Global survey of genetic variation in *CCR5*, *RANTES*, and *MIP-1α*: Impact on the epidemiology of the HIV-1 pandemic

Enrique Gonzalez[†], Rahul Dhanda[‡§], Mike Bamshad[§¶], Srinivas Mummidi[†§], Reni Geevarghese[†], Gabriel Catano[†], Stephanie A. Anderson[‖], Elizabeth A. Walter[**], Kevin T. Stephan[**], Michael F. Hammer[††], Andrea Mangano[‡‡], Luisa Sen[‡‡], Robert A. Clark[†], Seema S. Ahuja[†], Matthew J. Dolan[§§], and Sunil K. Ahuja[†§§]**

[†]South Texas Veterans Health Care System, and University of Texas Health Science Center at San Antonio, San Antonio, TX 78229-3900; [‡]Purdue Pharmaceuticals, Norwalk, CT 06901; [¶]Department of Pediatrics, University of Utah, Salt Lake City, UT 84112; [‖]Henry M. Jackson Foundation, and [**]Infectious Diseases Service, Wilford Hall Medical Center, San Antonio, TX 78236; [††]University of Arizona, Tucson, AZ 85721; and [‡‡]Hospital de Pediatría ''J.P. Garrahan,'' CP1245 Buenos Aires, Argentina

Communicated by Seymour J. Klebanoff, University of Washington School of Medicine, Seattle, WA, February 5, 2001 (received for review December 11, 2000)

Expression of CC chemokine receptor 5 (CCR5), the major coreceptor for HIV-1 cell entry, and its ligands (e.g., RANTES and MIP-1α) is widely regarded as central to the pathogenesis of HIV-1 infection. By surveying nearly 3,000 HIV+ and HIV− individuals from worldwide populations for polymorphisms in the genes encoding RANTES, MIP-1α, and CCR5, we show that the evolutionary histories of human populations have had a significant impact on the distribution of variation in these genes, and that this may be responsible, in part, for the heterogeneous nature of the epidemiology of the HIV-1 pandemic. The varied distribution of *RANTES* haplotypes (AC, GC, and AG) associated with population-specific HIV-1 transmission- and disease-modifying effects is a striking example. Homozygosity for the AC haplotype was associated with an increased risk of acquiring HIV-1 as well as accelerated disease progression in European Americans, but not in African Americans. Yet, the prevalence of the ancestral AC haplotype is high in individuals of African origin, but substantially lower in non-Africans. In a Japanese cohort, AG-containing *RANTES* haplotype pairs were associated with a delay in disease progression; however, we now show that their contribution to HIV-1 pathogenesis and epidemiology in other parts of the world is negligible because the AG haplotype is infrequent in non-Far East Asians. Thus, the varied distribution of *RANTES*, *MIP-1α*, and *CCR5* haplotype pairs and their population-specific phenotypic effects on HIV-1 susceptibility and disease progression results in a complex pattern of biological determinants of HIV-1 epidemiology. These findings have important implications for the design, assessment, and implementation of effective HIV-1 intervention and prevention strategies.

Since the beginning of the HIV-1 pandemic, it has been apparent that the epidemiology of HIV-1 varies throughout the world (1). This variation is due, in part, to differences in the biological (e.g., level of viremia), behavioral (e.g., rate of partner change), demographic (e.g., proportion of sexually active individuals), and economic/political (e.g., health care system, war) determinants that affect the spread of HIV-1 throughout a population (1, 2).

Biological determinants that influence the probability of HIV-1 transmission include the infectiousness of a sexual partner, the characteristics of the infecting viral strain, and the susceptibility of the uninfected partner (1, 2). Within a population, the course of an epidemic such as HIV also depends, in part, on the rate of contact between susceptible and infectious individuals. Thus, the net rate of spread of infection is proportional to the product of the density of susceptible people times the density of infectious individuals, a concept known as the ''mass action principle'' (3, 4). The epidemiology of HIV-1 in a population is therefore dependent, in part, on factors (e.g., behavioral, social) that influence the rate of contact between infectious and susceptible individuals as well as determinants that influence an individual's infectiousness and susceptibility. Solid epidemiological evidence supports the notion that individuals are not equally susceptible to HIV-1 infection (5–7) and that polymorphisms in host genes that facilitate viral cell entry and modulate immune responses play critical roles in influencing an individual's infectiousness and susceptibility to HIV-1 (refs. 8–19; reviewed in refs. 20 and 21).

We (17, 18, 22) and others (8–14, 16, 20, 21) have found that single polymorphisms in *CCR5* and *CCR2*, different combinations of polymorphisms in linkage disequilibrium (i.e., haplotypes), and pairs of haplotypes influence the rates of disease progression in infected children and adults, and the risk of mother-to-child transmission of HIV-1 (17–19). But polymorphisms in *CCR5* are only part of the overall picture. Recently, Gallo and colleagues have argued persuasively that the levels of the ligands of CCR5 (i.e., MIP-1α, MIP-1β, and RANTES) are consistent and reproducible immunological parameters associated with disease progression and HIV-1 transmission (23). These chemokines exhibit anti-HIV-1 properties *in vitro*, and *in vivo* evidence supports strongly the notion that their expression levels may influence susceptibility to HIV-1 infection (refs. 24 and 25; reviewed in ref. 23). Recent HIV vaccine studies also suggest that production of these chemokines are a true correlate of protection (reviewed in ref. 23).

If variation in host genes contributes to interindividual heterogeneity in risk of HIV susceptibility and differential host immune responses, then the challenge is to step back from the issue of single genes/polymorphisms and ask a more difficult and complex question, ''What is the relationship, if any, between the genetic variables that determine HIV transmission and course of infection within an individual and the genetic variables that influence the pattern of infection within communities of people?'' This is a fundamental question to address because public health policy, by its very nature, deals with the relationship between infection in *individuals* and infection in *populations*.

