

RESEARCH

Open Access



A genome-wide association study of survival in patients with sepsis

Tamara Hernandez-Beeftink^{1,2}, Beatriz Guillen-Guio^{1,3}, Jose M. Lorenzo-Salazar⁴, Almudena Corrales^{1,5}, Eva Suarez-Pajes¹, Rui Feng⁶, Luis A. Rubio-Rodríguez⁴, Megan L. Paynton³, Raquel Cruz⁷, M. Isabel García-Laorden^{2,5}, Miryam Prieto-González⁸, Aurelio Rodríguez-Pérez^{9,10}, Demetrio Carriedo¹¹, Jesús Blanco^{4,12}, Alfonso Ambrós¹³, Elena González-Higueras¹⁴, Elena Espinosa¹⁵, Arturo Muriel¹², Eduardo Tamayo^{16,17}, María M. Martín¹⁸, Leonardo Lorente¹⁹, David Domínguez¹⁵, Abelardo García de Lorenzo²⁰, Heather M. Giannini²¹, John P. Reilly²¹, Tiffanie K. Jones²¹, José M. Añón²⁰, Marina Soro²², Ángel Carracedo^{7,23,24}, Louise V. Wain^{3,25}, Nuala J. Meyer²¹, Jesús Villar^{2,5,26} and Carlos Flores^{1,4,5,27*} on behalf of the Genetics of Sepsis (GEN-SEP) Network

Abstract

Background: Sepsis is a severe systemic inflammatory response to infections that is accompanied by organ dysfunction and has a high mortality rate in adult intensive care units. Most genetic studies have identified gene variants associated with development and outcomes of sepsis focusing on biological candidates. We conducted the first genome-wide association study (GWAS) of 28-day survival in adult patients with sepsis.

Methods: This study was conducted in two stages. The first stage was performed on 687 European sepsis patients from the GEN-SEP network and 7.5 million imputed variants. Association testing was conducted with Cox regression models, adjusting by sex, age, and the main principal components of genetic variation. A second stage focusing on the prioritized genetic variants was performed on 2,063 ICU sepsis patients (1362 European Americans and 701 African-Americans) from the MESSI study. A meta-analysis of results from the two stages was conducted and significance was established at $p < 5.0 \times 10^{-8}$. Whole-blood transcriptomic, functional annotations, and sensitivity analyses were evaluated on the identified genes and variants.

Findings: We identified three independent low-frequency variants associated with reduced 28-day sepsis survival, including a missense variant in *SAMD9* (hazard ratio [95% confidence interval] = 1.64 [1.37–6.78], $p = 4.92 \times 10^{-8}$). *SAMD9* encodes a possible mediator of the inflammatory response to tissue injury.

Interpretation: We performed the first GWAS of 28-day sepsis survival and identified novel variants associated with reduced survival. Larger sample size studies are needed to better assess the genetic effects in sepsis survival and to validate the findings.

Keywords: Sepsis, Genome-wide association study, Genomics, 28 days, Outcome

Introduction

Sepsis is defined as a severe systemic inflammatory response to infections that is accompanied by organ dysfunction [1]. It is recognized as a global priority and its incidence in adults is estimated at approximately 189 cases per 100,000 people per year [2, 3]. In the intensive

*Correspondence: cflores@ull.edu.es

¹ Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Carretera del Rosario S/N, Santa Cruz de Tenerife, Spain
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

care units (ICUs), sepsis is associated with an overall mortality rate of about 30% [4, 5] and with significant morbidity in survivors. The risk of death in patients with sepsis increases with hemodynamic instability (i.e., septic shock) or due to respiratory complications, such as the acute respiratory distress syndrome (ARDS) [6]. Given the lack of specific therapeutic options and the underlying etiological complexity, multiple studies have focused on improving prevention, diagnosis, and prognosis of sepsis [5].

Genetic variation influences the host immune response to microbial agents [7–10]. In this sense, the genome-wide association studies (GWAS) have an enormous potential to reveal genetic factors for disease susceptibility, severity, and/or survival, as it has been shown for several infectious diseases [10–14]. However, the number of GWAS of sepsis or its complications is limited. To date, only four GWAS have been completed for sepsis mortality, although the likelihood of death for each patient over time was not considered [15–18]. Specifically, Rautanen and colleagues analyzed 28-day mortality in patients with pneumonia, linking the FER tyrosine kinase (*FER*) gene variation with reduced risk for death from sepsis [15]. Nevertheless, another study was unable to replicate this finding in independent patients [19]. Additionally, Scherag and colleagues found that the top ranking variants associated with 28-day mortality from sepsis in their study were located in the Vacuolar Protein Sorting 13 Homolog A (*VPS13A*) gene [16]. Rosier and colleagues identified variants within the Cytokine Inducible SH2 Containing Protein (*CISH*) gene associated with mortality due to septic shock at day 7 or day 28 [17]. Finally, D'Urso and colleagues performed a GWAS of susceptibility and mortality in septic shock and polygenic risk score (PRS) analysis to assess the genetic overlap between septic shock risk/mortality with clinically relevant traits. In the 28-day mortality GWAS analysis, they found association of an intronic variant in the Collagen Type IV Alpha 2 Chain (*COL4A2*) gene [18].

Based on all the above-mentioned evidence, here we performed a GWAS of 28-day sepsis survival to identify novel genetic variants associated with sepsis outcome taking into account the probability estimates of death for each patient over time.

Methods

Study design and participants

We performed a GWAS of 28-day survival in patients with sepsis. This study was conducted in two stages. The first stage was based on two cohorts of patients from the GEN-SEP study [20], where association results were meta-analyzed and used to prioritize variants (Fig. 1). In the second stage, we followed up these variants in

independent sepsis patients of European and African-American ancestry from the MESSI study [21]. Finally, a meta-analysis of results of the 2750 patients (1121 non-survivors) from the two stages was done and genome-wide significance was established at $p < 5.0 \times 10^{-8}$.

