Genomewide Association Study of an AIDS-Nonprogression Cohort Emphasizes the Role Played by *HLA* Genes (ANRS Genomewide Association Study 02)

Sophie Limou,[1,2,4,5,a] Sigrid Le Clerc,[1,2,4,a] Cédric Coulonges,[1,2] Wassila Carpentier,[2] Christian Dina,[6] Olivier Delaneau,[1] Taoufik Labib,[1,4] Lieng Taing,[1] Rob Sladek,[8] ANRS Genomic Group,[2,b] Christiane Deveau,[2] Rojo Ratsimandresey,[1] Matthieu Montes,[1] Jean-Louis Spadoni,[1] Jean-Daniel Lelièvre,[4] Yves Lévy,[4] Amu Therwath,[3] François Schächter,[1] Fumihiko Matsuda,[9] Ivo Gut,[5] Philippe Froguel,[6,10] Jean-François Delfraissy,[2] Serge Hercberg,[7] and Jean-François Zagury[1,2,4]

[1]Chaire de Bioinformatique, Conservatoire National des Arts et Métiers, [2]Agence Nationale de Recherche sur le SIDA et les Hépatites Virales (ANRS) Genomic Group, and [3]Laboratoire d'Oncologie Moléculaire, Université Paris 7, Paris, and [4]Henri Mondor Hospital, Institut National de la Santé et de la Recherche Médicale (INSERM) U841, Créteil, [5]Commissariat à L'Énergie Atomique/Institut de Génomique, Centre National de Génotypage, Evry, [6]Unité Mixte de Recherche (UMR) Centre National de la Recherche Scientifique 8090, Institut Pasteur de Lille, Lille, and [7]UMR U557 INSERM/U1125 L'Institut National de la Recherche Agronomique/Conservatoire National des Arts et Métiers/Université Paris 13, Centre de Recherche en Nutrition Humaine Ile-de-France, Santé-Médecine-Biologie Humaine Paris 13, Bobigny, France; [8]Department of Human Genetics, Faculty of Medicine, McGill University, and Génome Québec Innovation Centre, Montreal, Canada; [9]INSERM U852, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; [10]Genomic Medicine, Hammersmith Hospital, Imperial College London, London, United Kingdom

To elucidate the genetic factors predisposing to AIDS progression, we analyzed a unique cohort of 275 human immunodeficiency virus (HIV) type 1–seropositive nonprogressor patients in relation to a control group of 1352 seronegative individuals in a genomewide association study (GWAS). The strongest association was obtained for *HCP5* rs2395029 ($P = 6.79 \times 10^{-10}$; odds ratio, 3.47) and was possibly linked to an effect of sex. Interestingly, this single-nucleotide polymorphism (SNP) was in high linkage disequilibrium with *HLA-B, MICB, TNF,* and several other HLA locus SNPs and haplotypes. A meta-analysis of our genomic data combined with data from the previously conducted Euro-CHAVI (Center for HIV/AIDS Vaccine Immunology) GWAS confirmed the *HCP5* signal ($P = 3.02 \times 10^{-19}$) and identified several new associations, all of them involving *HLA* genes: *MICB, TNF, RDBP, BAT1–5, PSORS1C1,* and *HLA-C.* Finally, stratification by *HCP5* rs2395029 genotypes emphasized an independent role for *ZNRD1,* also in the HLA locus, and this finding was confirmed by experimental data. The present study, the first GWAS of HIV-1 nonprogressors, underscores the potential for some *HLA* genes to control disease progression soon after infection.

After 25 years of intensive research, there is still no definitive cure or vaccine for AIDS, and innovative strategies to fight HIV-1 infection are needed. Nowadays, genotyping by high-density arrays scanning the whole genome allows discovery of unsuspected genetic risk factors that influence the pathogenesis of disease [1]. This systematic genetic approach should reveal new leads for strategies targeting AIDS, given that associations based on a candidate gene approach accounted for no more than 10% of the genetic risk factors influencing disease progression [2]. Recently, a genomewide association study (GWAS) based on a European multicenter seroconverter HIV-1 cohort, the Euro-CHAVI (Center for HIV/AIDS Vaccine Immunology) cohort, identified 2 alleles in *HCP5* and *HLA-C* that explained nearly 15% of the variation in the viral load set point [3]. Although genomic studies of AIDS usually rely on seroconverter patients who display all stages of disease, our rationale was that the extreme nonprogression phenotype of the

Received 9 October 2008; accepted 3 November 2008; electronically published 10 January 2009.

Potential conflicts of interest: none reported.

Financial support: Agence Nationale de Recherche sur le SIDA et les Hépatites Virales (ANRS); Innovation 2007 program of the Conservatoire National des Arts et Métiers; AIDS Cancer Vaccine Development Foundation; Neovacs SA; Vaxconsulting. S.L. benefits from a fellowship from the French Ministry of Education, Technology, and Research, and S.L.C. benefits from a fellowship from ANRS.

[a] S.L. and S.L.C. contributed equally to this work.

[b] The ANRS Genomic Group oversees the AIDS genomic projects of the ANRS; study group members are listed at the end of the text.

Reprints or correspondence: Dr. Jean-François Zagury, 292 rue Saint Martin, 75003 Paris, France (zagury@cnam.fr).

The Journal of Infectious Diseases 2009; 199:419–26

0022-1899/2009/19903-0018$15.00
DOI: 10.1086/596067

Table 1. Fifty best results obtained for the comparison between nonprogressors and control subjects.

