IRF5-TNOP3 POLYMORPHISMS ARE ASSOCIATED WITH ELITE CONTROL OF HIV INFECTION: A RETROSPECTIVE STUDY

Sepúlveda-Crespo D, Jiménez-Sousa MA, Fernández-Rodíguez A, et al. IRF5-TNOP3 polymorphisms are associated with elite control of HIV infection: a retrospective study. J Med Virol. 2023;95:e28841. doi:10.1002/jmv.28841

Title page

Type of Manuscript: Short Communication

Title: *IRF5-TNOP3* polymorphisms are associated with elite control of HIV infection: a retrospective study

Short title: *IRF5-TNOP3* SNPs and HIV elite control

Authors: Daniel SEPÚLVEDA-CRESPO, Ph.D. [1, 2, (§)]; María A JIMÉNEZ-SOUSA, Ph.D. [1, 2, (*, §)]; Amanda FERNÁNDEZ-RODÍGUEZ, Ph.D. [1,2]; María A MUÑOZ-FERNÁNDEZ, Ph.D. [3, 4, 5]; José Luis JIMÉNEZ, Ph.D. [5,6]; Santiago MORENO, M.D., Ph.D. [2,7]; Felipe GARCIA, M.D., Ph.D. [8]; Isidoro MARTÍNEZ, Ph.D. [1,2]; José M. BENITO, M.D., Ph.D. [9, 10, (¥)]; Norma RALLÓN, Ph.D. [9, 10, (¥)]; Salvador RESINO, Ph.D. [1, 2, (*, ¥)]

(§), MAJS and DSC contributed equally to this study; (¥), JMB, NR, and SR contributed equally to this study; (*), MAJS and SR are corresponding authors.

Current affiliations:

(1) Unidad de Infección Viral e Inmunidad, Centro Nacional de Microbiología (CNM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain.

(2) Centro de Investigación Biomédica en Red en Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain.

(3) Servicio de Inmunología, Laboratorio InmunoBiología Molecular, Hospital General Universitario Gregorio Marañón, Madrid, Spain.

(4) Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Madrid, Spain.

(5) Fundación para la Investigación Biomédica, Hospital General Universitario Gregorio Marañón, Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain.

(6) Plataforma de Laboratorio, Hospital General Universitario "Gregorio Marañón", Madrid, Spain.

(7) Servicio de Enfermedades Infecciosas, Hospital Universitario Ramón y Cajal, Madrid, Spain.

(8) Servicio de Enfermedades Infecciosas, Hospital Clínic, Barcelona, Spain.

(9) HIV and Viral Hepatitis Research Laboratory, Instituto de Investigación Sanitaria Fundación Jiménez Díaz, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

(10) Hospital Universitario Rey Juan Carlos, Móstoles, Spain

Corresponding authors:

- **Corresponding author**: Salvador Resino; Centro Nacional de Microbiología, Instituto de Salud Carlos III; Carretera Majadahonda- Pozuelo, Km 2.2; 28220 Majadahonda (Madrid); Telf.: +34 918 223 266; Fax: +34 918 223 269; e-mail: sresino@isciii.es

- **Co-corresponding author:** María A Jiménez-Sousa; Centro Nacional de Microbiología, Instituto de Salud Carlos III; Carretera Majadahonda- Pozuelo, Km 2.2; 28220 Majadahonda (Madrid); Telf.: +34 918 223 266; Fax: +34 918 223 269; e-mail: jimenezsousa@isciii.es

Character count of Title: 98
Character count of Running Head: 37
Word count of Abstract: 206
Word count of Keywords: 5
Word count of Text: 2089
Count of References: 20
Count of Tables: 3
Count of supplementary files: 4

Abstract

Introduction: *IRF5–TNPO3* polymorphisms have previously been related to immune response, and TNPO3 plays a role in HIV-1 infection after nuclear import. Therefore, we analyzed the genetic association between *IRF5–TNPO3* polymorphisms and the HIV elite control in long-term non-progressors (LTNPs).

Methods: We performed a retrospective cohort study on 183 LTNPs, who were antiretroviral therapy (ART)-naïve with $CD4^+$ ≥500 cells/mm^3, viral load ≤10,000 copies/mL, and asymptomatic over ten years after HIV seroconversion. The primary outcome variable was HIV elite control (undetectable viral load in at least 90% of the measurements for at least one year). Seven *IRF5–TNPO3* polymorphisms were genotyped using Agena Bioscience's MassARRAY platform.

Results: We found a significant association between specific *IRF5–TNPO3* genotypes and HIV elite control: rs2004640 TT (aOR=2.05; *p*-value=0.041), rs10954213 AA (aOR=1.95; *p*=0.035), rs2280714 TT (aOR=2.02; *p*=0.031), and rs10279821 CC (aOR=2.12; *p*=0.017). We also found a significant association between *IRF5-TNPO3* haplotype TATC composed of the favorable significant polymorphisms (rs2004640, rs10954213, rs2280714, and rs10279821) and the HIV elite control (aOR=1.59; *p*=0.048).

Conclusions: *IRF5–TNPO3* rs2004640, rs10954213, rs2280714, and rs10279821 polymorphisms were related to HIV elite control in LTNPs. Our data provide new knowledge about the impact of *IRF5–TNPO3* polymorphisms on HIV pathogenesis to understand the phenomenon of natural HIV control.