The GEN-SEP study, used for the first stage, included 475 (146 non-survivors) European ancestry patients considered for a previous GWAS of sepsis-associated ARDS [20] recruited between 2002 and 2015 for which follow-up records of 28-day survival were available (1st GEN-SEP period). In a second round of patient recruitment, between 2016 and 2019, another 212 (35 non-survivors) patients were included in the GEN-SEP study (2nd GEN-SEP period) (Fig. 1) (see supplementary material for further details). Sepsis diagnosis was clinically defined according to the Third International Consensus Definitions for Sepsis [1].

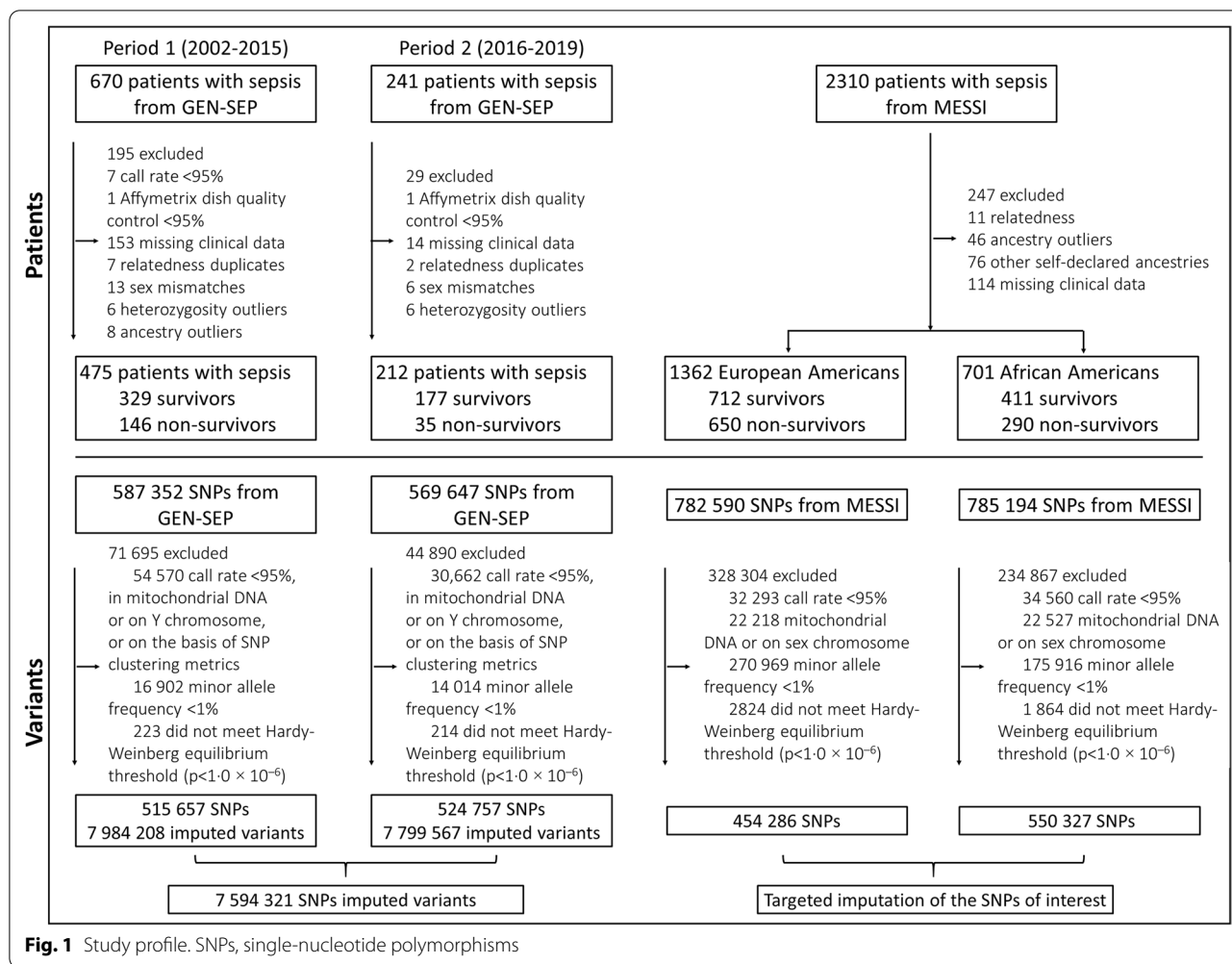
The MESSI study [21], used for the second stage, included a total of 1362 (650 non-survivors) unrelated European Americans and 701 (290 non-survivors) African-Americans with a period of recruitment from 2008 to 2019 (see supplementary material for further details).

All participating studies were done according to The Code of Ethics of the World Medical Association (Declaration of Helsinki), and written informed consent was obtained from all participants or their representatives. The Research Ethics Committees at all participating centers approved this study.

Genotyping and statistical analyses

Genotyping in the GEN-SEP was performed using the Axiom Genome-Wide Human CEU 1 array (Affymetrix, Santa Clara, CA, USA). Genotyping quality control procedures are detailed in supplementary material (Fig. 1). We also calculated the main axis of genetic variation using principal component (PC) analyses (Additional file 1: Figure S1). In MESSI, SNPs were genotyped using the Affymetrix Axiom TxArray v.1 (Affymetrix) (see supplementary material for further details).

In the first stage, we used Cox proportional hazards model to take into account time-to-event in the analysis of the genetic associations adjusting for age, sex, and the first two main PCs. Results were obtained for a total of 7,872,728 (1st GEN-SEP period) and 7,829,916 (2nd GEN-SEP period) SNPs. An intersection of 7,682,187 SNPs was considered for fixed-effect model meta-analysis. Variants were prioritized for the next stage if they satisfied having the same effect direction, a $p < 0.05$ in both GEN-SEP recruitment periods, and a $p < 5.0 \times 10^{-7}$ after the meta-analysis of both periods. Association results of this first stage were also inspected to evaluate whether the variants or genes previously associated



with sepsis mortality by other studies were also associated in GEN-SEP [15–18].

In the second stage, the prioritized independent variants were tested for association in the MESSI European Americans and African-Americans, separately, using Cox regression models, also considering sex, age, and the first two main PCs as covariates.