The table is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

GRIV (Genomics of Resistance to Immunodeficiency Virus) cohort could bring even more contrast to the attempt to identify genetic effects. Previous gene-candidate analyses have shown the power of this unique design, notably for the *HLA* and *CCR5* genes [4, 5].

METHODS

The GRIV cohort. The GRIV cohort was established in France in 1995 to generate a large collection of DNA for genetic studies to identify host genes associated with nonprogression to AIDS [4, 6]. Only white people of European descent living in France were eligible for enrollment to reduce confounding by population substructure. These criteria limit the influence of ethnic and environmental factors (all subjects live in a similar environment and are infected by B strains) and emphasize the genetic makeup of each individual in determining the various patterns of progression. Nonprogressors were included on the basis of the main clinical outcomes, CD4 T cell count and time to disease progression; inclusion criteria were asymptomatic HIV-1 infection for >8 years, no receipt of treatment, and a CD4 T cell count consistently remaining >500 cells/mm^3. Viral load was not part of the GRIV inclusion criteria; however, the values at inclusion were obtained and used to assess potential correlations with genotypes. DNA was obtained from fresh peripheral blood mononuclear cells or from Epstein-Barr virus–transformed cell lines. The nonprogressors group ($n = 275$) was composed of 201 men and 74 women whose ages at inclusion ranged from 19 to 62 years (median, 35 years). At inclusion, the median CD4 T cell count was 706 cells/mm^3 among the nonprogressors (minimum and maximum values, 501 and 2298 cells/mm^3). All patients provided written informed consent before enrollment in the GRIV genetic association study.

The seropositive control population. To determine whether positive signals corresponded either to an association with nonprogression or to an association with HIV-1 infection, we needed a group of seropositive control subjects who were not nonprogressors. For that, we used 86 white French subjects who qualified as rapid progressors to AIDS (i.e., a CD4 T cell count decreasing to <300 cells/mm^3 within 3 years of seroconversion). This control group was composed of 74 men and 12 women aged from 21 to 55 years (median, 32 years). The median CD4 T cell count of this seropositive control population was 230 cells/mm^3 (minimum and maximum values, 20 and 297 cells/mm^3). Viral loads were not available.

The SU.VI.MAX control group. The SU.VI.MAX (Supplémentation en Vitamines et Minéraux Antioxydants) study was a randomized, double-blind, placebo-controlled and primary-prevention trial designed to test the efficacy of daily supplements of antioxidant vitamins and minerals at nutrition-level doses in reducing the frequency of several major health problems in industrialized countries, especially the main causes of premature death, cancers and cardiovascular diseases. This cohort study was started in 1994 in France and was composed of 12,735 subjects [7]. The control group genotyped in the present study comprised 1352 representative SU.VI.MAX participants, all white persons living in France who were HIV-1 seronegative. This control cohort was composed of 525 men and 827 women, with a mean age of 53.1 and 48.5 years, respectively.

Genotyping method. Genotyping was performed for the GRIV cohort and the control groups by means of Infinium II HumanHap300 BeadChips (Illumina). The genomic DNA (750 ng) was whole-genome amplified, fragmented, denatured, and hybridized on prepared HumanHap300 BeadChips for a minimum of 16 h at 48°C. Nonspecifically hybridized fragments were removed by washing, and the remaining specifically hybridized DNA was fluorescently labeled by a single base-extension reaction and was detected using a BeadArray scanner (Illumina). Normalized bead-intensity data obtained for each sample were loaded into BeadStudio software (version 3.1; Illumina), which converted the fluorescence intensities into single-nucleotide polymorphism (SNP) genotypes.

Quality control. Using the BeadStudio software, we analyzed the crude genotyping data, and SNPs were filtered according to the following parameters. First, samples with a call rate (percentage of SNPs genotyped by sample) <95% in the Illumina clusters were deleted. Second, the SNPs having a call frequency (percentage of samples genotyped by SNP) <99% were reclustered. Third, after reclustering, samples with a call rate <97% were deleted. The clustering step can create SNP genotyping errors, which can be prevented by following the Illumina quality-control procedure (see http://www.illumina.com/downloads/GTDataAnalysis_TechNote.pdf). This method evaluates the quality of the newly created clusters according to several criteria, which can be manually checked and corrected as necessary. By this Illumina procedure, 1300 SNPs were excluded. Finally, after all the quality-control steps, the 15,731 SNPs with a call frequency <98% (>2% of missing data) were excluded. This quality-control procedure ensures reliable genotyping data with few missing data.

Hardy-Weinberg equilibrium analysis was performed for each SNP in each group using an exact statistical test [8] implemented in PLINK software (available at: http://pngu.mgh.harvard.edu/~purcell/plink/) [9]. Deviation from Hardy-Weinberg equilibrium in a group of patients suggests that the

The figure is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

Figure 1. Quantile-quantile plot for expected *(red)* vs. observed *(black)* *P* values from the comparison of nonprogressors with control subjects.

