Genetic Restriction of AIDS Pathogenesis by an SDF-1 Chemokine Gene Variant

Cheryl Winkler, William Modi, Michael W. Smith, George W. Nelson, Xueyun Wu, Mary Carrington, Michael Dean, Tasaku Honjo, Kai Tashiro, D. Yabe, Susan Buchbinder, Eric Vittinghoff, James J. Goedert, Thomas R. O'Brien, Lisa P. Jacobson, Roger Detels, Sharyne Donfield, Anne Willoughby, Edward Gomperts, David Vlahov, John Phair, ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), Stephen J. O'Brien*

Stromal-derived factor (SDF-1) is the principal ligand for CXCR4, a coreceptor with CD4 for T lymphocyte cell line–tropic human immunodeficiency virus–type 1 (HIV-1). A common polymorphism, *SDF1-3′A*, was identified in an evolutionarily conserved segment of the 3′ untranslated region of the SDF-1 structural gene transcript. In the homozygous state, *SDF1-3′A/3′A* delays the onset of acquired immunodeficiency syndrome (AIDS), according to a genetic association analysis of 2857 patients enrolled in five AIDS cohort studies. The recessive protective effect of *SDF1-3′A* was increasingly pronounced in individuals infected with HIV-1 for longer periods, was twice as strong as the dominant genetic restriction of AIDS conferred by *CCR5* and *CCR2* chemokine receptor variants in these populations, and was complementary with these mutations in delaying the onset of AIDS.

HIV-1 strains isolated from recently infected individuals are predominantly macrophage-tropic (M-tropic) and non–syncytium-inducing, and they co-opt CC-chemokine receptor proteins as entry ports in combination with CD4 molecules (*1*, *2*). Over the course of HIV-1 infection, viral phenotype and coreceptor use broaden to include the appearance of T lymphocyte cell line–tropic (T-tropic) variants near the time when AIDS symptoms are first observed (*1*, *3*, *4*). T-tropic strains induce the formation of syncytia in $CD4^+$ cell lines in vitro, infect peripheral blood mononuclear cells (PBMCs) faster, and replicate more aggressively than do the early M-tropic isolates (*1*, *4*). The occurrence of T-tropic isolates usually precedes a precipitous drop in CD4 T cells, which suggests that these viruses may contribute to T cell depletion. T-tropic HIV-1 enters target cells by means of CD4 and CXCR4 as a coreceptor complex, although T-tropic strains can also use CCR5 (*4*). Stromal-derived factor (SDF-1, also called pre–B cell growth stimulating factor), a powerful chemoattractant cytokine, is the natural ligand for CXCR4. Recent experiments have shown that SDF-1α (one of two transcriptional splice variants of the *SDF1* gene) is capable of down-regulating CXCR4 on cells by induction of endocytosis, effectively blocking infection by T-tropic but not M-tropic HIV-1 strains (*5*, *6*).

The use of available CXCR4 coreceptors by viral strains that emerge during late stage HIV-1 infection, together with the demonstration that SDF-1 effectively inhibits HIV-1 replication, prompted a polymorphism search for SDF-1 structural gene variants that might influence HIV-1 transmission or pathogenesis. We screened 1354 of the 3526 base pairs (bp) represented in human SDF-1β transcripts with a series of polymerase chain reaction (PCR) primers and single-strand conformation polymorphism (SSCP) heteroduplex assays (*7*) in a subgroup of 144 patients enrolled in five epidemiologic cohorts assembled to monitor HIV-1 infection and AIDS (*8*–*11*). Sequence analysis of a common variant revealed a G → A transition at position 801 (counting from the ATG start codon) in the 3′ untranslated region (3′UTR) of the reference sequence (GenBank accession number L36033). The polymorphism (designated *SDF1-3′UTR-801G-A* and abbreviated *SDF1-3′A* below) is represented in the SDF-1β transcript but not in the SDF-1α transcript. Because this variant eliminated an Msp I restriction site, a PCR–restriction fragment length polymorphism (RFLP) assay was used for rapid detection of genotypes (*7*). The allele and genotype frequencies of *SDF1-3′A* were determined in 2857 individuals from five AIDS cohorts (*8*–*11*). The following *SDF1-3′A* allele frequencies were found: Caucasians, 0.211 (n = 1835); Hispanics, 0.160 (n = 131); African Americans, 0.057 (n = 859); and Asians, 0.257 (n = 37) (*12*).

A role for *SDF1-3′A* in HIV-1 infection was investigated by genotyping 2419 HIV-1–infected patients and 435 HIV-1–exposed uninfected individuals. No significant differences in *SDF1* allele or genotype frequencies were observed in initial comparisons of exposed (or at risk) uninfected versus infected individuals in separate or combined cohort analyses [Fisher's exact test (FET), P = 0.16 to 1.0] (*12*). However, a group of 79 high-risk exposed uninfected individuals from MACS (those with extremely high-risk sexual practices) (*10*, *13*) showed a highly significant elevation in *SDF1-+/3′A* heterozygotes [50.6% among high-risk uninfected individuals compared to 31.1% among infected patients; FET, P = 0.002], suggesting a protective effect against HIV-1 infection for this genotype.