Keywords

LTNPs; HIV elite controllers; Single nucleotide polymorphisms; Interferon regulatory factor 5; Transportin 3

Introduction

In the absence of antiretroviral therapy (ART), most human immunodeficiency virus (HIV)-infected patients progress to acquired immunodeficiency syndrome (AIDS) within 8-10 years. There is a population of patients (1-5%) known as long-term non-progressors (LTNPs) whose progression of HIV infection is slower. LTNPs are clinically asymptomatic and does not develop AIDS-related events for at least 8-10 years without ART [1]. Moreover, HIV-1-elite controllers (ECs) represent a rare group of HIV-infected patients (0.15-1%) who can maintain an undetectable plasma viral load (pVL) in the absence of ART [2]. ECs represent about 50% of the total LTNPs (LTNP-ECs) who, in addition to the LTNP characteristics, can maintain undetectable pVL [3]. LTNP-ECs show slower AIDS progression and higher survival than LTNP-non-EC. Therefore, LTNP-ECs represent the most favorable phenotype of HIV-infected patients and a potential model of functional cure [3]. This variability in the spontaneous control of HIV infection is determined by a complex interaction between HIV and the host, where the genetic background is crucial [4,5].

The interferon (IFN) regulatory factors (IRFs) are a family of transcription factors vital for regulating the immune system and antiviral response [6]. IRF5 mediates IFN-I production downstream of pattern recognition receptors (PRRs) activation, such as Toll-like receptors 7 and 9 (TLR7 and TLR9) [6]. The *transportin 3 (TNPO3)* gene is located adjacent to *IRF5* gene. TNPO3 is a protein involved in HIV-1 nuclear import by binding to the serine/arginine-rich motifs in splicing factors or RNA recognition motifs. TNPO3 directly participates in the HIV life cycle through the import of HIV-1 across the nuclear membrane and the formation of the HIV pre-integration complex [7]. IRF5 and TNPO3 are involved in the antiviral response and the integration of the HIV provirus into cellular DNA, respectively [6,7].

The *IRF5-TNPO3* region (mapped to chromosome 7q32) contains several single-nucleotide polymorphisms (SNPs) associated with immune system diseases, such as multiple sclerosis and systemic lupus erythematosus [8-10]. The *IRF5–TNPO3* SNPs have previously been related to the modulation of the immune system. Specifically, rs2004640, rs10954213, and rs2280714 SNPs have been associated with increased expression of *IRF5* and excessive production of IFN-I [11-13]. Previous studies have shown that the downregulation of TNPO3 with siRNAs reduces nuclear import, HIV-1 integration, and infectivity [14-16]. However, there is no previous information about the effect of *IRF5–TNPO3* SNPs on *TNPO3* expression.

IRF5-TNPO3 SNPs could be crucial in HIV infection since IRF5 and TNPO3 are associated with the early stages of HIV-1 replication [6,7]. To our knowledge, there is no information on whether *IRF5–TNPO3* SNPs could be related to spontaneous HIV control. This study aimed to analyze the relationship between *IRF5–TNPO3* SNPs and the HIV elite control in LTNPs.

Methods

Study Design

We performed a retrospective cohort study in 183 HIV-infected patients from the Spanish Cohort of LTNPs with one peripheral blood sample deposited in the Spanish HIV HGM Biobank. LTNPs were defined as HIV-seropositive patients for at least ten years without ART with an average CD4+ count ≥500 cells/mm3, low plasma viral load levels (pVL ≤10,000 HIV-1 RNA copies/mL), and asymptomatic. Besides, 110 healthy blood donors seronegative for HIV, hepatitis B virus (HBV), and hepatitis C virus (HCV) from the Transfusion Center of the Community of Madrid were used as a control group.

This study was approved by the Research Ethics Committee of the Instituto de Salud Carlos III (CEI PI_2010-v3). This study was conducted under the Declaration of Helsinki, and all patients signed an informed consent form.

The Strengthening the Reporting of Genetic Association Studies (STREGA) guidelines were used for this study (**Supplementary Table 1**).

Outcomes

The primary outcome variable was HIV elite control (HIV-1-ECs), defined as maintaining undetectable pVL (<50-75 copies/mL) in the absence of ART in at least 90% of the measurements for at least one year [2]. The whole population of LTNPs was classified into two groups: i) LTNP-EC: those LTNP meeting the above criterion (n=86); ii) LTNP-non-EC: those LTNP that did not meet the above criterion (n=97).

Exposure variables

SNPs selection was carried out according to the following criteria: i) SNPs located around the *IRF5-TNPO3* region; ii) previously related to the immune system [9,10]; iii) a minor allelic frequency (MAF)>0.1 in the population of the Iberian Peninsula; iv) compatibility with other selected SNPs in the multiplex genotyping assay. Finally, six *IRF5–TNPO3* SNPs (rs4728142, rs2004640, rs10954213, rs2280714, rs10279821, and rs12537284).

DNA Genotyping

Blood samples were collected, sent to the Spanish HIV HGM Biobank, processed, and frozen after reception. Total DNA was extracted using Wizard® SV Genomic DNA Purification System (Promega, Madison, WI, USA). DNA genotyping of the selected SNPs was performed at Spanish National Genotyping Center (Cegen, http://www.usc.es/cegen/), using the MassARRAY platform and iPLEX® Gold technology (Agena Bioscience, San Diego, CA, USA). All SNPs had a DNA genotyping success rate greater than 95%.