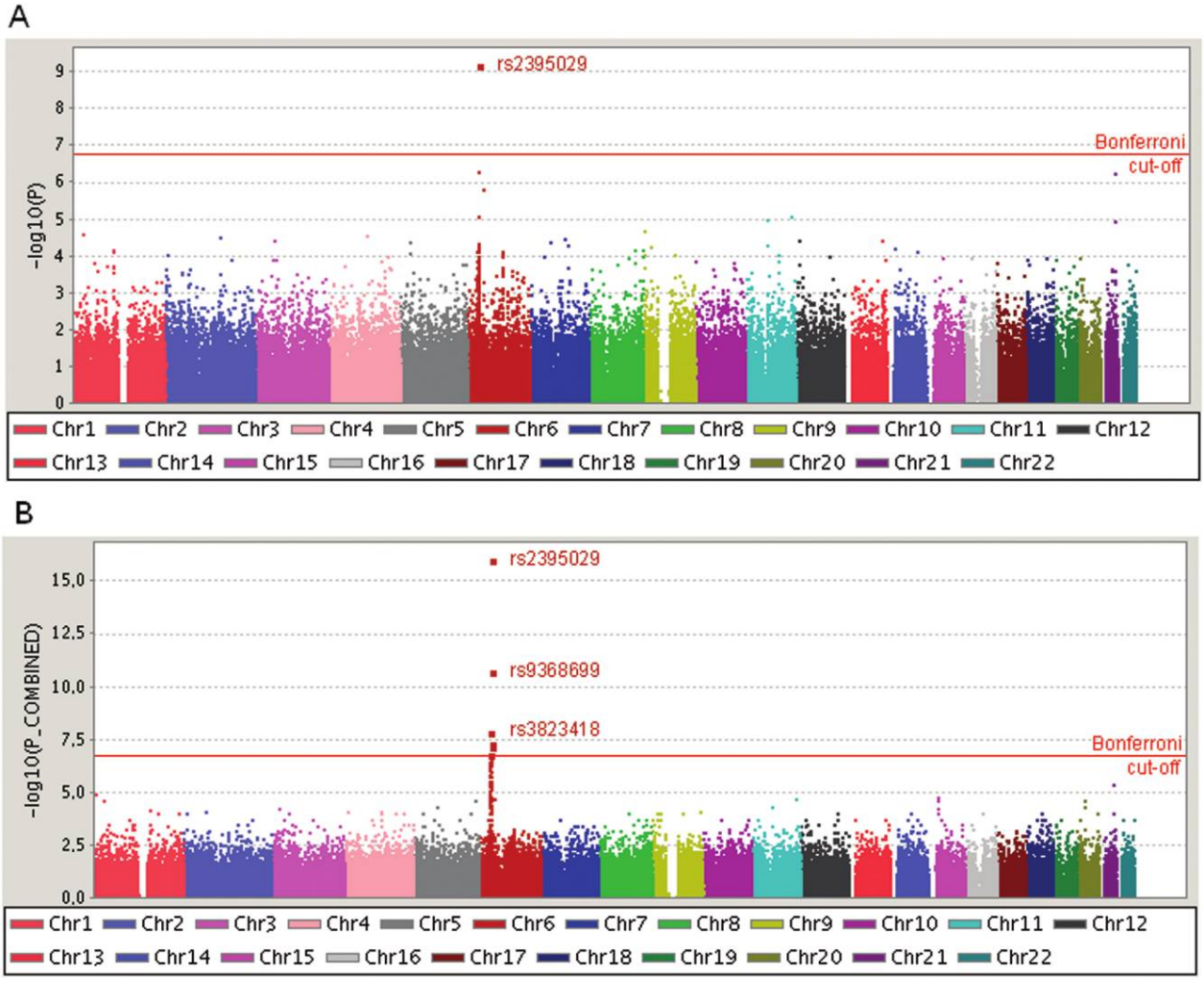

Figure 2. Distribution along the human autosomes of $-\log_{10}(P$ values) obtained for the comparison of nonprogressors with control subjects *(A)* and for the meta-analysis of the GRIV and Euro-CHAVI studies *(B)*. For the latter plot, we used the classical Fisher method, which allows combining *P* values obtained in 2 independent studies. The red line marks the Bonferroni threshold. Chr, chromosome.

SNP has a biological effect, while deviation in the control group or in all groups suggests a systematic error in genotyping. The 1475 SNPs that were not in Hardy-Weinberg equilibrium in the control group ($P < 1.0 \times 10^{-3}$) were rejected in this way.

A total of 235 SNPs with a low minor allelic frequency (<1%) in the global population were also filtered.

Linkage disequilibrium. For each SNP exhibiting a significant association, we looked for the other SNPs in linkage disequilibrium ($r^2 \geqslant 0.8$) in the HapMap population of Western European ancestry (CEU, HapMap data Release 21a/phase II January 2007, on NCBI B35 assembly, dbSNP125; available at: http://www.hapmap.org) to identify the genes possibly involved with the associations. A SNP was assigned to a gene if it was located in the gene or in the 2-kb flanking regions (potential regulatory sequence); otherwise, it was considered intergenic.

Statistical analysis. For each SNP, we performed a standard case-control analysis using Fisher's exact test (with PLINK software) to compare allelic distributions between the nonprogression group and the control group. To take into account the multiple comparisons, we computed the Bonferroni corrections. For all the SNPs meeting the statistical threshold (table 1), the quality of genotyping was individually rechecked with the BeadStudio software. We also checked that the allelic frequencies in the seropositive control population were similar to those in the seronegative SU.VI.MAX control population for those

Table 2. Comparative analysis of the single-nucleotide polymorphisms (SNPs) found to be highly significant by the GRIV genomewide association study (GWAS) and at least 1 other GWAS.

