Funded by the European Union

UNDINE	
Project number:	101057100
Project name:	The human genetic and immunological determinants of the clinical manifestations of SARS-CoV-2 infection: Towards personalised medicine
Topic:	HORIZON-HLTH-2021-DISEASE-04-07
Type of action:	HORIZON Research and Innovation Actions (RIA)
Starting date of action:	1 June 2022
Project duration:	48 months
Project end date:	31 May 2026
Deliverable number:	D2.3
Deliverable title:	List of genetic variants influencing immune responses to SARS-CoV-2
Document version:	Ver1.0
WP number:	WP2
Lead beneficiary:	IP
Main author(s):	Lluis Quintana-Murci (IP), Stefan Pöhlmann (DZP)
Internal reviewers:	Trine Mogensen (AU)
Nature of deliverable:	DATA
Dissemination level:	PU
Delivery date from Annex 1:	M48
Actual delivery date:	M48

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Executive Summary

The COVID-19 pandemic has revealed striking clinical variability among individuals infected with SARS-CoV-2, shaped by age, sex, comorbidities, and human genetics. However, the extent and determinants of variation in immune responses across individuals and populations remain incompletely understood. To investigate the genetic basis of immune response variability to SARS-Cov-2 across populations, ages, and sexes, we integrated single-cell transcriptomics with eQTL mapping across multiple immune cell types. We identify thousands of regulatory variants, including many influencing antiviral responses, and show that their effects are highly cell- and viral stimulus-dependent. Most regulatory variants governing responses to SARS-CoV-2 are shared with those involved in responses to IAV, with the notable exception of the myeloid compartment, which displays a distinct virus-specific regulatory architecture. We further show that SARS-CoV-2-specific regulatory variants are enriched for signatures of positive selection in East Asian populations, suggesting historical exposure to coronavirus-related selective pressures. Finally, we identify sex-linked mechanisms contributing to immune variability, including enhanced antiviral activity in females associated with escape from X-chromosome inactivation, and altered IFN signalling in males linked to mosaic loss of the Y chromosome with age.

Abbreviations

IAV	Influenza A virus
eQTL	Expression quantitative trait loci
reQTL	Response expression quantitative trait loci
XCI	X-chromosome inactivation
mLoY	Mosaic loss of the Y-chromosome

Contents

List of Figures

1 List of genetic variants influencing immune responses to SARS-CoV-2

1.1 Background

The COVID-19 pandemic has revealed considerable clinical variation among individuals infected with SARS-CoV-2, ranging from asymptomatic or mild disease to fatal outcome (1). Multiple factors have been shown to contribute to this variability, including advanced age, male sex, comorbidities, and human genetics (2, 3). Specifically, inborn errors of or auto-antibodies against type I interferons (IFNs) have been identified as strong contributors to the severity of COVID-19 pneumonia (4). In parallel, genome-wide association studies (GWAS) have identified more than 50 loci associated with infection susceptibility to SARS-CoV-2 or COVID-19 severity, further highlighting the central role of human genetic variation in the pathophysiology of the disease (3, 5). Despite the unprecedented progress made by immunological and genetic research during the pandemic, several important questions remain unresolved. These include the identification of the genetic determinants underlying variation in immune responses to SARS-CoV-2, the extent to which these determinants are shared with or specific to other respiratory RNA viruses, the evolutionary pressures that may have shaped this genetic variation across human populations, and the mechanisms through which such variation contributes to differences in disease susceptibility and severity.

1.2 Results

The objective of this deliverable was to characterise the impact of human genetic variation, both within and between populations, on transcriptional responses to SARS-CoV-2 ex vivo through single-cell expression quantitative trait loci (eQTL) mapping. To achieve this, we leveraged: (i) the dataset generated from individuals of Central African, West European, and East Asian ancestry (n=222) (see Deliverable D2.1), and (ii) the dataset obtained from individuals of European ancestry stratified by age and sex (n=384) (see Deliverable D2.2). In both cohorts, PBMCs were exposed to SARS-CoV-2 and, for comparative purposes, to IAV.