The influence of *SDF1* genotypes (+/+, +/3′A, and 3′A/3′A) on disease progression among HIV-1–infected individuals was analyzed by means of Cox proportional hazards models (*14*) (Fig. 1 and Table 1). A subgroup of 639 seroconverters from four cohorts was included in the analysis (*8*–*11*). These participants had well-characterized dates of seroconversion, with a first positive HIV test no more than 3 years after the last negative test, or, in the case of some SFCC participants, before the end of 1980 (*8*, *9*, *15*). Three AIDS endpoints reflecting advancing morbidity were evaluated: (i) AIDS-1993, as defined by the U.S. Centers for

C. Winkler, W. Modi, M. W. Smith, G. W. Nelson, X. Wu, M. Carrington, Science Applications International Corporation (SAIC), National Cancer Institute, Frederick, MD 21702, USA.
M. Dean and S. J. O'Brien, Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702, USA.
T. Honjo and D. Yabe, Department of Medical Chemistry, Kyoto University, Kyoto, Japan.
K. Tashiro, Center for Molecular Biology and Genetics, Kyoto University, Kyoto, Japan.
S. Buchbinder and E. Vittinghoff, AIDS Office, Department of Public Health, 25 Van Ness Avenue, Suite 500, San Francisco, CA 94102, USA.
J. J. Goedert and T. R. O'Brien, Viral Epidemiology Branch, National Cancer Institute, Executive Plaza North, 6130 Executive Boulevard, Bethesda, MD 20892, USA.
L. P. Jacobson, CAMACS, Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, Room E-7006, 615 North Wolfe Street, Baltimore, MD 21205, USA.
R. Detels, UCLA School of Public Health, Department of Epidemiology, Center for Health Sciences, Room 71-267, Box 951772, Los Angeles, CA 90095, USA.
S. Donfield, Rho Inc., 121 South Estes Drive, Suite 100, Chapel Hill, NC 27514, USA.
A. Willoughby, Center for Research for Mothers and Children, National Institute of Child Health and Human Development, Building 61E, Room 4B11H, Bethesda, MD 20892, USA.
E. Gomperts, Children's Hospital of Los Angeles, 4650 Sunset Boulevard, Los Angeles, CA 90027, USA.
D. Vlahov, Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, Room E-6008, 615 North Wolfe Street, Baltimore, MD 21205, USA.
J. Phair, Northwestern University Medical School, Comprehensive AIDS Center, 6980 North Lake Shore Drive, Suite 1106, Chicago, IL 60611, USA.
*To whom correspondence should be addressed.

Disease Control (*16*) (that is, HIV-1 infection plus AIDS-defining illness or decline of CD4 T lymphocytes to <200 cells/mm^3) or death; (ii) the more stringent AIDS-1987 definition (*16*) (HIV-1 infection plus AIDS-defining illness) or death; and (iii) death during follow-up for an HIV-1–infected patient (97% of these had AIDS-1993).

For the combined and separate cohort analyses, the *SDF1*-+/+ and *SDF1*-+/3′*A* individuals were indistinguishable in the pattern of progression to the three AIDS endpoints (Fig. 1). However, there was a marked slowing in progression to AIDS for individuals with the *SDF1*-3′*A*/3′*A* genotype [that is, relative hazards (RH) < 0.65; Table 1]. The delay was statistically significant with MACS and combined cohorts for AIDS-1987 and death, and with SFCC for death, for Caucasians or all ethnic groups (*12*, *13*, *17*). The extent of observed protection from AIDS progression associated with the *SDF1*-3′*A*/3′*A* genotype follows a gradation in combined and SFCC cohorts across increasingly severe AIDS endpoints (Fig. 1 and Table 1). The RH value for the combined Caucasian cohort sample was 0.65 for AIDS-1993, 0.36 for AIDS-1987, and 0.24 for death (lower values indicate increased protection; Table 1). The tendency to display increased protection in later stages of HIV-1 infection was also seen in MACS (RH = 0.59 > 0.22 > 0.1, respectively) and SFCC (RH = 0.83 > 0.30 > 0.00) cohorts (*12*). This gradation was extended when time to CD4 <200 cells/mm^3 (alone without AIDS disease or death) was used as an endpoint (Caucasians: RH = 0.67, P = 0.18; all races: RH = 0.63, P = 0.13) (*18*). The gradation suggests that *SDF1*-3′*A*/3′*A* protection is more pronounced in later stages of HIV-1 infection and is possibly related to interference with the appearance of T cell–tropic HIV-1 strains.

The protective effects of *SDF1*-3′*A*/3′*A* homozygotes were also apparent in defined disease category analyses, which allow the inclusion of seroprevalent patients (those whose seroconversion date is unknown because they were HIV-1 antibody–positive at the time of enrollment) in the slow/nonprogressor category (9). A significant elevation in the frequency of *SDF1*-3′*A*/3′*A* was observed among slow/nonprogressors within the combined cohorts for every AIDS outcome (Fig. 2A). The relative risk for AIDS avoidance, estimated by a case/control odds ratio, ranged from 2.4 to 4.1 for the three AIDS endpoints. The results of both the survival (Fig. 1 and Table 1) and disease category analyses (Fig. 2A) reveal a strong recessive *SDF1*-3′*A* association with delayed clinical outcomes of HIV-1 infection.

Variant alleles within the coding regions of the chemokine receptors CCR5 and CCR2, which are coreceptors for M-tropic HIV, have been shown to delay the rate of progression to AIDS (*8*, *9*, *19*, *20*). The mutant alleles *CCR5*-Δ*32* and *CCR2*-*64I* are dominant, genetically independent, and equally protective (*8*, *9*, *21*). An estimated 25 to 30% of long-term survivors who remain AIDS-free for >16 years can be attributed to a protective genotype for either *CCR5*-Δ*32* or *CCR2*-*64I* (9). A survival analysis of the relative contributions of *CCR5*-Δ*32*, *CCR2*-*64I*, and *SDF1*-3′*A* genotypes (Fig. 1 and Table 1) reaffirms the protective effects of *CCR2*, *CCR5*, and *SDF1* variant genotypes on progression to AIDS when the influence of the other protective loci are considered as covariables (*14*, *21*). For AIDS-1987 and death endpoints, the extent of *SDF1*-3′*A*/3′*A* protection in combined cohorts (Caucasian and all ethnic groups) is approximately twice that seen with *CCR2* or *CCR5* protection (that is, RH for *SDF1* versus CCR from Table 1 equals 0.37:0.64 for AIDS-1987 and 0.24:0.56 for death; P = 0.03) (*22*). In addition, CCR and *SDF1* protection are additive in AIDS cohorts, because patients with both *SDF1* and CCR protective genotypes avoid AIDS outcomes longer than do patients with only single-gene protection (P = 0.05 for AIDS-1993, P < 0.01 for AIDS-1987 and for death; Cox model log