The table is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

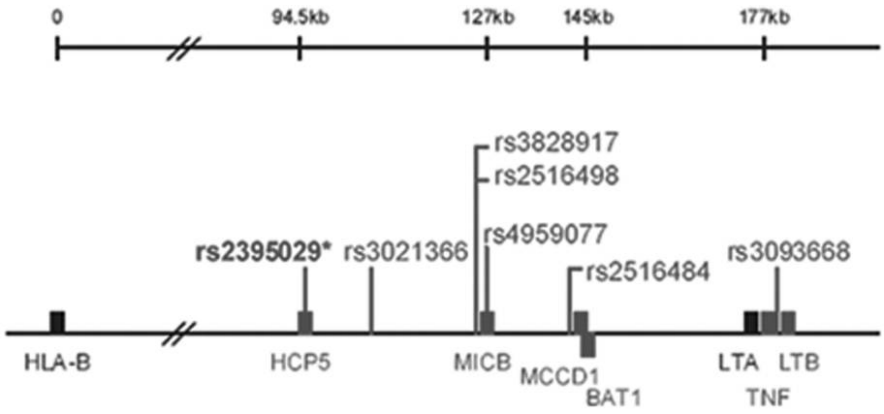

Figure 3. Complexity of the *HCP5* association. Shown is a genetic map of the HLA region. Table 3 lists single-nucleotide polymorphisms (SNPs) and haplotypes of several genes in strong linkage disequilibrium with the *HCP5* rs2395029 SNP, which is marked by an asterisk in this figure.

SNPs of interest, confirming that the observed associations were indeed linked to nonprogression.

Meta-analysis of the GRIV and Euro-CHAVI studies. A total of 286,529 SNPs were found to be in common between the GRIV GWAS and the previous GWAS of AIDS, the Euro-CHAVI study [3]. The genotypes obtained by the Euro-CHAVI study were not directly available, but the *P* values obtained for each SNP for the viral load end point could be obtained from the supporting online material for Fellay et al. [3] (available at: http://www.sciencemag.org/cgi/content/full/1143767/DC1). The *P* values obtained for each SNP in our study and in the Euro-CHAVI study were combined to provide a single probability value using the classical Fisher method [10]. We could not adjust the combined *P* values for opposite allelic effects (i.e., assign $P = 1$ if the odds ratios went in opposite directions), because the detailed allelic information for the Euro-CHAVI study was not available. We could compute this meta-analysis only for the Euro-CHAVI viral load end point data, because the *P* values for the progression to AIDS end point were not available.

Identification of population stratification. To correct for possible population stratification at the intercontinental level,

Table 3. Single-nucleotide polymorphisms (SNPs) and haplotypes of several genes in strong linkage disequilibrium ($r^2 \geqslant 0.8$ for HapMap data on white individuals) with the *HCP5* rs2395029 SNP (see figure 3).

Gene	Allele	Location	r^2
HCP5	rs2395029	Exon (Val112Gly)	. . .
HLA-B	HLA-B*5701	. . .	1.00
MICB	rs2516498	5'LR	0.83
	rs3828917	5'LR	1.00
	rs4959077	Intron	1.00
	rs2534654, rs2246626 (G-C)	Intron/3'LR	1.00
	rs3828917, rs1051788 (T-G)	5'LR/exon (Asp136Asn)	1.00
TNF	rs3093668	3'LR	0.83
	rs3093661	Intron	0.83
	rs3093661, rs4645843 (A-C)	Intron/exon (Pro84Leu)	0.83
	rs1799964, rs1800630, rs1800750 (C-C-G)	5'LR/5'LR/5'LR	0.83
MCCD1	rs2516484	5'LR	0.83
BAT1	rs2516484	3'LR	0.83
LTB	rs3093668	3'LR	0.83
	rs3093559, rs3093553 (C-G)	3'LR/intron	0.83
	rs3093726, rs3093559 (C-G)	3'LR/3'LR	0.83

NOTE. The *HCP5* rs2395029 SNP is marked by an asterisk in figure 3. To compute the linkage disequilibrium between the haplotypes and the *HCP5* SNP, we limited ourselves to haplotypes composed of 2 or 3 SNPs derived from the known HapMap SNPs in this HLA region. The list is only a small sample of the numerous haplotypes with $r^2 \geqslant 0.8$ identified. 5'LR corresponds to the 5' part within 2 kb of the gene; 3'LR corresponds to the 3' part within 0.5 kb of the gene.

Table 4. Influence of sex on the *HCP5* and *C6orf48* associations with nonprogression.

The table is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

Table 6. Fifty best combined *P* values obtained by the meta-analysis of the GRIV and Euro-CHAVI studies.

The table is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

case and control genotypes were analyzed using STRUCTURE software (version 2.2) [11]. We selected a set of 328 SNPs that were informative for ancestral origin (*F* statistics fixation index > 0.2) on the basis of the Perlegen data set and that were separated by 5 Mb to avoid linkage disequilibrium. We also included genotypes obtained from unrelated individuals representing the 3 populations studied by the HapMap project, to better separate the nonprogressor and control individuals according to their continental origin. All case and control subjects fell within the range of the white individuals from HapMap.

To avoid spurious associations resulting from possible population stratification or genotyping errors, a quantile-quantile plot was produced by plotting the ranked values of the test statistics against the approximated expected order statistic (figure 1). We also computed the genomic inflation factor λ [12]. The result ($\lambda = 1.064$), along with the quantile-quantile plot, suggested little overall effect of stratification.

RESULTS AND DISCUSSION

Using the Illumina HumanHap300 BeadChips, we performed a GWAS by comparing our nonprogression group ($n = 275$) with a control group ($n = 1352$) from the SU.VI.MAX cohort. After the different quality-control tests (see Methods), a total of 291,119 autosomal SNPs were tested for association with nonprogression. For each SNP, Fisher's exact test was performed comparing the allelic frequencies in the case group versus those in the control group, and the resulting *P* values were adjusted by the Bonferroni correction. Figure 2*A* depicts the distribution of the *P* values along the chromosomes, and table 1 presents the most significant signals.