In recent years, a growing amount of work has been aimed at understanding the influence of social, behavioral, and other heterogeneities on the epidemiological characteristics of host populations, and on the design and evaluation of HIV-1 immunization programs (1, 26). However, the potential impact of genetic heterogeneity within and among populations on these issues has been understudied. Indeed, in most mathematical models and trial simulations used for planning phase III HIV

Abbreviations: CCR5, CC chemokine receptor 5; *AF*, attributable fraction.

[§]R.D., M.B., and S.M. contributed equally to this work.

[§§]To whom reprint requests should be addressed. E-mail: ahujas@uthscsa.edu or matthew.dolan@59mdw.whmc.af.mil.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

vaccine efficacy trials, the impact of varying genetic susceptibility among populations remains an unmeasured factor (1, 26). This is understandable because we do not yet have a clear picture of the distribution of HIV susceptibility or resistance polymorphisms in worldwide populations.

In this study, we tested the hypothesis that the haplotypes of *CCR5* and its ligands that are associated with differences in interindividual susceptibility to HIV-1 infection and disease progression can lead to population-level differences in the biological correlates of HIV-1 transmission and clinical outcome. To this end, we surveyed world-wide populations for polymorphisms in *RANTES* (n = 2,508), *MIP-1α* (n = 1,784), and *CCR5* (n = 3,043), and determined the association between haplotype variation and HIV-1 susceptibility.

Methods

Cohorts. Control unlinked European-, African-, and Hispanic-American blood donors as well as blood donors from worldwide populations were genotyped, and their origins are shown in Table 3, which is published as supplemental data on the PNAS web site, www.pnas.org. A total of 1,120 HIV-seropositive patients were evaluated, including 507 seroconverting individuals. The demographic background of this cohort is 55% European American (EA), 36% African American (AA), 6% Hispanic American (HA), and 3% ''other.'' Additional epidemiologic features of the HIV+ cohort studied are in the supplemental data and as described (17, 18).

Polymorphisms and Genotyping. The SNPs in *MIP-1α* (+113 and +459) and *RANTES* (−96 and −471) were identified by comparing sequences available in GenBank and by bulk sequencing as described in Fig. 5, which is published as supplemental data. The methods for genotyping are in Fig. 5. While this work was in progress the identical *RANTES* SNPs were reported by Liu *et al.* (27). Positions −96 and −471 were designated as positions −28 and −403, respectively by Liu *et al.* (27) because their numbering system considers +1 as the first nucleotide of the RANTES mRNA; we have used the adenine (A) in ATG, the initiator Met codon as +1. The nomenclature of *CCR5* haplotypes is as described (18, 22) and is also summarized in Fig. 6, which is published as supplemental data. *CCR5* polymorphisms were genotyped as described (18, 22). The ancestral state for each polymorphism in the human genes was inferred by sequencing and genotyping the chimpanzee gene for the homologous position.

Statistical Analysis. The association between possession of a haplotype and risk of acquiring HIV was evaluated as described (19) using Fisher's exact test or χ^2 and multivariable logistic models (SAS, version 6.12; SAS Institute, Cary, NC). Time curves for progression to AIDS (1987 criterion) and survival were prepared by the Kaplan–Meier method as described (17, 18). Only individuals who had a minimum follow-up time of at least 6 months were included in the analysis (n = 1,120). Between-group analyses were completed by using the log-rank test.

Results

Global Survey of the *RANTES* Locus and Its Impact on the Epidemiology of HIV-1. The distribution of genetic variation at *MIP-1α* and *RANTES* differs substantially within and between populations (Fig. 1 and data not shown). Sequencing of *RANTES* in chimpanzees confirmed that the ancestral *RANTES* haplotype (AC) is defined by the presence of nucleotides A and C at positions −471 and −96, respectively (Fig. 1*a*). In addition, by PCR–restriction fragment length polymorphism (RFLP), the genotypes in 60 unrelated chimpanzees were *RANTES* −471A/A and −96C/C. The frequency of these SNPs in U.S.-based control blood donors and HIV-infected subjects is shown in Table 4, which is published as supplemental data. None of these frequencies deviated from Hardy–Weinberg equilibrium in European, African, or Hispanic Americans.

The *RANTES* −471G and −96G were in complete linkage disequilibrium with *RANTES* −96C and −471A, respectively (see Table 5, which is published as supplemental data). Along with sequencing data, these SNPs defined three RANTES haplotypes: AC, GC, and AG (Fig. 1*b* and data not shown). Among worldwide populations, the frequency of the ancestral haplotype was highest in individuals of African origin, intermediate in Asians, and lowest in those of European origin (Fig. 1*b*). The frequency of the AC haplotype is especially high in some isolated aboriginal human populations, such as African pygmies (Fig. 1*b*) and Nicobarese islanders (data not shown).

The C→G mutation at −96 is found almost exclusively in Asians, and is not found in any native Africans (Fig. 1*b*). Consequently, the *RANTES* AG haplotype is much more frequent in East Asians and in the Pacific Rim, specifically in Japan, than any other population (Fig. 1*b*). In contrast, the *RANTES* GC haplotype is most prevalent in Europeans and less frequent in Asians and Africans (Fig. 1*b*). Because of these differences in haplotype frequencies, the proportion of AG vs. GC-containing haplotype pairs in East Asians vs. non-Asian populations is significantly different (Fig. 1*c*). Liu *et al.* showed that HIV-positive Japanese homozygous or heterozygous for the *RANTES* AG-haplotype had a slower disease course compared with those who lacked this haplotype (27). However, the extremely low frequency of the AG haplotype in other non-Far East Asian populations (Fig. 1*b*), suggests that the contribution of this haplotype to the epidemiology of HIV-1 infection in these populations is small.