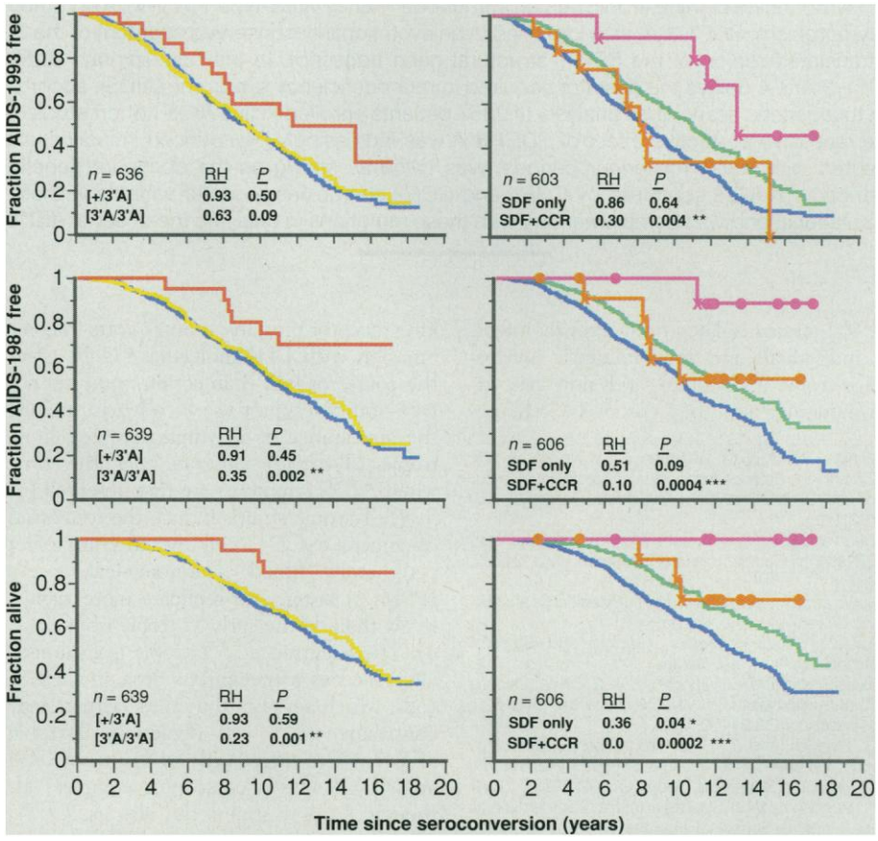

Fig. 1. Kaplan-Meier survival curves of seroconverters, showing relation of *SDF1*-3′*A*/3′*A* recessive protection to AIDS endpoints (*15*). Left panels: *SDF1*-3′*A*/3′*A* genotype survival (red) is compared with survival of *SDF1*-+/3′*A* (yellow) and *SDF1*-+/+ (blue) genotypes. Caucasians in the combined (ALIVE, MACS, MHCS, and SFCC) cohorts (*8*–*11*) [n = number of patients, RH = relative hazard, P = log likelihood P value based on the Cox proportional hazards model (*14*) for *SDF1*-3′*A*/3′*A* and +/3′*A* survival compared to *SDF1*-+/+ survival]. The value of n for AIDS-1993 is smaller than for AIDS-1987 or death because several subjects had CD4 T lymphocyte counts below 200 before HIV infection; for these subjects AIDS-1993 was impossible to define. Right panels: Survival curves for protective genotypes for *SDF1*, *CCR2*, and *CCR5* versus +/+ at the three loci. The protective genotypes are *SDF1*-3′*A*/3′*A*, *CCR2*-+/*64I* or *64I*/*64I*, and *CCR5*-+/Δ*32* and Δ*32*/Δ*32*. The four curves represent the following genotypes: +/+ at *SDF1*, *CCR2*, and *CCR5* (blue); one or more *CCR2*/*CCR5* protective genotypes and *SDF1*-+/+ (green); *SDF1*-3′*A*/3′*A* and +/+ at *CCR2*/*CCR5* (orange); and *SDF1*-3′*A*/3′*A* and protection by one or more *CCR2*/*CCR5* protective genotypes (pink). × indicates single events; ● indicates patient censoring. Summary statistics for the combined cohort analyses are shown in Table 1 (*12*). Log-log survival time versus log time plots were examined for proportionality with the combined cohort analysis. The plots were parallel and did not intersect, as assumed in the Cox proportional hazards model (*14*).

likelihood test) (*14*, *23*). For example, only 1 of the 10 seroconverter patients who were genotypically *SDF1-3′A/3′A* and either *CCR2-* or *CCR5*-protected has progressed to AIDS-defining conditions (AIDS-1987), while 5 of 13 *SDF1-3′A/3′A*, *CCR2-+/+*; *CCR5-+/+* patients did. However, 11 of 23 dual-protected seroprevalent patients ultimately succumbed to AIDS, although their time interval to AIDS was unknown.

The cumulative effects of the *SDF1-3′A/3′A* protective genotype, separately or in combination with *CCR* protective genotypes, were assessed over six intervals after HIV-1 seroconversion (Fig. 2B). The results reveal a significant increase of *SDF1-3′A/3′A* genotypes among patients who avoid AIDS for longer periods ($P = 0.02$). There was a complete absence of a dual *CCR/SDF1* composite variant genotype among patients who developed AIDS-1987 or died within the first 10 years after HIV-1 infection, and only a single individual with *SDF1* plus *CCR* protective genotypes developed AIDS-1993 during this interval. Combined with the survival analyses, these data emphasize the protective effect of the *SDF1-3′A/3′A* genotype and suggest that its effect is more than additive with the protection provided by *CCR2* and *CCR5* variant alleles (*23*).

