

Behavioral Risk Exposure and Host Genetics of Susceptibility to HIV-1 Infection

Sadeep Shrestha,^{1,2,3,a} Steffanie A. Strathdee,^{3,4} Noya Galai,³ Taras Oleksyk,^{1,2} M. Daniele Fallin,³ Shruti Mehta,³ Daniel Schaid,⁵ David Vlahov,⁶ Stephen J. O'Brien,¹ and Michael W. Smith^{1,2}

¹Laboratory of Genomic Diversity, National Cancer Institute at Frederick, National Institutes of Health, and ²Basic Research Program, Science Applications International Corporation, National Cancer Institute at Frederick, Frederick, and ³Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland; ⁴Division of International Health and Cross Cultural Medicine, University of California School of Medicine at San Diego, La Jolla; ⁵Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota; ⁶Center for Urban Epidemiologic Studies, New York Academy of Medicine, New York

(See the editorial commentary by Telenti and Ioannidis, on pages XXX–XX.)

Background. Some individuals are readily infected with low human immunodeficiency virus type 1 (HIV-1) exposure, whereas others appear less susceptible, suggesting that host genetics plays a role in the viral entry pathway. The matched case-control study design with measured risk exposures provides an avenue for discovering genes involved in susceptibility to infection.

Methods. We conducted a nested case-control study of African Americans (266 HIV-1 seroconverter cases and 532 seronegative controls from the AIDS Link to Intravenous Experience cohort), to examine the association between 50 single-nucleotide polymorphisms (SNPs) in 9 candidate genes (*CCR5*, *CCR2*, *RANTES*, *MPIA*, *MCP2*, *IL10*, *IFNG*, *MCSE*, and *IL2*) and susceptibility to HIV-1 infection. To account for differential exposure propensities, risk behavior self-reported during semiannual visits was used to estimate a standardized cumulative risk exposure (SCRE). Individual SNPs were evaluated using conditional logistic-regression models, and the inferred haplotypes were assessed in the haplotype trend regression analyses after adjusting for age and SCRE.

Results. Four SNPs (*CCR2*–*V64I*, *CCR5*–*2459*, *MPIA*+*954*, and *IL2*+*3896*) and specific haplotypes in the *IL2* and *CCR2/CCR5* regions were significantly associated with HIV-1 infection susceptibility in different genetic models.

Conclusions. Our results suggest that genetic variants in associated host genes may play an important role in susceptibility to HIV-1 infection.

There were 4.9 million new cases of HIV-1 infection worldwide during 2004 alone, and ~39.4 million people are currently living with HIV [1]. The primary risk factors for HIV-1 infection are unprotected sexual in-

tercourse, sharing of syringes, and being an infant born to an infected mother. In most cases, behavioral modification remains a foremost priority with regard to prevention of infection. The importance of biological and genetic differences between individuals in explaining differential susceptibility to HIV-1 infection is largely unknown. The course of HIV-1 susceptibility varies widely even among individuals with similar risk exposure levels [2–4]. For example, some sex workers and homosexual men have remained uninfected despite repeatedly engaging in unprotected sexual intercourse with HIV-1-infected partners or constantly engaging in high-risk behavior [4–6].

A possible role of host genetics in determining susceptibility to HIV-1 exposure is also suggested by the differential immunological responses that individuals have during the course of infection. Population-based genetic studies of HIV-1 infection susceptibility have been limited, since, in many cases, the seroconversion

Received 11 May 2005; accepted 29 June 2005; electronically published 22 November 2005.

Presented in part: 53rd Annual Meeting of the American Society of Human Genetics (poster 1259).

Potential conflicts of interest: none reported.

Financial support: This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Cancer Institute (NCI), Center for Cancer Research. This publication has been funded, in part, with federal funds from the NCI, NIH (contract N01-CO-12400), and National Institute on Drug Abuse (grants DA09225, DA8009, and DA12568).

^a Present affiliation: Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham.

Reprints or correspondence: Dr. Michael W. Smith, SAIC-Frederick, National Cancer Institute at Frederick, Bldg. 560, Rm. 21-74, Frederick, MD 21702 (smithm@ncifcrf.gov).

The Journal of Infectious Diseases 2006;193:000–000

© 2005 by the Infectious Diseases Society of America. All rights reserved. 0022-1899/2006/19301-00XX\$15.00

time is not known and the nongenetic risk exposure data are seldom available to evaluate disease infectivity related to host genetics. Most studies have focused on high-risk exposed uninfected or highly exposed but persistently seronegative individuals, such as discordant couples who have unprotected sex [2, 7, 8] and commercial sex workers [4, 9–12]. Immunologic and genetic studies of exposed yet uninfected individuals have helped to elucidate protective mechanisms for HIV-1 infection [13–15].

Previous studies have identified 14 genetic polymorphisms that show associations with disease progression [16]. However, apart from the HLA class I and II genes, only 4 other genes (*CCR5*, *IL10*, *RANTES*, and the *MCP1-MCP3-Eotaxin* gene clusters) show associations with infection in European Americans [16–21]. The most significant polymorphism associated with HIV-1 infection is a 32-bp deletion in the coding region of the *CCR5* gene [16, 22], where homozygotes (*CCR5-Δ32/Δ32*) show protection [14, 15, 23] and nearly resistance [4, 23, 24]. The frequency of *CCR5-Δ32/Δ32* is essentially zero in African Americans, and no distinct genetic polymorphisms have been associated with HIV-1 infection in this ethnic group.

We investigated 9 genes involved in the complex pathway of HIV-1 entry and replication, focusing on the *CCR5* coreceptor, which is predominantly used during primary infection [25–27]. Specifically, we examined polymorphisms in *CCR2*, *CCR5*, *RANTES*, *MIP1A*, *MCP2*, *IL10*, *IFNG*, *MCSF*, and *IL2*. The natural ligands of the *CCR5* coreceptor—*RANTES*, macrophage inflammatory protein 1 α (MIP1- α), and monocyte chemoattractant protein 2 (MCP2)—inhibit HIV-1 entry and down-regulate *CCR5* expression [28–32]. The cytokines interleukin (IL)–10 [33, 34] and IL-2 [35, 36] up-regulate the expression of *CCR5* in vitro and induce other cytokines that are potentially involved in the *CCR5* viral entry mechanism [37]. Macrophage colony-stimulating factor (MCSF) [38, 39] up-regulates *CCR5* expression and also enhances HIV-1 replication. Interferon (IFN)- γ [40, 41] has been shown to enhance HIV-1 transcription but also to inhibit viral entry. Although not exhaustive, this list of genes includes those among the major cellular participants in primary HIV-1 infection.

We examined polymorphisms in these host genes, using a nested case-control study within a cohort of African American injection drug users (IDUs), to investigate the genes' potential roles in susceptibility to HIV-1 infection. We assessed both individual single-nucleotide polymorphism (SNP) and haplotype associations with HIV-1 seroconversion in the *CCR2/CCR5* region and 7 other candidate gene regions, using a conditional logistic model framework accounting for differential risk exposure. To our knowledge, this is the first study to quantify HIV-1 risk exposure as a measurement and to explore its confounding effects in a controlled genetic analysis setting.

SUBJECTS, MATERIALS, AND METHODS

Subjects. AIDS Link to Intravenous Experience (ALIVE) is a prospective cohort study of primarily African American IDUs in Baltimore. Participants have been studied for natural history and risk factors of HIV-1 infection and progression to AIDS. Details of the cohort, along with materials and methods for the genetic study design, have been described elsewhere [42–44]. Only African American participants were included in the present study, to avoid the confounding effects associated with racial differences and the lack of statistical power for other groups in addressing HIV infection within the ALIVE cohort. This study was approved by the Committee on Human Research at the Johns Hopkins Bloomberg School of Public Health and the National Cancer Institute.

