Sex-biased expression of the *TLR7* gene in severe COVID-19 patients: Insights from transcriptomics and epigenomics

A. Gómez-Carballa [a,b,c], J. Pardo-Seco [a,b,c], S. Pischedda [a,b,c], I. Rivero-Calle [a,b,d], G. Butler-Laporte [e,f], J.B. Richards [e,f], S. Viz-Lasheras [a,b,c], F. Martinón-Torres [a,b,d], A. Salas [a,b,c,*], GENDRES network

[a] *Genetics, Vaccines and Infections Research Group (GENVIP), Instituto de Investigación Sanitaria (IDIS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain*
[b] *Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBER-ES), Madrid, Spain*
[c] *Unidade de Xenética, Instituto de Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, and GenPoB Research Group, Instituto de Investigación Sanitaria (IDIS), Hospital Clínico Universitario de Santiago (SERGAS), Galicia, Spain*
[d] *Translational Pediatrics and Infectious Diseases, Department of Pediatrics, Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain*
[e] *Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, Canada*
[f] *Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Québec, Canada*

ARTICLE INFO

ABSTRACT

There is abundant epidemiological data indicating that the incidence of severe cases of coronavirus disease (COVID-19) is significantly higher in males than females worldwide. Moreover, genetic variation at the X-chromosome linked *TLR7* gene has been associated with COVID-19 severity. It has been suggested that the sex-biased incidence of COVID-19 might be related to the fact that *TLR7* escapes X-chromosome inactivation during early embryogenesis in females, thus encoding a doble dose of its gene product compared to males. We analyzed *TLR7* expression in two acute phase cohorts of COVID-19 patients that used two different technological platforms, one of them in a multi-tissue context including saliva, nasal, and blood samples, and a third cohort that included different post-infection timepoints of long-COVID-19 patients. We additionally explored methylation patterns of *TLR7* using epigenomic data from an independent cohort of COVID-19 patients stratified by severity and sex. In line with genome-wide association studies, we provide supportive evidence indicating that *TLR7* has altered CpG methylation patterns and it is consistently downregulated in males compared to females in the most severe cases of COVID-19.

1. Introduction

It has been reported that women have a stronger immune system than men, which generally leads to a more efficient response to infections and vaccines (Zuk, 2009; Cordoba-Aguilar and Munguia-Steyer, 2013; Fink and Klein, 2018); conversely, women also suffer significantly much more from autoimmune diseases, such as rheumatoid arthritis, multiple sclerosis, etc (Youness et al., 2021). A complex interaction of chromosome X-linked genetic factors coupled with sex-specific hormones and environmental factors, may be involved in this phenomenon. For instance, it has been claimed that altered (X-linked) Toll-Like Receptor (TLR) 7 gene (*TLR7*) expression is involved in several autoimmune disorders (Brown et al., 2022). In addition, deficiencies in *TLR7*, together with other inborn errors of type 1 interferons (IFNs), have been found in several studies on infectious diseases (Asano et al. 2021).

TLR7 has also received much attention in the context of the coronavirus disease 2019 (COVID-19). This gene is essential for host defense against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, as its protein plays a fundamental role in pathogen recognition and activation of innate immunity (Petes et al., 2017). *TLR7* encodes for an endosomal innate immune sensor capable to detect single-stranded RNA (ssRNA) of the virus, inducing the production of type 1 IFN and the release of other proinflammatory cytokines. It has

been suggested that insufficient type 1 IFN immunity in the respiratory tract during the initial days of the infection may be critical for the virus spread, potentially constituting a decisive factor leading to pulmonary and systemic inflammation (Zhang et al., 2022).

At the same time, it is also known that severe COVID-19 has a marked male sex bias, as also found in other viral infections (Guerra-Silveira and Abad-Franch, 2013; Wang et al., 2022). Thus, mortality and hospitalization rates are higher in male COVID-19 patients (Spiering and de Vries, 2021). It has been suggested that the *TLR7* gene might play a role in this differential sex incidence of COVID-19. Most of the evidence comes from genomic studies. The first studies pointing to this link were small-scale both family- and population-based genetic association studies carried out in selected severe COVID-19 patients (van der Made et al., 2020; Asano et al. 2021; Fallerini et al., 2021). The role of *TLR7* in COVID-19 pathogenesis has found further support in more recent large-scale Whole-Exome and Whole-Genome Sequencing studies (WES and WGS, respectively). Thus, for instance, the WES association study of >580K COVID-19 patients of the COVID-19 Host Genetics Initiative revealed several genes of marginal genome-wide significance. One of these genes was *TLR7*, which emerged as a good association candidate when contrasting COVID-19 severe patients *versus* COVID-19 negative or unknown donors (P-value $= 4.28 \times 10^{-8}$ in a burden association test) (Kosmicki et al., 2021). The most definitive evidence of this association has been reported in a recent study that combined WES and WGS from 624K COVID-19 patients including a much larger number of severe cases (Butler-Laporte et al., 2022); these authors showed that carrying a rare predicted loss of function or missense variant in the *TLR7* gene was associated with a 5.3-fold increase in odds of severe disease (P-value $= 5.41 \times 10^{-7}$); this study indicated that deleterious variation at this gene could potentially affect both, males and females. This finding found strong evidence with exome-wide significance in the discovery phase, and solid replication in a large independent cohort (Butler-Laporte et al., 2022).