The sole association remaining after the Bonferroni adjustments was the *HCP5* rs2395029-G allele ($P = 6.79 \times 10^{-10}$; odds ratio, 3.47 [95% confidence interval, 2.39–5.04]) (table 1) located on chromosome 6 (figure 2*A*). This *HCP5* SNP was previously identified by Fellay et al. [3], who hypothesized that, because *HCP5* encodes a human endogenous retrovirus with sequence homology to HIV-1 *pol* [13], it may act as antisense RNA interfering with HIV-1 replication. This SNP is also in absolute linkage disequilibrium with the HLA-B*5701 allele [14], which has been associated with the control of HIV-1 replication and disease progression [15]. Interestingly, this SNP was also shown to be a major signal in a GWAS of psoriasis and psoriatic arthritis [16] (table 2). Figure 3 and table 3 show that this SNP could be tracking through linkage disequilibrium causal alleles in other major genes of the HLA locus, including *MICB, BAT1, LTB,* and *TNF*. MICB is a ligand for CD8 T cells and natural killer cells, which are key players in the anti–HIV-1 immune response. BAT1 is an essential component for splicing and RNA export [17] but is also known as a negative regulator of the inflammatory cytokines tumor necrosis factor (TNF), interleukin (IL)–1, and IL-6 [18]. Lymphotoxin β (LTB) is an inflammatory

Table 5. Best *P* values obtained by the meta-analysis of the GRIV and Euro-CHAVI studies.

SNP	Chr	Chr position	A1	A2	Allelic frequency (A1), % NP	CTR	SCP	Fisher $P_{NP\text{-}CTR}$	Fisher $P_{Euro\text{-}CHAVI}$	$P_{combined}$	Gene(s)/LD
rs2395029	6	31539759	G	T	8.9	2.7	2.3	6.79×10^{-10}	9.36×10^{-12}	3.02×10^{-19}	*HCP5,* intergenic, *MICB, MCCD1, BAT1, LTB, TNF*
rs9368699	6	31910520	C	T	8.2	3.2	3.5	5.27×10^{-07}	1.20×10^{-06}	1.84×10^{-11}	*C6orf48, RDBP, TNXB, BAT2, BAT3, LY6G5C, BAT5*
rs3823418	6	31208921	A	G	24.2	16.8	16.9	5.72×10^{-05}	1.11×10^{-05}	1.40×10^{-08}	*PSORS1C1*
rs2248462	6	31554775	A	G	31.4	24.2	25.8	5.61×10^{-04}	3.61×10^{-06}	4.26×10^{-08}	Intergenic, *MICB*
rs2516509	6	31557973	G	A	31.1	24.1	25.4	6.56×10^{-04}	3.61×10^{-06}	4.95×10^{-08}	Intergenic
rs10484554	6	31382534	T	C	18.3	13.3	11.1	3.78×10^{-03}	8.06×10^{-07}	6.27×10^{-08}	Intergenic, *HLA-C*
rs3815087	6	31201566	T	C	27.5	20.9	19.8	1.04×10^{-03}	7.09×10^{-06}	1.46×10^{-07}	*PSORS1C1,* intergenic

NOTE. This table presents the *P* values obtained by the classical Fisher method that met the Bonferroni threshold. For each SNP, the chromosome (Chr), the chromosome position, the allelic frequencies in the various populations (nonprogressors [NP], control subjects [CTR], and seropositive control population [SCP]), the *P* values obtained in each study, and the combined *P* value are shown. The seropositive control population is a group of 86 HIV-1–seropositive patients who are not nonprogressors and allows genetic associations with nonprogression to be distinguished from associations with HIV-1 infection. The gene or genes corresponding to the SNP or SNPs in linkage disequilibrium (LD; $r^2 \geqslant 0.8$ for HapMap data on white individuals) are also informed. A SNP was assigned to a gene if it was located in the gene or in the 2-kb flanking regions (potential regulatory sequence).

modulator essential for the development of lymphoid, dendritic, and natural killer cells [19]. TNF is a key proinflammatory cytokine that has been widely investigated in HIV-1 infection [20]. From a biological standpoint, all these genes are critical for immunity and, as such, are good candidates to intervene in the pathogenesis of HIV-1 infection. Indeed, they have all been associated with various immune-related diseases [21–24]. Overall, the complex genetic pattern of this region makes it difficult to discriminate between a specific signal alone or one in combination (i.e., haplotypes): as shown in figure 3 and table 3, several SNPs or haplotypes in the genes *HLA-B, MICB, BAT1, LTB, TNF,* and *MCCD1* are in high or full linkage disequilibrium with *HCP5* rs2395029.

To complete our analysis of the *HCP5* SNP association, we explored the influence of covariables, such as *CCR5-Δ32* and *CCR5-P1* haplotypes [5], the HIV-1 infection mode (mucosal or parenteral), and sex. No effect was observed except for a sex influence: the rs2395029-G frequency was 4.05% in nonprogressor women versus 10.70% in nonprogressor men ($P = 1.71 \times 10^{-2}$), whereas, in control subjects, the frequency was close to 3% in men and in women (table 4). Such interaction between genetic factors and sex have been previously described for both HLA and non-HLA genetic associations with other pathologies [25]. The lack of association for *HCP5* in women requires confirmation in a cohort containing a sufficiently large number of women, because the alleles of interest have a frequency of only 3%.