Cases and controls. A nested case-control study within the ALIVE cohort was used to investigate genetic variants of specific candidate genes associated with HIV-1 seroconversion, after adjusting for measured risk exposure. Cases were participants who underwent seroconversion during clinical follow-up. The date of HIV-1 seroconversion was estimated to be the midpoint between the last HIV-1–seronegative test and the first documented HIV-1–seropositive test. Two controls per case (cases, $n = 266$; controls, $n = 532$) were randomly chosen from among HIV-1–seronegative individuals and were matched according to the duration of follow-up. Incidence density sampling [45, 46] was used to select the controls from the pool of HIV-1–seronegative IDUs in active follow-up within a 6-month period (± 3 -month window) of the seroconversion date for cases. Under this scenario, 5 subjects were initially controls but later became cases, resulting in a total sample size of 793.

SNP selection and genotyping. Fifty SNPs in the candidate genes were identified and assessed for association with HIV-1 susceptibility (table A1 in the appendix, which is available only in the electronic edition of the *Journal*). SNPs were genotyped by Taqman assay [47], using an Applied Biosystems 7900 genetic analyzer. *CCR5-Δ32* was the only non-SNP polymorphism, and the variant was determined by an agarose gel sizing assay [48].

Risk exposure assessment. The matched follow-up time between cases and controls partially adjusts for unknown confounding factors; however, in addition, we quantified the cumulative risk exposure on the basis of self-reported risk behavior and subsequently adjusted for its effect. Previously, Nelson et al. [43] defined independent risk behavior associated with HIV-1 seroconversion within the ALIVE cohort [43]. On the basis of these documented risk factors, a measure of standardized cumulative risk exposure (SCORE) for each individual was estimated. The log of the estimated relative incidence values for each risk factor in the model were used as weights in the summary risk measure for all visits up to the seroconversion point for each individual, accounting for the person-time ex-

posure. CRE_i is the cumulative risk exposure for person i and is calculated as follows:

$$CRE_i = \sum_{j=1}^m \sum_{k=1}^n \beta_k R_{kj} t_j, \quad (1)$$

where β_k is the log relative incidence associated with risk variable k in the seroconversion model, R_{kj} is the indicator for risk behavior k at visit j ($1 = \text{yes}$; $0 = \text{no}$), and t_j is the number of days of follow-up at visit j .

Although the CRE values within the matched trios we examined were comparable, a measure that has a uniform interpretation among all individuals regardless of the differential follow-up time—an SCORE—was deemed appropriate for the analyses and was calculated as follows:

$$SCORE_i = \frac{CRE_i}{TFT_i}, \quad (2)$$

where CRE_i is the cumulative risk exposure (eq. [1]) and TFT_i is the total follow-up time for individual i . First, the proportions of cases were explored in different quantile groups of the SCORE distribution. Since the values of SCORE were based on risk factors for HIV-1 seroconversion, the proportion of cases in the risk groups was expected to increase with incremental values of SCORE. High, medium, and low risk levels were categorized on the basis of the proportion of seroconverters in the SCORE quantile groups.

Statistical analyses. Cases and controls were first compared in terms of demographic and risk behavior data. χ^2 tests were used to compare categorical variables (e.g., sex), whereas the Mann-Whitney U test was used to compare continuous variables (e.g., SCORE values and age). A relatively stringent threshold of $r^2 > 0.85$ was used to identify redundant SNPs. Analyses were conducted using the statistical packages SAS (version 9.0; SAS Institute), STATA (version 7.0; STATA), and S-Plus (version 6.0; MathSoft).

Conformity of the genotype proportions to Hardy-Weinberg equilibrium (HWE) was examined for each polymorphism in cases and controls. Pairwise linkage disequilibrium (LD) between SNPs at each gene was measured and viewed in the Graphical Overview of Linkage Disequilibrium program [49], using Lewontin's D' [50] and the square of the correlation coefficient, r^2 [51]. Initial exploratory analyses for differences in allele frequencies and genotype distribution between the cases and controls were performed using Fisher's exact test. Analyses comparing genotypes at each individual SNP between cases and controls were performed using conditional logistic regression. In our analyses, P values $\leq .05$ were considered to be statistically significant, and P values between .05 and .1 were deemed to indicate trends. Indicator variables for all genotypes were created by using the most common genotype as the ref-

erence category. In the case of the dominant model, heterozygotes and less common homozygotes were combined. Age and the SCORE risk levels were modeled as covariates.

The expectation/maximization (EM) algorithm [52], modeled after SNP HAP software (available at: <http://www-gene.cimr.cam.ac.uk/clayton/software/>) using an efficient progressive-insertion algorithm, was applied to estimate haplotype frequencies in a combined pool of cases and controls. On the basis of all possible haplotypes given the genotype, a haplotype matrix of posterior probabilities for each individual was estimated, and these probabilities were used in a haplotype trend regression (HTR) model as independent variables [53, 54]. The HTR model has been expanded to the conditional logistic-regression framework [46], allowing adjustment for matched case-control status, age, and risk levels. All rare haplotypes with frequencies $< 1\%$ were collapsed into 1 haplotype group, and the most frequent haplotypes were considered the references in the analyses. Global tests were conducted to assess the significance for the whole effect of haplotypes and individual tests for each haplotype. Further, stepwise regression was used to obtain the most parsimonious model. Haplotype-pair regression analysis was further performed for the significant haplotypes. When haplotypes were significantly associated with HIV-1 infection, the odds ratios (ORs) and 95% confidence intervals (CIs) were calculated. The haplotype analyses were implemented using the S-PLUS version (MathSoft) of the HaploStats statistical package (available at: <http://mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm>).

RESULTS

The case and control groups did not differ by sex, but the cases were slightly younger than the controls (median age, 36.1 vs. 38.5 years; $P < .0001$) (table 1). As expected, the median SCORE value was higher among cases than controls (0.62 vs. 0.49; $P < .0001$). SCORE values ranged from 0 to 1.86 (figure 1), and the distribution of values for all subjects was stratified into quintiles. The proportions of HIV-1 seroconverters (cases) in each quantile were calculated and are informative on a relative basis. The first and second quintiles (24% and 26% cases,

Table 1. Demographic characteristics of the study participants.

Demographic characteristic	Cases ^a ($n = 266$)	Controls ($n = 532$)	P
Male sex, %	74	75	.13
Age, median, years	36.1	38.5	$< .0001$
SCORE, median	0.63	0.49	$< .0001$

NOTE. Age and standardized cumulative risk exposure (SCORE) are median values at ascertainment time that are not corrected for the matching status of cases and controls.

^a Seroconverters with a < 2 -year interval between the last seronegative and first seropositive tests were considered to be cases.