HIV-Acquisition and Disease-Modifying Effects of *RANTES* Haplotypes. In HIV-positive European Americans, lack of a *RANTES* GC haplotype (the most common *RANTES* haplotype in worldwide populations) was associated with a rapid progression to AIDS and death (Fig. 2 *a* and *b*). Because the AG haplotype is restricted to individuals of Far East Asia, homozygosity for the ancestral *RANTES* haplotype (i.e., AC/AC) is the only non-GC-containing haplotype pair in European Americans (Fig. 1*c*). In HIV-positive European Americans, possession of the ancestral *RANTES* haplotype pair AC/AC was associated with a worse clinical outcome (Fig. 2 *c* and *d*).

Because possession of AC/AC was associated with a worse clinical outcome, we next determined whether possession of this haplotype pair in European Americans was also associated with an increased risk of disease acquisition. European Americans lacking the ancestral AC haplotype had a lower risk of acquiring HIV-1 {P = 0.05, OR = 0.78 [confidence interval (CI) = 0.61–1.00]}, and compared with the ancestral haplotype pair AC/AC the haplotype pairs AC/GC, GC/GC, and GC/AG were associated with a lower risk of acquiring HIV-1 (Table 1). Taken together, the findings shown in Fig. 2 and Table 1 indicate that possession of AC/AC is associated with an increased susceptibility to acquiring HIV-1 and disease progression. A disease- or transmission-modifying effect of the *RANTES* haplotype pairs was not found in seropositive African Americans.

HIV-Acquisition and Disease-Modifying Effects of *MIP-1α* Haplotypes. Population-specific HIV-1 transmission/disease-modifying effects were also observed for *MIP-1α* haplotypes. By PCR-RFLP and sequencing, the ancestral *MIP-1α* haplotype (CC) is defined by the presence of nucleotide C at positions +113 and +459 in chimpanzees (Fig. 1*d*). *MIP-1α* +113T was in complete linkage with *MIP-1α* +459T (see Table 6, which is published as supplemental data). This linkage pattern along with sequencing data defined three haplotypes: CC, CT, and TT (Fig. 1*e*). The frequency of these SNPs in control U.S.-based blood donors and HIV-infected subjects is shown in Table 7, which is published as supplemental data, and in European, African, and Hispanic Americans none of these frequencies deviated from Hardy–Weinberg equilibrium.

