**:2552–60.
87. Wang X-L, Yang L, Chan K-H, *et al*. Age and sex differences in rates of influenza-associated hospitalizations in Hong Kong. *Am J Epidemiol* 2015;**182**:335–44.
88. Steeg LG, Klein SL. Sex and sex steroids impact influenza pathogenesis across the life course. *Semin Immunopathol* 2019;**41**:189–94.
89. Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. *J Leukoc Biol* 2012; **92**:67–73.
90. Ramsuran V, Ewy R, Nguyen H, Kulkarni S. Variation in the untranslated genome and susceptibility to infections. *Front Immunol* 2018;**9**:2046.
91. Shirahama S, Miki A, Kaburaki T, Akimitsu N. Long noncoding RNAs involved in pathogenic infection. *Front Genet* 2020;**11**:454.
92. Oscanoa J, Sivapalan L, Gadaleta E, *et al*. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res* 2020; **48**:W185–92.
93. Dayem Ullah AZ, Oscanoa J, Wang J, *et al*. SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res* 2018;**46**:W109–13.
94. Dayem Ullah AZ, Lemoine NR, Chelala C. A practical guide for the functional annotation of genetic variations using SNPnexus. *Brief Bioinform* 2013;**14**:437–47.
95. Dayem Ullah AZ, Lemoine NR, Chelala C. SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic Acids Res* 2012;**40**:W65–70.
96. Chelala C, Khan A, Lemoine NR. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics* 2009;**25**:655–61.
97. Poliseno L, Salmena L, Zhang J, *et al*. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010;**465**:1033–8.
98. Yoon J-H, Abdelmohsen K, Gorospe M. Functional interactions among microRNAs and long noncoding RNAs. *Semin Cell Dev Biol* 2014;**0**:9–14.
99. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;**40**:D930–4.
100. Sudlow C, Gallacher J, Allen N, *et al*. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779.
101. Sun Q, Graff M, Rowland B, Wen J, Huang L, Miller-Fleming TW, *et al*. Analyses of biomarker traits in diverse UK biobank participants identify associations missed by European-centric analysis strategies. *J Hum Genet* 2021; 1–7.
102. Schurz H, Kinnear CJ, Gignoux C, *et al*. A sex-stratified genome-wide association study of tuberculosis using a multi-ethnic genotyping Array. *Front Genet* 2018;**9**:678.
103. Newport MJ, Finan C. Genome-wide association studies and susceptibility to infectious diseases. *Brief Funct Genomics* 2011;**10**:98–107.
104. Sen K. Relevance of Pseudogenes to Human Genetic Disease. In: *ELS*. John Wiley & Sons Ltd, Chichester, 2013.
105. Miles C, Wayne M. Quantitative trait locus (QTL) analysis. *Nature Education* 2008;**1**:208.
106. Le KTT, Matzaraki V, Netea MG, *et al*. Functional annotation of genetic loci associated with sepsis prioritizes immune and endothelial cell pathways. *Front Immunol* 2019;**10**: 1949.
107. Kukurba KR, Parsana P, Balliu B, *et al*. Impact of the X chromosome and sex on regulatory variation. *Genome Res* 2016;**26**:768–77.
108. Alasoo K, Rodrigues J, Mukhopadhyay S, *et al*. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat Genet* 2018;**50**:424–31.
109. 500 Functional Genomics Project. *Human Functional Genomics Project*. http://www.humanfunctionalgenomics.org/site/?page_id=82 (24 January 2021, date last accessed).

110. Yoshida K, Sugano S. Identification of a novel protocadherin gene (PCDH11) on the human XY homology region in Xq21.3. *Genomics* 1999;**62**:540–3.
111. Williams NA, Close JP, Giouzeli M, Crow TJ. Accelerated evolution of Protocadherin 11X/Y: a candidate gene-pair for cerebral asymmetry and language. *Am J Med Genet B Neuropsychiatr Genet* 2006;**141B**:623–33.
112. Veerappa AM, Saldanha M, Padakannaya P, Ramachandra NB. Genome-wide copy number scan identifies disruption of PCDH11X in developmental dyslexia. *Am J Med Genet B Neuropsychiatr Genet* 2013;**162B**:889–97.