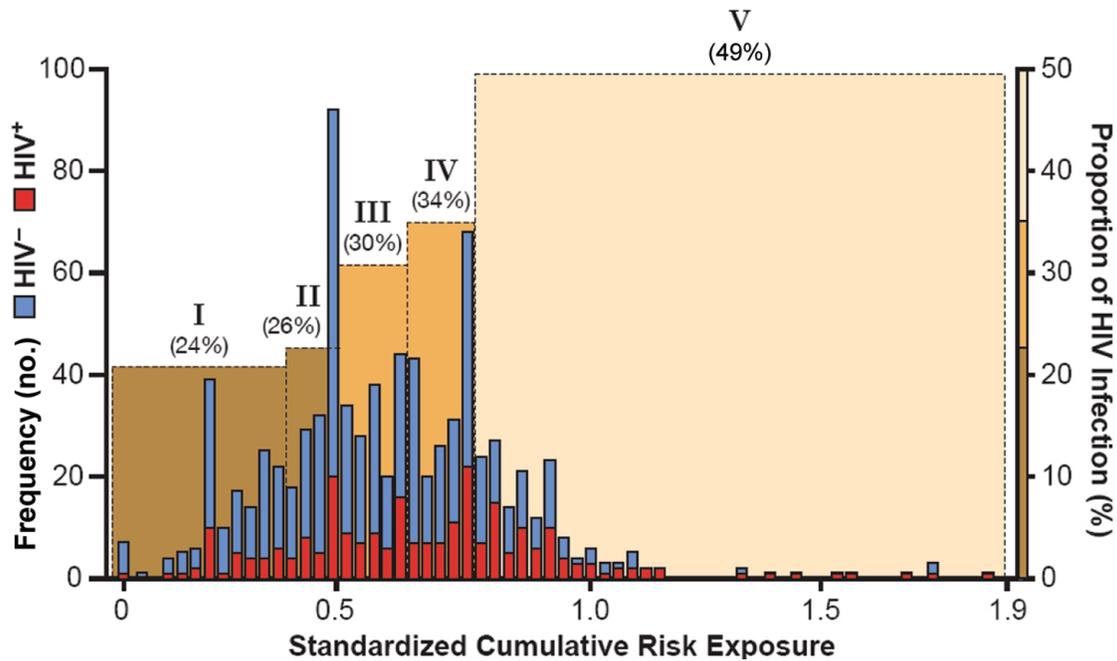


Figure 1. Risk level stratification based on distribution of standardized cumulative risk exposure (SCRE) values. Nelson et al. [43] showed that injection cocaine use (relative risk [RR], 1.61), homosexual activity (RR, 7.03), visiting “shooting galleries” (RR, 1.47), having any sexually transmitted diseases (RR, 1.81), and an interaction between drug injection and heterosexual behaviors (heterosexual sex and no drug injection: RR, 1.59; heterosexual sex and drug injection: RR, 1.77; heterosexual sex and drug injection more than once daily: RR, 3.14; no sex and drug injection less than once daily: RR, 4.28; no sex and drug injection more than once daily: RR, 5.06) were independently associated with seroconversion in a multivariate analysis. These estimates and measures were used to calculate SCRE values according to equations (1) and (2), from semiannual visits that queried the risk behaviors of all participants. The length of follow-up for exposure at each visit was determined as the no. of days between each visit and the previous one (maximum, 365 days). The first visit was considered to represent 180 days. The frequencies of SCRE values among subjects ($n = 798$) is labeled on the left vertical axis. The SCRE distribution was divided into quintiles, and the proportion of cases (seroconverters) in each quintile was calculated (*right vertical axis*). Quintiles I and II have 24% and 26% cases, respectively (low risk), quintiles III and IV have 30% and 34% cases, respectively (medium risk), and quintile V has 49% cases (high risk).

respectively) were categorized as low risk, the third and fourth quintiles (30% and 34% cases, respectively) were categorized as medium risk, and the fifth quintile (49% cases) was categorized as high risk.

A graphical representation of pairwise LD for 50 SNPs in 8 gene regions among controls is shown in figure 2. The distribution of alleles and genotypes in cases and controls at 41 SNPs (9 were excluded because of low frequency or high LD) was explored (table 2). Six (*CCR2*-V64I, *CCR5*-2459, *MIP1A*+954, *IL2*+161, *IL2*+3896, and *IFNG*+2112) trended toward statistical significance ($P < .10$) in both the allelic and genotypic frequency analyses. Matched case-control analysis using conditional logistic regression showed the effects of particular genotypes, both with and without (data not shown) adjustment for age and risk measured as SCRE; 4 of these 6 SNPs showed modest effects (figure 3). For the *CCR2*-V64I SNP, the adjusted OR (OR_{adj}) for heterozygotes was 0.69 (95% CI, 0.48–0.98), and that for A/A homozygotes was 0.39 (95% CI, 0.10–1.50), relative to the most common homozygotes. The rare minor-allele homozygote at *CCR2*-V64I was not statistically significant, but its effect

was similar to that of the heterozygote comparison, suggesting a dominant model (combining G/A and A/A: OR_{adj} , 0.66 [95% CI, 0.46–0.94]). Likewise, at *CCR5*-2459, the OR_{adj} values for heterozygotes and minor-allele homozygotes, compared with G/G homozygotes, were 0.71 (95% CI, 0.51–0.97) and 0.62 (95% CI, 0.38–1.00), respectively, whereas, when combined, a 0.69 (95% CI, 0.50–0.95) infection-protective effect was seen.

Other candidate genes within the CCR5 infection pathway also showed some signals of association with HIV infection. Effects seen for the *MIP1A*+954 T allele suggest a recessive model (OR_{adj} for T/A heterozygotes, 1.10 [95% CI, 0.79–1.56]; OR_{adj} for T/T homozygotes, 2.25 [95% CI, 1.03–12.29]) (figure 3). The support for the T/T homozygote effect is drawn from 11 individuals, and similar results are seen at the 2 highly linked loci ($r^2 \geq 0.85$; data not shown). None of the SNPs in *RANTES* or *MCP2* showed any significant association with susceptibility to HIV-1 infection. Since the genes encoding the CCR5 coreceptor ligands RANTES, MIP1- α , and MCP2 are localized to a 1.8-Mb region on 17q11-12, haplotype analysis was considered for all of the SNPs in the 3 ligand genes, but no significant

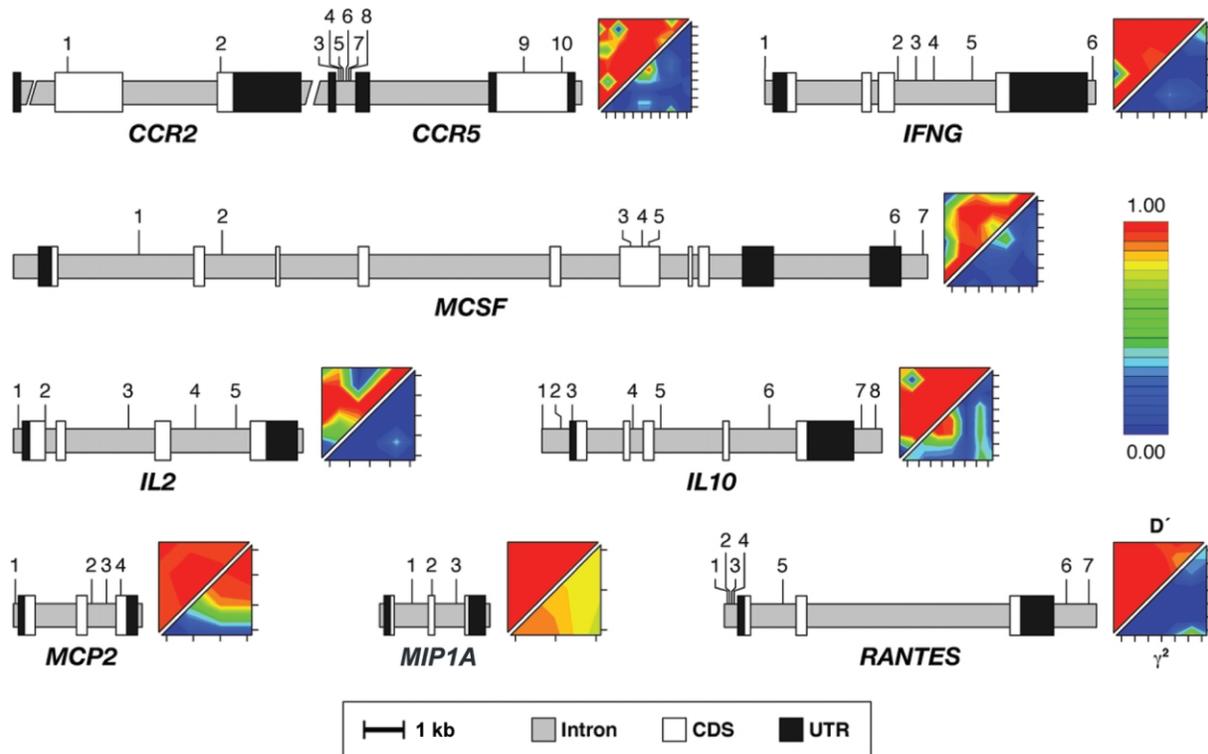


Figure 2. Linkage disequilibrium (LD) between single-nucleotide polymorphisms (SNPs) examined in 8 candidate gene regions. The physical positions of introns, coding regions (CDS), and untranslated regions (UTRs) are illustrated for each gene region. The SNPs examined are shown in the same order as in table 2 and the disequilibrium plots. D' (top left triangle) and r^2 (bottom right triangle) in controls are shown for very little (blue) to modest (green) to very strong (red) LD, ranging from 0 to 1.