It has been proposed that the rationale behind the role of *TLR7* in the differential sex incidence of COVID-19 may derive from the phenomenon of escape from X-chromosome inactivation, already described in the early 1960s (Lyon, 1962). To avoid double dosage of X-linked chromosomal genes in females compare to males, one of the female X-chromosomes is randomly silenced during early steps of embryonic development, thereby generating a mosaicism of activated X-chromosomes in females (X-chromosome inactivation or XCI). However, it seems that some X-linked genes, especially those in the pseudoautosomal regions (PARs) of pairing between the X and Y chromosomes, escape from this inactivation (escape from XCI). This phenomenon is not necessarily complete, it has inter-individual variability, and it also can differ between tissues (Spiering and de Vries, 2021).

However, little effort has been so far devoted to understanding the gene expression patterns of *TLR7* and their connection with different COVID-19 severe outcomes, as well as its presumed relationship with the reported sex-bias in this disease.

To shed light on this issue, we analyzed the sex differential transcription pattern of the *TLR7* gene in three recently published COVID-19 cohorts. Samples from two of these cohorts were selected to study *TLR7* expression during acute phase of infection: one dataset originated from a RNAseq study (Jackson et al., 2022), and the other one used the *n*-Counter (Nanostring) platform (Gómez-Carballa et al., 2022). The first study was carried out on blood samples extracted from COVID-19 patients (total $n = 69$), while the second one was carried out on blood, saliva, and nasopharyngeal samples (total $n = 52$); the two datasets overlap partially ($n = 23$ blood samples from the same timepoint), with this overlap serving as a technical validation. The third independent gene expression cohort, also stratified by sex and severity, was used to investigate the *TLR7* expression during COVID-19 convalescence and comes from a follow-up study also including long-COVID-19 patients (Ryan et al., 2022). In addition, we investigated the contribution of *TLR7* whole blood methylation patterns to the severity sex-bias observed

in COVID-19 patients by exploring a recent epigenomic study ($n = 473$) (Barturen et al., 2022).

2. Methods

First, we used gene expression data from two previous studies carried out on acute phase COVID-19 patients stratified by sex and severity. We have selected these two cohorts due to the accessibility to full raw expression data as well as to availability of complete clinical data; therefore allowing a homogeneous and acurate classification of patients by severity (ICU non-ICU and WHO score). The first one used RNAseq data (Jackson et al., 2022) while the second one used *n*-Counter data (Gómez-Carballa et al., 2022); the two cohorts overlap in 23 patients. Samples were collected during acute phase of infection. Processing and normalization of raw data generated were carried out as in the original articles (Gómez-Carballa et al., 2022; Jackson et al., 2022). More details about both cohorts, samples sizes, and clinical data associated are provided in Table S1.

The classification of patients into ICU and non-ICU categories or the computation of the WHO score of severity refers to the time of sample collection. We used a Wilcoxon test to assess statistical significance between severity groups, and a Pearson test for the correlation indices (r) and P-values (after checking the data for normality). Log$_2$FC was calculated by subtracting the arithmetic mean of Log$_2$ counts of gene A in the patient sample, and the arithmetic mean of Log$_2$ counts of gene A in the reference sample.

Next, raw counts data and metadata from the COVID-19 post infection RNAseq study (12-, 16-, and 24-weeks post infection) were downloaded from the link provided by the authors in the original publication (Ryan et al., 2022); only convalescence samples from critical patients (severity classification during the acute phase) were used in the present study. This cohort was the only one available including follow-up blood whole transcriptome data from COVID-19 patients stratified by severity in the acute phase of the infection. Data normalization was carried out using Deseq2 R package (Love et al., 2014) with default parameters and after removing low expressed genes (genes with sum of counts lower than five in all the samples).

In addition, DNA methylation data were obtained from a recent epigenomic study (Barturen et al., 2022) including a total of 473 COVID-19 positive patients. Whole blood samples were recruited during the acute phase of the infection, and according to the clinical outcome, they were subdivided into two groups: severe ($n = 113$) and mild patients ($n = 360$). Severity classification was made by their authors considering a WHO scale (mild: 1–4 scale, severe: 5–8 scale). Their WHO score division was comparable to the one used in the mentioned gene expression studies (WHO 1–3 *vs.* 4–8); a full unification of the scores was not possible due to the unavailability of clinical data in the later study (in contrast to the data in the expression studies that were generated by our group). Raw data were downloaded, filtered, and normalized using *minfi* (Aryee et al., 2014) package. To identify differentially methylated positions (DMPs) between the phenotype groups, the *limma* package (Ritchie et al., 2015) was used assuming a linear model adjusted for age.