The *SDF1-3′A* variant is located in a segment of the 3′UTR of the SDF-1β transcript (*6*) that is highly conserved in sequence (69% sequence between human and mouse SDF-1β 3′UTR sequence with no gaps). This extent of conservation within the segment suggests that it may serve as a target for cis-acting factors influencing transcript abundance, synthesis, transport, stability, or splice product abundance (*24*). The simplest hypothesis for *SDF1-3′A/3′A* action involves up-regulation of the quantity of SDF-1 protein available to bind CXCR4 and stem the appearance of late stage T-tropic HIV-1 strains in infected patients (*25*). This mechanism would be consistent with the gradation in survival outcome whereby *SDF1* protection is more pronounced in late stage AIDS outcomes than for earlier stages (Fig. 1).

The frequency of *SDF1-3′A* homozygous recessive individuals is low (<5%) in the study populations, but the effect is quite strong. Although *SDF1* protection is more apparent in later stages of HIV-1 infection, the principal effect may actually involve a strong protection against rapid progression to AIDS (Fig. 2B). The extent of observed *SDF1-3′A/3′A* protection is twice as strong overall as the influence of either *CCR5-Δ32* or *CCR2-64I*. This extent of protection mediated by a potential regulatory region of *SDF1*, effective in a test population and without obvious clinical cost, offers an attractive opportunity for therapeutics that could mimic the action of the variation.

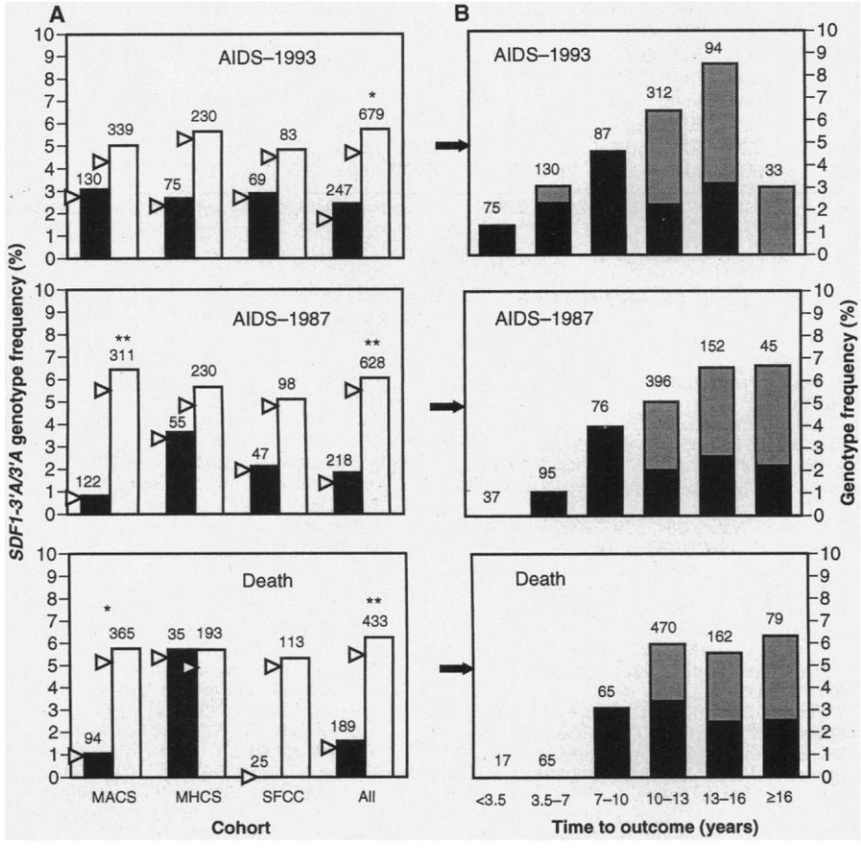

Fig. 2. (**A**) Defined disease category analysis of *SDF1-3′A/3′A* genotype frequencies for each cohort and combined cohorts for the three endpoints: AIDS-1993, AIDS-1987, and death (see text). Seroconverters who progressed to the designated outcomes before the cutoff time (defined below) were compared to seroconverters plus seroprevalents who survived outcome-free for at least that length of time. Imputed seroconversion dates for the seroprevalent subgroup for MHCS, HGDS, and ALIVE were provided by the cohort investigators (*10*, *11*). For MACS, date of enrollment was used as the starting date. Cutoffs (in years) were chosen as the time when approximately half of all seroconverters had progressed to the outcomes. The following times for the outcomes and cutoffs were used. For AIDS-1993: MACS, 6 years; MHCS, 9 years; SFCC, 12.5 years; and combined cohorts, 8 years. For AIDS-1987: MACS, 7.5 years; MHCS, 11.5 years; SFCC, 14.5 years; and combined, 10 years. For death: MACS, 8 years; MHCS, 11.5 years; SFCC, 14 years; and combined, 12 years (*9*). The number of individuals in each disease category is at the top of the bars. $^*P < 0.05$, $^{**}P < 0.01$ (FET for the null hypothesis of no *SDF1-3′A/3′A* protection compared with *SDF1-+/+* plus *SDF1-+/3′A*). Bars are for Caucasians; triangles indicate *SDF1-3′A/3′A* frequencies for all ethnic groups. The relative risk of rapid progression for unprotected (*SDF1-+/+* or *SDF1-+/3′A*) patients (relative to *SDF1-3′A/3′A* patients) was estimated for Caucasians in the combined cohorts by calculating the case/control odds ratios with slow progressors as controls; that is, the risk for each category is the ratio of the percentage of rapid progressors to the percentage of slow progressors. Estimate relative risks (with 95% confidence intervals and FET *P* values in parentheses) were as follows: for AIDS-1993, 2.4 (1.0 to 7.2, $P = 0.02$); for AIDS-1987, 3.4 (1.2 to 13.4, $P = 0.007$); for death, 4.1 (1.2 to 21.5, $P = 0.007$). (**B**) Frequencies of the protective *SDF1-3′A/3′A* genotype alone (black bars) or in combination with at least one *CCR2/CCR5* protective genotype (*CCR5-+/Δ32*, *CCR2-+/64I*, or *CCR2-64I/64I*; shaded bars) in six intervals of increasing survivorship from midpoint (seroconverters) or imputed (seroprevalents) seroconversion dates in Caucasians. Genotypic frequencies were determined separately for time to AIDS-1993, to AIDS-1987, and to death (see text) using seroconverters progressing to the outcomes in less than 3.5 years, 3.5 to <7 years, and 7 to <10 years, and including seroconverters and seroprevalents progressing to the outcomes in 10 to <13 years, 13 to <16 years, and ≥16 years. The number of individuals observed in each category is shown above the bars. The average frequency of the protective genotype for Caucasians is shown as an arrow. There is a statistically significant trend (Mantel-Haenzel χ^2) toward enrichment of *SDF1-3′A/3′A* genotypes at increasing survival intervals for AIDS-1993 ($P = 0.04$) and AIDS-1987 ($P = 0.01$), and toward enrichment of composite *SDF1-3′A/3′A* plus *CCR5* genotypes at increasing survival intervals for AIDS-1993 ($P = 0.005$), AIDS-1987 ($P = 0.01$), and death ($P = 0.05$).