LD was seen in this extended region (analysis not shown). For *IL2*+3896, the genotypes A/– and A/A did not separately indicate a significant association, compared with the common homozygous deletion genotype, but they showed a significant protective effect together (unadjusted OR, 0.74 [95% CI, 0.55–0.99]; OR_{adj}, 0.71 [95% CI, 0.52–0.97]) (figure 3).

Analysis of haplotypes (table 3) indicated that haplotypes in the *CCR2*/*CCR5* ($P = .03$) and *IL2* ($P = .02$) regions were significantly associated with HIV-1 infection. In the HTR analysis, when an additive approach was used, the CAAGAAC+ haplotype in the *CCR2*/*CCR5* region showed significant association with HIV-1 infection in the overall model (OR, 0.71 [95% CI, 0.57–0.85]), as well as in a more parsimonious model comparing it with all other haplotypes (OR, 0.77 [95% CI, 0.61–0.98]). In the analyses of the *IL2* haplotypes, TGCAA (OR, 0.79 [95% CI, 0.69–0.89]) and TGT-A (OR, 0.53 [95% CI, 0.25–0.81]) showed significant associations in the overall model. A more parsimonious model contrasting these 2 haplotypes with all others showed similar associations (ORs, 0.78 [95% CI, 0.71–0.86] and 0.52 [95% CI, 0.39–0.68], respectively). The results of the HTR analysis were consistent with the single-SNP results, in which the SNPs in LD with the significant haplotypes were also associated with HIV-1 infection (*CCR5*–2459A in LD with

the CAAGAAC+ haplotype and *IL2*+161A in LD with the TGCAA haplotype).

DISCUSSION

Our study suggests that some host genes may play a role in susceptibility to HIV-1 infection. The independent associations of specific SNPs and haplotypes were modest but statistically significant after age and differential risk exposure were adjusted for (figure 3 and table 3). We observed associations between susceptibility to HIV-1 infection and variants in the gene regions of *CCR2* and *CCR5* that encode the coreceptors; *MIP1- α* , a natural ligand of the coreceptor *CCR5*; and *IL-2*, a cytokine involved in immune regulation that has widely been used in therapy and vaccine trials. In contrast to previous studies that focused on only highly exposed uninfected individuals [14, 60–62], our study included a nested case-control design in which we focused on seroconverters and adjusted for measured risk behaviors in genetic analyses of HIV-1 susceptibility.

Exposure levels were measured in our study as SCRE, on the basis of self-reported risk behavior. Humans significantly influence their level of exposure to infectious agents like HIV-1 through their behavior, and such behavior is generally difficult

Table 2. Single-locus allele and genotype frequency distributions between cases and controls.

Gene, SNP	NCBI ID	Minor-allele frequency, %			Genotype frequency, %						P
		Cases	Controls	P	-/-		+/-		+/+		
					Cases	Controls	Cases	Controls	Cases	Controls	
<i>CCR2/CCR5</i>											
V64I (C/T)	rs1799864	12	16	.06	1	2	21	28	78	70	.04
N260N (G/A)	rs1799865	33	33	.65	11	10	45	45	45	45	.95
-2733 (A/G)	rs2856758	7	6	.35	0	1	14	11	86	89	.32
-2554 (G/T)	rs2734648	37	34	.10	16	11	43	45	41	44	.18
-2459 (G/A)	rs1799987	38	43	.01	16	18	45	51	40	31	.03
-2086 (A/G)	rs1800023	12	13	.60	1	2	21	22	78	77	.75
-1835 (C/T)	rs3181036	16	18	.54	3	3	26	29	71	67	.53
Δ32 (+/Δ)	rs333	2	2	.99	0	0	3	3	97	97	.86
<i>MIP1A</i>											
+954 (C/T)	rs1130371	18	15	.06	5	2	26	26	69	72	.01
<i>MCP2</i>											
-516 (G/A)	rs3138035	14	14	.45	3	2	23	22	74	76	.64
+1080 (C/A)	rs3138036	7	6	.76	0	0	14	12	86	88	.73
<i>IL10</i>											
-2507 (T/A)	rs1800890	28	26	.43	6	8	44	37	50	55	.12
-598 (T/C)	rs1800896	34	34	.92	17	17	51	46	32	37	.48
-14 (C/A)	rs1800872	42	40	.46	11	12	48	44	41	44	.58
+1165 (C/T)	rs1518111	41	40	.63	15	17	52	46	33	37	.37
+1575 (G/A)	rs1554286	40	38	.46	14	16	52	45	34	39	.16
+3399 (G/A)	rs3024495	4	4	.96	1	1	7	7	93	93	.86
+3948 (T/C)	rs3024496	41	42	.65	16	18	51	48	33	34	.58
+4283 (A/G)	rs3024498	14	13	.51	1	2	25	22	74	76	.75
<i>RANTES</i>											
-403 (C/T)	rs2107538	43	44	.54	19	18	46	51	35	30	.35
+363 (A/G)	rs2280789	21	20	.52	4	4	34	33	62	64	.54
+7725 (A/T)	rs3817655	41	44	.19	19	19	43	51	38	31	.08
+8812 (A/G)	rs1065341	22	24	.62	4	6	36	37	60	57	.56
<i>IL2</i>											
-338 (G/T)	rs2069762	12	11	.60	2	1	17	19	80	80	.23
+161 (T/G)	rs2069763	9	6	.02	1	0	16	13	83	87	.07
+1508 (C/T)	rs2069772	2	3	.34	1	0	3	6	96	94	.43
+3896 (-/A)	rs2069778	27	32	.06	8	11	38	43	54	46	.04
+4510 (A/G)	rs2069771	5	4	.73	1	1	8	12	91	88	.40
<i>MCSF</i>											
+3603 (A/G)	rs915357	34	31	.29	12	11	44	40	44	50	.29
+4274 (G/A)	rs3768484	8	9	.57	1	1	14	16	85	83	.57
+12882 (C/A)	rs333970	16	19	.24	3	5	25	29	72	66	.24
+13010 (A/G)	rs1058885	47	45	.43	22	20	51	48	27	32	.44
+13355 (C/A)	rs3738760	39	40	.82	15	16	49	47	36	37	.88
+18450 (G/A)	rs3093037	11	8	.11	1	0	19	14	80	85	.12
+19500 (C/A)	rs2050462	33	37	.33	12	15	43	44	46	41	.82
<i>IFNG</i>											
-179 (G/T)	rs2069709	3	3	.45	0	0	6	5	94	95	.57
+2112 (A/G)	rs2234685	18	15	.09	3	3	32	24	65	73	.06
+2477 (C/A)	rs2069714	5	5	.89	0	0	10	10	90	90	.88
+3289 (G/-)	rs2069733	6	5	.87	0	0	11	10	88	90	.56
+3360 (T/C)	rs2069718	38	40	.35	13	16	50	49	37	35	.53
+5299 (A/G)	rs2069727	20	22	.62	4	5	33	33	63	62	.58