Sexual chromosomes representation and locations of immune-related genes was generated using *KaryoploteR* R package (Gel and Serra, 2017). We used the R package *Enrichmentbrowser* (Geistlinger et al., 2016) to collect genes involved in immune response processes from the Gene Ontology (GO; http://geneontology.org) database. Graphing and statistical analyses were carried out using R software v4.1.1 (www.r-project.org).

3. Results

3.1. Systemic TLR7 gene expression and methylation patterns

We analyzed the transcription patterns of *TLR7* in COVID-19 patients stratified by sex and severity. The RNAseq data from Jackson et al.

(2022) indicate that *TLR7* expresses almost equally in male and females in the non-ICU group of patients, but it is downexpressed in males with respect to females in the ICU group (although the difference is not statistically significant; Fig. 1A). This pattern is also suggested when stratified patients by WHO severity score (Fig. 1A). Notably, there is a very weak positive correlation (not statistically significant) between *TLR7* expression and the WHO score in samples within the score range 1–3 (r = 0.11; *P*-value = 0.59) but this correlation is strongly negative and highly statistically significant for the more severe patients with WHO score values ≥ 4 (r = −0.47; *P*-value = 0.0071); Fig. 1B.

The data generated in the *n*-Counter study by Gómez-Carballa et al. (2022) showed that, for the overlapping samples with the RNAseq dataset, there is a significant correlation of gene expression values for the *TLR7* gene (r = 0.82; *P*-value = 1.9×10^{-6}; Fig. 1C). In agreement with the RNAseq data above, the *n*-Counter data also indicate that *TLR7* expresses less in males than females in the most severe patients. Although this difference is more subtle and non-statistically significant when considering the WHO score, it is clearly statistically significant when using the ICU *vs.* non-ICU criterion (*P*-value = 0.004); Fig. 1A.

Finally, in a comparison of ICU *vs.* non-ICU males, *TLR7* is downexpressed in severe male patients in both the RNAseq (*P*-value = 0.0054) and the *n*-Counter (*P*-value = 0.0045) datasets (Fig. S1A).

DNA methylation analysis of *TLR7* on blood samples revealed the presence of nine CpGs within the gene, three of them located within the gene body, while the remaining six were annotated to its promoter region (TSS200, TSS1550, 5′UTR). The nine positions did not show the same trend of methylation for the whole gene. When we interrogated the methylation status of these nine positions in females and males, separately, we observed not significant DMPs in females when comparing the severe group (WHO ≥ 5) against the mild group (WHO < 5); however, when the same analysis was performed in males, a statistically significant difference was observed in four out of the nine positions (Fig. S2). Three of these DMPs (cg08029608: *P*-value = 0.005; cg09542796: *P*-value = 0.0001; cg22859180: *P*-value = 0.003) showed a hypermethylation pattern in severe males, while one DMP (cg24735671: *P*-value = 0.045) presented an opposite methylation status, with higher methylation levels in mild patients.

3.2. *TLR7* gene expression in nasal and saliva samples

TLR7 expression has also been analyzed in saliva and nasal samples from COVID-19 patients (Gómez-Carballa et al., 2022) stratified by severity and sex.

It is again remarkable that *TLR7* is always downregulated in males with respect to females in both nasal and saliva samples and independently of the severity criteria used (Fig. 1D).

Thus, the *TLR7* expression in the comparison males *vs.* females is statistically significant in saliva samples in both non-ICU (*P*-value = 6.7×10^{-6}) and ICU patients (*P*-value = 0.022). This difference is also notable and in the same direction when using the WHO score, although it is only statistically significant in non-ICU patients (*P*-value = 0.036, Fig. 1D).

Females and males do not differ significantly in *TLR7* expression from nasal samples, but the data always show a decreasing expression pattern in males when compared to females, and this difference is more marked when considering the WHO score (although this comparison is particularly affected by a low sample size).