Finally, a fixed-effect model meta-analysis from the GEN-SEP and MESSI studies was performed, and the genome-wide significance was declared at $p < 5.0 \times 10^{-8}$. A sensitivity analysis was conducted for the genome-wide significant variants, adjusting the models for different clinical and demographic variables and the index event bias. We also used the Nagelkerke’s R^2 to estimate the proportion of variance explained by the independent most significant (sentinel) variants by separate or when combined into a PRS. More details are included in supplementary material.

Association studies in the HLA genes

Given the importance of the major histocompatibility complex (MHC) in inflammatory and immunological diseases, we performed association testing with 28-day sepsis survival of genetic variation in the HLA region. Association analyses were performed only on the GEN-SEP cohort by using Cox regressions adjusting for sex, age, and the main two PCs. This analysis was restricted to 207 classic HLA alleles and 1034 amino acids that had a frequency >1%. Considering the multiple tests adjustment, significance thresholds for the HLA analysis were set at $p < 2.49 \times 10^{-4}$ and $p < 4.83 \times 10^{-5}$, respectively.

Annotation of the functional effects of associated variants and gene expression

The functional effects of the sentinel variants, their linkage disequilibrium proxies ($r^2 > 0.7$), and related genes were assessed based on empirical data from different

integrated software tools and datasets (see supplementary material for further details). To assess differential gene expression of the genes near the sentinel variants, we accessed the public gene expression datasets GSE54514, GSE65682, and GSE32707, containing data for sepsis survival and sepsis-associated ARDS.

Polygenic risks of sepsis and effects on 28-day sepsis survival

We examined whether the polygenic component of sepsis risk was associated with 28-day sepsis survival through PRS (see supplementary material for further details).

First, we obtained a model of the genetic risk for sepsis by a GWAS of all available sepsis cases from GEN-SEP and population controls (Additional file 1: Table S1) using logistic regressions adjusted by sex, age, and the two main PCs.

Then, we constructed the PRS for sepsis risk by including in the score those variants that met a *p* value threshold in the sepsis risk GWAS and varied this threshold to investigate the effect of including more variants in the score.

Finally, we tested if the score was associated with 28-day survival among GEN-SEP patients, adjusting for sex, age, and two main PCs. For this, we used Cox regression and established the *p* < 0.001 threshold for defining significant associations of the risk score.

Additionally, we performed sensitivity analyses to assess the sepsis risk score association with 28-day sepsis survival after (1) excluding variants that deemed

significantly associated with sepsis in candidate gene studies (Additional file 1: Table S2), and (2) excluding variants significantly associated with sepsis mortality in previous GWAS (see supplementary material for further details).

Results

Demographic and clinical features of patients from the first stage are described in Additional file 1: Table S3. After association testing, 11 independent variants were prioritized in the first stage (Fig. 2; Additional file 1: Table S4). The genomic inflation factor of the results from this stage ($\lambda = 1.06$) did not indicate major systematic deviations from the null hypothesis of no association (Additional file 1: Figure S2).

In the second stage, we were able to follow up on the association of 10 of the 11 prioritized variants (Additional file 1: Table S4). Three of these variants reached the genome-wide significance threshold in the meta-analysis (Table 1): a missense variant in the Sterile Alpha Motif Domain containing 9 (*SAMD9*) (rs34896991; p.Ala1556Thr), an intergenic variant to Solute Carrier family 5 member 12 (*SLC5A12*) and Fin Bud Initiation Factor Homolog (*FIBIN*) genes (rs146257041), and an intergenic variant between two non-coding RNAs (LINC00378 and MIR3169) (rs138347802). These three SNPs were of low frequency, with a MAF of ~1% in the study population. Regional association results and Kaplan–Meier survival plots for the three genome-wide significant

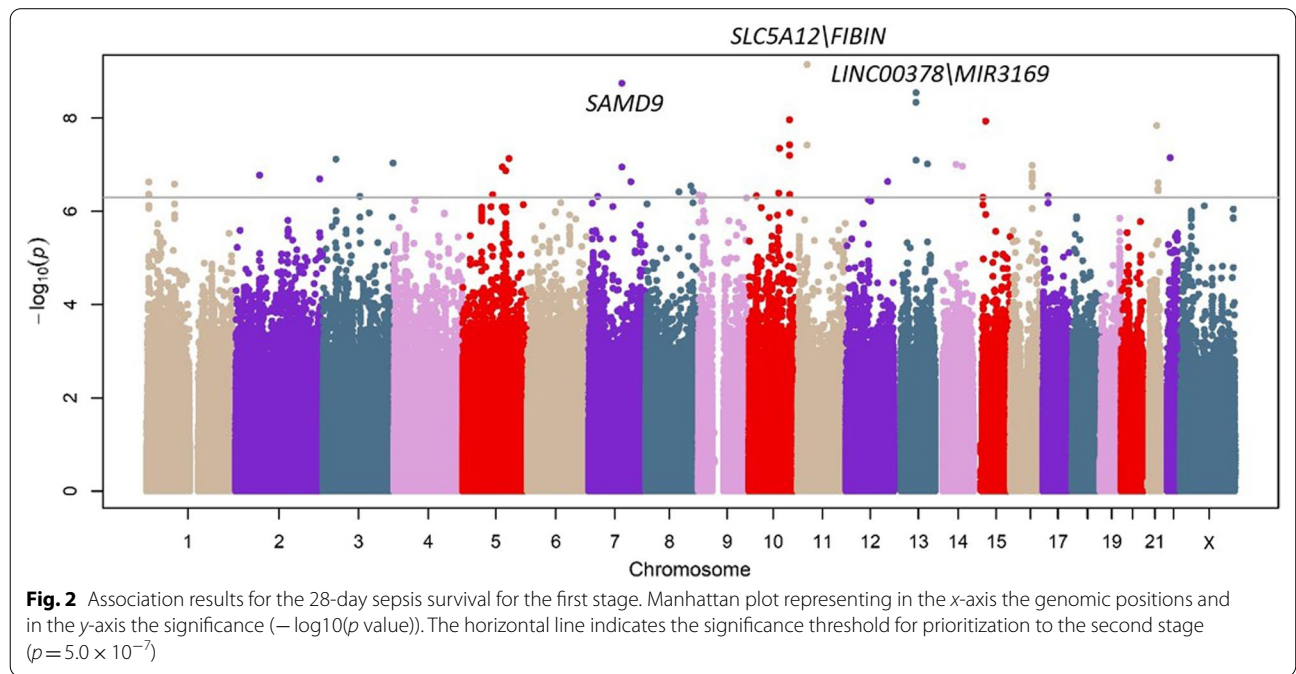


Table 1 Prioritized independent SNPs from the first stage of the GWAS of 28-day sepsis survival