The ancestral *MIP-1α* haplotype was the most common haplo-

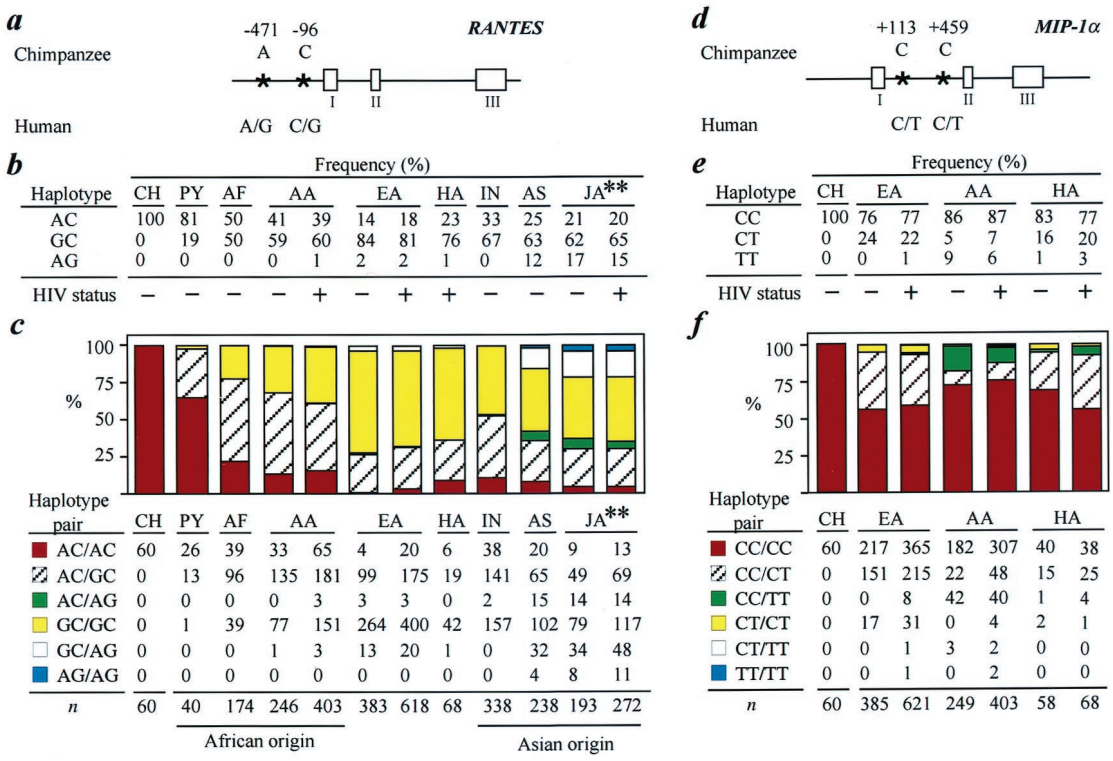

Fig. 1. (*a–c*) *RANTES* and (*d–f*) *MIP-1α* SNPs, haplotypes, and haplotype pairs. (*a* and *d*) Schematic illustration of the genomic organization and SNPs (*) in human *RANTES* and *MIP-1α*. Open boxes represent exons. The sequence at the position corresponding to the human SNPs in chimpanzees (ancestral state) is also shown. The designation of each haplotype is based on the 5′ > 3′ arrangement of the SNPs. Thus, for example, the haplotype that contains *RANTES* −471A and −96C is designated as the *RANTES* AC haplotype. (*b* and *e*) Frequencies of the *RANTES* and *MIP-1α* haplotypes in chimpanzee (CH) and control blood donors (−) and infected (+) worldwide populations, including European (EA), African (AA), and Hispanic Americans (HA). PY, pygmy; AF, non-Pygmy Africans; IN, Asian Indians; AS, non-Indian Asians; JA, Japanese. The ancestral haplotype for each gene is listed first in panels *b* and *e*. (*c* and *f*): (*Upper*) The frequency (%) of the *RANTES* and *MIP-1α* haplotype pairs in the different populations. (*Lower*) The total number of individuals with a given haplotype pair is shown. The color codes shown adjacent to each haplotype pair listed in the *Lower* correspond to the colored bars shown in the *Upper*. The ancestral haplotype pair for each gene is in red. **, data listed under the column heading designated as JA is derived from the study by Liu *et al.* (27), and shows the frequency of the *RANTES* haplotype (*b*) and haplotype pairs (*c*) in Japanese HIV− and HIV+ individuals; the data for the HIV+ and HIV− Japanese hemophiliacs and nonhemophiliacs was combined. The codes shown for the HIV status in *b* and *e* are also applicable for the data shown in *c* and *f*, respectively.

type in the three U.S.-based populations examined (Fig. 1*e*). However, the frequency of the other two *MIP-1α* haplotypes did vary substantially among populations. For example, the *MIP-1α* haplotype, TT, is three times more common in African Americans compared with Hispanics and is observed in less than 1% of all Caucasians (Fig. 1*e*). Thus, the *MIP-1α* haplotype pair, CC/TT is much more common in African Americans than any other group tested.

In European Americans, the time to develop AIDS was shorter in individuals who lacked the ancestral *MIP-1α* CC haplotype (Fig. 3 *a* and *b*). The majority of the European Americans who lacked the *MIP-1α* CC haplotype had the CT/CT haplotype pair (Fig. 1*f*). This suggested that the disease-accelerating effect observed for individuals lacking the CC haplotype was due primarily to the CT/CT haplotype pair. The disease-accelerating effect associated with CT/CT in European Americans was confirmed by a direct comparison of the three most prevalent *MIP-1α* haplotype pairs (CC/CC, CC/CT, and CT/CT; Fig. 1*f*). There was no difference in the disease course of individuals with the CC-containing haplotype pairs (CC/CC or CC/CT), but homozygosity for CT/CT was associated with a more rapid disease course (Fig. 3 *c* and *d*). There were too few European Americans with CC/TT or the other haplotype pairs for evaluation. A transmission-modifying effect for the *MIP-1α* haplotype pairs was not found in European Americans.

No disease-modifying effect was discovered for any of the *MIP-1α* haplotypes in African Americans. In African Americans, an overall significant effect for HIV-1 acquisition was observed for the proportion of individuals who possessed at least one CC, CT, or TT *MIP-1α* haplotype among HIV+ and control groups ($P = 0.027$). Possession of the TT *MIP-1α* haplotype was associated with a significantly lower risk of HIV-1 acquisition compared with the ancestral haplotype CC [$P = 0.029$, OR = 0.623 (CI = 0.406–0.956)]. This prompted us to determine whether there was in African Americans an overall difference in the proportion of individuals who possessed different *MIP-1α* haplotype pairs. The overall difference was significant ($P = 0.027$), and consistent with the finding that the *MIP-1α* TT haplotype was associated with a lower risk of acquiring HIV-1 in African Americans, possession of the haplotype pair CC/TT was associated with a lower risk of acquiring HIV compared with the ancestral haplotype pair CC/CC (Table 1). Notably, the distribution of CC/TT is restricted to African Americans (Fig. 1*f*).

Global Distribution of CCR5 Haplotypes and Its Impact on the Epidemiology of HIV-1. To develop a better understanding of the ramifications of the genetic diversity of *CCR5* on the HIV-1 epidemiological characteristics of different populations, we conducted a significantly more detailed analysis of the distribution of *CCR5* haplotypes in worldwide populations than we reported previously (ref. 18; Table 2, Fig. 4, and Table 8, which is published as supplemental data). This analysis further illuminated the striking divergence in the distribution of *CCR5* haplotypes within and between different populations (Table 2). For

