Reumatismo, 2024; 76 (2): 96-106

Association between *COMMD1* gene polymorphism rs11125908 and rheumatoid arthritis in the Cuban population

M. Carpio Alvarez[1], A. Cintado Benitez[2], T. Diaz Argudin[2], H. Nodarse Cuni[3], M.d.C Dominguez Horta[1], J.R. Fernández Massó[1]

[1]Pharmaceutical Department, Center for Genetic Engineering and Biotechnology, Havana, Cuba;
[2]Pharmacogenomics Department, Center for Genetic Engineering and Biotechnology, Havana, Cuba;
[3]Clinical Research Direction, Center for Genetic Engineering and Biotechnology, Havana, Cuba

SUMMARY

Objective. To evaluate the association of the rs11125908 polymorphism in the *COMMD1* gene in the Cuban population with rheumatoid arthritis (RA).

Methods. In this case-control study, 161 RA patients and 150 control subjects were genotyped for rs11125908 by the allele-specific polymerase chain reaction method. DNA sequencing was used to verify the assignation of the polymorphism. The odds ratios (OR) and their 95% confidence interval were calculated by logistic regression to determine the associations between genotypes and RA using the SNPStats software.

Results. An association of the single nucleotide polymorphism with the disease was found in the overdominant model (p=0.025; OR=1.91) for the AG genotype. Our analyses revealed an association between rs11125908 and the subgroup of patients with swollen joints < median under the codominant model for AG (p=0.034; OR=2.30) and GG genotype (p=0.034; OR=0.82) and with the overdominant model (p=0.01; OR=2.38). The subgroup of patients with an age of onset lower than the mean and AG genotype showed an association in the overdominant model (p=0.027; OR=2.27). Disease activity score 28 with erythrocyte sedimentation rate and disease duration variables were not associated with the rs11125908 polymorphism.

Conclusions. rs11125908 was associated with RA and with the number of swollen joints and age of onset subgroup analyses. We provide concepts for treatments for RA, based on pharmacological management of *COMMD1* expression.

Key words: *COMMD1, Cuban population, rheumatoid arthritis, single nucleotide polymorphism.*

Reumatismo, 2024; 76 (2): 96-106

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory joint disease with a worldwide prevalence of approximately 5 per 1000 adults (1). The prevalence of RA is 2-3 times higher in women than in men, but this ratio decreases with age (2). It is characterized by inflammation in the synovial membrane, cartilage destruction, bone erosion, joint deformity, and functional disability in the affected patients (3). The goal of RA therapy is to obtain a prolonged state of health associated with improved quality of life by controlling symptoms, preventing structural damage to the joints, and normalizing the patient's physical condition (4). Genome-wide association studies (GWAS) have uncovered the highly polygenic etiology of RA. More than 2127 unique single nucleotide polymorphisms (SNP) have been identified as associated with RA or with response to disease-modifying treatments in the GWAS catalog (5, 6). Genetic polymorphisms may serve as useful markers of RA disease prognosis and treatment (7). Most of GWAS on RA has been conducted in Caucasian, African American, and Asian populations (8-10). However, the genetic structure of the Cuban population differs from most studied populations and has been molded by the history of admixture between indigenous Americans, Europeans, and Africans (11). In addition, no association studies have been conducted in Cuban patients with RA.

Corresponding author:
Julio Raúl Fernández Massó
Pharmaceutical Department,
Center for Genetic Engineering and Biotechnology,
Cubanacan, P.O. Box 6162, Havana, Cuba
E-mail: julio.fernandez@cigb.edu.cu

In Japanese RA patients, the minor allele G at rs11125908 showed a significant cis expression quantitative trait loci association with increased *COMMD1* expression. Except for the publication by Murata et al. in 2017 (12), no reference has been made to its association or clinical relevance with any other disease. Murata et al. documented the relation between the high expression levels of *COMMD1* protein and a decrease in bone erosion of hand joints in Japanese patients with RA. Additionally, the rs11125908 G alleles demonstrated a strong suppressive association with finger joint destruction, independent of the presence and levels of rheumatoid factor. *COMMD1* is a recognized pleotropic protein that plays an important role in inflammation and hypoxic response through the negative regulation of transcription factors nuclear factor κB (NK-κB) and hypoxia-inducible factor-1 (13). In RA, according to the results of Murata et al., *COMMD1* is suppressed by hypoxia preventing its function as a suppressor of hypoxia-induced pathways under normoxic conditions and as a negative regulator of inflammatory gene expression, osteoclastogenesis and pathologic bone resorption (12).