113. Lopes AM, Ross N, Close J, *et al*. Inactivation status of PCDH11X: sexual dimorphisms in gene expression levels in brain. *Hum Genet* 2006;**119**:267–75.
114. Xie W, Agniel D, Shevchenko A, *et al*. Genome-wide analyses reveal gene influence on HIV disease progression and HIV-1C Acquisition in Southern Africa. *AIDS Res Hum Retrovir* 2017;**33**:597–609.
115. Laurikkala J, Pispa J, Jung H-S, *et al*. Regulation of hair follicle development by the TNF signal ectodysplasin and its receptor Edar. *Development* 2002;**129**:2541–53.
116. Redler S, Brockschmidt FF, Tazi-Ahnini R, *et al*. Investigation of the male pattern baldness major genetic susceptibility loci AR/EDA2R and 20p11 in female pattern hair loss. *Br J Dermatol* 2012;**166**:1314–8.
117. Chang B, Punj V, Shindo M, Chaudhary PM. Adenoviral-mediated gene transfer of ectodysplasin-A2 results in induction of apoptosis and cell-cycle arrest in osteosarcoma cell lines. *Cancer Gene Ther* 2007;**14**:927–33.
118. Gene: AL031073.1 (ENSG00000250625) - Summary - Homo_sapiens. *Ensembl genome browser 102*. http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000250625;r=X:141282138-141283276;t=ENST00000503029 (9 December 2020, date last accessed).
119. Song X, Hao J, Wang J, *et al*. The cancer/testis antigen MAGEC2 promotes amoeboid invasion of tumor cells by enhancing STAT3 signaling. *Oncogene* 2017;**36**:1476–86.
120. Gene: AC110995.1 (ENSG00000236120) - Summary - Homo_sapiens - Ensembl genome browser 102. *Ensembl genome browser 102*. http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000236120;r=X:5653418-5726326 (9 December 2020, date last accessed).
121. Laumonnier F, Bonnet-Brilhault F, Gomot M, *et al*. X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet* 2004;**74**:552–7.
122. Jamain S, Radyushkin K, Hammerschmidt K, *et al*. Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci U S A* 2008;**105**:1710–5.
123. Nguyen TA, Wu K, Pandey S, *et al*. A cluster of autism-associated variants on X-linked NLGN4X functionally resemble NLGN4Y. *Neuron* 2020;**106**:759–768.e7.
124. Libbey JE, Sweeten TL, McMahon WM, Fujinami RS. Autistic disorder and viral infections. *J Neurovirol* 2005;**11**:1–10.
125. Vigli D, Palombelli G, Fanelli S, *et al*. Maternal immune activation in mice only partially recapitulates the autism Spectrum disorders symptomatology. *Neuroscience* 2020;**445**:109–19.
126. Arango M-T, Shoenfeld Y, Cervera R, Anaya J-M. *Infection and Autoimmune Diseases*. El Rosario University Press, Bogotá, Colombia, 2013.
127. HaploReg v4.1 variant detail view rs5916341. http://pubs.broadinstitute.org/mammals/haploreg/detail_v4.1.php?query=&id=rs5916341 (14 January 2021, date last accessed).
128. Lei X, Han N, Xiao X, *et al*. Enterovirus 71 3C inhibits cytokine expression through cleavage of the TAK1/TAB1/TAB2/TAB3 complex. *J Virol* 2014;**88**:9830–41.
129. Thornbrough JM, Hundley T, Valdivia R, Worley MJ. Human genome-wide RNAi screen for host factors that modulate intracellular salmonella growth. *PLoS One* 2012;**7**:e38097.
130. Tian H, Lian R, Li Y, *et al*. AKT-induced lncRNA VAL promotes EMT-independent metastasis through diminishing Trim16-dependent Vimentin degradation. *Nat Commun* 2020;**11**:5127.
131. LINC01546 Gene Expression. *NCBI Genes*. https://www.ncbi.nlm.nih.gov/gene/100129464/?report=expression&bioproject=PRJEB4337 (11 December 2020, date last accessed).
132. HaploReg v4.1 variant detail view rs1823897. http://pubs.broadinstitute.org/mammals/haploreg/detail_v4.1.php?query=&id=rs1823897 (11 December 2020, date last accessed).
133. Franco B, Meroni G, Parenti G, *et al*. A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embryopathy. *Cell* 1995;**81**:15–25.