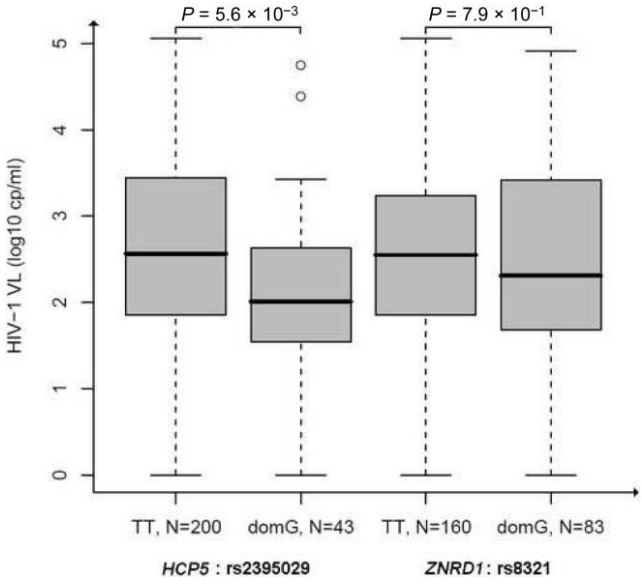

Figure 4. Correlation between genotypes and viral load. The box plots represent the viral load at inclusion of nonprogressor subjects carrying various *HCP5* and *ZNRD1* genotypes and were compared using Student's *t* test. A significantly lower viral load was found in the nonprogressor subjects carrying the *HCP5* rs2395029-GT genotype, compared with that in the ones with *HCP5* rs2395029-TT. No significant difference was observed between *ZNRD1* rs8321 genotypes (TT vs. dominant G [domG]) or for genotypes of *ZNRD1* (rs9261290, rs1245371, and rs259940; data not shown). rs8321 and rs9261290 were identified by our combined analysis with the Euro-CHAVI study, and rs1245371 and rs259940 were identified by our analysis of *HCP5*-independent signals. VL, viral load.

The figure is available in its entirety in the online edition of the *Journal of Infectious Diseases*.

Figure 5. Linkage disequilibrium map presenting single-nucleotide polymorphisms (SNPs) of the chromosome 6 HLA locus exhibiting strong *P* values in the meta-analysis of the GRIV and Euro-CHAVI studies.

The data from the sole AIDS GWAS of the Euro-CHAVI cohort published to date were available [3], and thus we performed a meta-analysis by combining their *P* values with ours using the classical Fisher method (see Methods). Figure 2*B* presents the distribution of the combined *P* values along the autosomes and several associations surpassing the Bonferroni threshold were found, all located on chromosome 6 (table 5) (see table 6 for an extended list of *P* values).

As expected, the strongest signal was obtained for the *HCP5* rs2395029 SNP ($P_{combined} = 3.02 \times 10^{-19}$). The second strongest association was obtained for the *C6orf48* rs9368699 SNP ($P_{combined} = 1.84 \times 10^{-11}$), which was in linkage disequilibrium with *HCP5* rs2395029 ($r^2 = 0.68$) and with several SNPs located in *HLA* genes, such as *TNXB, BAT2, BAT3,* and *RDBP,* suggesting a wide range of possible biological effects that could explain this association. For instance, tenascin XB (TNXB) is an extracellular matrix protein previously associated with systemic lupus erythematosus [26], HLA-B–associated transcript (BAT) 2 is a potential splicing factor previously associated with rheumatoid arthritis [27], BAT3 is an essential regulator of apoptosis and p53-mediated responses to genotoxic stress [28], and *RDBP* encodes a subunit of the negative elongation factor complex known to repress HIV-1 transcription elongation driven by Tat [29]. Then, the *PSORS1C1* gene exhibited 2 significant SNPs, rs3823418 and rs3815087 ($P_{combined} = 1.4 \times 10^{-8}$ and $P_{combined} = 1.46 \times 10^{-7}$, respectively), in partial linkage disequilibrium with each other ($r^2 = 0.66$ for HapMap data on white individuals). *PSORS1C1* is a psoriasis-susceptibility candidate gene [30]. The intergenic SNP rs2248462, which could not be assigned to a specific gene, exhibited a strong association ($P_{combined} = 4.26 \times 10^{-8}$) that could be explained by the linkage disequilibrium with the *MICB* gene discussed above. Finally, the combined analysis underlined the *HLA-C*–related SNP rs10484554 ($P_{combined} = 6.27 \times 10^{-8}$), which was also identified in the GWAS of psoriasis and psoriatic arthritis [16] (see table 2). This gene was widely discussed in the Euro-CHAVI GWAS for the association found with the rs9264942 SNP, which is unfortunately not present in the Illumina HumanHap300 BeadChip [3]. The sex dependence was still observed for the *C6orf48* rs9368699 SNP but not for the following SNPs (table 4).