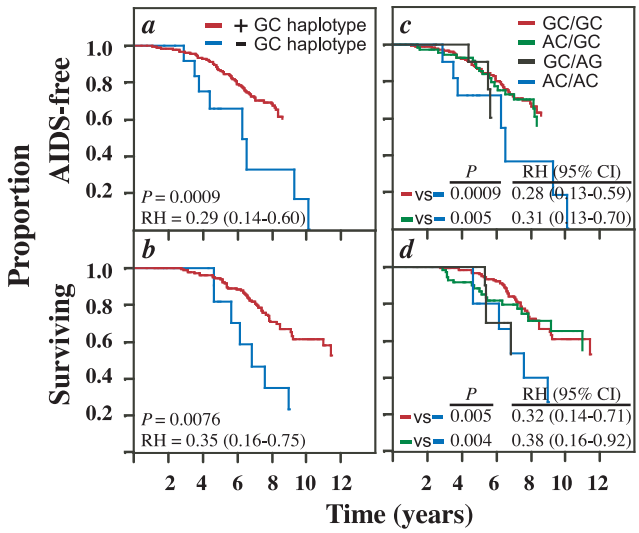

Fig. 2. Disease-modifying effects of *RANTES* haplotypes in seroconverting European Americans. (*a* and *b*) The Kaplan–Meier (KM) curves of the development of AIDS or death in the seroconverting European American portion of the cohort that lack (blue curve) or possess (red curve) the most common *RANTES* GC haplotype. Notably, most individuals who lack the GC haplotype have the haplotype pair AC/AC (Fig. 1*c*), and for this reason the pattern of the blue KM curves in *a* and *b* is very similar to those of the blue KM curve for AC/AC in *c* and *d*. *P* and RH [95% CI (confidence limits)] indicate the significance value by log-rank test and the relative hazard with respect to the reference group, respectively. (*c* and *d*) KM curves comparing the clinical course of seroconverting European Americans with the following haplotype pairs: GC/GC (red); AC/GC (green); GC/AG (black line); and AC/AC (blue). The reference group for the survival analyses is individuals that are homozygous for the ancestral *RANTES* haplotype AC.

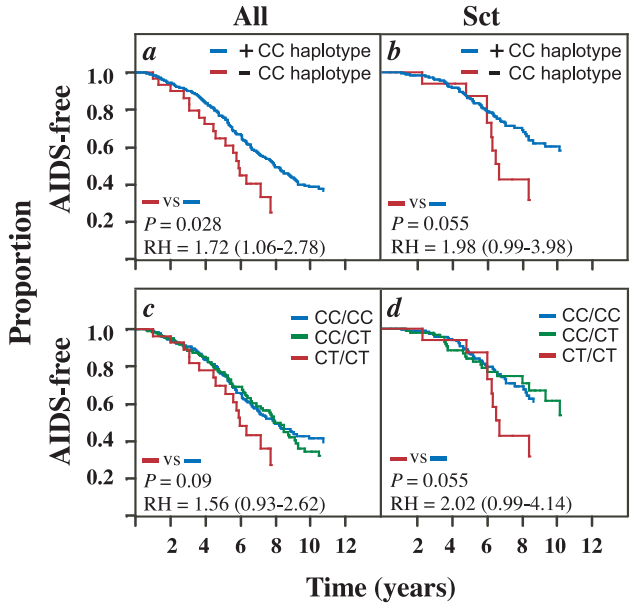

Fig. 3. Disease-modifying effects of *MIP-1α* haplotypes in the European American portion of the cohort. (*a* and *b*) The KM curves of the development of AIDS in European Americans that possess (blue) or lack (red) the ancestral *MIP-1α* haplotype (CC). All, the entire portion of the cohort and Sct, indicates seroconverting individuals. The reference group are individuals with the ancestral (CC) haplotype. (*c* and *d*) KM curves comparing the clinical course of European Americans with the following haplotype pairs: CC/CC (blue); CC/CT (green); and CT/CT (red). The reference group for the survival analyses is individuals that are homozygous for the ancestral *MIP-1α* haplotype CC (blue).

example, among Asians there are significant differences between Indian and non-Indian populations with respect to the distribution of the ancestral *CCR5* haplotype HHA or the CCR2-64I-containing haplotype, HHF*2 (Table 2). Indeed, in Indians the prevalence of these two haplotypes approximates that found in African populations. We previously reported that the distribution of *CCR5* HHD haplotypes was specific to individuals of African origin (18), and in the current analysis we identified that HHB is another *CCR5* haplotype that is specific to individuals of African origin. Notably, the prevalence of HHB is higher in Africans than African Americans. Similarly, the prevalence of HHF*1, the genetic context in which the CCR2-64I mutation arose (22), is highest in African populations.

This varied distribution of *CCR5* haplotypes results in an uneven distribution of *CCR5* haplotype pairs in global populations (Fig. 4). For example, 75% of the *CCR5* haplotype pairs in Europeans or Asian Indians can be accounted by only eight or six *CCR5* haplotype pairs, respectively, and six of these haplotype pairs are common to both populations. In contrast, in individuals of African descent (e.g., in African Americans), the eight most common *CCR5* haplotype pairs account for less than 50% of the haplotype pairs and, of these, only three are among the most common haplotype pairs found in Europeans or Asian Indians (Fig. 4).

Table 1. Risk of acquiring HIV-1 infection associated with *RANTES* and *MIP-1α* haplotype pairs (HP)

		European American			African American		
Gene	HP	OR	CI	*P*	OR	CI	*P*
RANTES	AC/AC†	1.0					
	AC/GC	0.35	0.12–1.06	0.064			
	GC/GC	0.30	0.10–0.90	0.031			
	GC/AG	0.31	0.01–1.11	0.071			
MIP-1α	CC/CC†				1.0		
	CC/TT				0.57	0.35–0.90	0.017

OR, odds ratio; CI, 95% confidence interval limit.
†Reference group.

Discussion

Within the global pandemic of HIV-1 infection there are many different epidemics, each with its own dynamics and each influenced by many factors, including time of introduction of the virus, population density, and cultural and social forces (1, 2). Although significant attention has been placed on understanding the influence of these factors in the spread and maintenance of HIV-1 in a given population, the contribution of host genetic determinants (i.e., polymorphisms) to the heterogeneous nature of the epidemiology of HIV-1 within and among populations has not been fully addressed. It is in this context that we will discuss the two major findings of this study: First, genetic variation in *MIP-1α* and *RANTES* is associated with interindividual differences in the risks for acquiring HIV-1 and rates of disease progression. Second, there are significant interpopulation differences in the distribution of *CCR5*, *MIP-1α*, and *RANTES* haplotypes that are associated with transmission/disease-modifying effects.