The rs11125908 polymorphism in the dbSNP Short Database National Center for Biotechnology Information (NCBI) Genetic Variations is described as a polymorphism located in the intronic region of *COMMD1*. Genetic variations located in the intronic regions contribute to disease development through alterations in transcription factor binding, enhancer activity, long-range enhancer-promoter interactions, post-translational histone modifications, and/or RNA polymerase function (14). Intronic SNPs generate a large expansion of transcriptomic diversity and promote or disrupt splicing, and affect long noncoding RNA function. It is likely that many of the SNPs are regulatory SNPs, with variations that modify gene expression levels (15).

The rs11125908 polymorphism was selected due to its relation with the NK-κB signaling pathway (12). Methotrexate is the gold-standard medication for RA. By modulating the NF-κB signaling pathway, methotrexate inhibits important pro-inflammatory properties of major cell lineages involved in RA pathogenesis (16). The rs11125908 polymorphism could become a prognostic pharmacogenomics marker of methotrexate response.

The frequency of the rs11125908 polymorphism and its association with the Cuban population are unknown. Our study aimed to evaluate the association of the rs11125908 polymorphism with RA and its clinical variables. This article is the first report on the allele and genotype frequency of the rs11125908 polymorphism in the Cuban population and its association with RA.

MATERIALS AND METHODS

Patients

A non-family-based case-control study was designed. A total of 161 RA patients recruited in Havana hospitals for the Jusvinza study (RPCEC00000404) who fulfilled the 2010 classification criteria of the American College of Rheumatology for RA were enrolled as the case group (17). The following data was collected for all RA study participants: sex, age, age at disease onset, age at diagnosis, presence or absence of erosions, and, extra-articular manifestations. Disease activity was assessed using the disease activity score 28 with ESR (DAS28) (18, 19). The control group consisted of 150 subjects from the general population who were strictly unrelated and without a personal history of chronic inflammatory or autoimmune diseases. The control population was selected and matched by sex and age from the DNA bank collection of the Center for Genetic Engineering and Biotechnology (CIGB) Pharmacogenomics group.

Ethics approval and consent to participate

The study was approved by the ethics committee of the respective medical centers and was performed according to the principles of the Helsinki Declaration. Informed consent was obtained from all of the participants for conducting genetic studies.

Blood collection and sample preparation

Venous blood was drawn from all subjects and collected in tubes containing ethylenediaminetetraacetic acid. DNA from peripheral leukocytes was extracted using a Wizard® Genomic DNA Purification Kit from Promega (Promega, Madison, WI, USA).

Oligonucleotides

The oligonucleotides used for the detection of the rs11125908 polymorphism for Sanger sequencing and allele-specific polymerase chain reaction (AS-PCR) methods were supplied by CIGB (CIGB, Havana, Cuba). Oligos were designed using the NCBI Primer-BLAST program (https://www.ncbi.nlm.nih.gov/tools/primer-blast). The forward primer was common for all reactions: 5´TCCCATTCTACCCTGTTCAATGT3´. The reverse oligo used for the assessment and corroboration of genotyping through the Sanger method was 5´TGAAGTCCATAATTCCCAGCTCT3´. To detect the presence of the polymorphism A by AS-PCR, we use the reverse primer 5´CAGCTCTTTGATTCAACCTATCACT3´ and for the polymorphism G, the primer 5´CAGCTCTTTGATTCAACCTATCACC3´.

Genotyping of samples by the Sanger method

The amplification conditions were as follows: 95ºC for 5 minutes, 95ºC for 10 seconds, 66ºC for 10 seconds, 72ºC for 20 seconds, for 35 cycles. Each PCR reaction was performed in 20 µL of reaction volume, containing 50 ng of the DNA sample, 10 pmol of oligonucleotides and 10 µL of 2X master mix. The purification of the amplified samples was carried out with the kit of the QIAGEN firm, "QIAquick® PCR Purification kit" with the QIAcubeTM equipment (QIAGEN, Hilden, Germany). Fifteen samples from healthy individuals were sequenced by automated sequencing based on the Sanger enzymatic method. The results of the sequences were analyzed using the bioinformatics program Finch TV version 1.4.0 (Informer Technologies, Inc., Los Angeles, CA, USA).

Genotyping of samples by allele-specific polymerase chain reaction methodology

AS-PCR assays were performed on the LightCycler® 480 real-time PCR system (Roche Diagnostics, Mannheim, Germany). Each sample was analyzed in duplicate under the same conditions. The amplification conditions were as follows: denaturation at 95ºC for 5 minutes, 95ºC for 10 seconds, hybridization for 10 seconds at 66ºC, and extension at 72ºC for 15 seconds, with 35 amplification cycles. Each 20 µL PCR reaction contained 10 µL of 2X Master mix, 50 ng of the DNA sample and 10 pmol of corresponding oligonucleotides. To determine the crossing point values (Cp), the LightCycler® 480 software was used with the absolute quantification protocol of the maximum of the second derivative. The genotype assignment was established as follows: for $|\Delta Cp|<1$, the heterozygous genotype was assigned, $|\Delta Cp|>2$ for AA or GG homozygotes. A positive control for each genotype (AA, GG and AG) and a negative control was used in each experimental run.