3.3. *TLR7* expression, admission days, days of symptoms and convalescence

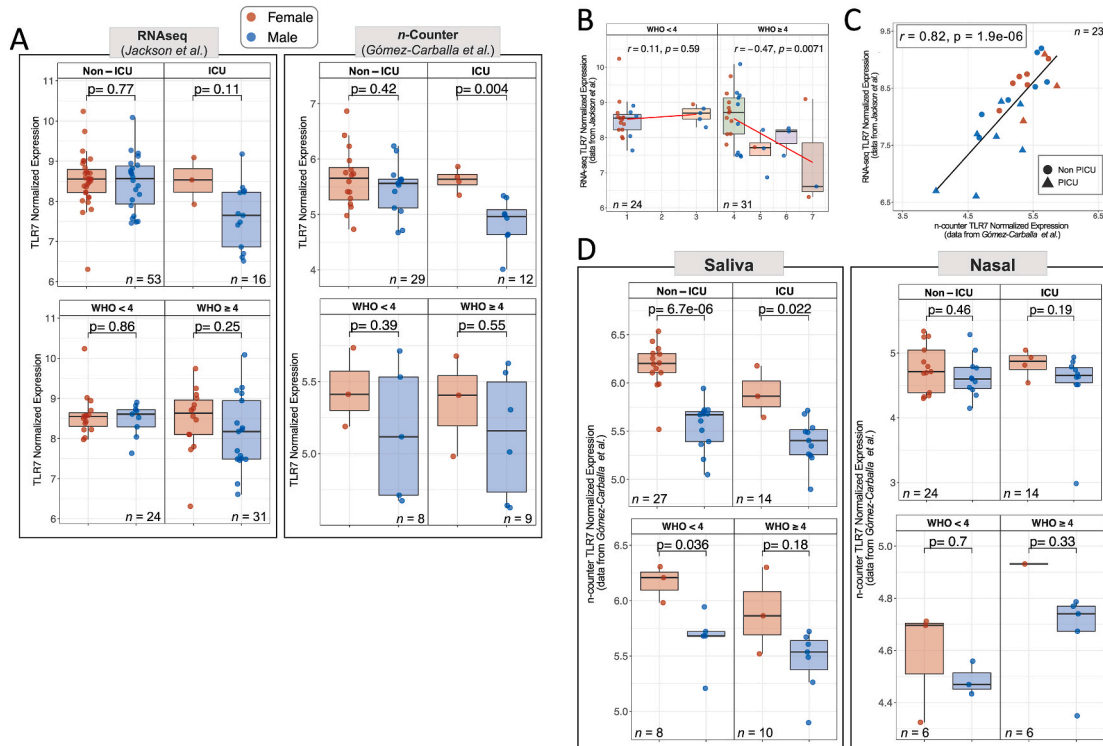We also analyzed the expression of *TLR7* in the context of

Fig. 1. (A) *TLR7* gene expression in blood samples from COVID-19 patients using two datasets, RNAseq from (Jackson et al., 2022) and *n*-Counter from (Gómez-Carballa et al., 2022), and stratifying by sex and severity: (*i*) ICU *vs.* non-ICU (*n* = 69 from (Jackson et al., 2022), and *n* = 41 from (Gómez-Carballa et al., 2022)) and (*ii*) WHO score (*n* = 55 from (Jackson et al., 2022), and *n* = 17 (Gómez-Carballa et al., 2022)). (B) Correlation between *TLR7* expression (*n* = 55) and the WHO score using RNAseq data (Jackson et al., 2022); *P*-values are also provided. (C) Correlation between *TLR7* expression in the RNAseq (Jackson et al., 2022) and the *n*-Counter (Gómez-Carballa et al., 2022) datasets (*n* = 23). (D) *TLR7* gene expression in nasal and saliva samples in the *n*-Counter dataset (Gómez-Carballa et al., 2022) stratified by sex and severity: ICU *vs.* non-ICU (*n* = 41 in saliva and *n* = 38 in nasal samples) and WHO score (*n* = 18 in saliva and *n* = 12 in nasal samples).

hospitalization days and days of symptoms to sample collection. It is most remarkable that expression of *TLR7* is negatively correlated with days from the onset of symptoms (whether hospitalized or not) to sample collection but only in ICU patients (r = −0.43; Fig. 2A). This negative correlation is even more clear when considering only those patients for which there is information on the WHO score (r = −0.66; Fig. S1B). A similar correlation is observed when using the WHO score as severity parameter, although yielding a more discrete value (r = −0.14; Fig. 2B).

TLR7 expression is also negatively correlated with admission days until sample collection (Fig. 2C); despite being a moderate correlation value, it is in the limit of the nominal statistical significance (r = −0.32; *P*-value = 0.059).

Additionally, we investigated how *TLR7* expression changes at different convalescence times after infection in patients classified as critical in the acute phase of the disease. Even though results are not statistically significant (probably owing to the small sample size), we observed that *TLR7* is again downexpressed in the acute phase of infection of the most critical males, and this *TLR7* reduction is maintained over time in convalescence samples, even 24 weeks post infection (Fig. 2 D).

4. Discussion

It has been reported that several genes involved in the regulation of the innate and adaptive immune response are strategically placed on the X-chromosome (Fig. 3A), including pathogen-related receptors (PRRs) such as the *TLR7*, *ACE2* and *TMPRSS2*. These and other genes have been found to be associated in COVID-19 in family-based, genome-wide and NGS population-based association studies. We examined expression patterns of *TLR7* because of its singularity of being located at a XCI escape region and, therefore, having (theoretically) a predominantly 'biallelic' behavior in females but 'monoallelic' in males (Fig. 3A).

Although genetic variation at this gene has been noted several times in genomic COVID-19 association studies, its gene expression patterns in whole blood, saliva and nasal samples has not been analyzed so far.

