TITLE: Association of Complement Receptor 2 polymorphisms with innate resistance to HIV-1 infection

RUNNING TITLE: Complement genetics and HIV-1 infection

AUTHORS: Rocio Herrero[1], Luis Miguel Real[2], Antonio Rivero-Juárez[3], Juan Antonio Pineda[2], Juan Macías[2], Ángela Camacho[3], Marina Laplana[4], Piotr Konieczny[1], Francisco J. Márquez[1], Juan-Carlos Souto[5], Jose-Manuel Soria[5], Mara Biasin[6], Irma Saulle[6], Sergio Lo Caputo[7], Antonio Rivero[3], Joan Fibla[4], Antonio Caruz[1*]

[1]Immunogenetics Unit, Department of Experimental Biology, University of Jaen, Jaen, Spain. [2]Infectious Diseases and Microbiology Clinical Unit. Valme Hospital, Seville, Spain. [3]Maimonides Institute for Research in Biomedicine of Cordoba (IMIBIC)/Reina Sofia University Hospital, Cordoba, Spain. [4]Human Genetics Unit, Department of Basic Medical Sciences, University of Lleida IRBLleida, Lleida, Catalonia, Spain, [5] Institut d'Investigació Biomèdica Sant Pau (IIB-Sant Pau), Hospital de la Santa Creu i de Sant Pau, Barcelone, Spain. [6] Department of Biomedical and Clinical Sciences, University of Milan, Milan, Italy. [7] Maria Annunziata Hospital, Florence, Italy.

[*]Corresponding author.

Corresponding author:

Antonio Caruz, PhD. Immunogenetics Unit, University of Jaen, Campus Las Lagunillas SN, 23071, Jaén, Spain. Telephone: +34 606352046. Fax: +34 953211875. e-mail: caruz@ujaen.es

Abstract

HIV-1 induces activation of complement through the classical and lectin pathways. However, the virus incorporates several membrane-bound or soluble regulators of complement activation (RCA) that inactivate complement. HIV-1 can also use the complement receptors (CRs) for complement-mediated antibody-dependent enhancement of infection (C´-ADE). We hypothesize that polymorphisms in RCA or CRs that reduce the ability of complement inhibition may protect from HIV-1 infection. For this purpose, 139 SNPs located in 19 RCA and CRs genes have been genotyped in a population of 201 Spanish HIV-1 exposed seronegative individual (HESN) and 250 HIV-1 infected patients. Two SNPs were associated with infection susceptibility, rs1567190 in *CR2* (OR = 2.27, $P= 1x10^{-4}$) and rs2842704 in *C4BPA* (OR2.11, $P= 2x10^{-4}$). To replicate this finding, we analyzed a cohort of Italian, sexually HESN individuals. Although not significant (P=0.25, OR=1.57), similar genotypic proportions were obtained for the *CR2* marker rs1567190. The results of the two association analyses were combined through a random effect meta-analysis, with a significant P value of $2.6x10^{-5}$ (OR=2.07). Furthermore, we found that the protective *CR2* genotype is correlated with lower levels *CR2* mRNA as well as differences in the ratio of the long and short CR2 isoforms.

Key words: HIV-1, HIV-1 exposed seronegative, injection drug users, complement, CR2, C4BPA

Introduction

HIV-1 is able to activate the classical pathway of complement in the absence of antibodies, such activation depends on direct binding of C1q envelope protein [1] and the attachment of C3b on the viral membrane. This fact is enhanced in the presence of specific anti-HIV-1 antibodies [2]. In addition, HIV-1 interacts with Mannose Binding Lectin (MBL) through direct binding to gp120, triggering the activation of the complement lectin pathway [3]. However, the virus appears to be resistant to complement [4,5]. The biological model to explain this fact is based in the active viral exploitation of the regulators of complement activation (RCA) protein family. The exploitation of RCA proteins by several families of viruses is a clear example of adaptive convergence [6]. In the case of HIV-1, it has been demonstrated the absorption of membrane-bound RCA proteins in virions [4,7,8]. These proteins enhance virus resistance to complement-mediated lysis [4,9,10]. This incorporation contributes to the survival of HIV-1 in blood or other tissues as both virions and infected cells are resistant to neutralizing antibodies [10]. Other soluble proteins such as CFH interacts with viral envelope [11]. Depletion of CFH in human serum increased destruction of cells infected by HIV-1 and free virus *in vitro* [11,12]. Similarly, viruses produced in cells defective for any of the RCA proteins are extremely sensitive to killing by serum in the presence or absence of neutralizing antibodies. Interestingly *Streptococcus spp*. has developed a similar strategy of resistance to serum by addition of CFH to its surface by the M protein [13]. Both complement genes and the RCA are extremely polymorphic and in some cases a functional effect of the polymorphisms has been found [14].

Complement-coated virions are concentrated in several tissues; especially the lymphatic system and can interact with complement receptors (CRs) that bind C3 fragments. Although viral inactivation through opsonisation and lysis have been observed [15], CRs expression in T /B lymphocytes or erythrocytes can also boost viral infectivity in a process known as complement-mediated antibody-dependent enhancement of infection (C´-ADE) [16], the CRs implicated in C´-ADE include CR1, CR2 and CR3. Both *cis* and *trans* C´-ADE have been observed on cell lines as well as on primary cells and with primary isolates of HIV-1[17].

In view of the above interactions of HIV-1 with RCA and complement receptors, we hypothesized that individuals carrying polymorphisms that reduce the functional capacity of the CRs or RCA proteins or their site of interaction with gp120/gp41 could display a natural resistance to HIV-1 infection mediated by an increased probability of virolisis mediated by full complement activation or lower complement-mediated enhancement of infection. To address this hypothesis we have performed a genetic association analysis as well as a phenotypic characterization of RCA and complement receptors in two independent cohorts of HESN and HIV-1 positive individuals.