Quality control

Sequencing of 19 samples with values of $|\Delta Cp|$ between one and two was performed to verify the assignation of the polymorphism. The agreement rate between the assignation by AS-PCR and sequencing was 100%.

Statistical analysis

A statistical analysis of genotype and allelic frequencies, Hardy-Weinberg equilibrium (HWE) and association analysis was performed using SNPStats (https://www.snpstats.net/start.htm, accessed on 2 May 2023) (20). Results with $p<0.05$ were considered statistically significant. The statistical analysis determined the risk of an event [odds ratio (OR)] and the confidence interval (95% CI) with the use of a linear regression model. Power calculations were conducted using Quanto 1.2.4 software (Informer Technologies, Inc., Los Angeles, CA, USA) (21).

Sample power

Quanto 1.2.4 power calculator was used to detect the power of the study considering

the total sample size (both cases and controls) studied (n=311), genotype relative risks of two for rs11125908 and the significance level equal to 0.05. We had 81.14% power, which indicates that the study sample size had sufficient power to detect the association (22, 23).

RESULTS

Demographic and clinical data

The age of RA patients ranged from 19 to 60 years, with a median of 50.29±7.92 years. The disease's duration was comprised between 2 and 42 years. The number of swollen joints ranged from 0 in one patient to 16. The estimated DAS28 values ranged from 3.25 to 10. Only six (0.03%) patients had high DAS activity (DAS28>5.1). Demographic and clinical data of patients included in the case group, or RA patients, are shown in Table I.

Allele and genotypic frequencies of the rs11125908 polymorphism in the Cuban population

The genotyping of 150 control samples and 161 from patients with RA was performed using the AS-PCR method. 12% of the control samples and 6% of the cases were in the range of |ΔCp| between 1 and 2, which was expected to be an ambiguous range, and genotype assignment was resolved by visual inspection of PCR amplification curves and Sanger sequencing with a 100% concordance between both methods. All study participants were genotyped for the rs11125908 polymorphism. There were no differences in the allele frequency values for cases and controls ($\chi 2$=1.17; p=0.28). However, in the distribution of genotypic frequencies, we observed differences ($\chi 2$=11.59; p=0.003) (Table II).

Table I - Demographic and clinical data of the patients with rheumatoid arthritis.

Parameters	Values
Age (mean±SD), years	50.29±7.92
Whites to non-whites ratio	0.87
Female to male ratio	6.31
Age of onset (mean±SD), years	41.03±11.52
Disease duration (mean±SD), years	9.26±8.25
Number of swollen joints	5.52±3.57
DAS28 (mean±SD)	4.76±0.63

SD, standard deviation; DAS28, disease activity score 28.

Association between rs11125908 polymorphism and rheumatoid arthritis

The observed genotypic frequencies in the control group were in HWE (p>0.05) and could be used for the following analysis. Departure from HWE resulted in genotype distribution in cases (P<0.0001) which is indicative of a strong association of this SNP with RA. This study used logistic regression adjusted by sex and age to test the association between SNP and RA under different genetic models. In our study, association analyses were performed under several genetic models, specifically codominant, dominant, recessive, overdominant, and log additive, to avoid possible biases in finding and reporting significant associations (20). In the models analyzed (Table III), only the association with RA is estimated in the AG genotype in the overdominant model (p=0.025; OR=1.91; 95% CI=1.08-3.35).

Association analysis of variables related to rheumatoid arthritis

Finally, we studied the possible association between the presence of the *COMMD1* allele and variables related to the course and outcome of the disease. Therefore, for the association analysis, we divided the RA cohort into two groups using the median values of the following variables: disease duration, number of inflamed joints, DAS28 and age of onset. With the number of inflamed joints lower than the average, the codominant and overdominant models showed an association. The results obtained are shown in Table IV. Our analyses revealed significant association of rs11125908 with patients with swollen joints < median under codominant model for AG (P=0.034; OR=2.30; 95% CI=1.14-4.62) and the GG genotype (p=0.034; OR=0.82; 95% CI=0.34-1.97) and with the overdominant model (p=0.01; OR=2.38; 95% CI=1.20-

Table II - Allele and genotypic frequencies of the rs11125908 polymorphism of the COMMD1 gene in the Cuban population.

Alleles/genotypes	Controls		Cases		OR (95% CI)	p value
	#	Frequency	#	Frequency		
A	211	0.7033	239	0.7422	1.21 (0.85-1.73)	0.28
G	89	0.2967	83	0.2578		
AA	76	0.5067	102	0.6335	1.00	
AG	59	0.3933	35	0.2174	*0.44 (0.26-0.73)*	*0.003**
GG	15	0.1000	24	0.1491	1.19 (0.59-2.42)	
HWE (p value)	>0.05		<0.0001*			

HWE, Hardy-Weinberg equilibrium; OR, odds ratio; CI, confidence interval.