Relative risk/hazard (or odds ratios) are used to indicate the strength of an association between a risk factor (e.g., genetic or environmental) and the presence or absence of disease. However, the attributable fraction (28, 29) is a more useful measure than the relative risk (or odds ratio) in quantifying the role of a specific risk factor in disease etiology and the overall public health impact of this factor. The public health relevance of this measure lies in estimating the proportion of cases of disease in a population that would not have occurred had the risk factor

Table 2. *CCR5* haplotype frequencies (%) in worldwide populations

Population	*N*	A	B	C	D	E	F*1	F*2	G*1	G*2
Africa	143									
Pygmy	36	71	3	1	0	13	6	6	1	0
Non-Pygmy	107	24	5	9	17	20	8	16	2	0
African American	646	21	1	15	20	19	4	15	4	3
European Origin	959	10	0†	35	1	32	1	8	4	8
Hispanic American	86	9	0	36	3	28	2	15	3	5
Argentinean	751	7	0†	35	1	31	4	15	5	3
Asia	458									
India	224	18	0†	36	1	29	2	18	1	0†
Non-Indian	234	5	0	42	0	25	3	5	0†	1

N, number of individuals; A–G*2 refers to *CCR5* human haplogroups (HH) HHA–HHG*2 (22).
†Frequency > 0, but $< 0.5\%$.

been absent or if its frequency varied within and between populations. By using the frequency of the risk factor in a population, the attributable fraction (*AF*) can be computed as (29, 30): $AF = f(R - 1) \div [1 + f(R - 1)]$, where f is the frequency of the risk factor in the population and R is the measure of relative risk/hazard (or odds ratios).

Thus, in host genotype-HIV-1 transmission/disease phenotype epidemiological studies, the *AF* of a given haplotype pair is related to its prevalence within a population as well as the magnitude of its influence on transmission or disease progression expressed as odds ratios or relative risk/hazard. In this context, the striking population-level differences in the distribution of transmission- and disease-modifying *RANTES* or *MIP-1α* haplotypes/haplotype pairs indicates that their attributable risks are not uniformly distributed among different populations. For example, the disease-modifying effects associated with the *RANTES* AG haplotype are confined to populations of the Far East, such as the Japanese (Fig. 1; ref. 27). Similarly, the transmission- and disease-modifying effects associated with the *MIP-1α* CC/TT and CT/CT haplotype pairs are restricted to African and European Americans, respectively. By using the transmission-modifying effects associated with the *MIP-1α* CC/TT haplotype pair as an example, and assuming that the associated effects (odds ratios) are similar in African and European Americans, one can estimate the *AF* of this haplotype pair (f = haplotype pair frequency) in preventing transmission in these two populations. The *AF* for the *MIP-1α* CC/TT haplotype pair in providing protection to African Americans is about 7.8%, whereas in European Americans it is $<0.1\%$, a 71-fold difference.

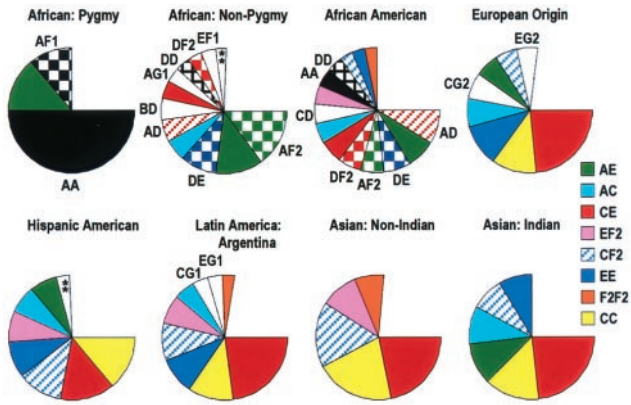

Fig. 4. Pie charts depicting the contrasting distribution of *CCR5* haplotype pairs in African Pygmies, non-Pygmy Aficans, African Americans (HIV+ and HIV−), Europeans (HIV+ and HIV−; combined data from European Americans and non-American Europeans), Hispanic Americans, Argentineans (HIV+ and HIV− children exposed perinatally to HIV; ref. 19), non-Indian Asians, and Asian Indians. The composition of each of the populations shown is in Table 3. The pie charts depict the frequencies of the most prevalent haplotype pairs that together account for 75% of the population studied. Thus, for example, the frequencies of six haplotype pairs in Asian Indians (C/E, C/C, A/E, A/C, C/F*2, and E/E) together account for 75% of all *CCR5* haplotype pairs in this population. The color codes for the haplotype pairs common to individuals of both non-African and African origin are shown to the right of the figure. A pie-slice that is not colored denotes a haplotype pair that is unique to the population indicated with respect to its prevalence among those haplotype pairs that together account for 75% of the population indicated. For example, the haplotype pair HHE/HHG*2 is one of the eight haplotype pairs that together comprise 75% of the European population studied, but it is not among the common haplotype pairs found in other populations. The detailed analysis of the haplotype pairs found in each of these populations is shown in Table 8, which is published as supplemental data. **, denotes that several haplotype pairs had the identical prevalence in the population studied; their identities are not listed, but can be derived from Table 8.

