Distribution of HIV-1 resistance-conferring polymorphic alleles *SDF-1-3¢A, CCR2-64I* and *CCR5-D32* in diverse populations of Andhra Pradesh, South India

G. V. RAMANA[1], A. VASANTHI[2], M. KHAJA[2], B. SU[1,3], V. GOVINDAIAH[4], L. JIN[1,3], L. SINGH[5] and R. CHAKRABORTY[1,3]*

[1]*Human Genetics Center, School of Public Health, University of Texas, P.O. Box 20186, Houston, TX 77225, USA*
[2]*Department of Genetics, Owaisi Medical and Research Centre, Deccan College of Medical Sciences and Allied Hospitals, Kanchan Bagh, Santosh Nagar, Hyderabad 500 058, India*
[3]*Center for Genome Information, Department of Environmental Health, University of Cincinnati, Cincinnati, OH 46267, USA*
[4]*Centre for DNA Fingerprinting and Diagnostics, 4-87/1 ECIL Road, Nacharam, Hyderabad 500 076, India*
[5]*Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India*

Abstract

Polymorphic allelic variants of chemokine receptors CCR2 and CCR5, as well as of stromal-derived factor-1 SDF-1, the ligand for the chemokine receptor CXCR4, are known to have protective effects against HIV-1 infection and to be involved with delay in disease progression. We have studied the DNA polymorphisms at the loci that encode these proteins in 525 healthy individuals without any history of HIV-1 infection from 11 diverse populations of Andhra Pradesh, South India. The two protective alleles *SDF-1-3¢A* and *CCR2-64I* at the SDF-1 and CCR2 loci, respectively, are present in all populations studied, although their frequencies differ considerably across populations (from 17% to 35% for the *SDF-1-3¢A* allele, and from 3% to 17% for *CCR2-64I*). In contrast the *CCR5-D32* allele is observed only in three populations (Yamani, Pathan and Kamma), all in low frequencies (i.e. 1% to 3%). The mean number of mutant alleles (for the three loci together) carried by each individual varies from 0.475 (in Vizag Brahmins) to 0.959 (in Bohra Muslims). The estimated relative hazard values for the populations, computed from the three-locus genotype data, are comparable to those from Africa and Southeast Asia, where AIDS is known to be widespread.

[Ramana G. V., Vasanthi A., Khaja M., Su B., Govindaiah V., Jin L., Singh L. and Chakraborty R. 2001 Distribution of HIV-1 resistance-conferring polymorphic alleles *SDF-1-3¢A, CCR2-64I* and *CCR5-D32* in diverse populations of Andhra Pradesh, South India. *J. Genet.* **80**, 137–140]

Introduction

Infection by human immunodeficiency virus type 1 (HIV-1) usually requires a cascade of events for its entry into target cells. Initially the viral envelope protein binds to CD4, a protein that is ubiquitously expressed on the cell surface of T-helper cells and macrophages (Robey *et al.* 1985). The viral protein then binds to a second kind of cell-surface molecule known as chemokine receptors. The macrophage-tropic (M-tropic) HIV-1 strains utilize the chemokine receptor CCR5 in combination with CD4 for entering target cells while the T-tropic HIV-1 uses CXCR4 (Littman 1998). Stromal-derived factor-1 (SDF-1, also known as pre-B cell growth stimulating factor) is the natural ligand for CXCR4. Variants of SDF-1 down-regulate CXCR4, thereby blocking infection by T-tropic but not M-tropic HIV strains (Bluel *et al.* 1996; Feng *et al.* 1996). Other chemokine receptors, e.g. CCR2, also play a role in HIV pathogenesis, although their mechanisms are less clear (Littman 1998; Mummidi *et al.* 1998). Recently, Saha *et al.* (2001) reported data characterizing two HIV-1 isolates, which infect target cells using cell-surface protein CD8 instead of CD4 without presence of CCR5 or CXCR4 as coreceptor complex.

*For correspondence. E-mail: ranajit.chakraborty@uc.edu.

Keywords. HIV-1 resistance polymorphism; chemokine receptor; stromal-derived factor-1; Andhra Pradesh; AIDS.

Individuals with variants of the genes encoding the chemokine receptors CCR2 and CCR5 and the ligand SDF-1 were shown to have resistance to HIV-1 infection (Dean *et al.* 1996; Fauci 1996; Martinson *et al.* 1997; Mummidi *et al.* 1998; Winkler *et al.* 1998). The deletion mutant *CCR5-D32* is widely observed in Caucasian populations (Martinson *et al.* 1997), whereas *CCR2-64I* and *SDF-1-3¢A* are distributed in all population groups worldwide. The relative hazard (RH) indices for all genotypic combinations at these three loci were estimated in cohort studies as well as in different ethnic populations (Smith *et al.* 1997; Winkler *et al.* 1998; Su *et al.* 1999).