Table III - Analysis of the association between rheumatoid arthritis and single nucleotide polymorphism of *COMMD1* (rs11125908) with adjustment for age and gender.

Model	Allele/genotypes	Case (%)	Control (%)	OR (95% CI)	p value	AIC	BIC
Codominant	A/A	102 (63.4)	76 (50.7)	1.00	0.079	354	372.7
	A/G	35 (21.7)	59 (39.3)	1.88 (1.05-3.37)			
	G/G	24 (14.9)	15 (10)	0.93 (0.41-2.11)			
Dominant	A/A	102 (63.4)	76 (50.7)	1.00	0.11	354.5	369.5
	A/G-G/G	59 (36.6)	74 (49.3)	1.53 (0.91-2.56)			
Recessive	A/A-A/G	137 (85.1)	135 (90)	1.00	0.48	356.6	371.6
	G/G	24 (14.9)	15 (10)	0.75 (0.34-1.68)			
Overdominant	A/A-G/G	126 (78.3)	91 (60.7)	1.00	*0.025**	352.1	367
	A/G	*35 (21.7)*	*59 (39.3)*	*1.91 (1.08-3.35)*			
Log-additive	–	–	–	1.16 (0.81-1.66)	0.42	356.5	371.4

CI, confidence interval; OR, odds ratio; AIC, Akaike information criterion; BIC, Bayesian information criterion; p values were calculated by logistic regression analysis with adjustment for age and gender; *p<0.05 indicates statistical significance.

4.70). In the analysis, when the age of onset of the disease was lower than the mean, the AG genotype showed an association in the overdominant model (p=0.027; OR=2.27; 95% CI=1.07-4.82). DAS28 and disease duration variables were not associated with the rs11125908 polymorphism.

DISCUSSION AND CONCLUSIONS

RA is caused by the interaction of genetic and environmental factors (10). In recent years, the association between genetic polymorphisms and diseases has been the focus of attention. RA can occur at any age, but it most commonly begins in middle age. Age of onset and sex may influence the disease course, outcome, and treatment. The demographic characteristics of our cohort of RA patients are very similar to those of other cohorts of patients in terms of variables such as sex and age of onset (2, 24, 25). RA is more frequent in women, with a reported women-to-men ratio of 3:1 (24, 25). In our case-control study, the ratio of wom-

Table IV - Association of the various parameters of rheumatoid arthritis with the single nucleotide polymorphisms "rs11125908" of *COMMD1* gene in the subgroup tests with adjustment for age and gender.

Model	Genotype	Case (%)	Control (%)	OR (95% CI)	p value	AIC	BIC
Swollen joints < median (5.52)							
Codominant	A/A	58 (65.2)	76 (50.7)	1.00			
	A/G	16 (18)	59 (39.3)	*2.30 (1.14-4.62)*	*0.034**	274.6	292
	G/G	15 (16.9)	15 (1)	*0.82 (0.34-1.97)*			
Dominant	A/A	58 (65.2)	76 (50.7)	1.00	0.11	276.8	290.7
	A/G-G/G	31 (34.8)	74 (49.3)	1.62 (0.90-2.92)			
Recessive	A/A-A/G	74 (83.2)	135 (90)	1.00	0.3	278.3	292.2
	G/G	15 (16.9)	15 (10)	0.64 (0.27-1.50)			
Overdominant	A/A-G/G	73 (82)	91 (60.7)	1.00	*0.01**	272.8	286.7
	A/G	16 (18)	59 (39.3)	*2.38 (1.20-4.70)*			
Log-additive	–	–	–	1.14 (0.76-1.71)	0.53	279	292.9
Swollen joints > median (5.52)							
Codominant	A/A	44 (61.1)	76 (50.7)	1.00			
	A/G	19 (26.4)	59 (39.3)	1.43 (0.69-2.97)	0.6	226.6	243.6
	G/G	9 (12.5)	15 (10)	1.29 (0.45-3.67)			
Dominant	A/A	44 (61.1)	76 (50.7)	1.00	0.32	224.6	238.2
	A/G-G/G	28 (38.9)	74 (49.3)	1.39 (0.72-2.69)			
Recessive	A/A-A/G	63 (87.5)	135 (90)	1.00	0.8	225.5	239.2
	G/G	9 (12.5)	15 (10)	1.14 (0.41-3.15)			
Overdominant	A/A-G/G	53 (73.6)	91 (60.7)	1.00	0.38	224.8	238.4
	A/G	19 (26.4)	59 (39.3)	1.37 (0.68-2.79)			
Log-additive	–	–	–	1.22 (0.76-1.95)	0.41	224.9	238.5
DAS28 < median (4.76)							
Codominant	A/A	37 (64.9)	76 (50.7)	1.00			
	A/G	12 (21.1)	59 (39.3)	2.07 (0.96-4.47)	0.15	224.2	240.8
	G/G	8 (14)	15 (10)	0.99 (0.36-2.75)			
Dominant	A/A	37 (64.9)	76 (50.7)	1.00	0.14	223.7	237.1
	A/G-G/G	20 (35.1)	74 (49.3)	1.66 (0.85-3.24)			
Recessive	A/A-A/G	49 (86)	135 (90)	1.00	0.64	225.7	239.1
	G/G	8 (14)	15 (10)	0.79 (0.29-2.13)			
Overdominant	A/A-G/G	45 (79)	91 (60.7)	1.00	0.052	222.2	235.5
	A/G	12 (21.1)	59 (39.3)	2.07 (0.97-4.39)			
Log-additive	–	–	–	1.23 (0.76-1.98)	0.4	225.2	238.6
DAS28 > median (4.76)							
Codominant	A/A	65 (62.5)	76 (50.7)	1.00			
	A/G	23 (22.1)	59 (39.3)	1.68 (0.86-3.31)	0.29	273.1	290.8
	G/G	16 (15.4)	15 (10)	0.99 (0.40-2.47)			
Dominant	A/A	65 (62.5)	76 (50.7)	1.00	0.24	272.2	286.3
	A/G-G/G	39 (37.5)	74 (49.3)	1.43 (0.79-2.59)			
Recessive	A/A-A/G	88 (84.6)	135 (90)	1.00	0.7	273.4	287.6
	GG	16 (15.4)	15 (10)	0.84 (0.35-2.04)			