Methods

Patients and controls

We recruited 451 white males exposed to HIV-1 infection by injection drug use (IDU-Spain), both infected (HIV, 250 subjects) and at HIV-1 exposed seronegative (HESN, 201 subjects), enrolled in prospective cohort studies in Spain, Arnau de Vilanova Hospital

(Lleida), Valme Hospital (Sevilla) and Reina Sofía Hospital (Córdoba) who had shared needles for >3 months [18]. Concurrent markers of hepatitis C virus (HCV) infection, one of the most prevalent blood borne virus, was present in 100% of HESN and HIV-1 positive subjects. The main epidemiological and clinical characteristics of the series studied here were described previously [18]. Additionally, a second population exposed to HIV-1 infection by sexual route was analyzed as replication population; both HIV-1 infected (62 subjects) and HESN (80 subjects), enrolled in prospective cohort studies in Italy (Sexual-Italy). The main characteristics of this population has been previously described [19].

Ethical aspects

The study was designed and performed according to the Helsinki declaration and was approved by the Ethics Committee of the three participating hospitals. All patients and donors provided written informed consent to participate in this study.

Polymorphism selection and genotyping

SNPs from complement regulatory genes (table 1) were selected according to two criteria, first, genotype data from Hapmap-CEU population were directly downloaded from the Hapmap database (www.hapmap.org) and the Tag SNPs were indentified using Haploview's V4.1 software tagger option (https://www.broad.harvard.edu/haploview/haploview). Second, SNPs list were completed with potentially functional polymorphisms previously associated or not with disease. DNA was extracted from fresh peripheral blood mononuclear cells or frozen whole blood using the Quick Pure Blood DNA extraction Kit (Macherey-Nagel, Düren, Germany) or Magna Pure system (Roche, Basel, Switzerland). The SNPs were

genotyped using a custom Golden Gate Veracode genotyping assay (Illumina, San Diego, California USA). Original genotyping data are available upon request.

Sequencing of C4BPA, CD55, CR2 and CR1 chromosome region

DNAs from 33 HESN individuals, homozygous CC for the rs1567190 SNP associated with the resistant phenotype, were mixed in 3 independent pools. In parallel, DNA from 46 HIV positive individuals homozygous TT associated with the risk phenotype were mixed in 5 independent pools. Sequencing of each pool DNA was done by custom capture and New Generation Sequencing using Illumina HiSeq2000 sequencer in Otogenetics facilities (www.otogenetics.com). The resequencing strategy was focused on 515.844 bp from chromosome 1 region including the genes *C4BPA*, *CD55*, *CR2* and *CR1* (coordinates from 207.224.455 bp to 207.740.299 bp, NCBI buildGRCh37/hg19). Raw data generated were analyzed using the SNP discovery tool of DNAnexus cloud computing software (www.dnanexus.com). We calculate a score for each polymorphic position identified (frequency of reads equal to the consensus sequence/number of total reads). The original sequencing data are available upon request.

Biostatistics and bioinformatics

Hardy-Weinberg equilibrium and pairwise linkage disequilibrium (D′) were calculated using Haploview V4.1 software. Block structure was considered for marker pairs showing D′ > 0.8, following the solid-spine block definition implemented in Haploview. Single marker and haplotype frequencies were estimated and compared using PLINK software, complemented with the Java-based implementation software gPLINK (http://pngu.mgh.harvard.edu/ ~purcell/plink/gplink.shtml). Complementary analysis was performed by use of SPSS software (version 14.0). Single-marker association P

values were corrected for multiple testing following the SNP spectral decomposition approach [20], experiment-wide significance threshold required to keep type I error rate at 5% was $5\ 10^{-4}$. Haplotype association tests were adjusted using a Bonferroni´s correction for multiple testing by dividing the significance level (0.05) by the number of major haplotypes (i.e., >5%), a p value $< 5\ 10^{-3}$ was considered statistically significant.

CR2 and C4BPA mRNA quantification

PBMC from healthy blood donors (City of Jaen, Blood Bank) were purified by centrifugation in Ficoll gradient, DNA was extracted as previously described and genotyped for the SNPs rs1567190 of *CR2* and rs2842704 of *C4BPA* with Taqman probes. Nine individuals harboring CC genotype and nine with TT genotype for *CR2* and ten individuals carrying AA genotype and ten with AG/GG genotype for *C4BPA* were randomly selected gene expression analyses. Then, from each sample, total RNA from $5x10^6$ PBMCs was isolated using Tripure reagent (Roche), quantified by spectrophotometry and disolved at 1 µg/µl. Reverse transcription of RNA was performed using the Superscript III (Invitrogen), following manufacturer's instructions. All the RNA samples were also subjected to control reactions without reverse transcriptase to check for contamination. The normalized *CR2* expression was determined in a multiplexed, one tube PCR reaction using the *CD19* as endogenous control (Hs01047410_g1, Applied Biosystems) and *CR2* Gene Expression Assays (Hs00153398_m1 FAM probe, Applied Biosystems). Quantitative real-time PCRs were performed in a total volume of 12.5 µl containing 6.25 µl of Quantimix Easy probes master mix (Biotools, Madrid, Spain), 0.25 µl of each probe/primer set CD19 and 1 µl of probe/primer set for CR2, 2 µl of cDNA and 3 µl of water. Cycling conditions were:

95 °C for 10 min, and 45 cycles at 95 °C for 15 s, 60 °C for 1 min. Similarly C4BPA expression was quantified. In that case, CD14 probe (Hs001691220_g1 VIC Probe, Applied Biosystems) was used as endogenous control mixed in the same PCR reaction tube with *C4BPA* probe (Hs00426339_m1 HEX probe, Applied Biosystems). The real-time PCR reaction were prepared in a total volume of 12.5 µl containing 6.25 µl of Quantimix Easyprobes master mix (Biotools, Madrid, Spain), 0.5 µl of each probe/primer set for CD14 and 1.2 µl of probe/primer set for C4BPA, 2 µl of cDNA and 2.55 µl of water. Cycling conditions were: 95 °C for 10 min, and 45 cycles at 95 °C for 25 s, 60 °C for 1 min. Fluorescence was monitored at the end of each annealing phase on a MX3005P thermocycler (Stratagene, La Jolla, CA, USA). Each donor cDNA were amplified in duplicate wells, and repeated in two independent days. Parallel negative control experiments were done with reverse transcriptase negative cDNAs. The relative expression of *CR2* and *C4BPA* mRNAs compared to endogenous *CD19* (*CR2*) or *CD14* (*C4BPA*) were calculated using the algorithm -$\Delta\Delta$ CT implemented in the MX-pro V4.0 software (Stratagene) according to manufacturer's instructions. The Comparative quantitation of short and long isoform of *CR2* transcripts has been done as previously described [21]