rsID	Position (Hg19)	A1/A2	Func.Gene	Gene	GEN-SEP (N = 687)			MESSI (N = 1362) (European Americans)			MESSI (N = 701) (African-Americans)			Meta-analysis (N = 2750)		
					EAF	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value	
rs34896991	7:92,730,745	C/T	Exonic	SAMD9	0.018	4.75 (2.86–7.89)	1.77×10^{-9}	0.61 (0.29–1.28)	0.188	1.44 (0.46–4.50)	0.534	1.64 (1.37–6.78)	4.92×10^{-8}			
rs146257041	11:26,983,813	A/G	Intergenic	SLC5A12/FIBIN	0.015	5.14 (3.06–8.65)	7.04×10^{-10}	NA	NA	0.89 (0.12–6.39)	0.911	4.59 (2.77–7.59)	$3.00 \times 10^{-9*}$			
rs138347802	13:61,367,197	A/G	Intergenic	LINC00378\ MIR3169	0.014	8.33 (4.1–16.91)	4.56×10^{-9}	0.68 (0.30–1.52)	0.345	2.98 (0.95–9.38)	0.062	2.57 (1.82–13.03)	4.44×10^{-8}			

A1, Ref allele; A2, Effect allele, EAF effect allele frequency, HR hazard ratio, CI confidence interval, Position: Genomic coordinates

*Meta-analysis performed between two studies

Table 2 Sensitivity analyses in the GEN-SEP study of the three genome-wide significant variants. The models adjusted for the indicated variables plus gender, age, and the two main principal components

	N	rs34896991		rs146257041		rs138347802	
		HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value
BMI	364	6.68 (3.22–13.86)	3.47×10^{-7}	5.64 (2.10–15.18)	6.12×10^{-4}	5.72 (0.72–45.32)	0.098
SAPS	264	4.06 (1.54–10.72)	4.66×10^{-3}	9.95 (4.55–21.74)	8.32×10^{-9}	8.63 (1.86–40.01)	5.90×10^{-3}
APACHE II	672	3.48 (2.05–5.91)	3.66×10^{-6}	5.84 (3.43–9.94)	7.39×10^{-11}	8.65 (4.27–17.53)	2.07×10^{-9}
Comorbidities [§]	600	4.24 (2.46–7.33)	2.19×10^{-7}	5.39 (3.09–9.40)	2.91×10^{-9}	7.60 (3.14–18.36)	6.69×10^{-6}
SOFA	661	4.58 (2.72–7.69)	9.49×10^{-9}	5.10 (2.99–8.71)	2.40×10^{-9}	8.50 (4.16–17.39)	4.59×10^{-9}
Sepsis of pulmonary origin	681	4.32 (2.62–7.13)	9.74×10^{-9}	4.71 (2.78–8.00)	9.33×10^{-9}	8.08 (3.98–16.43)	7.54×10^{-9}
Pathogen, %							
Gram-positive	487	4.97 (2.68–9.20)	3.51×10^{-7}	5.36 (3.02–9.51)	9.82×10^{-9}	13.11 (5.01–34.32)	1.61×10^{-7}
Gram-negative	487	4.97 (2.68–9.21)	3.42×10^{-7}	5.32 (3.00–90.46)	1.18×10^{-8}	13.17 (5.04–34.43)	1.47×10^{-7}
Others*	487	4.91 (2.64–9.13)	4.73×10^{-7}	5.13 (2.89–9.12)	2.30×10^{-8}	12.91 (4.93–33.85)	1.96×10^{-7}

APACHE II acute physiology and chronic health evaluation II, BMI Body mass index, ICU intensive care unit, SAPS simplified acute physiology score II, SOFA sequential organ failure assessment

*Includes: mixed Gram-positive and Gram-negative infection, fungi, virus, and polymicrobial

[§] Includes: cancer, age > 80 years old, hepatopathy, valvular disease, immunodeficiency, severe brain damage, morbid obesity, chronic disease, autoimmune disease, pregnancy, myopathy, pneumonia, and serious recurrent infections

variants in the GEN-SEP are reported in supplementary material (Additional file 1: Figures S3 and S4). Results remained robust in the sensitivity analyses (Table 2; Additional file 1: Table S5). Based on Nagelkerke's R^2 , the models for the sentinel variants explain separately nearly 6% of variance, whereas a PRS combination of these three variants explains nearly 9% (Additional file 1: Table S6). These results must be taken with caution because of the potential overfitting of the models.

Finally, we tested previously associated variants with sepsis mortality in the results of the first stage. We found that leading variants of previous GWAS [15–18] were not replicated in the GEN-SEP study (*FER* rs4957796, Hazard Ratio [HR]: 1.09 [95% CI = 0.83–1.43], $p = 0.550$; *VPS13A* rs117983287, HR: 0.36 [95% CI = 0.09–1.48], $p = 0.158$; *CISH* rs143356980, HR: 1.42 [95% CI = 0.61–3.33], $p = 0.419$; and *COL4A2* rs368584, HR: 1.02 [95% CI = 0.83–1.26], $p = 0.817$) (Additional file 1: Table S7). Similarly, an assessment of the association results in and around (± 50 kb) the corresponding genes (*FER*, *VPS13A*, *CISH*, and *COL4A2*) did not reveal any significant finding (Supplementary Methods; Table S7). Likewise, none of classical HLA alleles ($p_{\text{lowest}} = 0.0169$) or amino acids ($p_{\text{lowest}} = 0.0169$) was significantly associated with sepsis survival after multiple testing adjustments (Additional file 1: Figure S5; Table S8), suggesting that common genetic variation at the HLA is not a major driver of sepsis survival or has a modest effect size that could not be detected with the current design.

