Biology Faculty Articles

Department of Biological Sciences

7-15-2006

Consistent Effects of TSG101 Genetic Variability on Multiple Outcomes of Exposure to Human Immunodeficiency Virus Type 1

Arman A. Bashirova
National Cancer Institute at Frederick

Gabriela Bleiber
University Hospital of Lausanne - Switzerland

Ying Qi
National Cancer Institute at Frederick

Holli Hutcheson
National Cancer Institute at Frederick

Traci Yamashita
Johns Hopkins School of Hygiene and Public Health

See next page for additional authors

Follow this and additional works at: https://nsuworks.nova.edu/cnso_bio_facarticles

Part of the Genetics and Genomics Commons, Medicine and Health Sciences Commons, and the Virology Commons

Authors

Arman A. Bashirova, Gabriela Bleiber, Ying Qi, Holli Hutcheson, Traci Yamashita, Randall C. Johnson, Jie Cheng, Galit Alter, James J. Goedert, Susan Buchbinder, Keith Hoots, David Vlahov, Margaret May, Frank Maldarelli, Lisa Jacobson, Stephen J. O'Brien, Amalio Telenti, and Mary Carrington

Consistent Effects of *TSG101* Genetic Variability on Multiple Outcomes of Exposure to Human Immunodeficiency Virus Type 1†

Arman A. Bashirova,[1]‡ Gabriela Bleiber,[2] Ying Qi,[3] Holli Hutcheson,[1] Traci Yamashita,[4]
Randall C. Johnson,[3] Jie Cheng,[3] Galit Alter,[5] James J. Goedert,[6] Susan Buchbinder,[7]
Keith Hoots,[8] David Vlahov,[4] Margaret May,[9] Frank Maldarelli,[10] Lisa Jacobson,[4]
Stephen J. O'Brien,[1] Amalio Telenti,[2] and Mary Carrington[3]*

*Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702[1]; Institute of Microbiology and Division of
Infectious Diseases, University Hospital of Lausanne, Lausanne, Switzerland[2]; Laboratory of Genomic Diversity, SAIC-Frederick,
Inc., National Cancer Institute, Frederick, Maryland 21702[3]; Department of Epidemiology, School of Hygiene and Public Health,
The Johns Hopkins University, Baltimore, Maryland[4]; Partners AIDS Research Center, Massachusetts General Hospital,
55 Fruit Street, Charlestown, Massachusetts 02129[5]; Viral Epidemiology Branch, Division of Cancer Epidemiology and
Genetics, NCI, Bethesda, Maryland[6]; University of California, San Francisco, California[7]; Gulf States Hemophilia Center,
University of Texas Health Science Center, Houston, Texas[8]; Department of Social Medicine, University of Bristol,
Bristol, United Kingdom[9]; and HIV Drug Resistance Program, National Cancer Institute, Bethesda, Maryland[10]*

**Tumor susceptibility gene 101 (*TSG101*) encodes a host cellular protein that is appropriated by human
immunodeficiency virus type 1 (HIV-1) in the budding process of viral particles from infected cells. Variation
in the coding or noncoding regions of the gene could potentially affect the degree of TSG101-mediated release
of viral particles. While the coding regions of the gene were found to lack nonsynonymous variants, two
polymorphic sites in the *TSG101* 5′ area were identified that were associated with the rate of AIDS progression
among Caucasians. These single-nucleotide polymorphisms (SNPs), located at positions −183 and +181
relative to the translation start, specify three haplotypes termed A, B, and C, which occur at frequencies of 67%,
21%, and 12%, respectively. Haplotype C is associated with relatively rapid AIDS progression, while haplotype
B is associated with slower disease progression. Both effects were dominant over the intermediate haplotype A.
The haplotypes also demonstrated parallel effects on the rate of CD4 T-cell depletion and viral load increase
over time, as well as a possible influence on HIV-1 infection. The data raise the hypothesis that noncoding
variation in *TSG101* affects the efficiency of TSG101-mediated release of viral particles from infected cells,
thereby altering levels of plasma viral load and subsequent disease progression.**

One fundamental feature of a successful virus is its ability to
utilize the host cellular machinery in order to support its prop-
agation. Understanding the interactions between host and viral
proteins provides the opportunity to identify means for con-
trolling outcomes of viral infection. Recent studies have dis-
closed a variety of cellular molecules that are exploited by
human immunodeficiency virus type 1 (HIV-1) at different
stages of its life cycle, one of which, the tumor suppressor gene
101 (TSG101) protein, is essential for budding of the virus
from infected cells (10, 12, 21, 35).