Results

Genotyping statistics

Four SNPs did not passed the genotyping quality controls of Illumina Veracode Golden Gate protocol rs7084554 (*MBL2*), rs12406509 (*CFHR4*), rs7046 (*CD59*) and rs9837104. Additionally, we found significative deviation from HWE (P=1.3 10^{-5}) in the marker rs9331949 (*CLUS*). Mean genotyping call rates for the rest of the SNPs were 99.6%.

Univariate genetic association analysis

We selected 139 polymorphic markers (Table 1) within regulators of complement activation and complement receptors genes for susceptibility to HIV-1 association in the IDU-Spain cohort. Two intronic SNPs positioned in the same region of the chromosome 1 (Figure 1), showed experiment-wide significant allelic or genotypic association with resistance to HIV infection. The G allele of rs2842704 in *C4BPA* was associated with decreased risk of infection under dominant model (P=2x10^{-4}; OR: 0.47 [0.31-0.70]). Subjects carrying GG+AG genotypes were significantly over-represented in HESN group (43%) compared to HIV-1 infected one (26%) (Table 2). Regarding *CR2* gene, the T allele of rs1567190 was associated with increased risk of HIV infection under a dominant model (P=1x10^{-4}; OR: 2.27 [1.46-3.51]) (Table 3). Subjects carrying TT+TC genotypes were significantly over-represented in the HIV+ group (81%) compared to HESN (65.5%). Three other SNPs showed a trend to association with HIV-1 infection protection but did not reach statistical significance after correction by multiple testing; rs419137 in *CFH* (P=6x10^{-3}), rs831618 in *CD59* (P=7x10^{-3}) and rs6690037 in *C4BPB* (P=8x10^{-3}) (supplementary table 1).

Additionally to replicate the previous findings, rs2842704 in *C4BPA* and rs1567190 in *CR2* were genotyped in a second population (Sexual-Italy) exposed to the virus through sexual route. Both genetic markers were in accordance with the HWE law. Genotype distribution of rs2842704 marker differs to that observed in the IDU-Spain sample not supporting the observed association for this SNP in the IDU-Spain cohort (Table 3).

Similarly to what was observed in the IDU-Spain sample, *CR2* rs1567190 TT+TC genotype was over-represented in HIV-1 (77.5%) compared to HIV-1 (68.5%), although

it did not reach statistical significance (*P*=0.25; OR: 1.57 [0.7-3.4]) Table 3). The results of the two association analyses were combined through a random effect metaanalysis, with a significant p value of $2.6x10^{-5}$ (Table 3). These observations suggest that the T allele of rs2074560 confers risk to HIV-1 infection with a dominant effect, independently infection route.

Haplotype association analysis in Spanish intravenous drug users

A haplotype-based association test was performed by multimarker regression analysis using Plink and Haploview software. Haplotype frequencies were estimated and compared between the two groups of the IDU-Spain cohort. This analysis identified 3 haplotypes associated with HIV-1 resistance with a significant multiple-testing p value $<5x10^{-3}$. Haplotypes from, *CFH*, *C4BPA* and *CR2* genes included SNPs previously identified as associated in the univariate analysis (Table 4). Using Plink software, we calculated the chromosome phases for HESN and HIV groups and performed haplotype copy number calculation in HIV-1 positive and HESN individuals. The presence of 2 copies of the haplotype CGGGG encompassing the promoter, and the first 12 exons of *CR2* (rs1567190, rs1507764, rs1048971, rs17615, rs4317805) showed the strongest association with protection under a recessive model (*P*=$8x10^{-5}$) (table 4). The other two haplotypes associated with innate resistance to HIV-1 infection followed a dominant model and included presence of one or two copies the haplotype GCT (rs2842704, rs4425986 and rs9943077) located in the fourth intron of *C4BPA* (*P*=$1.1x10^{-3}$) (Table 4). As well as the haplotype AAACTGCG (rs2019727, rs6695321, rs10737680, rs419137, rs395544, rs2284664, rs11582939, rs16840658) that spam from the intron 9 of *CFH* to the 5´ end of *CFHR1* gene (Table 4).

Sequencing of *C4BPA*, *CD55*, *CR2* and *CR1* chromosome region

In order to identify new functional SNPs in the candidate region and perform a sequence haplotying, a mean of 900,000 bp have been sequenced confidently with a 30X mean coverage. Spanning the chromosome 1 region 207,224,455-207,740,299 (GRCh37/hg19); including *C4BPB*, *C4BPA*, *CD55*, *CR2*, *CR1* and *CR1L* in 5 pools of HIV positive homozygous for the risk (T) allele of the SNP rs1567190 and 3 pools of HESN patients homozygous for the protective (C) allele of the SNP rs1567190. The main statistics for pool sequencing can be found in Supplementary Table 2. We observed a perfect match between the two genotyping systems Veracode Golden Gate and New Generation Sequencing. A total of 189 new polymorphic sites have been identified in the region (supplementary table 3), most of them located in introns or intergenic regions. Only one SNP produces a non-synonymous coding change in the *CR1* gene (supplementary table 3). The individuals homozygous for the C allele of rs1567190 are also homozygous for 52 polymorphisms spanning 45 Kb of the *CR2* gene (Figure 3, supplementary table 4). In particular these individuals are homozygous for the coding SNPs rs1048971, rs17615, rs4308977, rs17616 (G-G-T-G). On the contrary, individuals homozygous for the risk allele of rs1567190 displayed a lower level of homozygosity along the *CR2* gene, only 11 polymorphisms are also homozygous and only one of them (rs17258982) are located in the *CR2* coding region (Figure 3, Supplementary Table 4).