Statistical Analysis

Frequency and absolute count were used for categorical variables and median and interquartile ranges for continuous variables. For comparing groups, the chi-squared test or Fisher's exact test was used for categorical variables, and the Mann-Whitney U test for continuous variables. The Hardy-Weinberg equilibrium (HWE) was evaluated using Stata's *genhwi* package. The genetic association study among *IRF5–TNPO3* SNPs and the elite control of HIV infection was assessed using logistic regression analysis according to the dominant, recessive, and additive inheritance models, providing the odds ratio (OR) and their 95% confidence intervals (CIs). Next, the selected SNPs (p-value ≤0.05) were evaluated by logistic regression analysis adjusted (adjusted OR (aOR)) by age at the moment of HIV diagnosis, gender, and risk group for HIV acquisition (Injecting drug use, IDU). This data analysis used Stata 17.0 (StataCorp, Texas, USA). Moreover, Haploview 4.2 software (MIT/Harvard Broad Institute, Cambridge, MA, USA) was used to evaluate the pairwise linkage disequilibrium (LD), reporting the standardized D' and r^2 values. PLINK 1.9 software was used for haplotype-based association analysis by logistic regression. The statistical significance was p≤0.05, and all p-values were two-tailed.

In silico analysis

The role of *IRF5-TNPO3* SNPs in *IRF5* and *TNPO3* gene expression was analyzed using the GTEx PORTAL Release V8 (https://gtexportal.org/home/ accessed on 08 Mar 2023) [17].

Results

Study population

LTNPs were 64.3% male and had a median age of 49 years at the moment of study, like the control group. Besides, LTNPs had a median age of 40 at the moment of HIV diagnosis, and HIV infection was acquired mainly by IDU (71.4%). We did not find significant differences in clinical and epidemiological characteristics between LTNP groups according to HIV elite control (**Table 1**).

Table 1. Clinical and epidemiological characteristics of HIV-infected patients and healthy donors.

	Healthy control *vs.* all LTNPs			**LTNP groups**		
Characteristics	**Healthy control**	**All LTNPs (*)**	***p*-value (a)**	**LTNP-non-EC**	**LTNP-EC**	***p*-value(b)**
No.	110	183	-	97	86	-
Male	75 (68.2%)	117 (64.3%)	0.487	68 (70.1%)	49 (57.6%)	0.080
Age (years)	47 (42; 53)	49 (46; 52)	0.197	48 (46; 52)	49 (47; 52)	0.192
Age at HIV diagnosis	-	40 (34; 44)	-	39 (35; 44)	40 (34; 43)	0.769
Year of HIV diagnosis	-	1993 (1990; 1997)	-	1993 (1991; 1998)	1992 (1989; 1996)	0.137
Risk group for HIV acquisition	-		-			0.247
IDU	-	130 (71.4%)	-	67 (69.1%)	63(74.1%)	-
Homosexual	-	13 (7.1%)	-	10 (10.3%)	3 (3.5%)	-
Heterosexual	-	27 (14.8%)	-	12 (12.4%)	15 (17.6%)	-
Others	-	12 (6.5%)	-	8 (8.2%)	4 (4.8%)	-

Statistics: *P*-values were calculated by Chi-square or Fisher's exact tests and Mann-Whitney test: (a), differences between the control group and all HIV-infected patients; (b), differences among HIV groups. **Abbreviations**: HIV: human immunodeficiency virus; IDU: intravenous drug users; LTNP: long-term nonprogressors; LTNP-EC: LTNPs who were elite controllers; LTNP-non-EC: LTNPs who were not HIV elite controllers.

IRF5–TNPO3 polymorphisms characteristics

All *IRF5-TNPO3* SNPs had a MAF > 5%, were in HWE, and their genotypic distribution in LTNPs was similar to that found in healthy controls (**Supplementary Table 2**). In addition, *IRF5–TNPO3* allelic and genotypic frequencies were consistent with those of the European population from the 1000 Genomes Project Phase 3 database (https://www.ensembl.org/index.html). The LD values in LTNPs are shown in **Supplementary Figure 1**, where D' and r^2 values higher than 0.8 indicate a strong LD.

IRF5–TNPO3 polymorphisms and HIV elite control

We analyzed the association between *IRF5–TNPO3* SNPs and HIV elite control for the dominant, recessive, and additive inheritance models (**Supplementary Table 3**), finding that the dominant inheritance model was the one that best fit. **Table 2** shows the association for the homozygous favorable allele. By unadjusted regression analysis, we found that carriers of homozygosity for major alleles in four SNPs (rs2004640 TT, rs10954213 AA, rs2280714 TT, and rs10279821 CC) had a higher likelihood of being an LTNP-EC ($p \leq 0.05$). Afterward, we analyzed these four significant SNPs using adjusted regression models, finding that homozygous carriers of the major allele for each SNP were more likely to be elite controllers of HIV: rs2004640 TT (aOR=2.05; p=0.041), rs10954213 AA (aOR=1.95; p=0.035), rs2280714 TT (aOR=2.02; p=0.024), and rs10279821 CC (aOR=2.12; p=0.017) (**Table 2**).

We also analyzed the association between four *IRF5-TNPO3* haplotypes composed of these four significant SNPs (rs2004640, rs10954213, rs2280714, and rs10279821) and the HIV elite control (**Table 3**), and we found by adjusted logistic regression analysis a significant association between *IRF5-TNPO3* TATC haplotype, composed by the major alleles, and HIV elite control (aOR=1.59; p=0.048).

Table 2. Genetic association of *IRF5–TNPO3* polymorphisms with HIV elite control.