TSG101 is an evolutionarily conserved gene located on hu-
man chromosome 11p15. It encodes a 46-kDa multidomain
protein that contains an N-terminal ubiquitin-conjugating en-
zyme E2 variant (UEV) domain, a proline-rich domain, a
coiled-coil region including a leucine zipper, and a C-terminal
α-helical domain. The TSG101 protein has been detected in
the nucleus and cytoplasm, and its localization is cell cycle
dependent (38, 39).

A potential role for *TSG101* as a tumor suppressor that
was suggested in an early study of the gene (17) remains
controversial (22, 34). Identification of molecules with
which TSG101 interacts has suggested its involvement in
transcriptional regulation (5, 13, 24, 32) and cell cycle con-
trol (6, 18, 27). Targeted deletion of *tsg101* in mice results in
early embryonic death due to a defect in cellular prolifera-
tion (29), and reports of cell cycle arrest and death in
TSG101-deficient cells have further confirmed a critical role
for TSG101 in cell survival (6, 16).

Many recent studies of TSG101 have focused on its role in
endosomal trafficking. TSG101 and its yeast orthologue,
Vps23, belong to the so-called "class E" proteins whose func-
tions are essential for vacuolar protein sorting (1). TSG101/
Vps23, along with two other proteins, Vps28 and Vps37, form
a ~350-kDa complex named ESCRT-I (*e*ndosomal *s*orting
*c*omplex *r*equired for *t*ransport) (14). ESCRT-I is involved in a
series of protein-protein interactions that result in sorting of
ubiquitylated proteins from early endosomes into multivesicu-
lar bodies (MVBs). During this process, early endosomes car-
rying protein cargo bud into MVBs, organelles that eventually
fuse with lysosomes for subsequent protein degradation. This
process entails direct binding of the TSG101/Vps23 UEV do-
main to ubiquitin, resulting in the delivery of ubiquitylated

* Corresponding author. Mailing address: Laboratory of Genomic
Diversity, SAIC-Frederick, Inc., National Cancer Institute, Frederick,
Maryland 21702. Phone: (301) 846-1390. Fax: (301) 846-1909. E-mail:
carringt@mail.ncifcrf.gov.
‡ Present address: Johns Hopkins University School of Medicine,
Baltimore, MD 21231.

proteins (that are destined for degradation) into MVBs (2, 12, 14, 20, 28, 33).

Some viruses have developed an ingenious mechanism for budding from cells that involves the host endosomal sorting process. This became evident after the discovery of the direct interaction between the UEV domain of TSG101 and a highly conserved motif in the p6 region of the HIV-1 Gag protein, Pro-Thr/Ser-Ala-Pro (PTAP), an interaction that was shown to be critical for the release of HIV-1 particles from the cellular membrane (10, 12, 21, 35). The HIV-1 PTAP motif belongs to a family of late (L) domains, so named for their late involvement in the viral life cycle. The viral L domains are also characterized by conserved PPXY and YXXL motifs, which, along with PTAP, interact with host proteins involved in the MVB pathway (reviewed in reference 9). The TSG101–HIV-1 relationship has been studied intensively by several laboratories, resulting in the identification of additional host proteins involved in the process of HIV-1 budding (30, 31, 36). The TSG101 molecule has also been shown to be involved in the budding of HIV-2 (25), Ebola virus (21), human T-cell leukemia virus type 1 (4), and bluetongue virus (37).

Given the central role of TSG101 in release of HIV-1 from infected cells, we hypothesized that genetic variations in *TSG101* could potentially affect the functional activity of TSG101 protein in viral budding, thereby altering levels of circulating virus in the blood of infected individuals and the clinical course of AIDS. Here, we report two single-nucleotide polymorphism (SNP) variants, located at positions −183 and +181 relative to the translation start site, that associate with differences in viral load dynamics, in CD4 T-cell decline, and, correspondingly, with the rate of AIDS progression after infection. The −183 variant has been recently reported to be associated with faster CD4 decline in the Swiss HIV Cohort Study (SHCS) (3). Here, we demonstrate a more detailed analysis of the association between *TSG101* variation and several outcomes of HIV-1 infection, which are consistent in the SHCS and a large sample of U.S. AIDS cohorts.