CR2 and *C4BPA* phenotyping

Next, we validated whether the observed effect could be associated with differential expression of *CR2* mRNA in PBMC from healthy donors. Nine samples homozygous CC

for rs1567190 of *CR2*, and nine homozygous TT samples were examined by Taqman real-time quantitative PCR. The relative amount of the CR2 expression of different genetics backgrounds were measured normalized to CD19 (a specific B- lymphocyte marker). The normalized expression of *CR2* revealed a lower level of expression of the protective genotype CC (n=9, mean=0.39±0.08) compared to the risk genotype TT (n=9, mean=0.77±0.11; *P*=0.014, Student T-test, Figure 2). These data suggest that the global level of *CR2* expression is increased in B-lymphocytes from rs1567190 TT carriers. These data represent the mean of two independent experiments performed in duplicate. The levels of the short and long *CR2* mRNA splicing isoforms were also quantified by real time PCR in both genetic backgrounds. The CD19-normalized short *CR2* isoform expression were significatively reduced in PBMC from homozygous CC individuals compared with TT (n=9, mean=0.12±0.09 vs. 0.64±0.19; *P*=0.024, Student T-test, Figure 2). Finally, we have found that the expression of long *CR2* isoform in rs1567190 CC carriers is more elevated (although without reaching statistical signification) compared to the other genotype (n=9, mean=1.9±1.04 vs. 0.07±0.07; *P*=0.09, Student T-test, Figure 2).

The mRNA levels of C4BPA normalized by CD19 in PBMC from individuals harboring the protective genotype does not display significative differences with the risk genotype (n=9, mean=0.36±0.81 vs. 0.26±0.26; *P*=0.77, Student T-test).

Discussion

In this hypothesis-driven, low-scale, candidate gene association study we have genotyped 139 SNPs in 19 genes related to complement regulatory proteins or complement receptors. Our discovery population is highly uniform: all are male, white Caucasians of Spanish ancestry, HCV positive and intravenous drug users; including 201 HESN and 250 HIV-1 positive. We found two SNPs that are significantly associated after Bonferroni´s correction with resistance to HIV-1 infection; rs1567190 located in *CR2* and rs2842704 in *C4BPA*. A metanalysis including an independent cohort of HESN individuals from Italy at risk of infection through sexual route gave support only the association of *CR2* with the risk of infection.

This study has several limitations, first, the list of genes related to complement system is not exhaustive; polymorphisms in other genes not included in this study may also be considered as potential modulators of HIV-1 susceptibility, additionally some genes included in this study, as CFHR1 or CD55 are represented by only one to three SNPs. Second, our phenotypic analysis is limited to mRNA expression; inclusion of protein quantification as well as complement-mediated antibody-dependent enhancement of infection in the context of CR2 isoforms and C4BP overexpression may give us a more complete picture of the role these polymorphisms in HIV-1 susceptibility and it will generate new tools to counteract the viral subversion of the complement system. Third, the sizes of our HESN cohorts are very limited and do not allow us to perform epistatic or gene-gene interaction analysis. Four, the replication cohort from Italy is significatively smaller than the Spanish one and the genetic factors affecting the innate

292 resistance against HIV-1 infection through sexual route could be different compared to
293 the parenteral one.

294 The *CR2* gene encodes a membrane protein also known as CD21, which functions as a
295 receptor for five ligands: C3d/iC3b, Epstein-Barr Virus, CD23 [22], Interferon alpha [23] and
296 DNA [24]. CR2 together with CD19, CD81, and CD225, forms the B-cell coreceptor
297 complex in the surface of B lymphocytes, and binding of CR2 with immune complexes
298 coated with C3 results in B-lymphocyte enhanced proliferation, up regulation of co-
299 stimulatory molecules, and activation of intracellular signaling pathways including
300 calcium release or MAP kinase activation (Cambier et al., 1994; Matsumoto et al.,
301 1991). On Follicular dendritic cells (FDCs) CR2 captures C3 coated antigens, apparently
302 to hold these on the surface for long periods of time, facilitating a prolonged immune
303 response [25]. *CR2* gene contains 18 exons and there exist two main isoforms that
304 results from alternative splicing of exon 11: *CR2*-L and *CR2*-S. The first form is
305 expressed mainly in follicular dendritic cells and contains 16 short consensus repeats
306 (SCRs); the second isoform is expressed principally in B lymphocytes, some T activated
307 cells and endothelial cells and it contains 15 SCRs. Both SCRs form the extracellular
308 domain [21]. The specific functions of each isoform as well as its binding capacity to the
309 natural ligands are unknown.

310 CR2 function can be subverted by some viruses that use antibodies and/or
311 complement for a highly efficient entry into target cells; in a process known as
312 antibody-dependent enhancement (ADE) of viral infection. Two types of ADE have
313 been described in the context of HIV-1. The first type of ADE is dependent on the
314 interaction between antibody and FcR, specifically the FcγRIIIa appears to be the most