134. Piard J, Hu J-H, Campeau PM, *et al*. FRMPD4 mutations cause X-linked intellectual disability and disrupt dendritic spine morphogenesis. *Hum Mol Genet* 2018;**27**:589–600.
135. Erhardt A, Wegscheid C, Claass B, *et al*. CXCR3 deficiency exacerbates liver disease and abrogates tolerance in a mouse model of immune-mediated hepatitis. *J Immunol* 2011;**186**:5284–93.
136. Chami B, Yeung A, Buckland M, *et al*. CXCR3 plays a critical role for host protection against salmonellosis. *Sci Rep* 2017;**7**:10181.
137. UBE2V1P9 UBE2V1 pseudogene 9 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene?cmd=search&term=100131981 (11 December 2020, date last accessed).
138. RNU6-555P RNA, U6 small nuclear 555, pseudogene - Gene - GTR - NCBI. https://www.ncbi.nlm.nih.gov/gtr/genes/106479813/ (11 December 2020, date last accessed).
139. Gene: AL023574.1. *GWAS Catalog*. https://www.ebi.ac.uk/gwas/genes/AL023574.1 (11 December 2020, date last accessed).
140. Shih JC, Chen K, Ridd MJ. MONOAMINE OXIDASE: from genes to behavior. *Annu Rev Neurosci* 1999;**22**:197–217.
141. Holz N, Boecker R, Buchmann AF, *et al*. Evidence for a sex-dependent MAOA× childhood stress interaction in the neural circuitry of aggression. *Cereb Cortex* 2016;**26**:904–14.
142. Lee HM, Sia APE, Li L, *et al*. Monoamine oxidase a is down-regulated in EBV-associated nasopharyngeal carcinoma. *Sci Rep* 2020;**10**:6115.
143. Villamor E, Saathoff E, Mugusi F, *et al*. Wasting and body composition of adults with pulmonary tuberculosis in relation to HIV-1 coinfection, socioeconomic status, and severity of tuberculosis. *Eur J Clin Nutr* 2006;**60**:163–71.
144. Berardi E. Muscular dystrophies and cancer cachexia: similarities in chronic skeletal muscle degeneration. *Journal of Functional Morphology and Kinesiology* 2017;**2**:39.
145. Gene: CHTF8P1 (ENSG00000233543) - *Summary - Homo_sapiens - Ensembl genome browser 102*. http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000233543;r=X:44633692-44635232;t=ENST00000414660 (19 December 2020, date last accessed).
146. Ding W-X, Yin X-M. Mitophagy: mechanisms, pathophysiological roles, and analysis. *Biol Chem* 2012;**393**:547–64.
147. Liu L, Feng D, Chen G, *et al*. Mitochondrial outer-membrane protein FUNDC1 mediates hypoxia-induced

mitophagy in mammalian cells. *Nat Cell Biol* 2012;**14**: 177–85.
148. Roperto S, Russo V, De Falco F, *et al*. FUNDC1-mediated mitophagy in bovine papillomavirus-infected urothelial cells. *Vet Microbiol* 2019;**234**:51–60.
149. Yuan Q, Sun N, Zheng J, *et al*. Prognostic and immunological role of FUN14 domain containing 1 in pan-cancer: friend or foe? *Front Oncol* 2020;**9**:1502.
150. Zhang L, Qin Y, Chen M. Viral strategies for triggering and manipulating mitophagy. *Autophagy* 2018;**14**: 1665–73.
151. Forero A. Differential activation of the transcription factor IRF1 underlies the distinct immune responses elicited by type I and type III interferons. *Immunity* 2019;**51**: 451–64.
152. Gene: DIAPH2-AS1 (ENSG00000236256) - *Summary - Homo_sapiens - Ensembl genome browser 102*. http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000236256;r=X:97431286-97642589 (19 December 2020, date last accessed).
153. Luther HP. Role of endogenous antisense RNA in cardiac gene regulation. *J Mol Med (Berl)* 2005;**83**:26–32.
154. Alam A, Imam N, Ahmed MM, *et al*. Identification and classification of differentially expressed genes and network meta-analysis reveals potential molecular signatures associated with tuberculosis. *Front Genet* 2019;**10**:932.