MATERIALS AND METHODS

Subjects. The study group included patients from five U.S. cohorts and a Swiss cohort: the AIDS Linked to the Intravenous Drug Experience Study, the Hemophilia Growth and Development Study, the Multicenter AIDS Cohort Study (MACS), the Multicenter Hemophilia Cohort Study, the San Francisco City Clinic Cohort Study (26), and the SHCS (http://www.shcs.ch). The seroconversion date was estimated as the midpoint between the first positive and the last negative HIV-1 antibody test (mean interval, 0.79 years; range, 0.07 to 3.0 years). High-risk exposed uninfected subjects were defined as those who reported engaging in behavior that confers high risk for HIV-1 exposure (continual sharing of injection equipment or anal receptive sex with multiple partners) or transfusions with factor VIII known to be contaminated with HIV-1. Longitudinal CD4 count values were available for MACS and SHCS, whereas viral load measurements were available only for the MACS cohort.

Genotyping. The single-strand conformation polymorphism technique was performed as described by Cullen et al. (8). The *TSG101* SNPs were typed using TaqMan Genotyping Assays (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol. Haplotypes were estimated in compound heterozygotes using the expectation maximization algorithm (19) and SAS software (SAS Institute, Cary, NC).

Statistical analysis. Four end points reflecting disease progression (AIDS outcome) were evaluated: time to CD4 < 200 cells/mm³; progression to AIDS according to the 1987 definition by the Centers for Disease Control and Prevention; progression to AIDS according to the 1993 definition by CDC; and AIDS-related death. Defined categorical analysis was performed comparing genotype and haplotype frequencies between groups of patients using a two-tailed Fisher's

exact test. Analyses presented (see Tables 1 to 3, Fig. 2, and tables in the supplemental material) were carried out using the SAS software. Participants were stratified by ethnic group, sex, and age at seroconversion (0 to 20, 20 to 40, and over 40 years).

CD4 decline (see Fig. 3) and HIV RNA increase (see Fig. 4) over time were assessed with random effects linear models in which the intercepts were allowed to vary randomly for each participant. These models accounted for the correlation between serial measurements from each person and were implemented using the lme 4 package in R (http://www.r-project.org/). The three genotypic categories were included as indicator variables in the models, and both the intercepts and slopes were estimated. These models were adjusted by age, race, and injecting drug use reported at either seroconversion or when CD4 fell between 500 and 600 cells/mm³ for MACS and SHCS, respectively. All CD4 and HIV RNA data were collected prior to report of highly active antiretroviral therapy use. For MACS patients included in these analyses, the seroconversion date was identified within a 2-year window. For over 90% of the 380 men included in the CD4 decline analysis, the seroconversion was known within a 1-year interval (median [interquartile range] = 0.51 [0.48 to 0.58] years).

RESULTS

Identification of *TSG101* SNPs and haplotypes. An initial screening of the *TSG101* coding region in 50 healthy Caucasian blood donors using the single-strand conformation polymorphism technique indicated the highly conserved nature of the gene, since no nonsynonymous nucleotide changes were identified. Sequencing of 79 randomly chosen Caucasian seroconverter patients (U.S. AIDS cohorts) was then performed in the 5′ area of the gene, which is likely to contain regulatory sequences. The sequenced fragment consisted of ~2 kb around exon 1, where nine SNPs were identified (Fig. 1A). Analysis of the SNP genotypes revealed six haplotypes with estimated frequencies of >1% that could be defined by four haplotype-tagging SNPs (Fig. 1B) (positions −600 [rs3802966], −518 [rs1857909], −183 [rs2292179], and +181 [rs1395319] relative to the translation start site).

Analysis of *TSG101* variants in the AIDS cohorts. The four *TSG101* haplotype-tagging SNPs located in the 5′ area were typed in the U.S. AIDS cohorts. Haplotypes based on the four SNPs were estimated in a large population of 1,895 Caucasian individuals (Fig. 1C). The frequencies of the haplotypes in Caucasians differed only slightly from the initial estimation in 79 individuals (Fig. 1B).