Major allele homozygosity distribution		Unadjusted		Adjusted	
Genotypes	**LTNP-non-EC *vs.* LTNP-EC**	**OR (95%CI)**	***p*-value**	**aOR (95%CI)**	***p*-value**
rs4728142 GG	29.9% *vs.* 30.2%	1.02 (1.91 - 0.54)	0.961	-	-
rs2004640 TT	20.0% *vs.* 33.3%	2.00 (3.94 - 1.02)	**0.045**	2.05 (4.08 - 1.03)	**0.041**
rs10954213 AA	29.9% *vs.* 44.2%	1.86 (3.41 - 1.01)	**0.046**	1.95 (3.61 - 1.05)	**0.035**
rs2280714 TT	31.9% *vs.* 47.7%	1.94 (3.53 - 1.06)	**0.031**	2.02 (3.73 - 1.10)	**0.024**
rs10279821 CC	30.9% *vs.* 47.6%	2.04 (3.72 - 1.11)	**0.021**	2.12 (3.92 - 1.15)	**0.017**
rs12537284 GG	74.2% *vs.* 74.4%	0.99 (0.51 - 1.92)	0.976	-	-

Statistics: Data were calculated by binary logistic regression adjusted by gender, age at HIV diagnosis, and injecting drug use (IDU) as the risk factor for HIV acquisition. Significant differences are shown in bold. **Abbreviations**: 95%CI: 95% confidence interval; OR: odds ratio; aOR: adjusted odds ratio; HIV: human immunodeficiency virus; IRF5: interferon regulatory factor 5; LTNP: long-term non-progressors; LTNP-EC: long-term non-progressors who are elite controllers; LTNP-non-EC: long-term non-progressors who are not elite controllers; TNPO3: transportin 3.

Table 3. Association of *IRF5-TNPO3* haplotypes with HIV elite control.

	Unadjusted		Adjusted	
Haplotypes	**OR (95% CI)**	***p*-value**	**aOR (95% CI)**	***p*-value**
TATC	1.54 (0.97; 2.44)	0.061	1.59 (1.01; 2.56)	**0.048**
GACT	0.57 (0.14; 2.27)	0.419	0.71 (0.17; 2.94)	0.222
TGCT	0.89 (0.46; 1.69)	0.712	0.78 (0.39; 1.52)	0.452
GGCT	0.74 (0.49; 1.11)	0.142	0.74 (0.49; 1.12)	0.150

Statistics: Associations were calculated by logistic regression with only haplotypes composed of significant SNPs (rs2004640, rs10954213, rs2280714, and rs10279821). Significant differences are shown in bold. **Abbreviations**: 95% CI: 95% confidence interval; aOR: adjusted odds ratio; HIV: human immunodeficiency virus; IRF5: interferon regulatory factor 5; LTNP: long-term nonprogressors; LTNP-EC: LTNPs who were elite controllers; LTNP-non-EC: LTNPs who were not HIV elite controllers; TNPO3: transportin 3.

Discussion

TNPO3 is a member of the karyopherin β family of proteins implicated in the nuclear import of splicing factors and HIV replication [7]. IFN secreted during viral infection is mainly controlled by IRFs, which coordinate the host immunity and induce an antiviral state [6]. However, the disruption of IRF activity by the HIV infection renders the immune response ineffective because HIV hijacks the host cell machinery and evades IFN-mediated antiviral mechanisms, thus contributing to AIDS progression [6]. *IRF5* encodes IRF5, which is involved downstream of the TLR-MyD88 signaling pathway essential for the transcriptional activation of *IFN-I* genes and induction of proinflammatory cytokines [6]. The nearest gene to *IRF5* is *TNPO3*, which encodes TNPO3 that transports into the nucleus pre-mRNA and serine/arginine-rich proteins, such as splicing factors. During HIV infection, HIV pre-integration complex components, such as viral cDNA and integrase, cross the nuclear membrane to deliver them into the nucleus through the nuclear pore complex [7]. The exact mechanism of this transport is unknown, but everything points to TNPO3 as a relevant factor in this process [7].

In this study, homozygosity for the major alleles of the *IRF5–TNPO3* SNPs (rs2004640, rs10954213, rs2280714, and rs10279821) was associated with the HIV elite control in a large cohort of LTNPs. Besides, the *IRF5-TNPO3* haplotype TATC, formed by the four major significant alleles, was associated with a higher probability of being an HIV elite controller than the other haplotypes.

This study is the first to show a significant association between several SNPs at the *IRF5-TNPO3* gene (rs2004640, rs10954213, rs2280714, and rs10279821) and the spontaneous control of HIV infection, pointing to a possible functional role of these *IRF5-TNPO3* SNPs in the HIV viral cycle. The rs2004640 T allele creates a donor splice site in an alternate exon 1 of the *IRF5* gene (exon 1B), allowing the expression of several unique *IRF5* isoforms [18]. In addition, the rs2004640 TT genotype has been previously associated with higher levels of *IRF5* mRNA expression, likely leading to excessive IFN-I production-enhancing immune responses [11]. The rs10954213 SNP is in a conserved 3' untranslated region (3' UTR) of the *IRF5* gene. The rs10954213 A variant has previously shown a strong association with *IRF5* expression, even higher than other *IRF5* variants [12]. Its effect is possible because the rs10954213 A allele alters the length of the *IRF5* 3' UTR [19], contributing to a shorter and more stable *IRF5* mRNA and, thus, decreasing its degradation [20]. The rs2280714 SNP is located downstream of the *IRF5* gene and in the 3'UTR of the *TNPO3* gene. The rs2280714 TT genotype has also been associated with upregulating *IRF5* mRNA expression [13]. Finally, the rs10279821 polymorphism is located in an intronic region of the *TNPO3* gene. Consistent with the literature, using the GTEx Portal [17], a comprehensive public resource to study tissue-specific gene expression and regulation, we found that these four polymorphisms at *IRF5-TNPO3* (rs2004640, rs10954213, rs2280714, and rs10279821) increase the *IRF5* expression in whole blood.