HIV-1 infection has spread to all population groups in India and has reached epidemic proportions (Misra *et al.* 1998). In the present study, we investigated the distribution of *SDF-1-3¢A*, *CCR2-64I* and *CCR5-D32* alleles in 11 diverse population groups from Andhra Pradesh state, South India. Also, the distribution of the number of mutants carried by each individual was computed. The relative hazard (RH) index was estimated in these populations to evaluate the potential for natural resistance against HIV-1 infection and progression of AIDS in each population.

Materials and methods

DNA was extracted from a total of 525 blood samples from diverse populations of Andhra Pradesh comprising castes (Vizag Brahmins, Peruru Brahmins, Kammas), tribes (Bagata, Poroja, Valmiki), and Muslim isolates (Yamani, Pathan, Sunni, Shia, Bohra) following standard procedure (Sambrooke *et al.* 1992). The castes and tribes were sampled from Visakhapatnam and the Muslim groups from Hyderabad. For genotyping the three loci, SDF-1, CCR2 and CCR5, the published primer sequences were adapted (Dean *et al.* 1996; Martinson *et al.* 1997; Smith *et al.* 1997; Winkler *et al.* 1998). For SDF-1 PCR products are digested with *Msp*I at 37°C for 4 h and genotypes are scored as wild type (245 bp and 110 bp), heterozygote (355 bp, 245 bp and 110 bp) and mutant homozygote (355 bp, no digestion). For CCR2 polymorphism is detected by digesting the PCR products with *Bsa*I at 37°C for 4 h; wild-type product shows up as a single 128-bp band (undigested), heterozygote as 128-bp and 110-bp bands, and mutant homozygote as a 110-bp band. For CCR5 the amplified products can be scored as wild type (217 bp) or mutant (185 bp).

From the genotype distributions at each locus, allele frequencies were computed by the gene count method (Li 1976). Also, the distribution of the number of mutant alleles (at the three loci together) carried by each individual was derived from the three-locus genotype data.

To evaluate AIDS onset risk for the populations sampled, the relative hazard (RH) was calculated for each population on the basis of the three-locus genotype of each individual. There are 27 possible three-locus genotypes (adapted from Winkler *et al.* 1998 cohort studies) which can be grouped into four distinct values. The RH of a population can be estimated by $RH = \Sigma W_i P_i$, where W_i and P_i are the genotype-specific RH and frequencies summed over all four groups of genotypes. Three AIDS definitions, AIDS-1993, AIDS-1987 and Death, following the classification of Dean *et al.* (1996), were considered in the RH evaluations.

Table 1. Genotype counts of alleles at SDF1, CCR2 and CCR5 loci.

Population	Sample size	SDF-1			CCR2			CCR5	
		GG	AG	AA	GG	AG	AA	N	HT
Castes									
V. Brahmins	59	39	18	2	51	8	0	59	0
P. Brahmins	45	25	20	0	36	9	0	45	0
Kamma	51	29	22	0	46	3	2	49	2
Tribes									
Bagata	45	30	15	0	36	9	0	45	0
Poroja	35	21	14	0	31	4	0	35	0
Valmiki	42	25	14	3	39	3	0	42	0
Muslims									
Yamani	49	30	16	3	45	4	0	48	1
Pathan	50	20	25	5	45	5	0	47	3
Sunni	50	29	14	7	34	15	1	50	0
Shia	50	25	15	10	38	11	1	50	0
Bohra	49	18	29	2	33	15	1	49	0
Total	525	291	202	32	434	86	5	519	6

The A allele at the SDF-1 locus represents the mutation *SDF-1-3¢A*, the A allele at the CCR2 locus represents the mutation *CCR2-64I*, and the genotype HT represents the heterozygote for the mutation *CCR5-D32* (N = wild-type homozygote at CCR5 locus).

Results and discussion

The genotype distributions for each locus in the 11 populations are shown in table 1. Homozygous mutant bearers are rare for the CCR2 locus, and no homozygous *CCR5-D32* mutation was found in this study. In contrast, *SDF-1-3¢A* mutant homozygotes are found in appreciable frequencies in the Muslim isolates.

Table 2 presents a summary of the three-locus genotype data, in which individuals are grouped by the number of mutant alleles present at these three loci. Also shown in this table are the average number of mutant alleles per individual and the standard deviation of the distribution. In general, individuals belonging to the Muslim groups carry more mutations (highest in the Bohra Muslims, almost one mutation per individual), while the caste and tribe populations do not differ in terms of the number of mutations carried (averaging between 0.43 and 0.64 mutations per individual).