The association of the individual SNPs and the five related haplotypes with disease progression was tested using the Cox proportional hazards model (7) and categorical analysis in which frequencies of genotypes were compared in people who developed AIDS during certain time periods. The Cox model did not reveal any significant effect for the four individual SNPs or for the five haplotypes (see Tables S1 and S2 in the supplemental material). However, we did observe differences in the frequencies of the −600G and −183C variants among individuals who developed AIDS before 7 years compared to those who developed AIDS later (see Table S3 in the supplemental material). In similar analyses, haplotype 2 conferred a susceptible effect (odds ratio [OR] = 1.6 to 2.5; P = 0.002 to 0.01), haplotypes 3 and 5 were protective (OR = 0.6 to 0.8, P = 0.007 to 0.2, and OR = 0.5 to 0.8, P = 0.06 to 0.5), while haplotype 4 was relatively neutral (OR = 0.8 to 1.1; P = 0.2 to 0.8) compared to a reference haplotype 1 (see Table S4 in the supplemental material). Of note, the −600G variant, which is in strong linkage disequilibrium (LD) with the −183T and +181C variants, specifies haplotype 3, and the protective effect
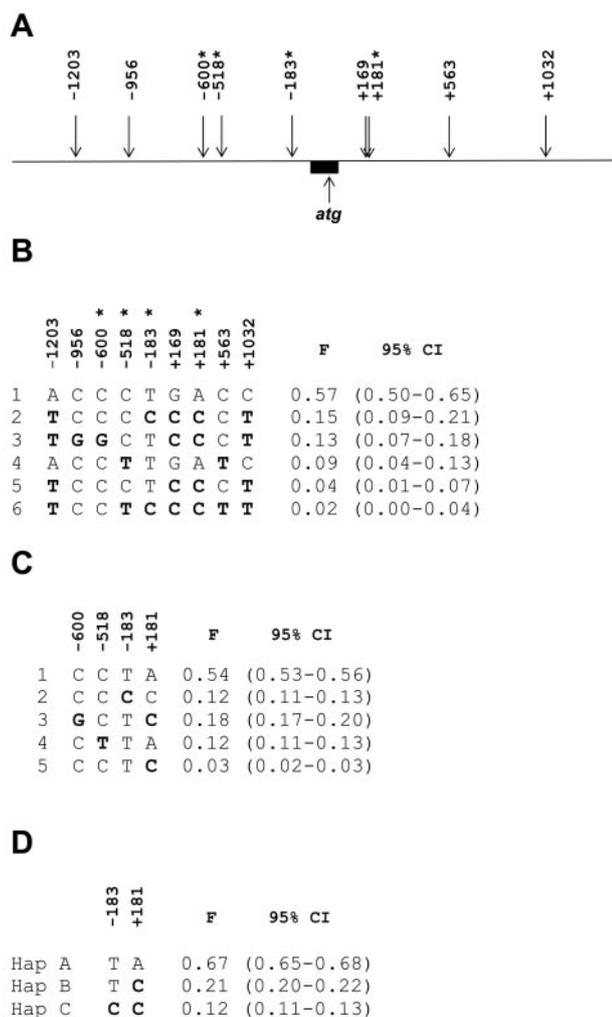
FIG. 1. SNPs and the corresponding haplotypes observed in the 5′ area of *TSG101*. (A) Schematic map of the nine SNPs identified in the 5′ area of *TSG101*. SNP positions are determined relative to the "A" nucleotide of the ATG start codon, which is shown within exon 1 (black box). (B) Haplotypes based on the nine SNPs. The haplotype frequencies were estimated based on sequencing data among 79 randomly chosen Caucasian seroconverters. Only haplotypes with frequencies of >0.01 are listed. Alleles with minor frequencies are shown in boldface. F, frequency; CI, confidence interval. (C) Haplotypes based on four SNPs estimated in 1,895 Caucasian individuals. (D) Structure and frequencies of haplotypes A, B, and C estimated in 2,071 Caucasian individuals.

of this haplotype corresponded to the effect of the individual SNP. Based on these data and the haplotype structures, we concluded that susceptibility and protection are associated with variation at positions −183 and +181. Three of the four possible haplotypes composed of these two SNPs were observed, indicating strong LD between the variants. The corresponding haplotypes, T-A, T-C, and C-C, were termed A, B, and C, respectively (Fig. 1D).

Haplotypes B and C are associated with different rates of AIDS progression. The frequencies of haplotypes A, B, and C among individuals who developed AIDS in one of six distinct time periods after seroconversion (≤3, 3 to 5, 5 to 7, 7 to 10, 10 to 12, and >12 years) were compared (data not shown). In

this analysis, seroprevalent individuals were included in the last three groups depending on the date of an AIDS-defining outcome after their first HIV⁺ visit. This analysis suggested that the effect of the *TSG101* haplotypes on HIV disease is not gradual over time, an observation that was further elucidated by the absence of a significant effect of these variants on AIDS progression in survival analysis using the Cox model, where time is a continuous variable starting from seroconversion (see Table S5 in the supplemental material). Rather, differences in susceptible (C) and protective (B) haplotypic frequencies appeared somewhat bimodal in that the frequency of the susceptible haplotypic group was significantly greater among individuals who progressed to AIDS within 7 years after seroconversion relative to those who remained AIDS-free for at least 7 years (dominant model; OR = 1.72 to 2.30; P = 0.0001 to 0.002) (Table 1). Conversely, the protective haplotypes were observed significantly more frequently among those who remained AIDS free for 7 years or longer after seroconversion compared to those who had progressed within 7 years (dominant model; OR = 0.56 to 0.64; P = 0.0005 to 0.01). Both the B and C haplotypes appeared to have dominant effects, since a codominant model did not fit the data as well as a dominant model. Haplotype A was relatively neutral (Table 1).