Potential biological effects of the 28-day sepsis survival-associated variants

We then explored the potential functional implications of the three genome-wide significant variants, rs34896991 in *SAMD9*, rs146257041 intergenic to *SLC5A12*\FIBIN, and rs138347802 intergenic to LINC00378\MIR3169. Based on distinct functional annotations, we observed a few regulatory activities linked to all three overlapped promoter or enhancer regions in multiple cell types including blood cells, and T cells (Additional file 1: Table S9). According to GTEx data, LINC00378 and MIR3169 are expressed only in testis, *SLC5A12* is mainly expressed in the kidney and in the small intestine, *FIBIN* is mainly expressed in arteries (aorta and tibial), tibial nerve, and vagina, while *SAMD9* is expressed broadly across many tissues, but mainly in the esophagus (mucosa), in transformed lymphocytes, and in whole blood. Based on the IPF Cell Atlas in control donors (information was absent for MIR3169), *SAMD9* is expressed ubiquitously across the different cell types, whereas the *SLC5A12*, *FIBIN*, and LINC00378 expression is restricted to a few cell types including vascular, muscular, epithelial, and immune (Additional file 1: Figure S6).

No significant expression quantitative trait loci (eQTLs) were identified in GTEx for rs34896991 and rs138347802. Nevertheless, four significant eQTLs were found in brain and testis for rs146257041. None of these three variants obtained a significant score predicted using DeepSEA. A scan for previously reported trait associations for the three variants based on PhenoScanner found that rs34896991 in *SAMD9* was also associated with the cause

of death in other specified degenerative diseases of the nervous system (Additional file 1: Table S9). Interestingly, other variants in *SAMD9*, *SLC5A12*, *FIBIN*, and in the non-coding RNA (LINC00378) have also been associated with different causes of death according to PhenoScanner results. Regarding the non-coding RNAs, LINC00378 has Cyclin-Dependent Kinase Inhibitor 1A (*CDKN1A*) as the main target and is linked to different types of cancers, while MIR3169 targets genes that are mainly involved in the p53 signaling pathway.

To further evaluate the biological implications of the genes near the identified GWAS loci, whole-blood transcriptomic array data from sepsis survivors and non-survivors were assessed. While information was only available for coding genes, an upregulation of *SAMD9* expression in non-surviving sepsis patients was observed in GSE54514 (log fold change: 0.545 adjusted FDR p value: 2.18×10^{-3}) (Additional file 1: Figure S7). Nevertheless, this gene expression difference among the sepsis patient groups was unrelated to the presence of ARDS (log fold change: 0.011; adjusted FDR p value: 0.996) (Additional file 1: Figure S8). A final assessment of transcriptomic array data from independent sepsis survivors and non-survivors from GSE65682 did not validate the association of *SAMD9* expression with sepsis survival (log fold change: -0.01 p value: 0.688) (Additional file 1: Figure S7).

Polygenic risks of sepsis and effects on 28-day sepsis survival

Finally, we used PRS to investigate whether the polygenic component of sepsis risk was associated with 28-day sepsis survival. We found that the sepsis risk PRS was not significantly associated with sepsis survival at any of the cutoffs (Additional file 1: Figure S9). These results were similar when the models excluded the variants significantly associated with sepsis in the previous candidate gene studies or in the sepsis mortality GWAS.

Discussion

To our knowledge, we report the first GWAS of 28-day sepsis survival conducted to date. In addition, given the importance of the MHC in inflammatory and immunological diseases, we also assessed for the first time the association of the classical HLA alleles and amino acids with 28-day sepsis survival. Our findings revealed three novel low-frequency variants associated with reduced 28-day survival among sepsis patients: one in *SAMD9* (the p.Ala1556Thr exonic variant), one intergenic to *SLC5A12\FIBIN*, and another intergenic to *LINC00378\MIR3169*. The functional annotation analyses revealed a modest regulatory activity of the sentinel variants.

Besides, we found inconsistent results in the association of *SAMD9* expression in whole blood in two independent cohort studies. We were unable to replicate the findings from previous GWAS of sepsis mortality as has been observed in other studies [18, 19]. We also found a lack of overlap between the polygenic component of sepsis risk and sepsis survival. Taken together, while these results are promising, they should be taken with caution given that they only explain a small proportion of the outcome (likely less than our estimates when tested in independent studies) and the studies will require hundreds of thousands of participants to optimally assess genetic effects in the entire allele frequency spectrum and to tackle the underlying complexities of sepsis.

SAMD9 could play a critical role in the inflammatory response during tissue injury and apoptosis [22, 23]. This gene encodes one of the SAM domain-containing proteins that has diverse roles for cellular processes via polymerization and participates in protein interactions and RNA binding [24–26]. It has been observed that the *SAMD9* upregulation triggered an accumulation of macrophages increasing low-grade glioma progression [27]. Linked to this, it was observed that *SAMD9* interacts with Ral Guanine Nucleotide Dissociation Stimulator Like 2 (RGL2) to decrease the expression of Early growth response protein 1 (EGR1) [28], which is a key regulator of inflammation in human macrophages [29]. *SAMD9* has been found to be significantly upregulated in vivo in peripheral blood mononuclear cells during inflammation and in vitro during T cell activation, and its expression is regulated at both the genetic and epigenetic levels [30]. Therefore, *SAMD9* could serve as a T cell activation marker acting as an anti-inflammatory factor [30]. The osmotic shock and interferon-gamma (IFN- γ) tightly regulate *SAMD9* expression [28, 31, 32]. In fact, Chefetz and colleagues observed that *SAMD9* was upregulated by tumor necrosis factor-alpha (TNF- α) through p38 mitogen-activated protein kinases (p38 MAPKs) and nuclear factor-kappa-B (NF- κ B) [31]. Moreover, mutations in *SAMD9* have been linked to immunodeficiency, neutropenia, impaired anti-cytomegalovirus response, and gastrointestinal disorder [33], and to severe multi-system disorders and complex phenotypes characterized by recurrent infection, dysphagia, and profound deafness [34]. *SAMD9* is also strongly associated with mean corpuscular hemoglobin or volume and red cell distribution width [35, 36]. Therefore, we could hypothesize that the *SAMD9* upregulation, which is inducible by various inflammatory, immunological, and stress factors, might activate T cells and produce the accumulation of macrophages through its interaction with RGL2, thus