Therefore, these 4 SNPs at *IRF5-TNPO3* (rs2004640, rs10954213, rs2280714, and rs10279821) promote increased IRF5 expression, which could be crucial for the spontaneous control of HIV infection. It cannot be ruled out an effect of these four SNPs in the TNPO3 expression level, an essential protein in the life cycle of HIV.

Limitations of the study

Our study has several limitations that must be considered for its correct interpretation. First, our study had limited statistical power because the sample size was small. This makes it difficult to reach statistical significance and/or increase the false positive rate. However, it should be noted that LTNPs are scarce, and reaching a high number is very difficult. Secondly, our study design is retrospective and could introduce biases. However, it should be noted that the two study groups (LTNP-non-EC vs. LTNP-EC) were similar in clinical-epidemiological characteristics. Thirdly, we did not analyze the association between IRF5-TNPO3 polymorphisms and comorbidities because we did not have detailed information. However, it should be noted that all patients were LTNPs, so a low rate of comorbidities is expected. Fourth, a functional analysis of the *IRF5-TNPO3* polymorphisms to confirm its effect on the HIV elite control would be desirable. However, it could not be performed due to the required samples' unavailability.

Conclusions

IRF5–TNPO3 rs2004640, rs10954213, rs2280714, and rs10279821 SNPs were associated with spontaneous control of HIV replication in LTNPs. Our data provide new knowledge about the impact of *IRF5–TNPO3* SNPs on HIV pathogenesis and contribute to a better understanding of the phenomenon of natural control of HIV.

Declarations

Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of the Instituto de Salud Carlos III (CEI PI_2010-v3).

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed during the current study may be available upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by Instituto de Salud Carlos III (grant number PI19/01638 to MAMF, PI20/01314 to JLJ, CP17CIII/00007, PI18CIII/00028 and PI21CIII/00033 to MAJS, PI17/00657, PI18CIII/00020 to AFR, and PI17CIII/00003 and PI20CIII/00004 to SR) and Ministerio de Ciencia e Innovación (grant number PID2021-126781OB-I00 to AFR). The study was also funded by the Spanish AIDS Research Network (grants RD16/0025/0013, RD16/0025/0019, and RD16CIII/0002/0002) and by the CIBER -Consorcio Centro de Investigación Biomédica en Red- (CB 2021), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación and Unión Europea – NextGenerationEU (grant numbers CB21/13/00044 to SR, CB21/13/00094 to FG, CB21/13/00086 to SM and CD22/01/00041 to MAMF). M.A.J-S is a Miguel Servet researcher supported and funded by ISCIII (grant no. CP17CIII/00007). N.R. is a Miguel Servet researcher from the ISCIII (grant no. CPII19/00025). D.S.-C. is a Sara Borrell researcher from ISCIII (grant no. CD20CIII/00001).

Acknowledgments

The authors thank the Spanish National Genotyping Center (CEGENPRB2-ISCIII) for providing SNP genotyping services (http://www.cegen.org). CEGEN is supported by grant PT13/0001, ISCIII-SGEFI/FEDER. We also acknowledge the patients in this study for their participation and the Transfusion Center of the Community of Madrid for the healthy donor blood samples provided.
This study would not have been possible without the collaboration of all the patients, medical and nursery staff, and data managers who have taken part in the project. We want to acknowledge the patients in this study for their participation and the HIV BioBank integrated into the Spanish AIDS Research Network and collaborating Centers (http://hivhgmbiobank.com/donor-area/hospitals-and-centres-transferring-samples/?lang=en) for the generous gifts of clinical samples used in this work. The HIV BioBank, integrated into the Spanish AIDS Research Network, is partially funded by the RD16/0025/0019 project as part of the Plan Nacional R + D + I and cofinanced by ISCIII- Subdirección General de Evaluación and el Fondo Europeo de Desarrollo Regional (FEDER).

Authors' contributions

Conceptualization: MAJS and SR.
Data curation: NR, JMB, SM, FG, MAMF, and JLJ.
Formal analysis: MAJS, DSC, and SR.
Funding acquisition: SR, NR, and JMB.
Investigation and methodology: MAJS, NR, and JMB.
Project Administration: MAJS and SR.
Supervision and visualization: SR, NR, and JMB.
Writing – original draft preparation: MAJS, DSC, and SR.
Writing – Review & Editing: AFR, IM, NR, and JMB.
All authors read and approved the final manuscript.

Authors' information (optional)

Not applicable.