The inclusion criteria for the Euro-CHAVI study were based on viral load during the asymptomatic set point period of infection, and for the GRIV study they were based on maintenance of CD4 T cell counts over time. The individuals with a low viral

load after infection are likely to be the ones with a stable high CD4 T cell count. Indeed, we found that, in the nonprogressors carrying the *HCP5* rs2395029-G allele, viral load was significantly lower ($P = 5.6 \times 10^{-3}$) than in the other nonprogressors (figure 4). It is thus not surprising to identify common genetic signals between these 2 studies, even though these cohorts were assembled independently. Notably, of the 50 best signals found in this meta-analysis, 46 originated from the HLA locus, emphasizing the massive role played by HLA in the nonprogression phenotype (table 6). The presence of these strong associations is a cross-validation of both cohorts and also emphasizes that the HLA locus is critical for the early control of HIV-1 replication and disease nonprogression. Reciprocally, in our GWAS alone, 31 of the 50 best signals were not from chromosome 6 (table 1) and were not found in the meta-analysis (table 6), suggesting that positive signals outside the HLA locus may be associated with the nonprogression phenotype without influencing viral load. This observation is in line with findings from a recent study by Mellors et al. [31], which showed that the viral load was predictive at a 34% level for the time to reach a CD4 T cell count <200 cells/mm^3.

Because the *HCP5* rs2395029-G allele was present in only 17.8% of the nonprogressor subjects, we reanalyzed the data in the nonprogressor and control individuals not carrying that allele in order to identify *HCP5*-independent signals. Dramatically, most of the signals from chromosome 6 disappeared because of the genetic linkage with the *HCP5* rs2395029-G. However, the strongest signals were still found in the HLA region, with 2 SNPs of the *ZNRD1/RNF39* region, rs1245371 and rs259940 ($P = 9.21 \times 10^{-7}$ and $P = 2.04 \times 10^{-6}$), in linkage disequilibrium. These 2 SNPs are genetically independent from the *HCP5* SNP ($r^2 = 0.06$) (figure 5). Interestingly, the *ZNRD1/RNF39* locus was also identified by our meta-analysis (for rs8321, $P = 4.66 \times 10^{-7}$; for rs9261290, $P = 5.11 \times 10^{-7}$) (table 6) and by the Euro-CHAVI study (for rs3869068, $P = 3.89 \times 10^{-7}$) with the progression-to-AIDS end point (defined as the time elapsed until treatment initiation or until reaching a CD4 T cell count <350 cells/mm^3). Unlike the *HCP5* rs2395029 SNP, none of the *ZNRD1/RNF39* SNPs alleles seemed to correlate with viral load (figure 4), suggesting that this locus influences disease progression. Functionally, the Genevar expression database identified an association between several *ZNRD1/RNF39* alleles and the differential expression of *ZNRD1* (table 2). Finally, we found that a recent genomewide RNA interference study identified zinc ribbon domain containing (ZNRD) 1 among the 273 proteins required for HIV-1 infection and replication [32], suggesting that this RNA polymerase I subunit is an active component in the *ZNRD1/RNF39* region.

In conclusion, the major novelty of the present GWAS of AIDS was the investigation of a cohort with the extreme HIV-1 nonprogression phenotype, in contrast to the usual seroconverter cohorts. We replicated the major role played by the *HCP5* gene of the HLA region in chromosome 6 previously reported in the Euro-CHAVI GWAS [3], a role that could be explained by linkage disequilibrium with other major genes of the HLA locus, such as *HLA-B, MICB, TNF, LTB,* and *BAT1*. The sex dependence of the *HCP5* SNP in our work is striking because it is in high linkage disequilibrium with *HLA-B*57,* which has been investigated for years. It was likely not observed before because most AIDS cohorts are deficient in women; however, this important observation needs confirmation. We then computed a meta-analysis with the previous GWAS and put forward new associations in the same locus: *C6orf48* (in linkage disequilibrium with *RDBP, TNXB,* and *BAT*), *PSORS1C1, MICB,* and *HLA-C*. The HLA region comes first in our AIDS-nonprogression genetic study; however, given that this region presents a complex pattern of high linkage disequilibrium and that all the genes identified display a strong relevance to immunology and AIDS, it is difficult to discriminate which one(s) is(are) the causal variant(s). More refined studies will be needed to discriminate which mechanisms and which HLA locus genes are at stake. Our study, however, has suggested an independent role for the *ZNRD1* gene in disease progression. Overall, our results underline the potential for controlling disease progression and/or viral replication by some *HLA* gene variants soon after infection. Notably, 2 major SNPs identified in our meta-analysis—*HCP5* and *HLA-C*—also had the strongest signals observed in the GWAS of psoriasis and psoriatic arthritis [16].

Because of the large amount of data generated in the present GWAS, statistical cutoffs were required to minimize false discovery, and many true positives with lower *P* values were likely missed but remain candidates of interest. As a reminder, allow us to state that the published *P* values from various cohorts for the widely recognized association between *CCR5-Δ32* and AIDS progression have all been in the range of 1×10^{-2} to 1×10^{-4} and would not be seen by the current genomewide studies. This latter observation emphasizes the need to analyze more patients and perform more meta-analyses to extract additional signals from the large pool of genes screened.

ANRS GENOMIC GROUP

The ANRS Genomic Group is composed of Prof. Jean-François Delfraissy (Agence Nationale de Recherche sur le SIDA et les Hépatites Virales, Paris), Dr. Laurence Meyer (Hôpital Kremlin-Bicêtre, France), Prof. Philippe Broët (Hôpital Kremlin-Bicêtre, France), Dr. Cyril Dalmasso (Hôpital Kremlin-Bicêtre, France), Dr. Wassila Carpentier (Hôpital La Salpêtrière, Paris), Prof. Patrice Debré (Hôpital La Salpêtrière, Paris), Dr. Ioannis Théodorou (Hôpital La Salpêtrière, Paris), Prof. Christine Rouzioux (Hôpital Necker, Paris), Cédric Coulonges (Conservatoire National des Arts et Métiers, Paris), Sigrid Le Clerc (Conservatoire National des Arts et Métiers, Paris), Sophie Limou (Conservatoire National des Arts et Métiers, Paris), and Prof. Jean-

François Zagury (Conservatoire National des Arts et Métiers, Paris).