conferring protection against the systemic dysregulation that occurs during sepsis. This would also be supported by its anti-tumoral, anti-inflammatory, and anti-viral activity. Because the T allele of rs34896991 predicts a missense change in *SAMD9*, this allele could act as a defective variant explaining its association with increased mortality risk among the patients with sepsis.

Our results also revealed two intergenic variants significantly associated with sepsis survival. One of them was located between *SLC5A12* and *FIBIN* genes. In particular, *SLC5A12* encodes an apical cell membrane protein that acts as a metabolite transporter involved in the nuclear factor E2-related factor 2 (NRF2) pathway [37, 38] which is key in response to oxidative stress [39–41]. *FIBIN* is a secreted protein acting downstream of retinoic acid and Wnt signaling [42, 43]. The other intergenic variant was located between a long non-protein coding RNA (LINC00378) and a microRNA (MIR3169). However, the functional information of these two non-coding genes is scarce, both linked to the cellular response to DNA damage.

We acknowledge some strengths and limitations of the study. Among the strengths, the results were supported by two independent geographically distinct studies with diverse ancestries. In addition, both cohorts were prospectively enrolled using consensus criteria for sepsis. Because of that, a robust sensitivity analysis was possible to control for potential confounders. Linked to this, although the significance and the effect direction of the associated variants were not affected by the index event bias correction, all the cases used in the survival analysis were also used to assess sepsis risk and this constitutes a limitation for the approach. The main weaknesses of the study are that we could not assess rarer or structural variants in the analyses. Other approaches such as exome or whole-genome sequencing are needed to analyze the role of these rare genetic variants. In addition, further functional characterization of the prioritized variants will be needed to further dissect the mechanistic connections with the pathophysiology of sepsis.

Conclusion

In conclusion, we have completed a GWAS of 28-day sepsis survival and have identified three novel variants associated with reduced survival, one of them involving a missense variant. Given that the three variants only explain a small proportion of the outcome, more studies with thousands of participants will be needed to optimally assess the genetic effects in sepsis survival and to further validate our findings.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13054-022-04208-5>.

Additional file 1. Supplementary methods, figures, and tables. The supplementary methods include a description of the GEN-SEP and MESSI study populations, genotyping and quality control, more details of the genome-wide association study of 28-day sepsis survival and the index event bias assessment, tools for the annotation of the functional effects of the sentinel variants and the gene expression of related genes, the association of polygenic risks of sepsis with the 28-day sepsis survival, the replication of genes from previous sepsis mortality GWAS, and related references.

The supplementary figures include a principal component analysis, the quantile-quantile plot of the GWAS, regional plots for the genome-wide significant variants, Kaplan-Meier 28-day survival plots, a manhattan plot of sepsis 28-day survival association study results in the HLA region, a plot of expression of related genes across cell types, a boxplot of *SAMD9* gene expression values in GSE54514 and GSE65682, a boxplot of *SAMD9* gene expression values in GSE32707, and polygenic risk score (PRS) model fitting.

The supplementary tables show the relevant demographic and clinical features of cases and controls used for the GWAS of sepsis risk and for the GEN-SEP patients, the list of previous candidate genes associated with sepsis risk, the prioritized independent SNPs, the index bias event results, the association results with 28-day mortality and the percentage of variation explained by the sentinel SNPs by separate or as part of a PRS, the association results of the sentinel variants from other sepsis mortality GWAS studies, the nominally significant results for the classical HLA alleles and amino acids, and the functional assessment of variants associated with 28-day sepsis survival.

Acknowledgements

We would like to thank all patients for their willingness to participate in the study. We thank all the investigators of the GEN-SEP Network and the nurses who assisted in the collection of donor samples. Similarly, we thank all members of the MESSI study groups for their willingness to support data sharing. TH-B, JML-S, ES-P, and LAR-R acknowledge the University of La Laguna for the training support during their PhD studies.

Author contributions

TH-B contributed to data analysis, figures and interpretation of data, and drafting the manuscript; BG-G performed experiments, analyzed the data, and revised the manuscript; JML-S, ES-P, RF, LAR-R, MP, and RC-G contributed to statistical analysis and revising the manuscript; ACo, MIG-L, MIP-G, AR-P, DC, JB, AA, EG-H, EE, AM, ET, MMM, LL, DD, AGL, HMG, JPR, TKJ, JMA, and MS performed experiments, collected sample and clinical data, and analyzed the data; ACa, LVW, and JV contributed to interpretation and critical manuscript review; NJM contributed to human subject enrollment, data collection and analysis, cohort funding; interpretation and critical manuscript review; CF contributed to the study concept and design, data analysis, interpretation of the data, drafting and critical revision of the manuscript, obtained funding, and conceived the project. All authors read and approved the final manuscript.