Our data indicate that *TLR7* is downregulated in blood from males with respect to females in the most severe cases of COVID-19 when stratifying the cohort in those attending ICU *vs.* non-ICU, and when using WHO scores for severity. Consistently, a statistically significant downregulation of *TLR7* occurs in males with a more severe disease (ICU) when compared to non-ICU males (who show similar *TLR7* expression levels as females). Moreover, it seems that *TLR7* expresses in a sex-biased manner in saliva and nasal samples. Note that, while in blood from ICU patients, *TLR7* expresses less in males than in females, in saliva *TLR7* follows this pattern independently of the severity classification used. A trend towards increased *TLR7* expression in females than in males is also observed when examaning Log$_2$FC values and median *TLR7* expression differences in females *vs.* males, in both, ICU and WHO ≥ 4 and in all tissues analyzed (Fig. 3B).

Therefore, the *TLR7* expression in severe females was always higher than in severe males, independently of the tissue and severity classification used, with blood being the tissue showing the largest expression differences. Even though not very large expression changes, our results indicate that *TLR7* may play a role in the sex-bias observed in the more severe outcomes of COVID-19.

Two of the three gene expression datasets used in the present study overlap in a few blood samples, and this fact offered an opportunity to assess the power of both RNAseq and *n*-Counter transcriptomic techniques to capture the real gene expression of the *TLR7* gene. The results of these two analyses correlate significantly, therefore providing a technical validation of the expression data.

In addition, we examined methylation patterns of *TLR7* in COVID-19 patients. We found four (statistically significant) DMPs in *TLR7* when comparing severe males (WHO ≥ 5) *vs.* males with mild disease (WHO < 5) but not in females. Interestingly, while three of them were
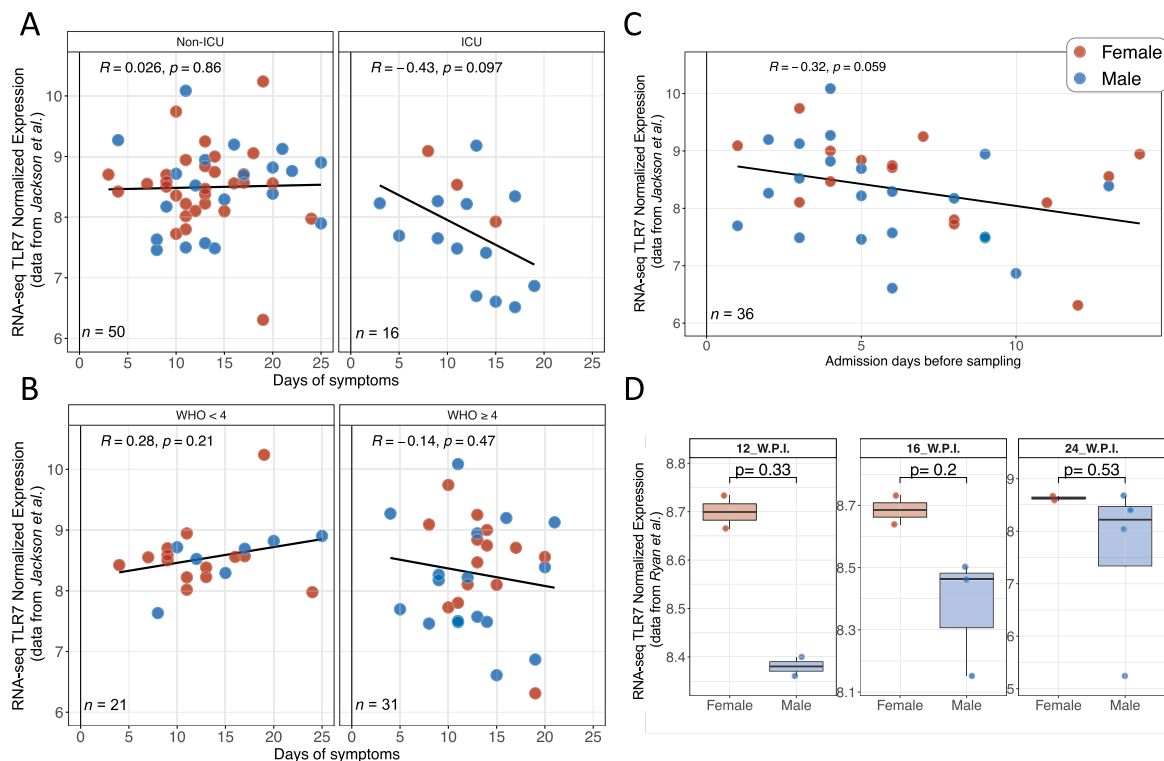