Estimates of mutant allele frequencies and relative hazard values are presented in table 3. The frequency of *SDF-1-3¢A* is quite high in the five Muslim populations (22% to 35%), compared to the castes and tribal groups (17% to 24%). *SDF-1-3¢A* occurs at exceptionally high frequency in Oceanic populations but is relatively low in African populations (Su *et al.* 1998). In contrast, the frequency of *CCR2-64I* allele shows a wide range of variation in the five Muslim groups (4% to 17%), while

Table 2. Individuals showing normal and mutant phenotypes at SDF-1, CCR2 and CCR5 loci and mean and standard deviation of number of mutant alleles.

Population	Sample size	No. of individuals with			No. of mutant alleles	
		0 mutant	1 mutant	2 mutants	Mean	S.D.
Castes						
V. Brahmins	59	31	26	2	0.475	0.568
P. Brahmins	45	20	21	4	0.644	0.645
Kamma	51	24	25	4	0.569	0.575
Tribes						
Bagata	45	24	18	2	0.533	0.625
Poroja	35	19	14	2	0.514	0.612
Valmiki	42	25	17	0	0.426	0.505
Muslims						
Yamani	49	26	22	1	0.49	0.545
Pathan	50	37	13	0	0.76	0.625
Sunni	50	21	21	8	0.74	0.723
Shia	50	19	25	6	0.74	0.664
Bohra	49	14	23	12	0.959	0.735

Table 3. Allele frequencies with standard errors for *SDF-1-3¢A*, *CCR2-64I* and *CCR5-D32* and relative hazard (RH) values in South Indian populations.

Population	Sample size	SDF-1	CCR2	CCR5	RH1	RH2	RH3
Castes							
V. Brahmins	59	0.19 (0.05)	0.07 (0.03)	0	0.94 (0.04)	0.93 (0.05)	0.92 (0.06)
P. Brahmins	45	0.22 (0.06)	0.10 (0.04)	0	0.93 (0.07)	0.93 (0.07)	0.92 (0.08)
Kamma	51	0.21 (0.06)	0.07 (0.04)	0.02 (0.02)	0.95 (0.06)	0.95 (0.06)	0.94 (0.07)
Tribes							
Bagata	45	0.17 (0.06)	0.10 (0.04)	0	0.93 (0.07)	0.93 (0.07)	0.92 (0.08)
Poroja	35	0.20 (0.07)	0.06 (0.04)	0	0.96 (0.05)	0.96 (0.05)	0.95 (0.06)
Valmiki	42	0.24 (0.06)	0.03 (0.03)	0	0.95 (0.04)	0.93 (0.06)	0.92 (0.07)
Muslims							
Yamani	49	0.22 (0.06)	0.04 (0.03)	0.01 (0.01)	0.95 (0.03)	0.93 (0.04)	0.92 (0.06)
Pathan	50	0.35 (0.07)	0.05 (0.03)	0.03 (0.02)	0.91 (0.05)	0.88 (0.07)	0.86 (0.08)
Sunni	50	0.28 (0.06)	0.17 (0.05)	0	0.85 (0.06)	0.83 (0.07)	0.80 (0.08)
Shia	50	0.35 (0.07)	0.13 (0.05)	0	0.86 (0.06)	0.81 (0.08)	0.78 (0.10)
Bohra	49	0.34 (0.07)	0.17 (0.05)	0	0.89 (0.08)	0.89 (0.08)	0.87 (0.09)

RH1, RH2 and RH3 were calculated under three different AIDS definitions, AIDS-1993, AIDS-1987 and Death, respectively.

in the tribes and castes this allele has a frequency not exceeding 10%. The worldwide data show complete absence of this allele in some Southeast Asian populations (Su *et al.* 2000). The frequency of the *CCR5-D32* deletion allele in the 11 populations studied here is either zero or negligible, which is consistent with data reported from other populations (Martinson *et al.* 1997; Su *et al.* 2000). The CCR5 deletion allele is observed mostly in European populations. The marginal presence of the allele seen in Asian populations could plausibly be due to gene flow from Caucasian populations. However, *CCR5-D32* is completely absent in populations from Africa, Oceania and the Americas.