Funding

This study was funded by Instituto de Salud Carlos III (CB06/06/1088; PI16/00049; PI17/00610; PI17/00177; PI19/00141; PI20/00876) and co-financed by the European Regional Development Funds, "A way of making Europe" from the European Union; agreement OA17/008 with Instituto Tecnológico y de Energías Renovables (ITER) to strengthen scientific and technological education, training, research, development and innovation in Genomics, Personalized Medicine and Biotechnology; and by Cabildo Insular de Tenerife (CGIEU0000219140). LVW holds a GSK/Asthma + Lung UK Chair in Respiratory Research (C17-1). The research was partially supported by the NIHR Leicester

Biomedical Research Centre; the views expressed are those of the author(s) and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health. BG-G is supported by Wellcome Trust grant 221680/Z/20/Z. The MESSI cohort is funded by NIH NHLBI awards HL137006, HL137915, and HL155804 (NJM). The genotyping service of GEN-SEP was done at CEGEN-PRB3-ISCIII; it is supported by grant PT17/0019, of the PE I + D + i 2013–2016, funded by Instituto de Salud Carlos III and European Regional Development Fund. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Availability of data and materials

The data used and analyzed during this study are available from the corresponding author on reasonable request.

Declarations

Ethical approval and consent to participate

The study was approved by the Ethics Committee from the Hospital Universitario de Canarias (Code: CHUNSC_2018-16). All participating studies were done according to The Code of Ethics of the World Medical Association (Declaration of Helsinki), and written informed consent was obtained from all participants or their representatives.

Competing interests

NJM reports funding to her institution from NIH related to this project; funding to her institution from Quantum Leap Healthcare Collaborative, Biomark Inc, and Athersys Inc for work unrelated to this manuscript; and participation on the Scientific Advisory Board for Endpoint Health, Inc. She is a member of the Data Safety Monitoring Board for the Careful Ventilation in ARDS (CAVIARDS) trial and the Observational Study Monitoring Board for the SPIROMICS II study. The rest of the authors declare no competing interests.

Author details

¹Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Carretera del Rosario S/N, Santa Cruz de Tenerife, Spain. ²Research Unit, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas de Gran Canaria, Spain. ³Department of Health Sciences, University of Leicester, Leicester, UK. ⁴Genomics Division, Instituto Tecnológico y de Energías Renovables (ITER), Santa Cruz de Tenerife, Spain. ⁵CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain. ⁶Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, USA. ⁷Genomic Medicine Group, Biomedical Research Center of Rare Diseases (CIBERER), University of Santiago de Compostela, Santiago de Compostela, Spain. ⁸Intensive Care Unit, Complejo Asistencial Universitario de Palencia, Palencia, Spain. ⁹Department of Anesthesiology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas de Gran Canaria, Spain. ¹⁰Department of Medical and Surgical Sciences, University of Las Palmas de Gran Canaria, Gran Canaria, Spain. ¹¹Intensive Care Unit, Complejo Hospitalario Universitario de León, León, Spain. ¹²Intensive Care Unit, Hospital Universitario Río Hortega, Valladolid, Spain. ¹³Intensive Care Unit, Hospital General de Ciudad Real, Ciudad Real, Spain. ¹⁴Intensive Care Unit, Hospital Virgen de la Luz, Cuenca, Spain. ¹⁵Department of Anesthesiology, Hospital Universitario N.S. de Candelaria, Santa Cruz de Tenerife, Spain. ¹⁶CIBER de Enfermedades Infecciosas, Department of Anesthesiology and Resuscitation, Hospital Clínico Universitario de Valladolid, Valladolid, Spain. ¹⁷Departamento de Cirugía, Facultad de Medicina, Universidad de Valladolid, Valladolid, Spain. ¹⁸Intensive Care Unit, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain. ¹⁹Intensive Care Unit, Hospital Universitario de Canarias, La Laguna, Tenerife, Spain. ²⁰Intensive Care Unit, Hospital Universitario La Paz, IdiPAZ, Madrid, Spain. ²¹Division of Pulmonary, Allergy, and Critical Care Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, USA. ²²Department of Anesthesiology, Hospital Clínico Universitario de Valencia, Valencia, Spain. ²³Genomic Medicine Group, CIMUS, University of Santiago de Compostela, Santiago de Compostela, Spain. ²⁴Galician Foundation of Genomic Medicine, Foundation of Health Research Institute of Santiago de Compostela (FIDIS), SERGAS, Santiago de Compostela, Spain. ²⁵Leicester Respiratory Biomedical Research, Centre, National Institute for Health Research, Glenfield Hospital, Leicester, UK. ²⁶Li Ka Shing Knowledge Institute, St. Michael's Hospital,

Toronto, ON, Canada. ²⁷Facultad de Ciencias de la Salud, Universidad Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain.

Received: 20 July 2022 Accepted: 17 October 2022

Published online: 05 November 2022

References

- Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. 2016;315:801–10.
- Fleischmann C, Scherag A, Adhikari NK, Hartog CS, et al. Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *Am J Respir Crit Care Med*. 2016;193(3):259–72.
- Fleischmann-Struzek C, Mellhammar L, Rose N, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. 2020;46(8):1552–62.
- Blanco J, Muriel-Bombín A, Sagredo V, et al. Incidence, organ dysfunction and mortality in severe sepsis: a Spanish multicentre study. *Crit Care*. 2008;12(6):R158.
- Levy MM, Evans LE, Rhodes A. The surviving sepsis campaign bundle: 2018 update. *Intensive Care Med*. 2018;44(6):925–8.
- Ranieri VM, Rubenfeld GD, ARDS Definition Task Force, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA*. 2012;307(23):2526–33.
- Lee MN, Ye C, Villani AC, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science*. 2014;343(6175):1246980.
- Zhang Q, Bastard P, Bolze A, et al. Life-threatening COVID-19: defective interferons unleash excessive inflammation. *Med (N Y)*. 2020;1(1):14–20.
- 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. *Sci Immunol*. 2021;6(62):eab4348.
- COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature*. 2021;600(7889):472–7.
- Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science*. 2007;317(5840):944–7.
- Tian C, Hromatka BS, Kiefer AK, et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat Commun*. 2017;8(1):599.
- Flatby HM, Rasheed H, Ravi A, et al. Risk of lower respiratory tract infections: a genome-wide association study with Mendelian randomization analysis in three independent European populations. *Clin Microbiol Infect*. 2022;28(5):732.e1–732.e7.
- Zhang Q, Bastard P, Cobat A, Casanova JL. COVID Human Genetic Effort. Human genetic and immunological determinants of critical COVID-19 pneumonia. *Nature*. 2022;603(7902):587–98.
- Rautanen A, Mills TC, Gordon AC, et al. Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. *Lancet Respir Med*. 2015;3(1):53–60.
- Scherag A, Schöneweck F, Kesselmeier M, et al. Genetic factors of the disease course after sepsis: a genome-wide study for 28day mortality. *EBioMedicine*. 2016;12:239–46.
- Rosier F, Brisebarre A, Dupuis C, et al. Genetic predisposition to the mortality in septic shock patients: FROM GWAS to the identification of a regulatory variant modulating the activity of a CISH enhancer. *Int J Mol Sci*. 2021;22(11):5852.
- D'Urso S, Rajbhandari D, Peach E, et al. Septic shock: a genomewide association study and polygenic risk score analysis. *Twin Res Hum Genet*. 2020;23(4):204–13.
- Schöneweck F, Kuhnt E, Scholz M, et al. Common genomic variation in the FER gene: useful to stratify patients with sepsis due to pneumonia? *Intensive Care Med*. 2015;41(7):1379–81.
- Guillen-Guio B, Lorenzo-Salazar JM, Ma SF, et al. Sepsis-associated acute respiratory distress syndrome in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med*. 2020;8(3):258–66.
- Reilly JP, Wang F, Jones TK, et al. Plasma angiopoietin-2 as a potential causal marker in sepsis-associated ARDS development: evidence from