Fig. 2. Correlation between *TLR7* gene expression using RNAseq data from (Jackson et al., 2022) and days of symptoms (asymptomatic patients were not included) since disease onset to sampling using (A) ICU (*n* = 66) and (B) WHO classification (*n* = 52); and (C) admission days since hospitalization to sampling (only hospitalized samples were included; *n* = 36). (D) *TLR7* expression in blood convalescence samples from critical COVID patients collected 12 w.p.i (*n* = 4), 16 w.p.i (*n* = 5) and 24 w.p.i (*n* = 6) (w.p.i = weeks post infection) from the study by (Ryan et al., 2022).
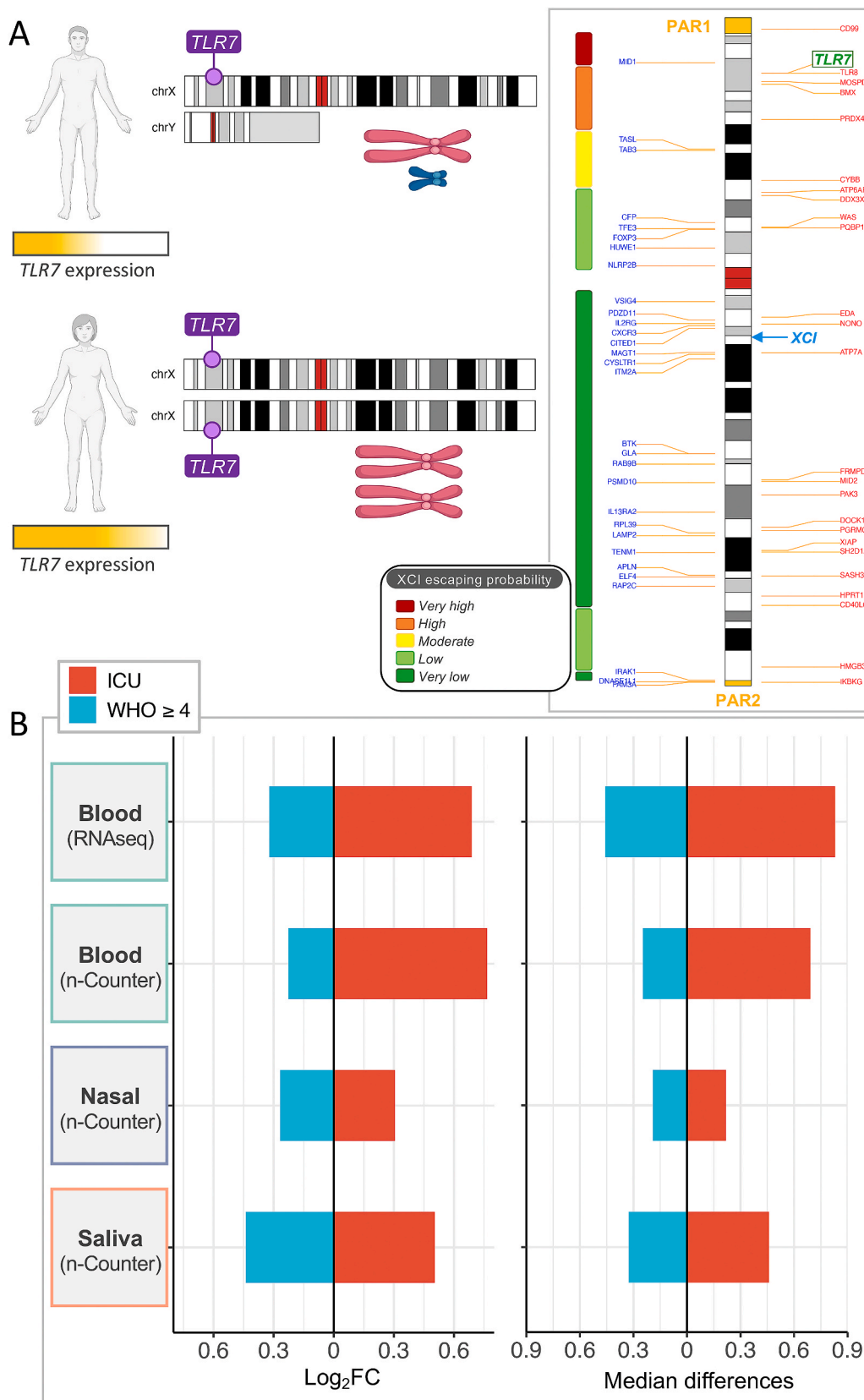
Fig. 3. (A) Schematic representation of sex chromosomes (left); X-chromosome location of *TLR7* and main immune-related genes is also displayed (gene names in red indicate location at positive strand, and in blue at negative strand); XCI escaping probability regions were obtained from (Schurz et al., 2019) (right). Male and female figures as well as illustrations of X- Y-chromosomes were downloaded from Biorender (https://biorender.com/). (B) *TLR7* expression changes in severe females with respect to severe males (for both criteria, ICU and WHO ≥ 4) in all tissues analyzed from RNA-seq (Jackson et al., 2022) and *n*-Counter (Gómez-Carballa et al., 2022) data expressed as Log$_2$FC and median differences. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