Acknowledgments

We are grateful to all the patients and medical staff who have kindly collaborated with the GRIV project.

References

1. Kingsmore SF, Lindquist IE, Mudge J, Gessler DD, Beavis WD. Genome-wide association studies: progress and potential for drug discovery and development. Nat Rev Drug Discov **2008**; 7:221–30.
2. O'Brien SJ, Nelson GW. Human genes that limit AIDS. Nat Genet **2004**; 36:565–74.
3. Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. Science **2007**; 317:944–7.
4. Flores-Villanueva PO, Hendel H, Caillat-Zucman S, et al. Associations of MHC ancestral haplotypes with resistance/susceptibility to AIDS disease development. J Immunol **2003**; 170:1925–9.
5. Winkler CA, Hendel H, Carrington M, et al. Dominant effects of CCR2-CCR5 haplotypes in HIV-1 disease progression. J Acquir Immune Defic Syndr **2004**; 37:1534–8.
6. Rappaport J, Cho YY, Hendel H, Schwartz EJ, Schachter F, Zagury JF. 32 bp CCR-5 gene deletion and resistance to fast progression in HIV-1 infected heterozygotes. Lancet **1997**; 349:922–3.
7. Hercberg S, Galan P, Preziosi P, et al. Background and rationale behind the SU.VI.MAX study, a prevention trial using nutritional doses of a combination of antioxidant vitamins and minerals to reduce cardiovascular diseases and cancers. SUpplementation en VItamines et Mineraux AntioXydants Study. Int J Vitam Nutr Res **1998**; 68:3–20.
8. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet **2005**; 76:887–93.
9. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet **2007**; 81:559–75.
10. Fisher R. Statistical methods for research workers. Edinburgh: Oliver & Boyd, **1932**.
11. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics **2000**; 155:945–59.
12. Devlin B, Roeder K. Genomic control for association studies. Biometrics **1999**; 55:997–1004.
13. Kulski JK, Dawkins RL. The P5 multicopy gene family in the MHC is related in sequence to human endogenous retroviruses HERV-L and HERV-16. Immunogenetics **1999**; 49:404–12.
14. de Bakker PI, McVean G, Sabeti PC, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet **2006**; 38:1166–72.
15. Stephens HA. HIV-1 diversity versus HLA class I polymorphism. Trends Immunol **2005**; 26:41–7.
16. Liu Y, Helms C, Liao W, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. PLoS Genet **2008**; 4:e1000041.
17. Reed R, Hurt E. A conserved mRNA export machinery coupled to pre-mRNA splicing. Cell **2002**; 108:523–31.
18. Allcock RJ, Williams JH, Price P. The central MHC gene, BAT1, may encode a protein that down-regulates cytokine production. Genes Cells **2001**; 6:487–94.
19. Rennert PD, Browning JL, Mebius R, Mackay F, Hochman PS. Surface lymphotoxin alpha/beta complex is required for the development of peripheral lymphoid organs. J Exp Med **1996**; 184:1999–2006.
20. Kedzierska K, Crowe SM. Cytokines and HIV-1: interactions and clinical implications. Antivir Chem Chemother **2001**; 12:133–50.
21. Kilding R, Iles MM, Timms JM, Worthington J, Wilson AG. Additional genetic susceptibility for rheumatoid arthritis telomeric of the DRB1 locus. Arthritis Rheum **2004**; 50:763–9.
22. Kula D, Jurecka-Tuleja B, Gubala E, Krawczyk A, Szpak S, Jarzab M. Association of polymorphism of LTalpha and TNF genes with Graves' disease. Folia Histochem Cytobiol **2001**; 39(Suppl 2):77–8.
23. Rodriguez-Rodero S, Rodrigo L, Fdez-Morera JL, et al. MHC class I chain-related gene B promoter polymorphisms and celiac disease. Hum Immunol **2006**; 67:208–14.
24. Vandenbroeck K. Cytokine gene polymorphisms in multifactorial conditions. Taylor & Francis Group, ed. Boca Raton: Taylor & Francis Group, **2006**.
25. Ordovas JM. Gender, a significant factor in the cross talk between genes, environment, and health. Gend Med **2007**; 4(Suppl B):S111–22.
26. Kamatani Y, Matsuda K, Ohishi T, et al. Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population. J Hum Genet **2008**; 53: 64–73.
27. Singal DP, Li J, Lei K. Genetics of rheumatoid arthritis (RA): two separate regions in the major histocompatibility complex contribute to susceptibility to RA. Immunol Lett **1999**; 69:301–6.
28. Sasaki T, Gan EC, Wakeham A, Kornbluth S, Mak TW, Okada H. HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. Genes Dev **2007**; 21:848–61.
29. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol Cell Biol **2004**; 24:787–95.
30. Holm SJ, Carlen LM, Mallbris L, Stahle-Backdahl M, O'Brien KP. Polymorphisms in the SEEK1 and SPR1 genes on 6p21.3 associate with psoriasis in the Swedish population. Exp Dermatol **2003**; 12:435–44.
31. Mellors JW, Margolick JB, Phair JP, et al. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 cell count slope for progression to AIDS and death in untreated HIV-1 infection. JAMA **2007**; 297:2349–50.
32. Brass AL, Dykxhoorn DM, Benita Y, et al. Identification of host proteins required for HIV infection through a functional genomic screen. Science **2008**; 319:921–6.