References

1. Kumar P. Long term non-progressor (LTNP) HIV infection. *Indian J Med Res.* 2013;138(3):291-293.
2. Navarrete-Munoz MA, Restrepo C, Benito JM, Rallon N. Elite controllers: A heterogeneous group of HIV-infected patients. *Virulence.* 2020;11(1):889-897.
3. Capa L, Ayala-Suarez R, De La Torre Tarazona HE, et al. Elite controllers long-term non progressors present improved survival and slower disease progression. *Sci Rep.* 2022;12(1):16356.
4. Lopez-Galindez C, Pernas M, Casado C, Olivares I, Lorenzo-Redondo R. Elite controllers and lessons learned for HIV-1 cure. *Curr Opin Virol.* 2019;38:31-36.
5. Gingras SN, Tang D, Tuff J, McLaren PJ. Minding the gap in HIV host genetics: opportunities and challenges. *Hum Genet.* 2020;139(6-7):865-875.
6. Marsili G, Perrotti E, Remoli AL, Acchioni C, Sgarbanti M, Battistini A. IFN Regulatory Factors and Antiviral Innate Immunity: How Viruses Can Get Better. *J Interferon Cytokine Res.* 2016;36(7):414-432.
7. Tabasi M, Nombela I, Janssens J, Lahousse AP, Christ F, Debyser Z. Role of Transportin-SR2 in HIV-1 Nuclear Import. *Viruses.* 2021;13(5).
8. Kottyan L, Kelly JA, Harley JB. 127 - Genetics of lupus. In: Hochberg MC, Silman AJ, Smolen JS, Weinblatt ME, Weisman MH, eds. *Rheumatology (Sixth Edition).* Philadelphia: Mosby; 2015:1045-1051.
9. Kottyan LC, Zoller EE, Bene J, et al. The IRF5-TNPO3 association with systemic lupus erythematosus has two components that other autoimmune disorders variably share. *Hum Mol Genet.* 2015;24(2):582-596.
10. Jarvinen TM, Hellquist A, Zucchelli M, et al. Replication of GWAS-identified systemic lupus erythematosus susceptibility genes affirms B-cell receptor pathway signalling and strengthens the role of IRF5 in disease susceptibility in a Northern European population. *Rheumatology (Oxford).* 2012;51(1):87-92.
11. Wang JM, Huang AF, Yuan ZC, Su LC, Xu WD. Association of IRF5 rs2004640 polymorphism and systemic lupus erythematosus: A meta-analysis. *Int J Rheum Dis.* 2019;22(9):1598-1606.
12. Graham RR, Kyogoku C, Sigurdsson S, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A.* 2007;104(16):6758-6763.
13. Hu W, Ren H. A meta-analysis of the association of IRF5 polymorphism with systemic lupus erythematosus. *Int J Immunogenet.* 2011;38(5):411-417.
14. Brass AL, Dykxhoorn DM, Benita Y, et al. Identification of host proteins required for HIV infection through a functional genomic screen. *Science.* 2008;319(5865):921-926.
15. Konig R, Zhou Y, Elleder D, et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell.* 2008;135(1):49-60.
16. Christ F, Thys W, De Rijck J, et al. Transportin-SR2 imports HIV into the nucleus. *Curr Biol.* 2008;18(16):1192-1202.
17. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. *Nature Genetics.* 2013;45(6):580-585.
18. Graham RR, Kozyrev SV, Baechler EC, et al. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet.* 2006;38(5):550-555.
19. Dideberg V, Kristjansdottir G, Milani L, et al. An insertion-deletion polymorphism in the interferon regulatory Factor 5 (IRF5) gene confers risk of inflammatory bowel diseases. *Hum Mol Genet.* 2007;16(24):3008-3016.
20. Liu HF, An XJ, Yang Y, et al. Association of rs10954213 polymorphisms and haplotype diversity in interferon regulatory factor 5 with systemic lupus erythematosus: a meta-analysis. *J Huazhong Univ Sci Technolog Med Sci.* 2013;33(1):15-21.

Appendix 1

Centers involved in Long Term Non-Progressors (LTNPs) cohort

Centro Sandoval - Madrid
Hospital 12 de Octubre - Madrid
Hospital Arnau de Vilanova - Lleida
Hospital Asturias
Hospital Bellvitge - Barcelona
Hospital Castellón
Hospital Clínic - Barcelona
Hospital Donostia - San Sebastián
Hospital Elche - Alicante
Hospital Germans Trias i Pujol - Badalona
Hospital Gregorio Marañón - Madrid
Hospital Joan XXIII - Tarragona
Hospital La Fe - Valencia
Hospital La Paz/Carlos III - Madrid
Hospital La Princesa - Madrid
Hospital Navarra - Pamplona
Hospital Parc Taulí- Sabadell
Hospital Ramón y Cajal - Madrid
Hospital San Cecilio - Granada
Hospital San Pedro - Logroño
Hospital Son Dureta - Mallorca
Hospital Virgen del Rocío – Sevilla

Supplementary Materials

Supplementary Table 1. STrengthening the REporting of Genetic Association studies (STREGA) reporting recommendations, extended from STROBE Statement

Item	Item no	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Page no
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract.		1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found.		3
Introduction				
Background rationale	2	Explain the scientific background and rationale for the investigation being reported.		4-5
Objectives	3	State specific objectives, including any pre-specified hypotheses	***State if the study is the first report of a genetic association, a replication effort, or both.***	5
Methods				
Study design	4	Present key elements of study design early in the paper.		6
Setting	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection.		6
Participants	6	(a) **Cohort study –** Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. **Case–control study –** Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. **Cross-sectional study –** Give the eligibility criteria, and the sources and methods of selection of participants.	***Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant.***	6
		(b) Cohort study – For matched studies, give matching criteria and number of exposed and unexposed. **Case–control study –** For matched studies, give matching criteria and the number of controls per case.		
Variables	7	(a) Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	***(b) Clearly define genetic exposures (genetic variants) using a widely –used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin).***	6
Data sources measurement	8*	(a) For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	***(b) Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory /centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the***	6-7