315 important receptor for FcR-dependent ADE, and polymorphisms in this gene were
316 associated with outcomes in a trial testing recombinant gp120 vaccination in
317 preventing sexually acquired HIV-1 infection [26]. Second, complement alone can boost
318 infection of dendritic cells as well as monocytes/macrophages and antibodies can
319 augment this process by a more efficient recruitment of complement on the viral
320 surface. This mechanism is dependent on the CR2/CR1 expression in the target cells
321 [17,27] and probably is mediated by and increased adhesion of the virus-antibody-
322 complement (C3d) complexes through binding with CR2 on the target surface [28] [29]. The
323 protective SNP identified in this study, belongs to an extended haplotype that display
324 strong linkage disequilibrium and include other 52 genetic polymorphisms that spam
325 45 Kb from the promoter to 3´ downstream region. The protective *CR2* haplotype
326 displays lower levels of total *CR2* mRNA as well as reduced short isoform expression; in
327 view of the data about the role of CR2 in complement-mediated ADE, we hypothesize
328 that the ratio of short/long *CR2* isoform expression and/or the global CR2 level, may
329 contribute to the likelihood of HIV-1 acquisition through the parenteral as well as
330 sexual routes. ADE is an important factor affecting other viruses as Dengue, Ross River
331 or Ebola [16]; we conjecture that the *CR2* haplotypes could have also an influence in the
332 susceptibility to these viruses. Although the main SNP identified in this study has not
333 been previously related with any pathology, the protective haplotype has been
334 associated with susceptibility to systemic lupus erythematosus [30] [21].

335 *C4BPA* gene encodes a plasma protein that forms an oligomer complex with *C4BPB*,
336 producing three isoforms: α7β1, α7β0 y α6β1. C4BP is a plasmid regulator of the
337 classical and alternative complement pathways. It hydrolyzes the complement
338 fragment C4b, accelerates the degradation of the C3 convertase and regulates as a

cofactor the alternative pathway by binding to Factor I [31]. Gene expression profiling of peripheral blood mononuclear cells from patients with acute HIV-1 infection has identify C4BPA as a protein capable of restrict HIV-1 replication [32]. Other viruses as Dengue, West Nile or yellow fever actively recruit C4BP through the NS1 protein, reducing the functional capacity of C4 and complement activation. Bacterial infections as *Streptococcus* [33] , *Yersinia* [34], *Lepstospira* [35] or *Neisseria* [36] among others, have also developed active protein systems to capture C4BP in the surface to escape complement mediated lysis. The observed association of *C4PBA* haplotypes with HIV-1 resistance in HESN at risk thought the parenteral route suggests that this gene could be a candidate for further functional phenotypic characterization. Several observations may underlie the absence of association of *C4BPA* marker in the Italian cohort at risk through the sexual route. First, this cohort is mainly constituted by women [19] and the Spanish one is exclusively constituted by males. There are significative gender differences in the innate and adaptive immune responses, the pathogenesis of infectious diseases, response to vaccines and the prevalence of autoimmune diseases [37,38]. In the case of C4BP, the synthesis appears to be influenced by sex hormones [39], pregnancy [40] or contraceptive treatment [41]. Finally, environmental factors as HCV infection (with 100% prevalence in the Spanish cohort) may modulate the genetic susceptibility in the drug users but not in the sexual cohort. HCV actively interferes with the complement system, inhibiting the transcription and protein production of C3 [42], C4 [43] and C9 [44].

CFH regulates the function of the alternative complement pathway in fluid phase and on cellular surfaces. It binds to C3b, accelerates the decay of the alternative pathway convertase C3bBb, and also acts as a cofactor for complement factor I, another C3b

363 inhibitor [45] [46]. The *CFH* gene consists of 20 short consensus repeats (SCR), each of
364 which encodes a functional domain containing a C3d/C3b binding site [47] [14]. CFH is
365 synthesized primarily by the liver and secreted into the circulation [14]. SNPs in the *CFH*
366 region that appears to be related with HIV-1 resistance at the haplotypic level, has
367 been previously associated with several autoimmune diseases [14]; as well as
368 susceptibility to meningitis [48] and Dengue [49]. High levels of CFH in human serum can
369 protect cells infected by HIV-1 and free virus *in vitro* [14,22,23,24], we hypothesize
370 that the protective haplotype of CFH may be related to lower levels of CFH expression
371 by the liver or altered splicing efficacy.

372 The discovery of genetic polymorphisms associated with innate resistance to HIV-1
373 infection can increase our understanding on the viral replication and spreading
374 strategies and can potentially lead to the identification of new antiviral targets acting
375 on the host proteins needed by the virus to achieve optimal replication or immune
376 escape. The results presented here if confirmed, could lead to develop alternative
377 strategies to combat the viral infection by the means of development of new
378 antagonists of CR2, C4BPA or CFH. This approach has formerly assayed using inhibitors
379 of CD59 [50], another membrane-bound RCA implicated in HIV-1 resistance to serum.
380 Finally, our results may have implications in the pharmacogenetics of HIV-1 vaccine.