1.2.1 Genetic determinants of population-level variation in immune responses

To investigate the contribution of human genetic diversity to variation in antiviral immune responses across populations, we mapped eQTLs across 22 immune cell types using single-cell transcriptomic

data (6). This analysis identified more than 12,000 eQTLs, approximately 10% of which displayed population-specific effects. Particular attention was given to response eQTLs (reQTLs), defined as variants modulating transcriptional responses following viral stimulation. In total, 1,505 reQTLs affecting 1,213 genes were identified. Comparative analyses of responses to SARS-CoV-2 and IAV revealed that the majority of reQTLs were shared across immune cell lineages, supporting the existence of a broadly conserved genetic architecture governing responses to RNA viruses (**Fig. 1a**). However, substantially greater virus specificity was observed within the myeloid compartment, where 49% of the identified reQTLs exhibited stimulus-specific effects. An illustrative example is provided by a strong SARS-CoV-2-specific reQTL detected in myeloid cells at the *MMP1* locus (**Fig. 1b**), a gene encoding a biomarker of COVID-19 severity (7). Collectively, these analyses reveal that the effects of virus-induced regulatory variants are highly cell-type dependent and highlight the virus specificity of the genetic basis of the myeloid response.

To further dissect the sources of population-level differences in immune responses, mediation analyses were performed to distinguish the relative contributions of immune-cell composition and genetic regulation. These analyses showed that variation in cellular composition constitutes a major determinant of population differences at the global transcriptional level, whereas genetic effects, although affecting a smaller number of genes, exerted stronger locus-specific impacts. Depending on the cellular context, genetic regulation accounted for population differences affecting between 141 and 433 genes, corresponding to 32-58% of the observed transcriptional variation. Together, these findings indicate that population variation in immune responses is driven largely by cellular heterogeneity, but genetic variants with marked allele frequency variation contribute to population differences at specific loci.

Additional analyses explored the evolutionary and clinical relevance of these regulatory variants. Variants associated with SARS-CoV-2-specific responses were significantly enriched for signatures of positive selection exclusively in individuals of East Asian ancestry (**Fig. 1c**). This finding supports the notion that populations of East-Asian ancestry have been exposed for longer to coronavirus-like infections and aligns with the proposed timing (25,000 years ago) of an epidemic that affected the evolution of host coronavirus-interacting proteins (8). Finally, we investigated the contributions of genetic variants altering responses to SARS-CoV-2 ex vivo to COVID-19 risk in vivo by determining whether eQTLs were more strongly associated with COVID-19 GWAS hits (5, 9) than random,

matched SNPs. Colocalization analysis identified 30 loci where disease-associated signals colocalized with regulatory variants influencing gene expression (**Fig. 1d**), supporting a direct role for regulatory genetic variation in modulating to COVID-19 susceptibility and severity.

1.2.2 Genetic determinants of age- and sex-related variation in immune responses

This objective aimed to characterize the mechanisms underlying age- and sex-associated differences in immune responses to SARS-CoV-2, with a particular focus on genetic variation involving the sex chromosomes. Specifically, we investigated the contribution of X-chromosome inactivation (XCI) escape in females and mosaic loss of the Y chromosome (mLoY) in males (**Fig. 2a,b**). Analyses of XCI escape revealed stronger escape patterns within lymphoid cell populations, with no significant association with age. Among X-linked genes displaying both XCI escape and induction following viral stimulation, expression levels were strongly correlated with interferon-stimulated gene (ISG) activity. Notably, genes such as *DDX3X* were associated with enhanced antiviral transcriptional responses, supporting the hypothesis that increased dosage of specific X-linked genes contributes to stronger antiviral immunity in females.

The analysis of mLoY revealed a markedly different cellular distribution. Mosaic Y chromosome loss was most frequently observed in myeloid cells and NK cells, reaching frequencies of up to 30% in monocytes from individuals older than 55 years. Although mLoY was generally uncommon within lymphoid populations, regulatory T cells (Tregs), particularly memory Tregs, exhibited unexpectedly elevated frequencies, including among younger individuals. Specifically, mLoY was detected in approximately 17% of memory Tregs and 12% of naïve Tregs. Functional analyses further showed that, within the same individual and cell type, cells exhibiting mLoY displayed enhanced IFN responses and enrichment for antiviral and innate immune defence pathways. These findings suggest that mLoY influence immune-cell function and contribute to age-related variation in antiviral responses.

Finally, genome-wide interaction analyses identified genetic variants whose regulatory effects varied according to age or sex. Across the different cell types and stimulation conditions examined, we identified 15 genotype-by-age (G×Age) interaction eQTLs and 4 genotype-by-sex (G×Sex) interaction eQTLs, further supporting the contribution of host genetic factors to age- and sex-dependent variation in immune responses to SARS-CoV-2.