- Mendelian randomization and mediation analysis. *Intensive Care Med.* 2018;44(11):1849–58.
22. Mekhedov SL, Makarova KS, Koonin EV. The complex domain architecture of SAMD9 family proteins, predicted STAND-like NTPases, suggests new links to inflammation and apoptosis. *Biol Direct.* 2017;12(1):13.
 23. Nounamo B, Li Y, O'Byrne P, et al. An interaction domain in human SAMD9 is essential for myxoma virus host-range determinant M062 antagonism of host anti-viral function. *Virology.* 2017;503:94–102.
 24. Schultz J, Ponting CP, Hofmann K, Bork P. SAM as a protein interaction domain involved in developmental regulation. *Protein Sci.* 1997;6(1):249–53.
 25. Aviv T, Lin Z, Lau S, et al. The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nat Struct Biol.* 2003;10(8):614–21.
 26. Meruelo AD, Bowie JU. Identifying polymer-forming SAM domains. *Proteins.* 2009;74(1):1–5.
 27. Ma W, Zhang K, Bao Z, et al. SAMD9 is relating with M2 macrophage and remarkable malignancy characters in low-grade glioma. *Front Immunol.* 2021;12:659659.
 28. Hershkovitz D, Gross Y, Nahum S, et al. Functional characterization of SAMD9, a protein deficient in normophosphatemic familial tumoral calcinosis. *J Invest Dermatol.* 2011;131(3):662–9.
 29. Trizzino M, Zucco A, Deliard S, et al. EGR1 is a gatekeeper of inflammatory enhancers in human macrophages. *Sci Adv.* 2021;7(3):eaaz8836.
 30. He P, Wu LF, Bing PF, et al. SAMD9 is a (epi-) genetically regulated anti-inflammatory factor activated in RA patients. *Mol Cell Biochem.* 2019;456(1–2):135–44.
 31. Chefetz I, Ben Amitai D, Browning S, et al. Normophosphatemic familial tumoral calcinosis is caused by deleterious mutations in SAMD9, encoding a TNF-alpha responsive protein. *J Invest Dermatol.* 2008;128(6):1423–9.
 32. Nagamachi A, Kanai A, Nakamura M, et al. Multiorgan failure with abnormal receptor metabolism in mice mimicking Samd9/9L syndromes. *J Clin Invest.* 2021;131(4):e140147.
 33. Formankova R, Kanderova V, Rackova M, et al. Novel SAMD9 mutation in a patient with immunodeficiency, neutropenia, impaired Anti-CMV response, and severe gastrointestinal involvement. *Front Immunol.* 2019;10:2194.
 34. Zhang Y, Zhang Y, Zhang VW, et al. Mutations in both SAMD9 and SLC19A2 genes caused complex phenotypes characterized by recurrent infection, dysphagia and profound deafness: a case report for dual diagnosis. *BMC Pediatr.* 2019;19(1):364.
 35. Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell.* 2016;167(5):1415–1429.e19.
 36. Chen MH, Raffield LM, Mousas A, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell.* 2020;182(5):1198–1213.e14.
 37. Srivastava S, Nakagawa K, He X, et al. Identification of the multivalent PDZ protein PDZK1 as a binding partner of sodium-coupled monocarboxylate transporter SMCT1 (SLC5A8) and SMCT2 (SLC5A12). *J Physiol Sci.* 2019;69(2):399–408.
 38. Pucino V, Certo M, Bulusu V, et al. Lactate buildup at the Site of chronic inflammation promotes disease by inducing CD4+ T cell metabolic rewiring. *Cell Metab.* 2019;30(6):1055–1074.e8.
 39. Buendia I, Michalska P, Navarro E, et al. Nrf2-ARE pathway: an emerging target against oxidative stress and neuroinflammation in neurodegenerative diseases. *Pharmacol Ther.* 2016;157:84–104.
 40. Ahmed SM, Luo L, Namani A, et al. Nrf2 signaling pathway: pivotal roles in inflammation. *Biochim Biophys Acta Mol Basis Dis.* 2017;1863(2):585–97.
 41. Saha S, Buttari B, Panieri E, et al. An overview of Nrf2 signaling pathway and its role in inflammation. *Molecules.* 2020;25(22):5474.
 42. Wakahara T, Kusu N, Yamauchi H, et al. Fibin, a novel secreted lateral plate mesoderm signal, is essential for pectoral fin bud initiation in zebrafish. *Dev Biol.* 2007;303(2):527–35.
 43. Lakner J, Seyer C, Hermsdorf T, Schöneberg T. Characterization of the expression, promoter activity and molecular architecture of fibin. *BMC Biochem.* 2011;12:26.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