			study simultaneously or in smaller batches.	
Bias	9	(a) Describe any efforts to address potential sources of bias.	***(b) For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias, and explain what approach was used to deal with this.***	7
Study size	10	Explain how the study size was arrived at.		6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.	***If applicable, describe how effects of treatment were dealt with.***	6-7
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding.	***State software version used and options (or settings) chosen.***	7
		(b) Describe any methods used to examine subgroups and interactions.		7
		(c) Explain how missing data were addressed.		
		(d) **Cohort study –** If applicable, explain how loss to follow-up was addressed. **Case–control study –** If applicable, explain how matching of cases and controls was addressed. **Cross-sectional study –** If applicable, describe analytical methods taking account of sampling strategy.		
		(e) Describe any sensitivity analyses.		
			(f) State whether Hardy-Weinberg equilibrium was considered and, if so, how.	7
			(g) Describe any methods used for inferring genotypes or haplotypes.	7
			(h) Describe any methods used to assess or address population stratification.	7
			(i) Describe any methods used to address multiple comparisons or to control risk of false positive findings.	7
			(j) Describe any methods used to address and correct for relatedness among subjects.	7
Results				
Participants	13*	(a) Report the numbers of individuals at each stage of the study – e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up and analysed.	***Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful.***	8
		(b) Give reasons for non-participation at each stage.		
		(c) Consider use of a flow diagram.		
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders.	***Consider giving information by genotype.***	8
		(b) Indicate the number of participants with missing data		

		for each variable of interest.		
		(c) **Cohort study –** Summarize follow-up time, e.g. average and total amount.		
Outcome data	15*	**Cohort study –** Report numbers of outcome events or summary measures over time.	***Report outcomes (phenotypes) for each genotype category over time***	8
		Case–control study – Report numbers in each exposure category, or summary measures of exposure.	***Report numbers in each genotype category***	
		Cross-sectional study – Report numbers of outcome events or summary measures.	***Report outcomes (phenotypes) for each genotype category***	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence intervals). Make clear which confounders were adjusted for and why they were included.		8
		(b) Report category boundaries when continuous variables were categorized.		
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.		
			(d) Report results of any adjustments for multiple comparisons.	
Other analyses	17	(a) Report other analyses done – e.g. analyses of subgroups and interactions, and sensitivity analyses.		8
			(b) If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken.	
			(c) If detailed results are available elsewhere, state how they can be accessed.	
Discussion				
Key results	18	Summarize key results with reference to study objectives.		9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.		10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.		9-10
Generalizability	21	Discuss the generalizability (external validity) of the study results.		9-10
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.		12

STROBE: STtrengthening the Reporting of Observational Studies in Epidemiology

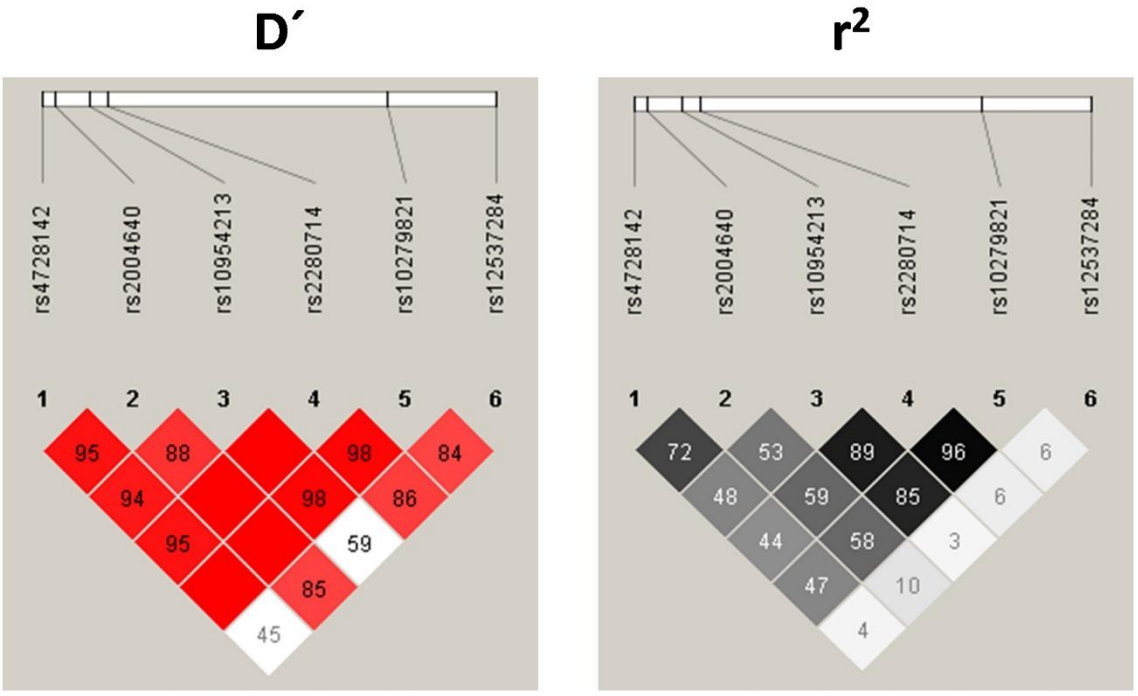
*Give information separately for cases and controls in case–control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Supplementary Table 2. Characteristics of *IRF5–TNPO3* polymorphisms in LTNPs and control group.