Table 1. Survival analysis of protection from progression to AIDS outcomes by *SDF1-3′A/3′A* variant, *CCR5*, or *CCR2* protective polymorphisms. *Analysis 1:* Survival analysis for progression to three AIDS endpoints among HIV-1–infected seroconverters for *SDF1-3′A/3′A* versus *SDF1-+/+* or *SDF-+/3′A* genotypes, as in Fig. 1. Seroconverters in combined cohorts including only Caucasians and for all ethnic groups were analyzed using the Cox proportional hazards model (*13, 14*). A log likelihood test (1 df) (LL), *P* value, relative hazard (RH), and 95% confidence interval (CI) were calculated for each variable in the analysis of AIDS outcomes (*14*). Times to AIDS-1993, AIDS-1987, and death (see text) were calculated from the midpoint of the last HIV-1–negative test date and the first HIV-1–positive test date. The value of *n* for AIDS-1993 is smaller than for AIDS-1987 or death because several subjects had CD4 T lymphocyte counts below 200 before HIV infection; for these subjects AIDS-1993 was impossible to define. Seroconverters with an interval greater than 3 years between last negative and first positive were excluded from the analysis. Analyses were adjusted for age, where age is a categorical variable with three categories: <30, 30 to 40, or >40 years old. *Analysis 2: SDF1-3′A/3′A* versus *SDF1-+/+* or *SDF1-+/3′A* controlling for the protective genotypes of *CCR2* and *CCR5* (*21*). *Analysis 3: CCR2-64I/64I* or *CCR2-+/64I* or *CCR5-+/Δ32* versus *CCR5-+/+* and *CCR2-+/+* (normal at two loci) controlling for the protective genotype of *SDF1*. The LL *P* values use a Bonferroni correction for multiple independent tests performed in each of the three analyses. An LL calculation for χ^2 was performed because of the small numbers of patients and few failures in *SDF1-3′A/3′A* individuals (*14*). [+/+] for *SDF1* includes *SDF1-+/+* and *SDF1-+/3′A*.

AIDS outcome	Caucasians					All races				
	n	Events	RH	95% CI	LL-*P*	*n*	Events	RH	95% CI	*P*
Analysis 1: SDF1-3′A/3′A versus SDF1-+/3′A and SDF1-+/+ genotypes										
AIDS-1993	636	410	0.65	0.39–1.1	0.10	844	514	0.61	0.37–1.0	0.06
AIDS-1987	639	315	0.36	0.18–0.71	0.003*	857	387	0.34	0.18–0.67	0.002*
Death	639	243	0.24	0.10–0.57	0.001*	857	294	0.23	0.10–0.55	0.001**
Analysis 2: SDF1-3′A/3′A versus [+/+] for SDF1, CCR2, and CCR5										
AIDS-1993	603	391	0.67	0.40–1.14	0.14	800	491	0.63	0.38–1.06	0.08
AIDS-1987	606	303	0.37	0.18–0.72	0.004*	806	371	0.35	0.18–0.68	0.002*
Death	606	234	0.24	0.10–0.58	0.002*	806	282	0.23	0.10–0.55	0.001**
Analysis 3: CCR2/5 protection versus [+/+] for SDF1, CCR2, and CCR5										
AIDS-1993	603	391	0.63	0.51–0.77	0.00001****	800	491	0.65	0.54–0.78	0.000007****
AIDS-1987	606	303	0.64	0.50–0.81	0.0002**	806	371	0.66	0.53–0.83	0.0002**
Death	606	234	0.56	0.42–0.73	0.00003***	806	282	0.60	0.47–0.77	0.00007***

*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$. Calculated with the Bonferroni correction for multiple tests (six tests).