The highly varied distribution of *CCR5* haplotypes/haplotype pairs among populations (Table 2) further highlights that the *AF* of these genetic factors will vary significantly from population to population. Assuming that homozygosity for HHG*2, (i.e., *CCR5*-Δ*32*-containing haplotype) is associated with complete protection, and that its prevalence is ≈1% in European populations, its *AF* in providing protection to this population is 1%, whereas its extremely low prevalence in African populations precludes precise estimates of the *AF*. Conversely, because HHD is African-specific (18), the *AF* of HHD/HHD in the occurrence of HIV-1 mother-to-child transmission is restricted to this population (14). Similarly, although HHE/HHE is associated with increased risk for mother-to-child transmission, and an accelerated disease course in infected European Americans and Argentinean children (Tables 9 and 10, which are published as supplemental data; refs. 18 and 19), the prevalence of HHE/HHE is significantly higher in non-African populations. Thus, even if one assumed that the phenotypic effects associated with HHE/HHE are equal among all populations, its *AF* is not.

Collectively, these observations and the findings of our previous work and those of other investigators permit two major conclusions. First, an individual's HIV-1 susceptibility and clinical outcome is likely to be determined by the collective effects (and possibly interactions) of at least several loci. However, the effects of haplotype pairs at several different loci can converge to produce a similar phenotype. For example, in European Americans each of the haplotype pairs *CCR5* HHE/HHE, *RANTES* AC/AC, and *MIP-1α* CT/CT are associated with an accelerated disease course. Conversely, haplotype pairs of the same gene may produce different phenotypes in distinct populations. For example, the haplotype pairs *RANTES* AC/AC or *MIP-1α* CT/CT influence clinical outcome in HIV-seropositive European Americans, but not African Americans.

Second, varied distribution of transmission- and disease-modifying CCR5-ligand haplotypes in populations and their

population-specific associations suggests that variation in these genes serves as an important biological determinant of interindividual risks of susceptibility, and the epidemiology of HIV-1 infection. These population-level differences could lead to the heterogeneous manner in which populations respond and adapt to the spread of HIV-1.

These findings also have important implications for studies designed to test vaccines or therapeutic agents. For example, although countries in sub-Saharan Africa and Asia (e.g., Thailand) may well prove to be suitable populations for vaccine trials, the distribution of genetic variation in *CCR5* and *RANTES* in these and other populations is extremely heterogeneous. Thus, the efficacy of a single vaccine candidate used in a trial against a predominant viral strain may differ among populations, in part because of the varied distributions of transmission- and/or disease-modifying haplotypes of *CCR5* and its ligands. Accounting for these interpopulation genetic differences either by stratifying study populations by their genetic risk or controlling for this variation in the statistical analysis of the efficacy data may be necessary. Failure to consider these effects may potentially undermine the validity of a result in a specific ethnic group by either masking a positive result or overlooking a negative result. For example, a vaccine could appear to be effective in a population because it has a high prevalence of protective haplotype pairs compared with the control population, thus attributing efficacy to an intervention that actually represents a preponderance of a protective haplotype pair in one study arm.

This study also has important implications for genotype–phenotype studies. Because genetic variation at susceptibility loci can be unevenly distributed among populations, it is emphasized that when comparing the associated phenotypic effects in individuals who possess a given genotype versus those that lack it, this latter control group in statistical analysis is not identical among different populations. For example, HHA/HHF*2 (ancestral haplotype paired with the CCR2-64I-containing haplotype) is the most common haplotype pair found in Africans (Fig. 4), and in our cohort is associated with disease retardation in African Americans (18). In contrast, HHA/HHF*2 is rare in European Americans (Fig. 4 and Table 8), and neither HHA nor HHF*2 were associated with phenotypic effects in the HIV-seropositive European American population that we studied (17, 18). Thus, the results of genotype–phenotype studies are likely to be highly dependent on the overall genetic context of the population studied.

It is possible that, similar to the evolutionary pressure of malaria on the human genome (30), HIV-1 might also drive the evolution of their hosts, especially in view of the fact that HIV-1 is infecting a relatively naïve host. Thus, as we watch the consequences of the HIV-1 epidemic unfold, this infection may shape the genetic architecture of the most heavily afflicted populations (e.g., selection of the phenotype reflected by HIV− children born to HIV+ mothers).

From a public health perspective, the insights from our studies highlight the importance of considering the genetic heterogeneity within and among populations in the design and evaluation of HIV-1 intervention and prevention strategies. The importance of considering this genetic heterogeneity in vaccine trials is illustrated further by evidence suggesting an important role for chemokine production in response to HIV vaccines (reviewed in ref. 23), and the general concept that genetic variation can influence vaccine responsiveness (31).

Note Added in Proof. While this paper was in review, McDermott *et al.* also showed that *RANTES* haplotypes can influence HIV transmission and rate of disease progression in a U.S.-based cohort (32).

We thank V. Telles, J. Barnes, and C. Hensler for superb technical assistance; A. S. Ahuja for forbearance; Drs. D. W. Bowden, B. I. Freedman, and K. Murthy for DNA samples; and two anonymous reviewers for their excellent suggestions. Because of space constraints, we deeply regret our inability to refer to additional papers that have also examined the HIV transmission/disease-modifying effects of coreceptor/chemokine polymorphisms. The Henry M. Jackson Foundation and the Military HIV Program at the Walter Reed Army Institute of Research contributed support for the Wilford Hall Medical Center patient cohort as part of the Tri-Service HIV Program. This work was also supported in part by a Veterans Affairs Hospitals Merit Award and National Institutes of Health grants (AI43279 and AI46326) to S.K.A., and by a National Institutes of Health minority supplement grant (AI43279) to E.G. S.K.A. is an Elizabeth Glaser Scientist.