381

References

1. Thielens NM, Tacnet-Delorme P, Arlaud GJ. Interaction of C1q and mannan-binding lectin with viruses. Immunobiology. 2002;205:563-574.
2. Prohaszka Z, Hidvegi T, Ujhelyi E, et al. Interaction of complement and specific antibodies with the external glycoprotein 120 of HIV-1. Immunology. 1995;85:184-189.
3. Ji X, Gewurz H, Spear GT. Mannose binding lectin (MBL) and HIV. Mol Immunol. 2005;42:145-152.
4. Stoiber H, Clivio A, Dierich MP. Role of complement in HIV infection. Annu Rev Immunol. 1997;15:649-674.
5. Banki Z, Stoiber H, Dierich MP. HIV and human complement: inefficient virolysis and effective adherence. Immunol Lett. 2005;97:209-214.
6. Tortorella D, Gewurz BE, Furman MH, Schust DJ, Ploegh HL. Viral subversion of the immune system. Annu Rev Immunol. 2000;18:861-926.
7. Saifuddin M, Hedayati T, Atkinson JP, Holguin MH, Parker CJ, Spear GT. Human immunodeficiency virus type 1 incorporates both glycosyl phosphatidylinositol-anchored CD55 and CD59 and integral membrane CD46 at levels that protect from complement-mediated destruction. J Gen Virol. 1997;78 (Pt 8):1907-1911.
8. Montefiori DC, Cornell RJ, Zhou JY, Zhou JT, Hirsch VM, Johnson PR. Complement control proteins, CD46, CD55, and CD59, as common surface constituents of human and simian immunodeficiency viruses and possible targets for vaccine protection. Virology. 1994;205:82-92.
9. Schmitz J, Zimmer JP, Kluxen B, et al. Antibody-dependent complement-mediated cytotoxicity in sera from patients with HIV-1 infection is controlled by CD55 and CD59. J Clin Invest. 1995;96:1520-1526.
10. Saifuddin M, Parker CJ, Peeples ME, et al. Role of virion-associated glycosylphosphatidylinositol-linked proteins CD55 and CD59 in complement resistance of cell line-derived and primary isolates of HIV-1. J Exp Med. 1995;182:501-509.
11. Stoiber H, Schneider R, Janatova J, Dierich MP. Human complement proteins C3b, C4b, factor H and properdin react with specific sites in gp120 and gp41, the envelope proteins of HIV-1. Immunobiology. 1995;193:98-113.
12. Stoiber H, Pinter C, Siccardi AG, Clivio A, Dierich MP. Efficient destruction of human immunodeficiency virus in human serum by inhibiting the protective action of complement factor H and decay accelerating factor (DAF, CD55). J Exp Med. 1996;183:307-310.
13. Haapasalo K, Vuopio J, Syrjanen J, et al. Acquisition of complement factor H is important for pathogenesis of Streptococcus pyogenes infections: evidence from bacterial in vitro survival and human genetic association. J Immunol. 2012;188:426-435.
14. Rodriguez de Cordoba S, Esparza-Gordillo J, Goicoechea de Jorge E, Lopez-Trascasa M, Sanchez-Corral P. The human complement factor H: functional roles, genetic variations and disease associations. Mol Immunol. 2004;41:355-367.
15. Wilflingseder D, Banki Z, Garcia E, et al. IgG opsonization of HIV impedes provirus formation in and infection of dendritic cells and subsequent long-term transfer to T cells. J Immunol. 2007;178:7840-7848.
16. Takada A, Kawaoka Y. Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. Rev Med Virol. 2003;13:387-398.
17. Willey S, Aasa-Chapman MM, O'Farrell S, et al. Extensive complement-dependent enhancement of HIV-1 by autologous non-neutralising antibodies at early stages of infection. Retrovirology. 2011;8:16.
18. de la Torre MS, Torres C, Nieto G, et al. Vitamin D receptor gene haplotypes and susceptibility to HIV-1 infection in injection drug users. J Infect Dis. 2008;197:405-410.

19. Sironi M, Biasin M, Cagliani R, et al. A Common Polymorphism in TLR3 Confers Natural Resistance to HIV-1 Infection. J Immunol. 2012;188:818-823.
20. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet. 2004;74:765-769.
21. Douglas KB, Windels DC, Zhao J, et al. Complement receptor 2 polymorphisms associated with systemic lupus erythematosus modulate alternative splicing. Genes Immun. 2009;10:457-469.
22. Frank MM. CD21 deficiency, complement, and the development of common variable immunodeficiency. J Allergy Clin Immunol. 2012;129:811-813.
23. Asokan R, Hua J, Young KA, et al. Characterization of human complement receptor type 2 (CR2/CD21) as a receptor for IFN-alpha: a potential role in systemic lupus erythematosus. J Immunol. 2006;177:383-394.
24. Asokan R, Banda NK, Szakonyi G, Chen XS, Holers VM. Human complement receptor 2 (CR2/CD21) as a receptor for DNA: implications for its roles in the immune response and the pathogenesis of systemic lupus erythematosus (SLE). Mol Immunol. 2013;53:99-110.
25. Cinamon G, Zachariah MA, Lam OM, Foss FW, Jr., Cyster JG. Follicular shuttling of marginal zone B cells facilitates antigen transport. Nat Immunol. 2008;9:54-62.
26. Forthal DN, Gabriel EE, Wang A, Landucci G, Phan TB. Association of Fcgamma receptor IIIa genotype with the rate of HIV infection after gp120 vaccination. Blood. 2012;120:2836-2842.
27. Huber G, Banki Z, Lengauer S, Stoiber H. Emerging role for complement in HIV infection. Curr Opin HIV AIDS. 2011;6:419-426.
28. Lund O, Hansen J, Soorensen AM, Mosekilde E, Nielsen JO, Hansen JE. Increased adhesion as a mechanism of antibody-dependent and antibody-independent complement-mediated enhancement of human immunodeficiency virus infection. J Virol. 1995;69:2393-2400.
29. Fust G. Enhancing antibodies in HIV infection. Parasitology. 1997;115 Suppl:S127-140.
30. Wu H, Boackle SA, Hanvivadhanakul P, et al. Association of a common complement receptor 2 haplotype with increased risk of systemic lupus erythematosus. Proc Natl Acad Sci U S A. 2007;104:3961-3966.
31. Martinez-Barricarte R, Heurich M, Valdes-Canedo F, et al. Human C3 mutation reveals a mechanism of dense deposit disease pathogenesis and provides insights into complement activation and regulation. J Clin Invest. 2010;120:3702-3712.
32. Chang HH, Soderberg K, Skinner JA, et al. Transcriptional network predicts viral set point during acute HIV-1 infection. J Am Med Inform Assoc. 2012;19:1103-1109.
33. Ermert D, Weckel A, Agarwal V, Frick IM, Bjorck L, Blom AM. Binding of complement inhibitor C4b-binding protein to a highly virulent Streptococcus pyogenes M1 strain is mediated by protein H and enhances adhesion to and invasion of endothelial cells. J Biol Chem. 2013;288:32172-32183.
34. Kirjavainen V, Jarva H, Biedzka-Sarek M, Blom AM, Skurnik M, Meri S. Yersinia enterocolitica serum resistance proteins YadA and ail bind the complement regulator C4b-binding protein. PLoS Pathog. 2008;4:e1000140.
35. Castiblanco-Valencia MM, Fraga TR, Silva LB, et al. Leptospiral immunoglobulin-like proteins interact with human complement regulators factor H, FHL-1, FHR-1, and C4BP. J Infect Dis. 2012;205:995-1004.
36. Ram S, Cullinane M, Blom AM, et al. Binding of C4b-binding protein to porin: a molecular mechanism of serum resistance of Neisseria gonorrhoeae. J Exp Med. 2001;193:281-295.
37. Fish EN. The X-files in immunity: sex-based differences predispose immune responses. Nat Rev Immunol. 2008;8:737-744.
38. Pennell LM, Galligan CL, Fish EN. Sex affects immunity. J Autoimmun. 2012;38:J282-291.