REFERENCES AND NOTES

1. M. P. D'Souza and V. A. Harden, *Nature Med.* **2**, 1293 (1996); B. A. Premack and T. J. Schall, *ibid.*, p. 1174; J. M. McNicholl *et al.*, *Emerg. Infect. Dis.* **3**, 261 (1997).
2. G. Alkhatib *et al.*, *Science* **272**, 1955 (1996); H. Deng *et al.*, *Nature* **381**, 661 (1996); T. Dragic *et al.*, *ibid.*, p. 667; H. Choe *et al.*, *Cell* **85**, 1135 (1996); B. J. Doranz *et al.*, *ibid.*, p. 1149; J. Rucker *et al.*, *ibid.* **87**, 437 (1996).
3. A. S. Fauci, *Nature* **384**, 529 (1996); R. A. Weiss, *Science* **272**, 1885 (1996).
4. R. I. Connor *et al.*, *J. Exp. Med.* **185**, 621 (1997); H. Schuitemaker *et al.*, *J. Virol.* **64**, 356 (1991); H. Schuitemaker *et al.*, *ibid.* **66**, 1354 (1992); B. Asjo, *Lancet* **ii**, 660 (1986); R. I. Connor, H. Mohri, Y. Cao, D. D. Ho, *J. Virol.* **67**, 1772 (1993); M. T. Roos *et al.*, *J. Infect. Dis.* **165**, 427 (1992); T. Zhu *et al.*, *Science* **261**, 1179 (1993).
5. Y. Feng, C. C. Broder, P. E. Kennedy, E. A. Berger, *Science* **272**, 872 (1996); C. C. Bluel *et al.*, *Nature* **382**, 829 (1996); E. O'Berlin *et al.*, *ibid.*, p. 833; C. C. Bluel *et al.*, *J. Exp. Med.* **184**, 1101 (1996); A. Amaru *et al.*, *ibid.* **186**, 139 (1997).
6. M. Shirozu *et al.*, *Genomics* **28**, 495 (1995); K. Tashiro *et al.*, *Science* **261**, 600 (1993).
7. PCR products were run out using two conditions: (i) 6% acrylamide (acrylamide:bisacrylamide, 37.5:1), 4°C at 50 W for 3.5 hours, and (ii) 5% acrylamide (acrylamide: bisacrylamide, 19:1), room temperature at 40 W for 4 to 6 hours. Both conditions used 1× tris-boric acid–EDTA buffer [M. White *et al.*, *Genomics* **12**, 301 (1992); M. Ravnik-Glavac *et al.*, *Hum. Mol. Genet.* **3**, 801 (1994); M. Orita *et al.*, *Genomics* **5**, 874 (1989)]. The following primers were used: 5′UTR (197 bp), GGCAGGTGGCGAGCTTGAGC (forward) and CTGGAGGGCGCTTATTGTC (reverse); exon 1 (130 bp), AGCCGCATTGCCCGCTCGGCGTC (F) and CGTCGCTGAGGCAGAGCGCGGTC (R); exon 2 (218 bp), CAAAATCTGACAGGGTAGTA (F) and TCGTTAGATGCAACTATGTTC (R); exon 3 (189 bp), AGCCGCGCTTTCCTCCTGTGC (F) and TAGTTTTCCTCGAGTGGGTC (R); exon 4 (318 bp), CTGTCCTGCTGGAGCTGGC (F) and TTTCAGAGCTGGGCTCCTAC (R); 3′UTR (302 bp), CAGTCAACCTGGGCAAAGCC (F) and AGCTTTGGTCCTGAGAGTCC (R).
8. M. Dean *et al.*, *Science* **273**, 1856 (1996).
9. M. W. Smith *et al.*, *ibid.* **277**, 959 (1997).
10. R. Kaslow *et al.*, *Am. J. Epidemiol.* **126**, 310 (1987); J. Phair *et al.*, *J. Acquired Immune Defic. Syndr.* **5**, 490 (1992); R. Detels *et al.*, *ibid.* **7**, 1263 (1994). MACS principal investigators include A. Saah, A. Munoz, and J. Margolick (Johns Hopkins School of Public Health), J. Phair (Northwestern University Medical School), R. Detels and J. V. Giorgi (University of California, Los Angeles), and C. Rinaldo (University of Pittsburgh). The study was funded by National Institute of Allergy and Infectious Diseases cooperative agreements Uo1-Ai-35042, 5-Mo1-Rr-00722(Gcrc), Uo1-Ai-35043, Uo1-Ai-37984, Uo1-Ai-35039, Uo1-Ai-35040, Uo1-A1-37613, and Uo1-Ai-35041.
11. S. P. Buchbinder, *AIDS* **8**, 1123 (1994); K.-J. Lui *et al.*, *Science* **240**, 1333 (1988); M. W. Hilgartner *et al.*, *Am. J. Pediatr. Hematol. Oncol.* **15**, 208 (1993); J. J. Goedert *et al.*, *N. Engl. J. Med.* **321**, 1141 (1989); M. M. Lederman *et al.*, *J. Infect. Dis.* **172**, 228 (1995); D. Vlahov *et al.*, *NIDA Research Monograph Series 103* (Public Health Service, Alcohol and Drug Abuse Administration, Washington, DC, 1991). The last of these sources describes the ALIVE cohort, including intravenous drug users and predominantly (94%) African Americans.
12. Additional analysis and data can be inspected at rex.ncl.nih.gov/RESEARCH/basic/lgd/Front_page.htm.
13. Y. Huang *et al.*, *Nature Med.* **2**, 1240 (1996); R. Detels *et al.*, *AIDS* **10**, 102 (1996).
14. D. R. Cox, *J. R. Stat. Soc. B* **34**, 187 (1972); Proportional Hazard Regression (PHREG), SAS Release 6.10 (Sas Institute Inc., Cary, NC); D. R. Cox and D. Oakes, *Analysis of Survival Data* (Chapman & Hall, London, 1984), p. 36; E. Marubini and M. G. Valsecchi, *Analysing Survival Data for Clinical Trials and Observational Studies* (Wiley, New York, 1995), p. 160.
15. Seroconverter patients included 867 subjects with a maximum interval of 3 years between an HIV-1 antibody–negative test date and their first HIV-1 antibody–positive test date. Seroconversion date was the midpoint between the last HIV-1 antibody–negative and first positive clinic visits. Ninety patients enrolled in the SFCC study before 31 December 1980 were included using imputed seroconversion dates based on their date of first HIV-1 antibody–positive test, because the likelihood of infection before 1 January 1978 (a 3-year window of infection) was extremely low (≤0.01). Seroconversion dates for the imputed SFCC subjects were set at 60 days, 120 days, and 180 days before the date for first antibody-positive visit for patients enrolled in 1978, 1979, and 1980, respectively (*11*).
16. U.S. Centers for Disease Control, *Morb. Mortal. Wkly. Rep.* **36** (suppl. 1) (14 August 1987); *ibid.*, **41** (18 December 1992). The latter publication contains a revised classification system for HIV-1 infection and an expanded surveillance case definition for AIDS among adolescents and adults. It went into effect in January 1993.
17. The MHCS cohort had 140 seroconverters, of whom four were *SDF1-3′A/3′A* homozygotes. Two developed AIDS at 8 and 10 years, while the other two have remained AIDS-free for 14 and 15 years after seroconversion. These small numbers render this cohort, when analyzed alone, less informative statistically than larger cohorts, for example, MACS (n = 332 seroconverters) or combined cohorts (n = 857). SFCC, a cohort of homosexual men with a preponderance of long-term survivors, had no deaths among six *SDF1-3′A/3′A* homozygous seroconverters, making estimates of relative hazard imprecise.
18. Additional evidence in support of an increasing appearance of *SDF1-3′A/3′A* protection in late stages of HIV-1 infection involves our finding a statistically significant difference in the frequency of homozygotes (FET, $P \leq 0.01$) in seroconverter (F = 3.5%; n = 669) versus seroprevalent (F = 6.2%; n = 743) cohort members. The enrichment of homozygotes in seroprevalent individuals is consistent with late-stage protection for two reasons: (i) The seroprevalent group was seropositive at enrollment and therefore included more patients with longer undefined intervals since infection; and (ii) studies may be biased to include more long-term survivors than rapid progressors [M. W. Smith *et al.*, *Nature Med.* **3**, 1052 (1997)]. MACS specifically excluded enrollment of individuals with AIDS-defining conditions, whereas SFCC selected for long-term survivors (*10, 11*).
19. R. Liu *et al.*, *Cell* **86**, 367 (1996); W. A. Paxton *et al.*, *Nature Med.* **2**, 412 (1996); M. Samson, *Nature* **382**, 722 (1996); N. L. Michael *et al.*, *Nature Med.* **3**, 338 (1997); P. A. Zimmerman *et al.*, *Mol. Med.* **3**, 23 (1997).
20. R. Biti *et al.*, *Nature Med.* **3**, 252 (1997); T. R. O'Brien *et al.*, *Lancet* **349**, 1219 (1997); I. Theodorou *et al.*, *ibid.*, p. 1219; N. L. Michael *et al.*, *Nature Med.* **3**, 1160 (1997); M. W. Smith *et al.*, *ibid.*, p. 1052; O. J. Cohen, *J. Clin. Invest.* **100**, 1581 (1997). These papers contain caveats about *CCR2* and *CCR5* genetic protection.
21. Protective genotypes include *CCR2-64I* protection (*CCR5-+/+*, *CCR2-+/64I* or *64I/64I*, and *SDF1-+/+*), *CCR5-Δ32* protection (*CCR5-+/Δ32*, *CCR2-+/+*, and *SDF1-+/+*), and *SDF1-3′A* protection (*SDF1-3′A/3′A*). Protective genotypes at either *CCR5* or *CCR2* are referred to as "*CCR* protection."
22. The null hypothesis of equality of *SDF1* and *CCR2/CCR5* protection was tested in a Cox model analysis by comparing a model with a single variable for *SDF1* or *CCR2/CCR5* protection, corresponding to equality of *CCR2/CCR5* and *SDF1* protection, with a model containing an additional variable for differential *SDF1* protection. For Caucasians in combined cohorts, relative hazards of <1 were shown for the differential *SDF1* protection variable for AIDS-1987 and death, indicating stronger protection by *SDF1* than by *CCR2* or *CCR5*, with P = 0.04 (AIDS-1987)