1. Anderson, R. M., May, R. M., Boily, M. C., Garnett, G. P. & Rowley, J. T. (1991) *Nature (London)* **352,** 581–589.
2. Carael, M., Buve, A. & Awusabo-Asare, K. (1997) *AIDS* **11,** S23–S31.
3. Dietz, K. & Hadeler, K. P. (1988) *J. Math. Biol.* **26,** 1–25.
4. Hamer, W. H. (1906) *Lancet* **i,** 733–739.
5. Plummer, F. A., Ball, T. B., Kimani, J. & Fowke, K. R. (1999) *Immunol. Lett.* **66,** 27–34.
6. Liu, S. L., Schacker, T., Musey, L., Shriner, D., McElrath, M. J., Corey, L. & Mullins, J. I. (1997) *J. Virol.* **71,** 4284–4295.
7. Operskalski, E. A., Busch, M. P., Mosley, J. W. & Stram, D. O. (1997) *J. Acquired Immune Defic. Syndr.* **15,** 145–150.
8. Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A. & Landau, N. R. (1996) *Cell* **86,** 367–377.
9. Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., Goedert, J. J., Buchbinder, S. P., Vittinghoff, E., Gomperts, E., *et al.* (1996) *Science* **273,** 1856–1862.
10. Samson, M., Libert, F., Doranz, B. J., Rucker, J., Liesnard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C., *et al.* (1996) *Nature (London)* **382,** 722–725.
11. Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., Goedert, J. J., O'Brien, T. R., Jacobson, L. P., Kaslow, R., *et al.* (1997) *Science* **277,** 959–965.
12. Michael, N. L., Louie, L. G., Rohrbaugh, A. L., Schultz, K. A., Dayhoff, D. E., Wang, C. E. & Sheppard, H. W. (1997) *Nat. Med.* **3,** 1160–1162.
13. Martin, M. P., Dean, M., Smith, M. W., Winkler, C., Gerrard, B., Michael, N. L., Lee, B., Doms, R. W., Margolick, J., Buchbinder, S., *et al.* (1998) *Science* **282,** 1907–1911.
14. Kostrikis, L. G., Neumann, A. U., Thomson, B., Korber, B. T., McHardy, P., Karanicolas, R., Deutsch, L., Huang, Y., Lew, J. F., McIntosh, K., *et al.* (1999) *J. Virol.* **73,** 10264–10271.
15. Kostrikis, L. G., Huang, Y., Moore, J. P., Wolinsky, S. M., Zhang, L., Guo, Y., Deutsch, L., Phair, J., Neumann, A. U. & Ho, D. D. (1998) *Nat. Med.* **4,** 350–353.
16. McDermott, D. H., Zimmerman, P. A., Guignard, F., Kleeberger, C. A., Leitman, S. F. & Murphy, P. M. (1998) *Lancet* **352,** 866–870.
17. Mummidi, S., Ahuja, S. S., Gonzalez, E., Anderson, S. A., Santiago, E. N., Stephan, K. T., Craig, F. E., O'Connell, P., Tryon, V., Clark, R. A., *et al.* (1998) *Nat. Med.* **4,** 786–793.
18. Gonzalez, E., Bamshad, S., Sato, N., Mummidi, S., Dhanda, R., Catano, G., Cabrera, S., McBride, M., Cao, X.-H., Merrill, G., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96,** 12004–12009.
19. Mangano, A., Gonzalez, E., Dhanda, R., Catano, G., Bamshad, M., Bock, A., Duggirala, R., Williams, K., Mummidi, S., Clark, R. A., *et al.* (2001) *J. Infect. Dis.* **184,** in press.
20. Kaslow, R. A. & McNicholl, J. M. (1999) *Proc. Assoc. Am. Physicians* **111,** 299–307.
21. O'Brien, S. J. & Moore, J. P. (2000) *Immunol. Rev.* **177,** 99–111.
22. Mummidi, S., Bamshad, M., Ahuja, S. S., Gonzalez, E., Feuillet, P. M., Begum, K., Galvis, M. C., Kostecki, V., Valente, A. J., Murthy, K. K., *et al.* (2000) *J. Biol. Chem.* **275,** 18946–18961.
23. Gallo, R. C., Garzino-Demo, A. & DeVico, A. L. (1999) *J. Clin. Immunol.* **19,** 293–299.
24. Garzino-Demo, A., Moss, R. B., Margolick, J. B., Cleghorn, F., Sill, A., Blattner, W. A., Cocchi, F., Carlo, D. J., DeVico, A. L. & Gallo, R. C. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11986–11991.
25. Zagury, D., Lachgar, A., Chams, V., Fall, L. S., Bernard, J., Zagury, J. F., Bizzini, B., Gringeri, A., Santagostino, E., Rappaport, J., *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95,** 3857–3861.
26. Boily, M. C., Masse, B. R., Desai, K., Alary, M. & Anderson, R. M. (1999) *Vaccine* **17,** 989–1004.
27. Liu, H., Chao, D., Nakayama, E. E., Taguch, H., Gotto, M., Xin, X., Takamatsu, J. K. Saito, H., Ishikawa, Y., Akaza, T., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96,** 4581–4585.
28. Miettinen, O. S. (1974) *Am. J. Epidemiol.* **99,** 325–332.
29. Adams, M. J., Jr., Khoury, M. J. & James, L. M. (1989) *J. Clin. Epidemiol.* **42,** 659–662.
30. Clegg, J. B. & Weatherall, D. J. (1999) *Proc. Assoc. Am. Physicians* **111,** 278–282.
31. Poland, G. A. (1998) *Am. J. Hum. Genet.* **62,** 215–220.
32. McDermott, D. H., Beecroft, M. J., Kleeberger, C. A., Al-Sharif, F. M., Ollier, W. E., Zimmerman, P. A., Boatin, B. A., Leitman, S. F., Detels, R., Hajeer, A. H., *et al.* (2000) *AIDS* **14,** 2671–2678.