hypermethylated in severe males, the other one showed higher methylation in male patients with a mild phenotype. This might suggest that the interaction between methylation and demethylation of *TLR7* gene could be contributing to COVID-19 severity condition in males. The coexistence of both hypo- and hypermethylated sites in the same gene region has not been fully understood (He et al., 2019). The 'dogma' defining the classical correlation between DNA methylation and gene expression asserts that when methylation occurs at the promoter region, the expression of the gene is inhibited or downregulated (Baylin, 2005). Actually, whole-genome studies have revealed that the effect of DNA methylation is not restricted to the promoter regions, but epigenetic changes occurring in low CpG density regions (such as gene bodies), could also impact on the regulation of gene expression (Jones, 2012). With the data available, it is not possible to assess to what extent, the different patterns of methylation observed in the *TLR7* gene (in the promoter region [three DMPs] and in the gene body [one DMP]), could contribute to a global downregulation of the gene in blood (as showed by the transcriptomic data).

The data examined in the present study are consistent in the different cohorts analyzed, using different technical procedures for gene expression measurement (RNAseq and *n*-Counter from Nanostring) and different '-omic' approaches (transcriptomics and epigenomics). The patterns observed are in line to expectations according to genomic association studies. However, many questions remain unanswered regarding the role of *TLR7* in COVID-19. For instance, as inferred from large-scale genome studies and the gene expression results of the present study, *TLR7* alone cannot fully explain severity in COVID-19 patients. For instance, this gene seems to be upregulated in a minor proportion of females who also have a severe outcome. This observation would be in good agreement with the fact that variation at the deleterious variants in *TLR7* is related to severity in both, males and females (Butler-Laporte et al., 2022).

There are three main limitations in the present study. First, the relatively small sample size of the studies, which might explain the lack of statistical significance of a few pairwise comparisons. Second, *TLR7* gene expression was measured in whole blood, nasal and saliva samples showing certain tissue variability with respect to sex differences and severity; therefore, its expression in other tissues that are directly involved in the immune response to SARS-CoV-2 infection (i.e. lower respiratory tract tissues) remains to be analyzed. Third, our study does not allow to investigate the level of *TLR7* expression in different cell populations in e.g blood. Finally, other external or internal factors might be modulating *TLR7* expression to some extent, including the different treatments given to patients with a more severe course of the disease (especially those from ICU) or sexual hormone levels, which were recently associated to an increased type 1 IFN response in females (Webb et al., 2018). Instead, the main strength of the study comes from the consistent trend seen in these comparisons, which basically indicates a more limited gene expression of *TLR7* in severely affected males (even in the convalescence phase) and a negative correlation with admission days and days of symptoms.

To the best of our knowledge, this is the first study investigating expression levels of *TLR7* in different COVID-19 scenarios and tissues. The findings suggest that its gene expression could (at least in part) explain the sex-biased incidence in severe patients. Further studies are necessary to investigate more thoroughly the relationship between *TLR7* expression, sex, and COVID-19 severity.

Funding sources

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data from four other studies were used. We did not generated new data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2022.114288.

References

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., Irizarry, R.A., 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30, 1363–1369.

Asano, T., Boisson, B., Onodi, F., Matuozzo, D., Moncada-Velez, M., Maglorius, Renkilaraj, MRL, Zhang, Meertens, P., Bolze, L., Materna, A., et al., 2021. X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. Sci Immunol 6.

Barturen, G., Carnero-Montoro, E., Martinez-Bueno, M., Rojo-Rello, S., Sobrino, B., Porras-Perales, O., Alcantara-Dominguez, C., Bernardo, D., Alarcon-Riquelme, M.E., 2022. Whole blood DNA methylation analysis reveals respiratory environmental traits involved in COVID-19 severity following SARS-CoV-2 infection. Nat. Commun. 13, 4597.

Baylin, S.B., 2005. DNA methylation and gene silencing in cancer. Nat. Clin. Pract. Oncol. 2 (Suppl. 1), S4–S11.

Brown, G.J., Canete, P.F., Wang, H., Medhavy, A., Bones, J., Roco, J.A., He, Y., Qin, Y., Cappello, J., Ellyard, J.I., et al., 2022. TLR7 gain-of-function genetic variation causes human lupus. Nature 605, 349–356.

Butler-Laporte, G., Povysil, G., Kosmicki, J.A., Cirulli, E.T., Drivas, T., Furini, S., Saad, C., Schmidt, A., et al., 2022. Exome-wide Association Study to Identify Rare Variants Influencing COVID-19 Outcomes: Results from the Host Genetics Initiative. *PLoS Genet* In press.

Cordoba-Aguilar, A., Munguia-Steyer, R., 2013. The sicker sex: understanding male biases in parasitic infection, resource allocation and fitness. PLoS One 8, e76246.