Continued >>>

Continued >>>

Model	Genotype	Case (%)	Control (%)	OR (95% CI)	p value	AIC	BIC
Overdominant	A/A-G/G	81 (77.9)	91 (60.7)	1.00	0.12	271.1	285.2
	A/G	23 (22.1)	59 (39.3)	1.69 (0.87-3.25)			
Log-additive	–	–	–	1.14 (0.76-1.72)	0.53	273.2	287.3
Disease duration < median (9.26)							
Codominant	A/A	65 (65)	76 (50.7)	1.00	0.16	291.1	308.7
	A/G	23 (23)	59 (39.3)	1.87 (0.98-3.55)			
	G/G	12 (12)	15 (10)	1.19 (0.47-3.02)			
Dominant	A/A	65 (65)	76 (50.7)	1.00	0.087	289.9	304
	A/G-G/G	35 (35)	74 (49.3)	1.65 (0.93-2.94)			
Recessive	A/A-A/G	88 (88)	135 (90)	1.00	0.95	292.8	306.9
	G/G	12 (12)	15 (10)	0.97 (0.39-2.42)			
Overdominant	A/A-G/G	77 (77)	91 (60.7)	1.00	0.059	289.3	303.4
	A/G	23 (23)	59 (39.3)	1.82 (0.97-3.41)			
Log-additive	–	–	–	1.29 (0.85-1.96)	0.23	291.4	305.5
Disease duration > median (9.26)							
Codominant	A/A	37 (60.7)	76 (50.7)	1.00	0.18	205.3	222.1
	A/G	12 (19.7)	59 (39.3)	1.86 (0.82-4.19)			
	G/G	12 (19.7)	15 (10)	0.71 (0.26-1.91)			
Dominant	A/A	37 (60.7)	76 (50.7)	1.00	0.43	206.2	219.6
	A/G-G/G	24 (39.3)	74 (49.3)	1.32 (0.66-2.62)			
Recessive	A/A-A/G	49 (80.3)	135 (90)	1.00	0.27	205.6	219
	G/G	12 (19.7)	15 (10)	0.58 (0.22-1.53)			
Overdominant	A/A-G/G	49 (80.3)	91 (60.7)	1.00	0.083	203.8	217.2
	A/G	12 (19.7)	59 (39.3)	1.99 (0.90-4.38)			
Log-additive	–	–	–	1.00 (0.63-1.59)	1	206.8	220.2
Age of onset < median RA							
Codominant	A/A	38 (64.4)	76 (50.7)	1.00	0.084	225.7	242.4
	A/G	12 (20.3)	59 (39.3)	2.22 (1.03-4.79)			
	G/G	9 (15.2)	15 (10)	0.87 (0.32-2.35)			
Dominant	A/A	38 (64.4)	76 (50.7)	1.00	0.13	226.4	239.7
	A/G-G/G	21 (35.6)	74 (49.3)	1.66 (0.86-3.21)			
Recessive	A/A-A/G	50 (84.8)	135 (90)	1.00	0.43	228	241.4
	G/G	9 (15.2)	15 (10)	0.68 (0.26-1.78)			
Overdominant	A/A-G/G	47 (79.7)	91 (60.7)	1.00	*0.027**	223.8	237.1
	A/G	12 (20.3)	59 (39.3)	*2.27 (1.07-4.82)*			
Log-additive	–	–	–	1.18 (0.74-1.89)	0.48	228.2	241.5
Age of onset > median RA							
Codominant	A/A	64 (62.8)	76 (50.7)	1.00	0.41	258.6	276.2
	A/G	23 (22.6)	59 (39.3)	1.59 (0.79-3.19)			
	G/G	15 (14.7)	15 (10)	0.99 (0.39-2.51)			
Dominant	A/A	64 (62.8)	76 (50.7)	1.00	0.32	257.4	271.5
	A/G-G/G	38 (37.2)	74 (49.3)	1.36 (0.74-2.52)			