Estimates of the average relative hazard values based on three-locus genotypes for each individual, shown in the last three columns of table 3, are consistent with the distribution of the number of mutations carried by individuals. For example, the RH values within caste and tribal groups and the Yamani Muslims are nearly the same, in parallel with the homogeneity of the number of mutations carried by individuals in these populations. In contrast, the four other Muslim groups (Pathan, Sunni, Shia and Bohra) exhibit somewhat lower RH values, owing to the fact that in these populations the number of mutations carried is also higher. Seven of the 11 populations exhibit RH values > 0.9 for all three criteria (AIDS-1993, AIDS-1997 and Death), indicating the higher propensity for susceptibility of these populations to HIV-1 infection. In contrast, Oceanic populations have the lowest RH values, which means they have the highest protection from AIDS onset or HIV-1 infection (Su *et al.* 1999). Worldwide data, published in Su *et al.* (2000), suggest that Southeast Asian populations and African populations exhibit the highest RH values, which indicates high prevalence of AIDS epidemic in these two regions. It is likely that intermediate frequencies of two mutant alleles (*SDF-1-3¢A* and *CCR2-64I*) and complete absence of *CCR5-D32* result in high RH values for Southeast Asian populations (Su *et al.* 2000). In this sense, our data are consistent with this observation, since we found comparatively higher RH values for the caste and tribal populations of Andhra Pradesh, together with absence of *CCR5-D32* allele in these populations.

Acknowledgements

All DNA samples analysed here were anonymized, and were collected with informed consent of the subjects. Research supported by the U.S. Public Health Services Research Grant GM 41399 from the National Institutes of Health.

References

Bluel C. C., Farzan M., Choe H., Parolin C., Clark-Lewis I., Sodraski J. and Springer T. A. 1996 The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusion and blocks HIV-1 entry. *Nature* **382**, 829–833.

Dean M., Carrington M., Winkler C., Huttley G. A., Smith M. W., Allikmets R., Goedert J. J. *et al.* 1996 Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CCR5 structural gene. *Science* **273**, 1856–1862.

Fauci A. S. 1996 Resistance to HIV-1 infection: it's in the genes. *Nature Med.* **2**, 966–967.

Feng Y., Broder C. C., Kennedy P. E. and Berger E. A. 1996 HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane G protein coupled receptor. *Science* **272**, 872–877.

Li C. C. 1976 *First course in population genetics*. Boxwood Press, Pacific Grove.

Littman D. R. 1998 Chemokine receptors: keys to AIDS pathogenesis? *Cell* **93**, 677–680.

Martinson J. J., Chapman N. H., Rees D. C., Liu Y. T. and Clegg J. B. 1997 Global distribution of the CCR5 gene 32-base pair deletion. *Nature Genet.* **16**, 100–103.

Misra S. N., Sengupta D. and Sathpathy S. K. 1998 AIDS in India: recent trends in opportunistic infections. *Southeast Asian J. Trop. Med. Public Health* **29**, 373–376.

Mummidi S, Ahuja S. S., Gonzalez E., Anderson S. A., Santiago E. N., Setphan K. T., Craig F. E., O'Connell P., Tryon V., Clark R. A., Dolan M. J. and Ahuja S. K. 1998 Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nature Med.* **4**, 786–793.

Robey W. G., Safai B., Oroszlen S., Arthur L. O., Gonda M. A., Gallo R. C. and Fischinger P. J. 1985 Characterization of envelope and core structural gene products of HTLV-III with sera from AIDS patients. *Science* **228**, 593–595.

Saha K., Zhang J., Gupta A., Dave R., Yimen M. and Zerhouni B. 2001 Isolation of primary HIV-1 that target $CD8^+$ T lymphocytes using CD8 as a receptor. *Nature Med.* **7**, 65–72.

Sambrooke J., Fritsch E. F. and Maniatis T. 1992 *Molecular cloning: a laboratory manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Smith M. W., Dean M., Carrington M., Winkler C., Huttley G. A., Lomb D. A., Goedert J. J. *et al.* 1997 Contrasting genetic influence of CCR2 and CCR5 variants of HIV-1 infection and disease progression. *Science* **277**, 959–965.

Su B., Chakraborty R., Jin L., Xiao J. H. and Lu D. R. 1998 Frequency of an HIV-resistant allele is exceptionally high in New Guinea highlanders. *J. Am. Med. Assoc.* **280**, 1830.

Su B., Jin L., Hu F., Xiao J., Luo J., Lu D., Zhang W., Chu J., Du R., Geng Z., Qiu X., Tan J., O'Brien S. J. and Chakraborty R. 1999 Distribution of two HIV-1 resistant polymorphisms (SDF1-3′A and CCR2-64I) in East Asia and world populations and its implication in AIDS epidemiology. *Am. J. Hum. Genet.* **65**, 1047–1053.

Su B., Sun G., Lu D., Hu F., Chakraborty R., Deka R. and Jin L. 2000 Distribution of three HIV-1 resistant polymorphisms SDF1-3′A, CCR2-64I, and CCR5-Δ32 in global populations. *Eur. J. Hum. Genet.* **8**, 975–979.

Winkler C., Modi W., Smith M. W., Nelson G. W., Wu X., Carrington M., Dean M. *et al.* 1998 Genetic restrictions of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science* **279**, 389–393.

Received 3 September 2001