NOTE. Several databases and sources were used to identify and select single-nucleotide polymorphisms (SNPs) within the candidate genes. SNPs in *IL2*, *IL10*, and *IFNG* were identified from the genotype data for the UW-FHCRC Variation Discovery Resource [55]. The Molecular Evolutionary Genetic Analysis program, Mega 2.0 [56], was used to create haplotype phylogeny trees from these genotype data, to generate an evolutionary perspective of the haplotypes. SNPs (minor-allele frequency, >2%) that differentiated the haplotype lineages were selected for genotyping. Previous studies have identified the sets of informative SNPs in the *CCR2/CCR5* and promoter regions [57, 58]. Two public databases, the dbSNP database, maintained by the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/projects/SNP/>), and the JSNP database, maintained by the Japan Biological Informatics Consortium (<http://snp.ims.u-tokyo.ac.jp/>), were used to select SNPs in the *RANTES*, *MIP1A*, *MCP2*, and *MCSF* genes. The minor allele is the second allele in the SNP column. Overall, among the 50 SNPs, there were 4 with minor-allele frequencies $\geq 2\%$ (*RANTES*: rs4239253, rs2280788, and rs1800825; *CCR2/CCR5*: rs1799863) and 5 that were in strong linkage disequilibrium (LD) ($r^2 \geq .85$) with another SNP in the gene region (*MIP1A*: rs1719130 and rs1719134 in LD with rs1130371; *MCP2*: rs3138038 and rs3138038 in LD with rs3138036; *CCR2/CCR5*: rs1799988 in LD with rs1799987). The most common genotype is represented by +/+, the heterozygous genotype by +/-, and the less common genotype by -/-.

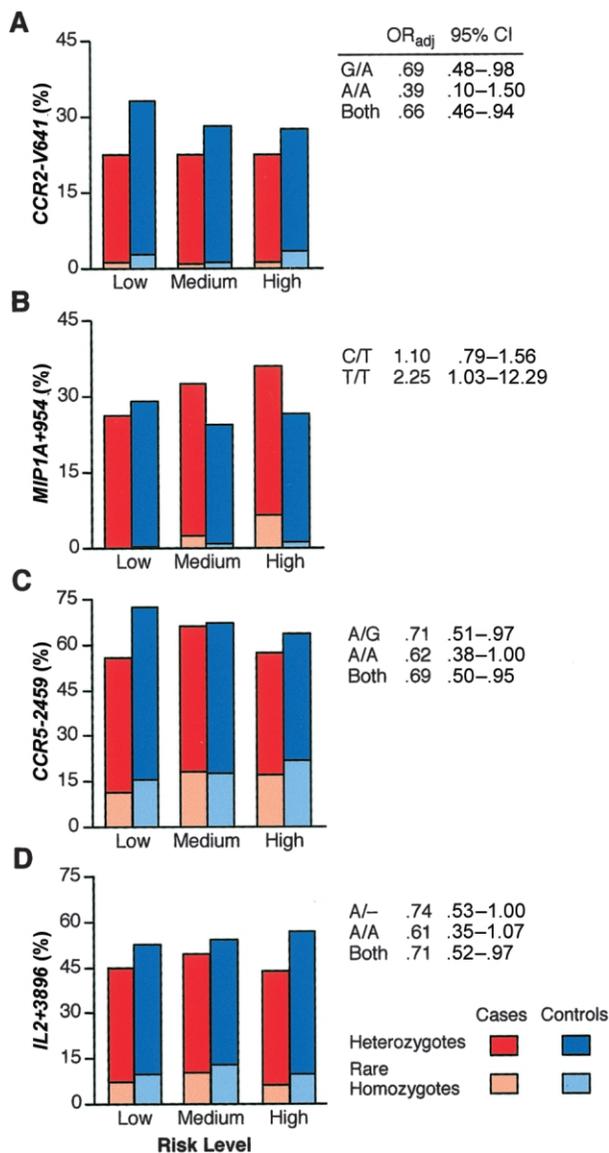


Figure 3. Genotypic frequencies and HIV infection adjusted odds ratios (OR_{adj}) at single-nucleotide polymorphisms (SNPs) *CCR2-V64I* (A), *MIP1A+954* (B), *CCR5-2459* (C), and *IL2+3896* (D). Adjustments in the conditional logistic regression models were made for risk levels (also see figure 2) and age. Common homozygotes at each SNP were used as the reference group (frequencies not shown).

to quantify or assess. The reliability and validity of behavior self-reported by IDUs in numerous publications [63–65], coupled with relatively low rates of abstinence from risk behavior reported at ALIVE clinical visits, suggests that these self-reports are robust indicators of true behavior. The SCORE we estimated from ALIVE self-reports of both injection and sexual behavior captured varying IDU behavior over several years. In contrast, most cross-sectional studies of IDUs [66, 67] represent a less reliable snapshot of varying risk behavior. Given the high seroprevalence of HIV-1 among IDUs in Baltimore, our study

subjects are at greater risk for HIV-1 infection than are other populations. The SCORE we developed quantifies the level of risk for each individual and estimates the differential propensity of risk exposure for genetic analyses.

Correcting for multiple comparisons, using methods such as Bonferroni or false discovery rate, did not result in significant associations with any of the SNPs or haplotypes. Given the biological plausibility of the importance of CCR5 pathway genes in primary HIV infection, some exploration of these results with $P < .05$ seems justified to guide later research on African Americans. We did not observe any *CCR5-Δ32/Δ32* homozygotes in our study sample, which is not surprising, given the rarity of the $\Delta32$ allele in African Americans. We found that individuals with the *64I* allele in *CCR2*, a gene encoding a β chemokine receptor, were less susceptible (dominant allele A model) to HIV-1 infection (figure 3). The G→A polymorphism in the coding region of the HIV-1 coreceptor CCR2 (*-64I* A allele frequency, 0.098 in white individuals and 0.151 in African Americans) causes a single amino acid change from Val to Ile in the first transmembrane domain but does not appear to alter the mechanism of the coreceptor [60]. This A variant has been found to be associated with delayed HIV-1 disease progression in several studies [60, 68–70], but no effect on HIV-1 transmission has yet been reported. In vitro, an isoform of CCR2 (*CCR2A*) binds to CCR5 in the cytoplasm and sequesters CCR5 receptor before it reaches the cell surface [71], providing a possible cellular mechanism underlying delayed progression to AIDS in individuals with the *64I* polymorphisms. The consistency of *CCR2-64I* decreasing surface availability of CCR5 for viral entry supports our association in which individuals with these polymorphisms are less susceptible to HIV-1 infection. We also observed that another variant in the promoter region of *CCR5*, *-2459A*, appeared to be protective in a dominant model. A protective effect of the *-2459A* allele against HIV infection, however, seems unlikely, because it has previously been shown to accelerate progression to AIDS [57, 72]. The SNP showing association with infection in *IL2* is in a noncoding region, and the one in *MIP1A* is a synonymous SNP. Very little is known regarding the relationship between these genes and HIV infection, which requires further examination—for example, the association between structural variation of *MIP1A* (*CCL3L1*) in the number of copies and HIV-1 susceptibility should be investigated [73].