Continued >>>

Continued >>>

Model	Genotype	Case (%)	Control (%)	OR (95% CI)	p value	AIC	BIC
Recessive	A/A-A/G	87 (85.3)	135 (90)	1.00	0.74	258.2	272.4
	G/G	15 (14.7)	15 (10)	0.86 (0.34-2.13)			
Overdominant	A/A-G/G	79 (77.5)	91 (60.7)	1.00	0.18	256.6	270.7
	A/G	23 (22.6)	59 (39.3)	1.59 (0.81-3.14)			
Log-additive	–	–	–	1.12 (0.73-1.71)	0.6	258.1	272.2

CI, confidence interval; OR, odds ratio; AIC, Akaike information criterion; BIC, Bayesian information criterion; DAS28, disease activity score 28, RA, arthritis rheumatoid; p values were calculated by logistic regression analysis with adjustment for age and gender; *p<0.05 indicates statistical significance.

en to men was 6:1. These observed differences between the sex incidence of RA in populations and our study have been described in other RA clinical investigations. A 2007 study, described by Galea and Tracy (26), suggested that women were more likely to participate in scientific research than their male counterparts. In addition, the underrepresentation of men in RA randomized clinical trials was also explained by the usually publicized increased prevalence of rheumatic diseases in women, which may have affected their inclusion in clinical trials (27).

The association of the rs11125908 polymorphism with RA has not been described previously in the Cuban population. We found the allele frequencies in the control group for rs11125908 were 0.7033 for the allele A and 0.2967 for the minor G allele. These allele frequencies are closest to the frequency values of the European population (A=0.665613, G=0.334387), the Japanese population (A=0.70758, G=0.29242), and the Latin American population (A=0.6648, G=0.3352). Interestingly, the polymorphism of the Cuban population is very different from that of the African population (A=0.9028, G=0.0972) (https://www.ncbi.nlm.nih.gov/snp/?term=rs11125908, *Supplementary Table 1*), taking into account the admixture of the Cuban population in Havana City (28). In addition, the frequency of this polymorphism is similar in the group of RA patients who self-classified as white or not white (*Supplementary Table II*).

Considering the results obtained by Murata *et al*. in 2017 (12), where in Japanese RA patients the minor G allele at rs11125908 showed a higher expression of *COMMD1* and was associated with less bone erosions of hand joints, a protective association of AG and GG genotype compared to AA on RA would be expected. We show in Table II the distribution of the rs11125908 alleles and genotypes in the Cuban population. We could not find differences in allele frequencies between cases and controls. However, the analysis of the genotypic frequencies ($\chi 2$=11.59; p=0.003) shows an association of the rs11125908 polymorphism with RA. The AG genotype shows a protective effect with an OR of 0.44 (0.26-0.73). In our sample, we could not find an association between the alleles GG and RA. An increase in the genotypic frequency for AA of 12.68% was found in these cases, which would imply an increase in bone erosion and number of swollen joints, making these patients more difficult to treat. It is striking that the frequency of the GG genotype increased by 4.9% in the cohort of RA patients with respect to the control group. This cohort of patients should show a moderate manifestation of symptoms with a slow progression of disease. Nevertheless, the complex genetic nature of the disease, in combination with the environmental factors that contribute to its course, makes it difficult to reach conclusions. Unfortunately, the article by Murata *et al.* does not publish the allele and genotypic frequency of the polymorphism, making it impossible to establish a comparison between the Japanese and Cuban populations (12).

The choice of genetic model in RA studies remains open, but investigators seem to test the association between SNPs under different genetic models (29). Moreover, sex and age are linked to RA (2). For this reason, we decided to estimate the association with ad-

justment for the covariates age and sex and with multiple inheritance models. Under these assumptions, the AG genotype continues to show a significant association of RA with the overdominant model [OR=1.91; (1.08-3.35 at 95% CI); p=0.025] (Table III). This model of inheritance has been observed in the meta-analysis of the interleukin-1 α +4845G/T (rs17561) SNP in RA (29).