and $P = 0.03$ (death). The corresponding values for all ethnic groups were $P = 0.03$ and $P = 0.02$.

23. To test for a significant additive effect (that is, for an advantage in having protective genotypes both at the *SDF1* locus and at one or both of the *CCR* loci), we performed a Cox model test with an interaction variable and a single covariable for *SDF1* or *CCR* protection. This test had significant log likelihood P values ($P < 0.01$) for AIDS-1987 and death for Caucasians in combined cohorts, with relative hazards of 0.31 (AIDS-1987) and 0.0 (death) (no deaths in doubly protected group), showing a significant advantage to having both protective genotypes. As an additional test of the additivity of the interaction between *SDF1* and *CCR*, a Cox model test was performed with separate variables for protection by *SDF1* and *CCR*, plus an interaction variable. The relative hazards of the interaction term were 0.55 (AIDS-1993), 0.31 (AIDS-1987), and 0.0 (death), with the P values falling short of significance ($P = 0.13$ to 0.31). These results indicate that the nonadditivity in the interaction between *SDF1* and *CCR* protective genotypes is not significant, but that the interaction tends toward being stronger than additive—that is, synergistic.
24. J. Ross, *Trends Genet.* **12**, 171 (1996); K. C. Tsai, V. V. Cansino, D. T. Kohn, R. L. Neve, N. I. Perrone-Bizzozero, *J. Neurosci.* **17**, 1950 (1997); K. M. Mcgowan, S. Police, J. B. Winslow, P. H. Pekala, *J. Biol. Chem.* **272**, 1331 (1997); G. Shaw and R. Kamen, *Cell* **46**, 659 (1986); D. Kube *et al.*, *Cytokine* **7**, 107 (1995).
25. The *SDF1* gene contains four exons over a 5.6-kb region of chromosome 10q11.1 (*6*). Two alternatively spliced transcripts that specify SDF-1α and SDF-1β are made from the gene; the isomers differ by the addition of four COOH-terminal amino acids from the fourth exon in SDF-1β (*6*). The two transcripts have completely different 3′UTRs, and the *SDF1-3′A* mutation is found in the SDF-1β transcript within a sequence block that is conserved between mouse and human SDF-1β UTR sequences. The possibility of linkage disequilibrium tracking of the *SDF1-3′A* mutation through linkage disequilibrium was investigated first by sequence determination of the four *SDF1* coding exons in eight *SDF1* 3′A/3′A homozygotes. No additional polymorphisms were detected. Further sequence analysis of two *SDF1*-+/+ homozygotes and two *SDF1-3′A/3′A* homozygotes for 3253 nucleotides (out of 3524 in the entire transcript) identified two variants (positions 1912 and 2950) in single *SDF1*-+/+ individuals and revealed no additional mutations tracking with *SDF1-3′A*.
26. We very gratefully acknowledge the patients, their families, and clinicians who participated in the ALIVE, MACS, MHCS, HGDS, and SFCC cohort studies. We thank E. Binns, R. Boaze, K. Boyd, S. Cevario, K. Gong, D. Hague, A. Houser, L. Kenefic, M. Konsovich, D. Lomb, M. McNally, M. Mylasky, and M. Weedon for excellent technical assistance and D. Lipmann, G. Huttley, G. Smythers, C. Stephens, and J. Wang for helpful discussions. We also thank L. Main for secretarial assistance and the Frederick Biomedical Supercomputing Center for computational resources.