We developed a regression-based approach to study the association between inferred haplotype probabilities and susceptibility to HIV-1 infection in a nested case-control analysis. The primary assumptions of the method are (1) that sampling of cases and controls is random and (2) that HWE conditions are met when posterior probabilities of all possible haplotype pairs are estimated [53, 54]. Unlike the methods used in previous studies with inferred haplotypes based on the likelihood-ratio

Table 3. Global and individual haplotype tests, based on regression analysis, and the frequencies of haplotypes, estimated using the expectation/maximization algorithm.

Gene, haplotype	Frequency, %		P	
	Cases	Controls	Global	Haplotype
<i>CCR2/CCR5</i> ^a			.03	
CGAGGAC+	23	24		...
CAATGAC+	26	21		.60
CAAGAAC ^b	15	19		.01
TAAGAAT+	12	15		.51
CGATGGC+	9	9		.13
CAGGAAC+	5	4		.31
CAATGGC+	3	3		.45
CAAGAAT+	4	2		.92
CAGGAAC-	2	2		.68
<i>IL10</i>			.93	
TTATAGTA	40	38		...
TTCCGGTA	12	13		.72
ACCCGGCG	13	12		.56
TTCCGGCA	7	9		.21
TCCCGGCA	6	8		.18
ACCCGGCA	6	6		.57
TCCCGGTA	4	4		.71
ATCCGGCA	4	4		.80
ACCCGACA	4	4		.81
<i>RANTES</i>			.89	
CAAA	56	55		...
TATG	20	23		.34
TGTA	19	19		.97
<i>MCP2</i>			.82	
GA	78	80		...
AA	14	13		.52
GC	8	7		.92
<i>IL2</i>			.02	
TGC-A	49	48		...
TGCAA ^c	27	33		.03
GGC-A	8	7		.39
TTC-G	4	6		.39
TTC-A	7	4		.31
TGT-A ^c	2	3		.02
<i>IFNG</i>			.37	
GACGTA	36	36		...
GACGCG	20	22		.46
GACGCA	18	19		.16
GGCGTA	12	10		.33
GGC-TA	6	5		.20
GAAGTA	5	5		.97
TACGTA	3	3		.99
<i>MCSF</i>			.51	
AGCGAGC	29	29		...
GGAACGA	9	11		.26
GGCACGC	11	8		.05
AGCACGC	8	7		.77
AGAACGA	6	7		.84
AACACGA	5	6		.30
GGCGCAC	7	5		.14
AGCACGA	4	6		.15
AGCAAGC	4	4		.68
AGCGAGA	4	4		.65
GGCACGA	4	3		.50

test [74, 75], our method has several advantages. First, haplotype-specific association can be easily computed, along with evaluation of individual haplotypes; second, nongenetic covariates can be adjusted; and third, computation time is much shorter than with the likelihood-based approach, especially when heterozygotes are frequent.

Applying the new analysis method to the 8 gene regions, we found associations between *IL2* and *CCR2/CCR5* haplotypes and HIV-1 infection. Beyond studies implicating *CCR5-Δ32* homozygosity, there have been no reports indicating associations between haplotypes in the *CCR2/CCR5* region and susceptibility to infection. Some studies have shown that, among European Americans, +.P1.+ haplotype homozygotes (table 3, footnote "a") have accelerated AIDS progression [57, 58]. Among African Americans, a weak association with AIDS progression (using a set of samples that included 164 in our analyses) was reported in one study [72], and no association was reported in another [58]. One might expect concordance between AIDS progression and HIV susceptibility genes, but we observed a protective effect for HIV-1 infection (OR, 0.71 [95% CI, 0.57–0.85]) with the AIDS progression–accelerating +.P1.+–containing haplotype, CAAGAAC+ (table 3). Haplotypic analysis also indicates that variants in the *IL2* gene could be involved in protection against infection. Single-SNP analyses indicated an association with *IL2+3896A* that is also seen with the TGCAA haplotype it defines; however, it missed the haplotype TGT-A, which has a relatively low frequency. The functions of these 2 haplotypes are not known, but the importance of *IL2* is clear. *IL2* has been successfully used in therapy for HIV/AIDS and

NOTE. All haplotypes with frequencies <1% were collapsed into 1 group for the analysis but are not listed in the table. All single-nucleotide polymorphisms in *MIP1A* were highly correlated, and therefore haplotype analysis was not performed. Wald's test was performed to test the significance of individual haplotypes in the conditional logistic-regression model. The most common haplotype was used as the baseline reference in the haplotype test and adjusted for sex and standardized cumulative risk exposure (SCRE). *P* values are based on the difference between a simpler model with fewer variables and a complex one with more variables, in which a χ^2 is calculated as Δ deviance and Δ degrees of freedom.

^a The relationship between these *CCR2/CCR5* haplotypes and those in other studies [58, 59] are as follows: CGAGGAC+ corresponds to HHA and to +.P2.1.+; +.P5.1.+; +.P8.1.+; +.P9.1.+; and +.P11.w1.+; CAATGAC+ corresponds to HHD and HHB and to +.P3.1.+; CAAGAAC+ corresponds to HHE and to +.P1.1.+; TAAGAAT+ corresponds to HHF*2 and to 64I.P1.w3.+; CGATGGC+ corresponds to HHC and to +.P4.1.+; CAGGAAC+ corresponds to HHG*1 and to +.P1.w2.+; CAATGGC+ corresponds to HHC and to +.P4.1.+; CAAGAAT+ corresponds to HHF*1 and to +.P1.w3.+; and CAGGAC– corresponds to HHG*2 and to +.P1.w2.Δ32.

^b Haplotype-pair analysis for CAAGAAC+/CAAGAAC+ homozygotes (OR, 0.50 [95% CI, 0.22–1.14]) and CAAGAAC+/* heterozygotes (OR, 0.82 [95% CI, 0.36–1.88]) compared with other pairs of haplotypes were not significant. A dominant model also was not significant (OR, 0.79 [95% CI, 0.61–1.02]) but trended toward a protective effect.

^c In the haplotype-pair analysis, both the homozygotes and heterozygotes for TGCAA (ORs, 0.76 [95% CI, 0.49–1.18] and 0.82 [95% CI, 0.64–1.05], respectively) and TGT-A (ORs, 0.22 [95% CI, 0.03–1.58] and 0.63 [95% CI, 0.32–1.23], respectively) were not significant. The dominant model trended toward a protective effect both for TGCAA (OR, 0.81 [95% CI, 0.64–1.02]) and TGT-A (OR, 0.53 [95% CI, 0.28–0.99]).

currently has been extended to phase 3 trials [76, 77]. A deficiency in IL-2 production is one of the first immunologic defects to be described in HIV-1-infected individuals. The consistency of these modest associations with both the allelic and haplotype analysis of the *CCR2/CCR5* and *IL2* gene regions awaits replication in other cohorts and studies.

Discovering the role of host genetics in susceptibility to HIV-1 infection can provide important insights into the growing pandemic. Comprehensive scans [78–80] of genes involved in innate immunity and, indeed, the entire genome are possible with the study design and samples we have developed. Human genetic variation is known to play a role in susceptibility to many infectious diseases, most notably malaria, schistosomiasis, and tuberculosis [81–83]. Most previous studies have focused on twin, adoptee, and family designs, to control for differential exposure and the environment. However, studies of such related individuals are difficult, since exposure to infectious agents such as HIV-1 is frequently not similar among family members. To address these limitations, we developed a promising case-control study that explicitly takes into account measured exposure levels, to study the association of host genetics with HIV-1 infection. On examination of polymorphic markers in 9 *CCR5* pathway genes, SNPs in 4 of these genes (*CCR2*, *CCR5*, *MIP1A*, and *IL2*) showed moderate associations with HIV-1 infection in single-SNP and haplotype analyses. One of these SNPs (*CCR2-V64I*) has a plausible biological role, limiting the availability of HIV-1 coreceptor (*CCR5*) in the host cells [84, 85]. Like studies of the genetics of progression to AIDS [16], our study suggests a complex association of multiple genes with small contributory effects to HIV-1 infection.