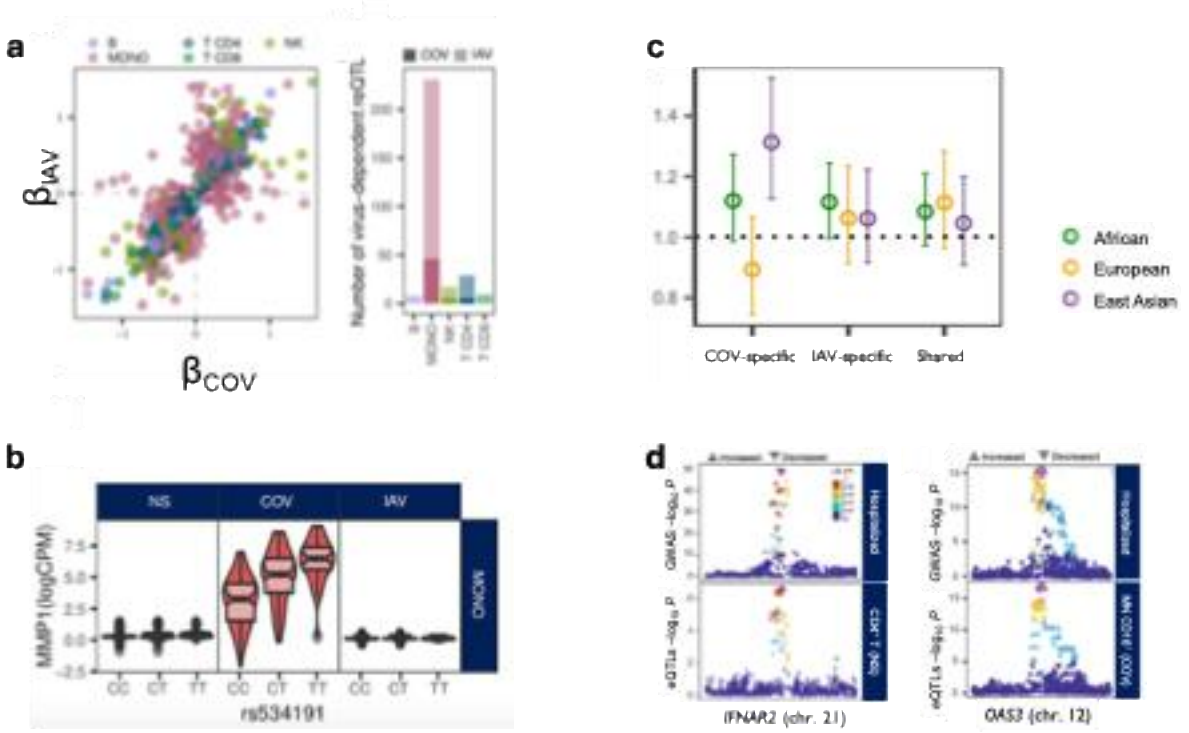

Figure 1. Genetic determinants of variation in transcriptional responses to SARS-CoV-2.
(a) Comparison of reQTL effect sizes (β) between SARS-CoV-2- and IAV-stimulated cells. Each dot represents a specific reQTL (defined by a SNP, gene, and immune lineage), and colours indicate the lineage in which the reQTL was detected (right panel).
(b) Example of a SARS-CoV-2-specific reQTL at the *MMP1* locus. The associated SNP shows no detectable effect on gene expression at baseline (NS) or following IAV stimulation, whereas a significant genotype-dependent effect is observed in monocytes following SARS-CoV-2 stimulation.
(c) Enrichment analyses of signatures of natural selection among reQTLs. Significant enrichment was observed only for SARS-CoV-2-specific reQTLs identified in individuals of East Asian ancestry.
(d) Colocalization of *IFNAR2* and *OAS3* eQTLs with loci associated with COVID-19 severity. Top panels show the −log10(P) association profiles for COVID-19-related hospitalization. Bottom panels show the −log10(P) association profiles for gene expression in non-stimulated CD4+ T cells (*IFNAR2*) and SARS-CoV-2-stimulated CD16+ monocytes (*OAS3*). Colours indicate the degree of linkage disequilibrium (LD; r^2) with the consensus SNP identified through colocalization analyses (purple). Arrows indicate the direction of the allelic effect for each SNP. Chr., chromosome.