39. Marcovina SM, Zoppo A, Vigano-D'Angelo S, Di Cola G, D'Angelo A. Determination of serum levels of complement component C4b-binding protein: influence of age and inflammation. Int J Clin Lab Res. 1991;21:171-175.
40. Comp PC, Thurnau GR, Welsh J, Esmon CT. Functional and immunologic protein S levels are decreased during pregnancy. Blood. 1986;68:881-885.
41. Melissari E, Kakkar VV. The effects of oestrogen administration on the plasma free protein S and C4b-binding protein. Thromb Res. 1988;49:489-495.
42. Mazumdar B, Kim H, Meyer K, et al. Hepatitis C virus proteins inhibit C3 complement production. J Virol. 2012;86:2221-2228.
43. Banerjee A, Mazumdar B, Meyer K, Di Bisceglie AM, Ray RB, Ray R. Transcriptional repression of C4 complement by hepatitis C virus proteins. J Virol. 2011;85:4157-4166.
44. Kim H, Meyer K, Di Bisceglie AM, Ray R. Hepatitis C virus suppresses C9 complement synthesis and impairs membrane attack complex function. J Virol. 2013;87:5858-5867.
45. Ault BH. Factor H and the pathogenesis of renal diseases. Pediatr Nephrol. 2000;14:1045-1053.
46. Perez-Caballero D, Gonzalez-Rubio C, Gallardo ME, et al. Clustering of missense mutations in the C-terminal region of factor H in atypical hemolytic uremic syndrome. Am J Hum Genet. 2001;68:478-484.
47. Sanchez-Corral P, Perez-Caballero D, Huarte O, et al. Structural and functional characterization of factor H mutations associated with atypical hemolytic uremic syndrome. Am J Hum Genet. 2002;71:1285-1295.
48. Davila S, Wright VJ, Khor CC, et al. Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. Nat Genet. 2010;42:772-776.
49. Pastor AF, Rodrigues Moura L, Neto JW, et al. Complement factor H gene (CFH) polymorphisms C-257T, G257A and haplotypes are associated with protection against severe dengue phenotype, possible related with high CFH expression. Hum Immunol. 2013;74:1225-1230.
50. Hu W, Yu Q, Hu N, et al. A high-affinity inhibitor of human CD59 enhances complement-mediated virolysis of HIV-1: implications for treatment of HIV-1/AIDS. J Immunol. 2010;184:359-368.

Conflict of interest: The authors declare no conflicts of interest.

Financial support: Spanish Health Ministry (Fondo de Investigaciones Sanitarias PI021476, PI051778, PI10/01232 and ISCIII-RETIC RD06/006), Fundació Marató TV3 (020730 and 020732) and Junta de Andalucía (PI-0335/2009). JAP is the receptor of an intensification grant from the Fundación Progreso y Salud of the Consejería de Salud de la Junta de Andalucía (Reference AI-0021).

Authorship contributions:

Conceived and designed the experiments: AC, JF
Performed the experiments: RH, PK, ML, AC
Contributed reagents/materials/analysis: AR, JAP, ARJ, JM, ACam, JCS, JMS, MB, IS, SLC. JF, FJM
Wrote the manuscript: AC, RH, JF

Figures legends

Figure 1: Scatter plot of genetic association results. Negative log10-transformed P-values are plotted against physical position in the chromosome 1. SNPs rs2842704 rs1567190 are statistically significant after Bonferroni´s correction.