We found very few published studies that have evaluated the association between radiographic joint damage and genetic analysis. Rodriguez-Rodriguez et al. reported a probable association between the prostaglandin E receptor 4 variant, rs76523431 and radiographic joint destruction in Caucasian patients with RA (30). Suzuki et al. reported an association between the PADI4 risk allele and radiographic joint destruction amongst Japanese patients with RA (31). In addition, rs2295926, belonging to the *GALNT12* gene and rs11958855 belonging to the *KCNN2* gene were strongly associated with rapid radiographic joint destruction (32).

Taking into consideration that rs11125908 G alleles demonstrated a strong suppressive effect on finger joint damage, independently of rheumatoid factor positivity (12), we decided to study the relation of this SNP with clinical variables such as age at onset, disease duration, number of inflamed joints, and DAS28 (Table IV), distributing the RA cohort into two groups using the median values of the variables. As expected, the presence of the allele G in the genotype showed an association with a lower number of swollen joints in the codominant (AG and GG) and overdominant models (A/G). It was surprising that the AG genotype showed an association with a lower age of onset in the overdominant model, as we expect an association with the subgroup of higher age onset. DAS28 in addition to swollen joint count, includes in the calculation other parameters such as tender joint count, erythrocyte sedimentation rate, and general health, and probably the influence of the other variables on DAS28 is relevant. No association with the SNP was found under any model.

The present study has some limitations. First, we did not measure the protein levels of *COMMD1* in the control or patient cohorts to be sure that the Murata et al. results were replicated in the Cuban population. Secondly, the stratification of the RA group decreases the power of the study from 0.7965 for the number of swollen joints to 0.5773 for the age of onset subgroup. We suggest increasing the size of the sample to increase the reliability of our results in stratified analyses. Finally, the cohorts were comprised entirely of Cuban patients; SNPs may differ in subjects of different ethnicities.

For the first time, the frequency of the rs11125908 polymorphism in the Cuban population has been described, and its association with RA has been analyzed. Considering the results of this case-control study, the rs11125908 polymorphism is associated with RA, and the GG genotype could be protective. However, the latter is just speculation, which may need further studies with a larger sample size. To the best of our knowledge, our study is the first to find evidence that this polymorphism is potentially associated with the development of RA and possible protective in the Cuban population. These results could provide new ideas for individualized treatment of RA based on the increase in *COMMD1* levels, to improve disease management and quality of life of patients (32).

Contributions

MCA, ACB, JRFM, conceptualization and design; MCA, ACB, HNC, MCDH, JRFM, analysis and interpretation of the data; MCA, JRFM, drafting of the article; TDA, HNC, provision of study patients and assembly of clinical data and samples; ACB, statistical expertise. All authors were involved in the revisions of the manuscript and approved the final version for submission.

Conflict of interest

The authors' salary is paid by the sponsoring institution of the Clinical Trial (CIGB).

Ethics approval and consent to participate

The study was approved by the ethics committee of the respective medical centers and

was performed according to the principles of the Helsinki Declaration.

Informed consent:
Informed consent was obtained from all of the participants for conducting genetic studies.

Funding
This work was supported by the *Ministerio de Ciencia, Tecnología y Medio Ambiente* (CITMA, Project No. PN385LH007-007).

Availability of data and materials
Primary data are available on request from the authors.

Acknowledgments:
The authors are grateful to RA patients and doctors who take care of them: Dinorah Prada Hernández and Jorge Alexis Gómez Morejón from the Cuban Rheumatology Institute and Araceli Chico Capote from "Hermanos Ameijeiras" Surgical Clinical Hospital.

Additional information
This article was already published as a non-peer-reviewed article in Research Square. Doi: https://doi.org/10.21203/rs.3.rs-3283108/v1.