Fallerini, C., Daga, S., Mantovani, S., Benetti, E., Picchiotti, N., Francisci, D., Paciosi, F., Schiaroli, E., Baldassarri, M., Fava, F., et al., 2021. Association of Toll-like receptor 7 variants with life-threatening COVID-19 disease in males: findings from a nested case-control study. Elife 10, e67569.

Fink, A.L., Klein, S.L., 2018. The evolution of greater humoral immunity in females than males: implications for vaccine efficacy. Curr Opin Physiol 6, 16–20.

Geistlinger, L., Csaba, G., Zimmer, R., 2016. Bioconductor's EnrichmentBrowser: seamless navigation through combined results of set- & network-based enrichment analysis. BMC Bioinf. 17, 45.

Gel, B., Serra, E., 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33, 3088–3090.

Gómez-Carballa, A., Rivero-Calle, I., Pardo-Seco, J., Gómez-Rial, J., Rivero-Velasco, C., Rodriguez-Nunez, N., Barbeito-Castineiras, G., Pérez-Freixo, H., Cebey-López, M., Barral-Arca, R., et al., 2022. A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity. Environ. Res. 210, 112890.

Guerra-Silveira, F., Abad-Franch, F., 2013. Sex bias in infectious disease epidemiology: patterns and processes. PLoS One 8, e62390.

He, L., Khanal, P., Morse, C.I., Williams, A., Thomis, M., 2019. Differentially methylated gene patterns between age-matched sarcopenic and non-sarcopenic women. J Cachexia Sarcopenia Muscle 10, 1295–1306.

Jackson, H., Rivero-Calle, I., Broderick, C., Habgood-Coote, D., D'Souza, G., Nichols, S., Vito, O., Gómez-Rial, J., Rivero-Velasco, C., Rodríguez-Núñez, N., et al., 2022. Characterisation of the blood RNA host response underpinning severity in COVID-19 patients. Sci. Rep. 12, 12216.

Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.

Kosmicki, J.A., Horowitz, J.E., Banerjee, N., Lanche, R., Marcketta, A., Maxwell, E., Bai, X., Sun, D., Backman, J.D., Sharma, D., et al., 2021. Pan-ancestry exome-wide association analyses of COVID-19 outcomes in 586,157 individuals. Am. J. Hum. Genet. 108, 1350–1355.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.

Lyon, M.F., 1962. Sex chromatin and gene action in the mammalian X-chromosome. Am. J. Hum. Genet. 14, 135–148.

Petes, C., Odoardi, N., Gee, K., 2017. The toll for trafficking: Toll-like receptor 7 delivery to the endosome. Front. Immunol. 8, 1075.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. *Limma* powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47.

Ryan, F.J., Hope, C.M., Masavuli, M.G., Lynn, M.A., Mekonnen, Z.A., Yeow, A.E.L., Garcia-Valtanen, P., Al-Delfi, Z., Gummow, J., Ferguson, C., et al., 2022. Long-term perturbation of the peripheral immune system months after SARS-CoV-2 infection. BMC Med. 20, 26.

Schurz, H., Salie, M., Tromp, G., Hoal, E.G., Kinnear, C.J., Moller, M., 2019. The X chromosome and sex-specific effects in infectious disease susceptibility. Hum. Genom. 13, 2.

Spiering, A.E., de Vries, T.J., 2021. Why females do better: the X chromosomal *TLR7* gene-dose effect in COVID-19. Front. Immunol. 12, 756262.

van der Made, C.I., Simons, A., Schuurs-Hoeijmakers, J., van den Heuvel, G., Mantere, T., Kersten, S., van Deuren, R.C., Steehouwer, M., van Reijmersdal, S.V., Jaeger, M., et al., 2020. Presence of genetic variants among young men with severe COVID-19. JAMA 324, 663–673.

Wang, C., Lashua, L.P., Carter, C.E., Johnson, S.K., Wang, M., Ross, T.M., Ghedin, E., Zhang, B., Forst, C.V., 2022. Sex disparities in influenza: a multiscale network analysis. iScience 25, 104192.

Webb, K., Peckham, H., Radziszewska, A., Menon, M., Oliveri, P., Simpson, F., Deakin, C. T., Lee, S., Ciurtin, C., Butler, G., et al., 2018. Sex and pubertal differences in the Type 1 Interferon pathway associate with both X chromosome number and serum sex hormone concentration. Front. Immunol. 9, 3167.

Youness, A., Miquel, C.H., Guery, J.C., 2021. Escape from X chromosome inactivation and the female predominance in autoimmune diseases. Int. J. Mol. Sci. 22.

Zhang, Q., Bastard, P., Effort, C.H.G., Cobat, A., Casanova, J.L., 2022. Human genetic and immunological determinants of critical COVID-19 pneumonia. Nature 603, 587–598.

Zuk, M., 2009. The sicker sex. PLoS Pathog. 5, e1000267.