28 August 1997; accepted 25 November 1997

Characterization of an Avian Influenza A (H5N1) Virus Isolated from a Child with a Fatal Respiratory Illness

Kanta Subbarao,* Alexander Klimov, Jacqueline Katz, Helen Regnery, Wilina Lim, Henrietta Hall, Michael Perdue, David Swayne, Catherine Bender, Jing Huang, Mark Hemphill, Thomas Rowe, Michael Shaw, Xiyan Xu, Keiji Fukuda, Nancy Cox

An avian H5N1 influenza A virus (A/Hong Kong/156/97) was isolated from a tracheal aspirate obtained from a 3-year-old child in Hong Kong with a fatal illness consistent with influenza. Serologic analysis indicated the presence of an H5 hemagglutinin. All eight RNA segments were derived from an avian influenza A virus. The hemagglutinin contained multiple basic amino acids adjacent to the cleavage site, a feature characteristic of highly pathogenic avian influenza A viruses. The virus caused 87.5 to 100 percent mortality in experimentally inoculated White Plymouth Rock and White Leghorn chickens. These results may have implications for global influenza surveillance and planning for pandemic influenza.

The introduction and subsequent spread in the human population of influenza A viruses with a novel hemagglutinin (HA) or a novel HA and neuraminidase (NA) subtype results from a sudden and major change in virus antigenicity, which is referred to as an antigenic shift. Lack of protective immunity in the human population against the new HA and NA proteins can result in rapid global spread of the virus, leading to widespread morbidity and mortality. Pandemic strains contain new HA or NA genes derived from animal influenza A viruses. Influenza A viruses of 15 recognized HA subtypes and 9 NA subtypes are known to circulate in birds and other animals, creating a reservoir of influenza A virus genes available for genetic reassortment with circulating human strains of influenza virus. However, on the basis of seroarchaeology and virus isolation since 1933, only viruses of the H1, H2, and H3 subtypes are known to infect and cause disease in humans (*1*).

In general, avian influenza A viruses, including those that are highly pathogenic in birds, do not appear to replicate efficiently or cause disease in humans. The only reported natural infections of humans by avian viruses are two cases of conjunctivitis associated with avian H7 viruses, one of which was an infection with a seal virus of avian origin (*2*). Serosurveys of farm workers in southern China by single radial hemolysis revealed a seroprevalence ranging from 1 to 38% for avian viruses of the H4 through H13 subtypes, including 7% seroprevalence for H5 viruses (*3*). In contrast, there were no documented infections in U.S. poultry workers exposed to strains of avian (H5) influenza A viruses that were highly pathogenic in poultry (*4*).

On 9 May 1997, a previously healthy 3-year-old boy, who was a resident of Hong Kong, developed a sore throat, dry cough, and fever. He was diagnosed with pharyngitis and prescribed antibiotics and aspirin. The child continued to be symptomatic and febrile and was hospitalized on 15 May. Upon admission, he was noted to be febrile (axillary temperature above 39°C) and irritable. His laboratory tests were most remarkable for leukopenia (2000 white blood cells per cubic millimeter). His chest x-ray was within normal limits. The next day, he was transferred to another hospital, where he developed progressive respiratory distress associated with hypoxemia, consistent with acute respiratory distress syndrome. He also became increasingly unresponsive. Computerized tomography of the head was unremarkable, and examination of his cerebrospinal fluid was not suggestive of an inflammatory process. Despite mechanical ventilation and broad antibiotic coverage, the child died on 21 May with several complications, including respiratory failure, renal failure, and disseminated intravascular coagulopathy. Postmortem liver and kidney biopsies showed evidence of microvascular fatty infiltration consistent with Reye's syndrome, which is a recognized complication of influenza.

A tracheal aspirate specimen was obtained on day 10 of illness and was cultured for respiratory viruses. A cytopathic effect was noted in mammalian Madin Darby canine kidney (MDCK) cells and rhesus monkey kidney (LLC-MK2) cells 2

K. Subbarao, A. Klimov, J. Katz, H. Regnery, H. Hall, C. Bender, J. Huang, M. Hemphill, T. Rowe, M. Shaw, X. Xu, K. Fukuda, N. Cox, Influenza Branch, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA.
W. Lim, Government Virus Unit, Queen Mary Hospital, Hong Kong, China.
M. Perdue and D. Swayne, Southeast Poultry Research Laboratory, Agriculture Research Service, U.S. Department of Agriculture, Athens, GA 30605, USA.
*To whom correspondence should be addressed.