REFERENCES

1. Finckh A, Gilbert B, Hodkinson B, Bae SC, Thomas R, Deane KD, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol 2022; 18: 591-602.
2. Nilsson J, Andersson ML, Hafström I, Svensson B, Forslind K, Ajeganova S, et al. Influence of age and sex on disease course and treatment in rheumatoid arthritis. Open Access Rheumatol 2021; 13: 123-38.
3. Petrelli F, Mariani FM, Alunno A, Puxeddu I. Pathogenesis of rheumatoid arthritis: one year in review 2022. Clin Exp Rheumatol 2022; 40: 475-82.
4. Jin S, Zhao J, Li M, Zeng X. New insights into the pathogenesis and management of rheumatoid arthritis. Chronic Dis Transl Med 2022; 8: 256-63.
5. Akhtar M, Ali Y, Islam ZU, Arshad M, Rauf M, Ali M, et al. Characterization of rheumatoid arthritis risk-associated SNPs and identification of novel therapeutic sites using an in-silico approach. Biology (Basel) 2021; 10: 501.
6. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS catalog: knowledgebase and deposition resource. Nucleic Acids Res 2023; 51: D977-85.
7. Aluko A, Ranganathan P. Pharmacogenetics of drug therapies in rheumatoid arthritis. Methods Mol Biol 2022; 2547: 527-67.
8. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nature Genetics 2010; 42: 508-14.
9. Lee YH, Bae S-C, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. Mol Biol Rep 2012; 39: 10627-35.
10. Okada Y, Eyre S, Suzuki A, Kochi Y, Yamamoto K. Genetics of rheumatoid arthritis: 2018 status. Ann Rheum Dis 2019; 78: 446-53.
11. Marcheco-Teruel B, Parra EJ, Fuentes-Smith E, Salas A, Buttenschon HN, Demontis D, et al. Cuba: exploring the history of admixture and the genetic basis of pigmentation using autosomal and uniparental markers. PLoS Genet 2014; 10: e1004488.
12. Murata K, Fang C, Terao C, Giannopoulou EG, Lee YJ, Lee MJ, et al. Hypoxia-Sensitive COMMD1 integrates signaling and cellular metabolism in human macrophages and suppresses osteoclastogenesis. Immunity 2017; 47: 66-79.e5.
13. Riera-Romo M. COMMD1: a multifunctional regulatory protein. J Cell Biochem 2018; 119: 34-51.
14. Vockley CM, Barrera A, Reddy TE. Decoding the role of regulatory element polymorphisms in complex disease. Curr Opin Genet Dev 2017; 43: 38-45.
15. Salih MH, Al-Azzawie AF, Al-Assie AHA. Intronic SNPs and genetic diseases: a review. Int J Res Appl Sci Biotechnol 2021; 8: 267-74.
16. Cronstein BN, Aune TM. Methotrexate and its mechanisms of action in inflammatory arthritis. Nat Rev Rheumatol 2020; 16: 145-54.
17. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010; 69: 1580-8.
18. Kow J, Lim GH, Tan YK. Applying 28-joint disease activity score (DAS28) and swollen joint count together as single time-point measures better depict rheumatoid arthritis disease status when compared to DAS28 alone: Perspectives from an ultrasound imaging study. Int J Rheum Dis 2023; 26: 581-4.
19. van Riel PL. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clin Exp Rheumatol 2014; 32: S-65-74.

20. Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. Bioinformatics 2006; 22: 1928-9.
21. Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 2002; 21: 35-50.
22. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample size calculations. Nephron Clin Pract 2011; 118: c319-23.
23. Politi C, Roumeliotis S, Tripepi G, Spoto B. sample size calculation in genetic association studies: a practical approach. Life (Basel) 2023; 13: 235.
24. Nishino T, Hashimoto A, Tohma S, Matsui T. Comprehensive evaluation of the influence of sex differences on composite disease activity indices for rheumatoid arthritis: results from a nationwide observational cohort study. BMC Rheumatol 2023; 7: 4.
25. van Vollenhoven RF. Sex differences in rheumatoid arthritis: more than meets the eye. BMC Med 2009; 7: 12.
26. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007; 17: 643-53.
27. Dong G, Huang B, Verbeeck J, Cui Y, Song J, Gamalo-Siebers M, et al. Win statistics (win ratio, win odds, and net benefit) can complement one another to show the strength of the treatment effect on time-to-event outcomes. Pharm Stat 2023; 22: 20-33.
28. Cintado A, Companioni O, Nazabal M, Camacho H, Ferrer A, De Cossio ME, et al. Admixture estimates for the population of Havana City. Ann Hum Biol 2009; 36: 350-60.
29. Padyukov L. Genetics of rheumatoid arthritis. Semin Immunopathol 2022; 44: 47-62.
30. Rodriguez-Rodriguez L, Ivorra-Cortes J, Carmona FD, Martin J, Balsa A, van Steenbergen HW, et al. PTGER4 gene variant rs76523431 is a candidate risk factor for radiological joint damage in rheumatoid arthritis patients: a genetic study of six cohorts. Arthritis Res Ther 2015; 17: 306.
31. Suzuki T, Ikari K, Yano K, Inoue E, Toyama Y, Taniguchi A, et al. PADI4 and HLA-DRB1 are genetic risks for radiographic progression in RA patients, independent of ACPA status: results from the IORRA cohort study. PLoS One 2013; 8: e61045.
32. Hayashi S, Matsubara T, Fukuda K, Maeda T, Funahashi K, Hashimoto M, et al. A genome-wide association study identifying the SNPs predictive of rapid joint destruction in patients with rheumatoid arthritis. Biomed Rep 2021; 14: 31.

Supplementary material available online