Acknowledgments

First and foremost, we thank the AIDS Link to Intravenous Experience cohort participants and staff. We thank Cheryl Winkler, Michael Dean, and Terri H. Beaty, for helpful insights in developing the ideas for this article, and David Thomas, for buffy coat samples. We are also grateful to Kai Zhao, Ann Truelove, Bailey Kessing, Michael Malasky, Mary Thompson, Mahboobeh Safaeian, and Joseph Bareta for their assistance. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

References

- UNAIDS. AIDS epidemic update. Geneva, Switzerland: UNAIDS/WHO, 2004.
- Mazzoli S, Trabattoni D, Lo Caputo S, et al. HIV-specific mucosal and cellular immunity in HIV-seronegative partners of HIV-seropositive individuals. *Nat Med* 1997; 3:1250–7.
- Goh WC, Markee J, Akridge RE, et al. Protection against human immunodeficiency virus type 1 infection in persons with repeated exposure: evidence for T cell immunity in the absence of inherited *CCR5* coreceptor defects. *J Infect Dis* 1999; 179:548–57.
- Plummer FA, Ball TB, Kimani J, et al. Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: what mediates protection and why does it develop? *Immunol Lett* 1999; 66:27–34.
- Shearer GM, Clerici M. Protective immunity against HIV infection: has nature done the experiment for us? *Immunol Today* 1996; 17:21–4.
- Detels R, Liu Z, Hennessey K, et al. Resistance to HIV-1 infection. Multicenter AIDS Cohort Study. *J Acquir Immune Defic Syndr* 1994; 7:1263–9.
- Clerici M, Giorgi JV, Chou CC, et al. Cell-mediated immune response to human immunodeficiency virus (HIV) type 1 in seronegative homosexual men with recent sexual exposure to HIV-1. *J Infect Dis* 1992; 165:1012–9.
- Beretta A, Furci L, Burastero S, et al. HIV-1-specific immunity in persistently seronegative individuals at high risk for HIV infection. *Immunol Lett* 1996; 51:39–43.
- Fowke KR, Kaul R, Rosenthal KL, et al. HIV-1-specific cellular immune responses among HIV-1-resistant sex workers. *Immunol Cell Biol* 2000; 78:586–95.
- Dorrell L, Hessel AJ, Wang M, et al. Absence of specific mucosal antibody responses in HIV-exposed uninfected sex workers from the Gambia. *AIDS* 2000; 14:1117–22.
- Promdej N, Costello C, Wernett MM, et al. Broad human immunodeficiency virus (HIV)-specific T cell responses to conserved HIV proteins in HIV-seronegative women highly exposed to a single HIV-infected partner. *J Infect Dis* 2003; 187:1053–63.
- Rowland-Jones SL, McMichael A. Immune-responses in HIV-exposed seronegatives—have they repelled the virus? *Curr Opin Immunol* 1995; 7:448–55.
- Paxton WA, Martin SR, Tse D, et al. Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposures. *Nat Med* 1996; 2:412–7.
- Dean M, Carrington M, Winkler C, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CCR5* structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 1996; 273:1856–62.
- Huang Y, Paxton WA, Wolinsky EM, et al. The role of mutant *CCR5* allele in HIV-1 transmission and disease progression. *Nature Med* 1996; 2:1240–3.
- O'Brien SJ, Nelson GW. Human genes that limit AIDS. *Nat Genet* 2004; 36:565–74.
- Modi WS, Goedert JJ, Strathdee S, et al. MCP-1-MCP-3-Eotaxin gene cluster influences HIV-1 transmission. *AIDS* 2003; 17:2357–65.
- Shin HD, Winkler C, Stephens JC, et al. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of *IL10*. *Proc Natl Acad Sci USA* 2000; 97:14467–72.
- Ping A, Nelson GW, Wang L, et al. Modulating influence on HIV/AIDS by interacting *RANTES* gene variants. *Proc Natl Acad Sci USA* 2002; 99:10002–7.
- O'Brien SJ, Gao X, Carrington M. HLA and AIDS: a cautionary tale. *Trends Mol Med* 2001; 7:379–81.
- Kaslow RA, Dorak T, Tang JJ. Influence of host genetic variation on susceptibility to HIV type 1 infection. *J Infect Dis* 2005; 191(Suppl 1): S68–77.
- Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ, Dean M. Novel alleles of the chemokine-receptor gene *CCR5*. *Am J Hum Genet* 1997; 61:1261–7.
- Liu R, Paxton WA, Choe S, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996; 86:367–77.
- O'Brien TR, Winkler C, Dean M, et al. HIV-1 infection in a man homozygous for *CCR5* Delta 32. *Lancet* 1997; 349:1219.
- Berson JF, Long D, Doranz BJ, Rucker J, Jirik FR, Doms RW. A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *J Virol* 1996; 70:6288–95.
- Ioannidis JPA, Rosenberg PS, Goedert JJ, et al. Effects of *CCR5-Δ32*, *CCR2-64I*, and *SDF-1 3A* alleles on HIV-1 disease progression: an