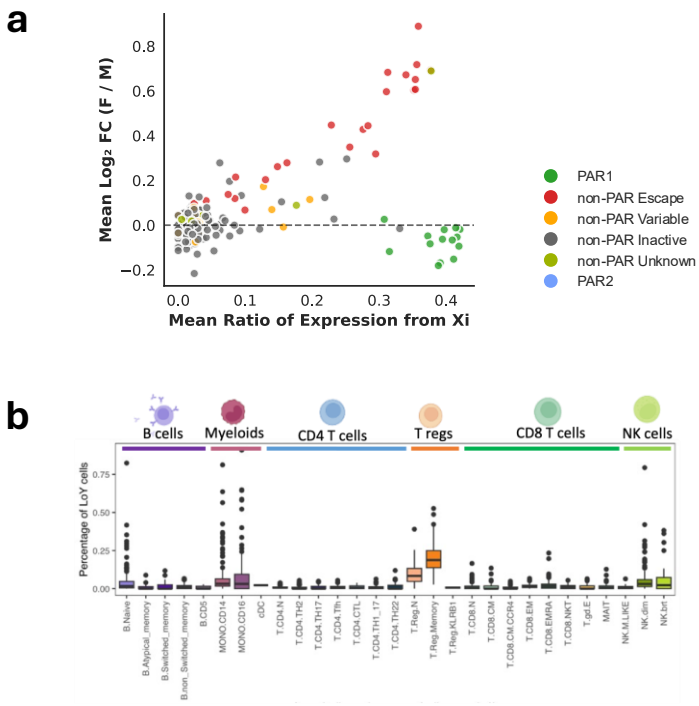

Figure 2. Contribution of sex chromosomes to variation in age- and sex-dependent immune responses to SARS-CoV-2.
(a) Expression levels of X-linked genes escaping X-chromosome inactivation (XCI) plotted against the fold-change difference in expression between females and males. A substantial proportion of genes located within the pseudoautosomal region 1 (PAR1; green) escape XCI but do not contribute substantially to sex differences in expression. In contrast, XCI-escaping genes located outside the pseudoautosomal regions (red) contribute to differential gene expression between males and females.
(b) Proportion of cells exhibiting mosaic loss of the Y chromosome (mLoY) across multiple immune cell types.

2 Conclusion

Using single-cell technologies and quantitative genetic approaches, we mapped the genetic basis of leukocyte responses to SARS-CoV-2 across populations of diverse ancestry, ages, and sexes. We demonstrate that population genetic variation has a substantial impact on immune response heterogeneity, alongside sex-chromosome–linked mechanisms that contribute to differences between males and females and across age groups.

These findings establish host genetic variation as a key driver of both inter-individual and population-level differences in immune responses to SARS-CoV-2. They advance our understanding of the genetic determinants underlying COVID-19 susceptibility and severity and provide a broader framework for investigating host genetic factors shaping responses to emerging viral pathogens.

3 References

1. O'Driscoll M, Ribeiro Dos Santos G, Wang L, Cummings DAT, Azman AS, et al. 2021. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* 590: 140-5
2. Zhang Q, Bastard P, Effort CHG, Cobat A, Casanova JL. 2022. Human genetic and immunological determinants of critical COVID-19 pneumonia. *Nature* 603: 587-98
3. Cobat A, Zhang Q, Effort CHG, Abel L, Casanova JL, Fellay J. 2023. Human Genomics of COVID-19 Pneumonia: Contributions of Rare and Common Variants. *Annu Rev Biomed Data Sci* 6: 465-86
4. Bastard P, Zhang Q, Zhang SY, Jouanguy E, Casanova JL. 2022. Type I interferons and SARS-CoV-2: from cells to organisms. *Curr Opin Immunol* 74: 172-82
5. COVID-19 Host Genetics Initiative. 2023. A second update on mapping the human genetic architecture of COVID-19. *Nature* 621: E7-E26
6. Aquino Y, Bisiaux A, Li Z, O'Neill M, Mendoza-Revilla J, et al. 2023. Dissecting human population variation in single-cell responses to SARS-CoV-2. *Nature* 621: 120-8
7. Syed F, Li W, Relich RF, Russell PM, Zhang S, et al. 2021. Excessive Matrix Metalloproteinase-1 and Hyperactivation of Endothelial Cells Occurred in COVID-19 Patients and Were Associated With the Severity of COVID-19. *J Infect Dis* 224: 60-9
8. Souilmi Y, Lauterbur ME, Tobler R, Huber CD, Johar AS, et al. 2021. An ancient viral epidemic involving host coronavirus interacting genes more than 20,000 years ago in East Asia. *Curr Biol* 31: 3704
9. COVID-19 Host Genetics Initiative. 2021. Mapping the human genetic architecture of COVID-19. *Nature* 600: 472-7