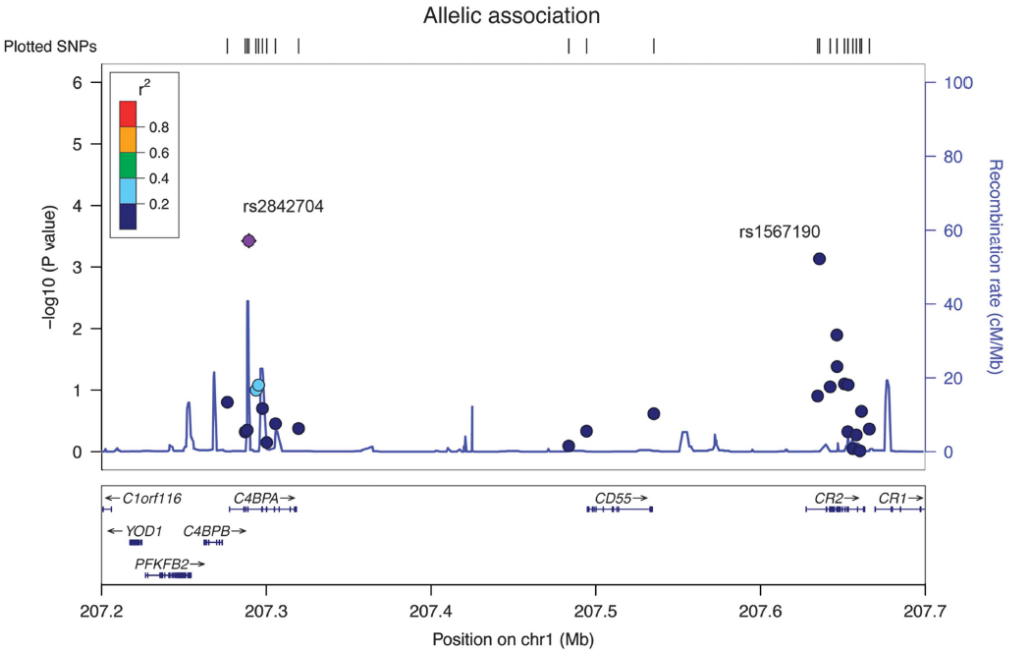
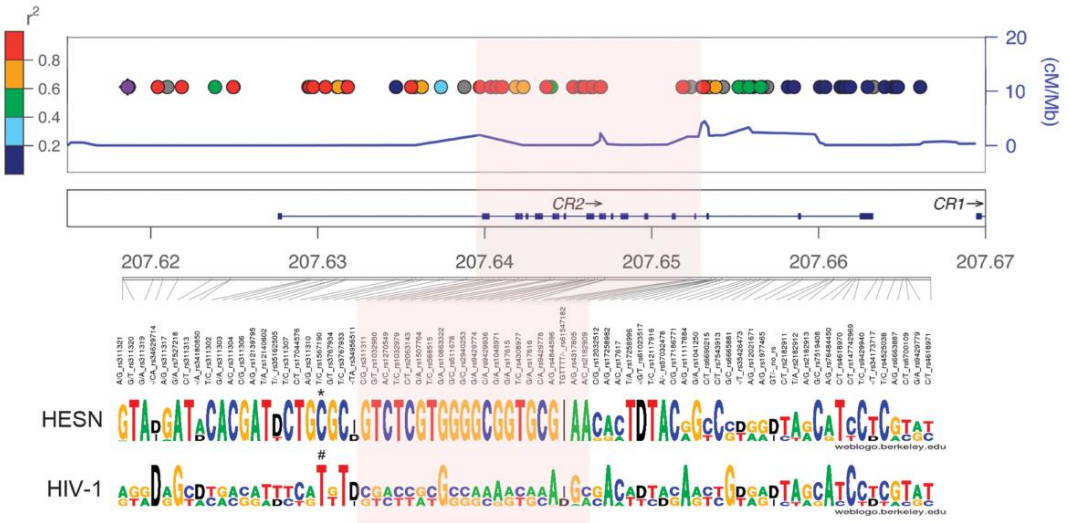
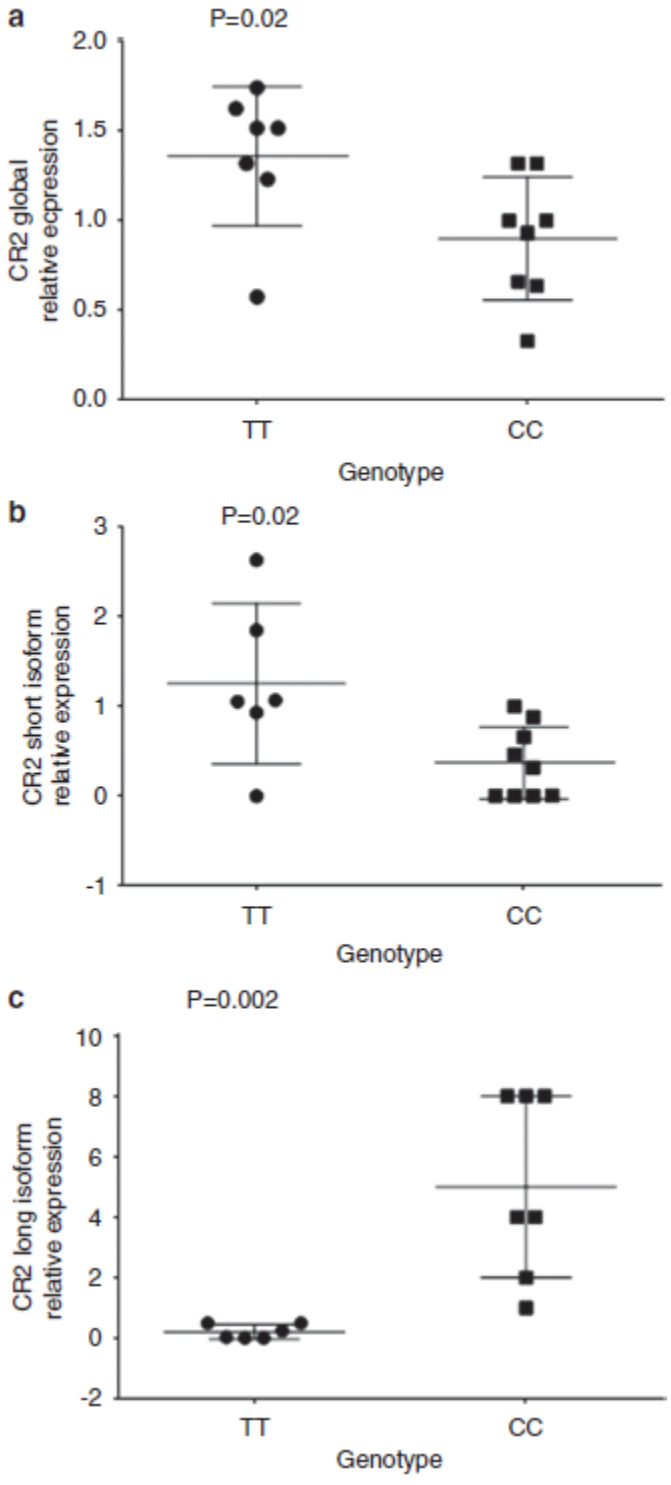


Figure 2: CR2 mRNA expression normalized by CD19 according to the genotype for rs1567190. a) Global CR2 expression levels (*P*=0.014) b) Short CR2 isoform (*P*=0.024) and c) Long CR2 isoform (*P*=0.77).



538 **Figure 3:** Sequence-inferred haplotypes of *CR2* gene in pools of HESN with protective
539 (a) or HIV-1 infected patients with risk (b) genotypes. The Y-axis represents the % of
540 reads with consensus sequence. The X-axis represents the relative position of CR2
541 SNPs along the gene.



542

543