		Distribution of genotypes (1)			HWE (2)	
SNPs	**Genotype**	**Control group**	**LTNPs**	***p*-value**	**Control group**	**LTNPs**
rs4728142	n=293					
	GG	33 (30.0%)	55 (30.1%)	0.933	0.451	0.658
	AG	51 (46.4%)	88 (48.1%)			
	AA	26 (23.6%)	40 (21.9%)			
rs2004640	n=289					
	TT	36 (32.7%)	47 (26.3%)	0.495	0.563	1.000
	GT	51 (46.4%)	90 (50.3%)			
	GG	23 (20.9%)	42 (23.5%)			
rs10954213	n=293					
	AA	46 (41.8%)	67 (36.6%)	0.482	0.529	0.755
	AG	53 (48.2%)	90 (49.2%)			
	GG	11 (10.0%)	26 (14.2%)			
rs2280714	n=293					
	TT	52 (47.3%)	72 (39.3%)	0.357	0.659	0.425
	CT	49 (44.5%)	90 (49.2%)			
	CC	9 (8.2%)	21 (11.5%)			
rs10279821	n=293					
	CC	52 (47.3%)	71 (38.8%)	0.316	0.498	0.198
	CT	50 (45.5%)	93 (50.8%)			
	TT	8 (7.3%)	19 (10.4%)			
rs12537284	n=292					
	GG	81 (74.3%)	136 (74.3%)	0.951	0.447	0.758
	AG	25 (22.9%)	43 (23.5%)			
	AA	3 (2.8%)	4 (2.2%)			

Statistic: (1), Differences in the genotype's distribution between HIV patients and controls; (2), Hardy-Weinberg equilibrium. P-values were calculated by chi-squared. P-values were calculated by chi-squared. **Abbreviations**: HIV: human immunodeficiency virus; HWE: Hardy-Weinberg equilibrium; IRF5: interferon regulatory factor 5; LTNP: long-term nonprogressor; SNP: single nucleotide polymorphism; TNPO3: transportin 3.

Supplementary Figure 1. Pairwise linkage disequilibrium (LD) pattern for *IRF5–TNPO3* polymorphisms. The intensity of red or grey decreases with decreasing D´ and r2 values. The genomic location of SNPs is indicated on top. The diagonal represents an SNP, and the square represents a pairwise comparison between two SNPs, indicating the magnitude of LD (D´ and r^2). D´ and r^2 vary from 0 (absence) to 100 (complete). **Abbreviations**: D´: D-prime or proportion of the possible LD that was present between two SNPs; HIV: human immunodeficiency virus; IRF5: interferon regulatory factor 5; LD: linkage disequilibrium; r^2: square of the correlation coefficient; SNP: single nucleotide polymorphism; TNPO3: transportin 3.



Supplementary Table 3. Genetic association of *IRF5–TNPO3* polymorphisms with HIV elite control.

SNPs	Inheritance	Genotypes	OR (95%CI)	p-value	aOR (95%CI)	p-value
rs4728142	Dominant	GG vs. AG/AA	0.98 (0.52 - 1.85)	0.961	0.99 (0.52 - 1.88)	0.963
rs4728142	Recessive	GG/AG vs. AA	1.72 (0.85 - 3.49)	0.134	1.80 (0.87 - 3.70)	0.112
rs4728142	Additive	nA	1.19 (0.79 - 1.79)	0.403	1.21 (0.80 - 1.83)	0.373
rs2004640	Dominant	TT vs. GT/GG	0.50 (0.25 - 0.98)	**0.045**	0.49 (0.25 - 0.97)	**0.041**
rs2004640	Recessive	TT/GT vs. GG	0.71 (0.35 - 1.43)	0.339	0.76 (0.37 - 1.55)	0.452
rs2004640	Additive	nG	0.67 (0.44 - 1.03)	0.068	0.68 (0.45 - 1.05)	0.084
rs10954213	Dominant	AA vs. AG/GG	0.54 (0.29 - 0.99)	**0.046**	0.51 (0.28 - 0.95)	**0.035**
rs10954213	Recessive	AA/AG vs. GG	0.80 (0.35 - 1.86)	0.606	0.74 (0.32 - 1.75)	0.495
rs10954213	Additive	nG	0.69 (0.44 - 1.06)	0.092	0.66 (0.42 - 1.03)	0.064
rs2280714	Dominant	TT vs. CT/CC	0.52 (0.28 - 0.94)	**0.031**	0.50 (0.27 - 0.91)	**0.024**
rs2280714	Recessive	TT/CT vs. CC	0.66 (0.26 - 1.69)	0.388	0.65 (0.25 - 1.67)	0.371
rs2280714	Additive	nC	0.63 (0.40 - 0.99)	**0.043**	0.61 (0.38 - 0.97)	**0.036**
rs10279821	Dominant	CC vs. CT/TT	0.49 (0.27 - 0.90)	**0.021**	0.47 (0.26 - 0.87)	**0.017**
rs10279821	Recessive	CC/CT vs. TT	0.80 (0.31 - 2.10)	0.652	0.79 (0.30 - 2.09)	0.632
rs10279821	Additive	nT	0.63 (0.39 – 0.99)	**0.05**	0.61 (0.38 - 0.98)	**0.041**
rs12537284	Dominant	GG vs. AG/AA	0.99 (0.51 - 1.92)	0.976	1.01 (0.51 - 1.97)	0.988
rs12537284	Recessive	GG/AG vs. AA	0.37 (0.04 - 3.61)	0.391	0.35 (0.04 - 3.51)	0.372
rs12537284	Additive	nA	0.92 (0.51 - 1.65)	0.772	0.92 (0.51 - 1.68)	0.793

Statistics: Data were calculated by binary logistic regression adjusted by gender, age at HIV diagnosis, and injecting drug use (IDU) as the risk factor for HIV acquisition. Significant differences are shown in bold. **Abbreviations**: 95%CI: 95% confidence interval; OR: odds ratio; aOR: adjusted odds ratio; HIV: human immunodeficiency virus; IRF5: interferon regulatory factor 5; LTNP: long-term non-progressors; LTNP-EC: long-term non-progressors who are elite controllers; LTNP-non-EC: long-term non-progressors who are not elite controllers; TNPO3: transportin 3.