- international meta-analysis of individual-patient data. *Ann Int Med* **2001**; 135:782–95.
27. Cohen OJ, Paolucci S, Bende SM, et al. CXCR4 and CCR5 genetic polymorphisms in long-term nonprogressive human immunodeficiency virus infection: lack of association with mutations other than CCR5-Delta32. *J Virol* **1998**; 72:6215–7.
 28. Hogan CM, Hammer SM. Host determinants in HIV infection and disease. Part 2: genetic factors and implications for antiretroviral therapeutics. *Ann Int Med* **2001**; 134:978–96.
 29. McDermott DH, Beecroft MJ, Kleeberger CA, et al. Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the multicenter AIDS cohort study. *AIDS* **2000**; 14:2671–8.
 30. Alkhatib G, Combadaiere C, Broder CC, et al. CC CKR5: a RANTES, MIP-1 α and MIP-1 β receptor as a fusion cofactor for macrophagic-tropic HIV-1. *Science* **1996**; 272:1955–8.
 31. Gong W, Howard OM, Turpin JA, et al. Monocyte chemotactic protein-2 activates CCR5 and blocks CD4/CCR5-mediated HIV-1 entry/replication. *J Biol Chem* **1998**; 273:4289–92.
 32. Ruffing N, Sullivan N, Sharmeen L, Sodroski J, Wu L. CCR5 has an expanded ligand-binding repertoire and is the primary receptor used by MCP-2 on activated T cells. *Cell Immunol* **1998**; 189:160–8.
 33. Sozzani S, Ghezzi S, Iannolo G, et al. Interleukin 10 increases CCR5 expression and HIV infection in human monocytes. *J Exp Med* **1998**; 187:439–44.
 34. Weissman D, Poli G, Fauci AS. IL-10 synergizes with multiple cytokines in enhancing HIV production in cells of monocytic lineage. *J Acquir Immune Defic Syndr Hum Retroviro* **1995**; 9:442–9.
 35. Weissman D, Dybul M, Daucher MB, Davey RT, Walker RE, Kovacs JA. Interleukin-2 up-regulates expression of the human immunodeficiency virus fusion coreceptor CCR5 by CD4⁺ lymphocytes in vivo. *J Infect Dis* **2000**; 181:933–8.
 36. Zou W, Foussat A, Houhou S, et al. Acute upregulation of CCR-5 expression by CD4⁺ T lymphocytes in HIV-infected patients treated with interleukin-2. ANRS 048 IL-2 Study Group. *AIDS* **1999**; 13:455–63.
 37. Alfano M, Poli G. Cytokine and chemokine based control of HIV infection and replication. *Curr Pharm Des* **2001**; 7:993–1013.
 38. Gendelman HE, Orenstein JM, Martin MA, et al. Efficient isolation and propagation of human immunodeficiency virus on recombinant colony-stimulating factor-1-treated monocytes. *J Exp Med* **1988**; 167: 1428–41.
 39. Kutza J, Crim L, Feldman S, et al. Macrophage colony-stimulating factor antagonists inhibit replication of HIV-1 in human macrophages. *J Immunol* **2000**; 164:4955–60.
 40. Biswas P, Poli G, Kinter AL, et al. Interferon γ induces the expression of human immunodeficiency virus in persistently infected promonocytic cells (U1) and redirects the production of virions to intracytoplasmic vacuoles in phorbol myristate acetate-differentiated U1 cells. *J Exp Med* **1992**; 176:739–50.
 41. Losana G, Bovolenta C, Rigamonti L, et al. IFN- γ and IL-12 differentially regulate CC-chemokine secretion and CCR5 expression in human T lymphocytes. *J Leukoc Biol* **2002**; 72:735–42.
 42. Strathdee SA, Galai N, Safaiean M, et al. Sex differences in risk factors for HIV seroconversion among injection drug users. *Arch Intern Med* **2001**; 161:1281–8.
 43. Nelson K, Galai N, Safaiean M, Strathdee SA, Celentano D, Vlahov D. Temporal trends in the incidence of human immunodeficiency virus infection and risk behavior among injection drug users in Baltimore, Maryland, 1988–1998. *Am J Epidemiol* **2002**; 156:641–53.
 44. Vlahov D, Anthony JC, Munoz A, et al. The ALIVE study, a longitudinal study of HIV-1 infection in intravenous drug users: description of methods and characteristics of participants. *NIDA Res Monogr* **1991**; 109:75–100.
 45. Szklo M, Nieto JF. *Epidemiology beyond the basics*. Gaithersburg, MD: Aspen Publication, **2000**.
 46. Breslow N, Day N. *Statistical methods in cancer research: the analysis of case-control studies*. Oxford: Oxford University Press, **1980**.
 47. De La Vega FM, Dailey D, Ziegler J, Williams J, Madden D, Gilbert D. New generation pharmacogenomic tools: a SNP linkage disequilibrium map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. *Biotechniques* **2002**(Suppl):48–50, 52, 54.
 48. Dean M, Stephens JC, Winkler C, et al. Polymorphic admixture typing in human ethnic populations. *Am J Hum Genet* **1994**; 55:788–808.
 49. Abecasis G, Cookson W. GOLD—graphical overview of linkage disequilibrium. *Bioinformatics* **2000**; 16:182–3.
 50. Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **1964**; 49:49–67.
 51. Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **1995**; 29:311–22.
 52. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* **1995**; 12: 921–7.
 53. Schaid D, Rowland C, Tines D, Jacobson R, Poland G. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* **2002**; 70:425–34.
 54. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* **2002**; 53:79–91.
 55. Carlson CS, Eberle MA, Rieder MJ, Yi Y, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* **2004**; 74:106–20.
 56. Kumar S, Tamura K, Jakobsen I, Nei M. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* **2001**; 17:1244–5.
 57. Martin MP, Dean M, Smith MW, et al. Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science* **1998**; 282:1907–11.
 58. Gonzalez E, Bamshad M, Sato N, et al. Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci USA* **1999**; 96:12004–9.
 59. Carrington M, Dean M, Martin MP, O'Brien SJ. Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences. *Hum Mol Genet* **1999**; 8:1939–45.
 60. Smith MW, Dean M, Carrington M, et al. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. *Science* **1997**; 277:959–65.
 61. Winkler C, Modi W, Smith MW, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science* **1998**; 279: 389–93.
 62. Duggal P, An P, Beaty TH, et al. Genetic influence of CXCR6 chemokine receptor alleles on PCP-mediated AIDS progression among African Americans. *Genes Immun* **2003**; 4:245–50.
 63. Darke S. Self-report among injection drug users: a review. *Drug Alcohol Depend* **1998**; 51:253–63.
 64. Goldstein MF, Friedman SR, Neaigus A, Jose B, Ildefonso G, Curtis R. Self-reports of HIV risk behavior by injecting drug-users—are they reliable. *Addiction* **1995**; 90:1097–104.
 65. Petry NM. Reliability of drug users' self-reported HIV risk behaviors using a brief, 11-item scale. *Subst Use Misuse* **2001**; 36:1731–47.
 66. Peters AD, Reid MM, Griffin SG. Edinburgh drug-users: are they injecting and sharing less? *AIDS* **1994**; 8:521–8.
 67. Lalota M, Kwan BW, Waters M, Hernandez LE, Liberti TM. The Miami, Florida, Young Men's Survey: HIV prevalence and risk behaviors among urban young men who have sex with men who have ever runaway. *J Urban Health* **2005**; 82:327–38.
 68. Mummidi S, Ahuja SS, Gonzalez E, et al. Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med* **1998**; 4:786–93.
 69. Kostrikis LG, Huang YX, Moore JP, et al. A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nat Med* **1998**; 4:350–3.
 70. Hendel H, Henon N, Lebuane H, et al. Distinctive effects of CCR5,

- CCR2 and SDF1 genetic polymorphisms in AIDS progression. *J Acquir Immune Defic Syndr Hum Retrovirol* **1998**;19:381–6.
71. Nakayama EE, Tanaka Y, Nagai Y, Iwamoto A, Shioda T. A CCR2-64I polymorphism affects stability of CCR2A isoform. *AIDS* **2004**;18:729–38.
 72. Ping A, Martin MP, Nelson GW, et al. Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *AIDS* **2000**;14:2117–22.
 73. Gonzalez E, Kulkarni H, Bolivar H, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **2005**;307:1434–40.
 74. Zhao JH, Sham PC. Faster haplotype frequency estimation using unrelated subjects. *Hum Hered* **2002**;53:36–41.
 75. Fallin D, Cohen A, Essioux L, et al. Genetic analysis of case/control data using estimated haplotype frequencies: application to APOE locus variation and Alzheimer's disease. *Genome Res* **2001**;11:143–51.
 76. Natarajan V, Lempicki RA, Sereti I, et al. Increased peripheral expansion of naive CD4+ T cells in vivo after IL-2 treatment of patients with HIV infection. *Proc Natl Acad Sci USA* **2002**;99:10712–7.
 77. Paredes R, Lopez B, Fernandez-Cruz E, Clotet B, Lane HC. The potential role of interleukin-2 in patients with HIV infection. *AIDS Rev* **2002**;4:36–40.
 78. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* **1996**;273:1516–7.
 79. The International HapMap Project. *Nature* **2003**;426:789–96.
 80. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* **2005**;6:109–18.
 81. Segal S, Hill A. Genetic susceptibility to infectious disease. *Trends Microbiol* **2003**;11:445–8.
 82. Abel L, Dessein A. Genetic epidemiology of infectious diseases in humans: design of population-based studies. *Emerg Infect Dis* **1998**;4:593–603.
 83. McNicholl J, Cuenco K. Host genes and infectious diseases: HIV, other pathogens, and a public health perspective. *Am J Prev Med* **1999**;16:141–54.
 84. Mellado M, Rodriguez-Frade J, Manes S, Martinez A. Chemokine signaling and functional responses: the role of receptor dimerization and TK pathway activation. *Annu Rev Immunol* **2001**;19:397–421.
 85. Mellado M, Rodriguez-Frade J, Vila-Coro A, de Ana A, Martinez A. Chemokine control of HIV-1 infection. *Nature* **1999**;400:723–4.