Protective, susceptible, and neutral *TSG101* haplotype groups. Haplotypes B and C exhibited opposite dominant effects over the neutral haplotype A (Table 1), so it follows that the haplogenotypes A/B and B/B would both be protective, and A/C and C/C would both confer susceptibility in terms of AIDS progression before or after 7 years after seroconversion. We hypothesized that the dominant effects of haplotypes B and C would result in a neutral phenotype among B/C heterozygotes similar to the A/A haplogenotype. Thus, the three nonoverlapping groups based on genotypic data were tested for their effects on AIDS progression. As expected, the A/C-C/C grouping showed a strong susceptible effect (OR = 1.74 to 2.61; P = <0.0001 to 0.003), A/B-B/B associated with protection (OR =

TABLE 1. Effects of *TSG101* haplotypes on AIDS progression among Caucasians from the combined U.S. AIDS cohorts (dominant model)

Haplotype	AIDS outcome	No.[a] of fast progressors[b] (F)	No.[a] of slow progressors[c] (F)	OR	P value
A	CD4 < 200	218 (0.88)	637 (0.89)	0.94	0.71
	AIDS-1993	245 (0.87)	775 (0.89)	0.91	0.54
	AIDS-1987	132 (0.86)	995 (0.88)	1.06	0.80
	Death	92 (0.90)	1169 (0.88)	1.19	0.41
B	CD4 < 200	74 (0.30)	286 (0.40)	0.64	0.005
	AIDS-1993	82 (0.29)	353 (0.41)	0.60	0.0005
	AIDS-1987	44 (0.29)	451 (0.40)	0.60	0.006
	Death	27 (0.26)	523 (0.39)	0.56	0.01
C	CD4 < 200	77 (0.31)	145 (0.20)	1.78	0.0006
	AIDS-1993	89 (0.32)	184 (0.21)	1.72	0.0005
	AIDS-1987	52 (0.34)	251 (0.22)	1.78	0.002
	Death	40 (0.39)	291 (0.22)	2.30	0.0001

[a] Numbers of individuals who had at least one of the corresponding haplotypes (dominant model).
[b] Seroconverter patients who progressed to an AIDS outcome in ≤7 years after HIV-1 infection.
[c] Patients who avoided an AIDS outcome for >7 years after HIV-1 infection (seroconverters and seroprevalent individuals).

TABLE 2. *TSG101* haplogenotypes and AIDS progression among Caucasians from the combined U.S. AIDS cohorts

Haplogenotype	AIDS outcome	No.[a] of fast progressors[b] (F)	No.[a] of slow progressors[c] (F)	OR	P value
Neutral (A/A or B/C)	CD4 < 200	133 (0.54)	357 (0.50)	1.17	0.30
	AIDS-1993	147 (0.52)	426 (0.49)	1.14	0.37
	AIDS-1987	72 (0.47)	560 (0.50)	0.89	0.55
	Death	47 (0.46)	668 (0.50)	0.85	0.47
Protective (A/B or B/B)	CD4 < 200	56 (0.23)	251 (0.35)	0.54	0.0003
	AIDS-1993	64 (0.23)	307 (0.35)	0.54	<0.0001
	AIDS-1987	37 (0.24)	384 (0.34)	0.61	0.01
	Death	21 (0.21)	447 (0.34)	0.51	0.006
Susceptible (A/C or C/C)	CD4 < 200	59 (0.24)	109 (0.15)	1.74	0.003
	AIDS-1993	71 (0.25)	137 (0.16)	1.80	0.0005
	AIDS-1987	45 (0.29)	183 (0.16)	2.13	0.0002
	Death	34 (0.33)	214 (0.16)	2.61	<0.0001

[a] Numbers of individuals who had at least one of the corresponding haplotypes (dominant model).
[b] Seroconverter patients who progressed to an AIDS outcome in ≤7 years after HIV-1 infection.
[c] Patients who avoided an AIDS outcome for >7 years after HIV-1 infection (seroconverters and seroprevalent individuals).

0.51 to 0.61; P = <0.0001 to 0.01), and the A/A-B/C grouping appeared neutral (Table 2). The protective and susceptible haplotype groups were further compared to the neutral group by the Cox proportional-hazards model (Fig. 2). The relative shapes of the curves in the Kaplan-Meier plots confirm that the effects of the *TSG101* variants are not constant over time, but rather occur most obviously between about 4 and 12 years after seroconversion.
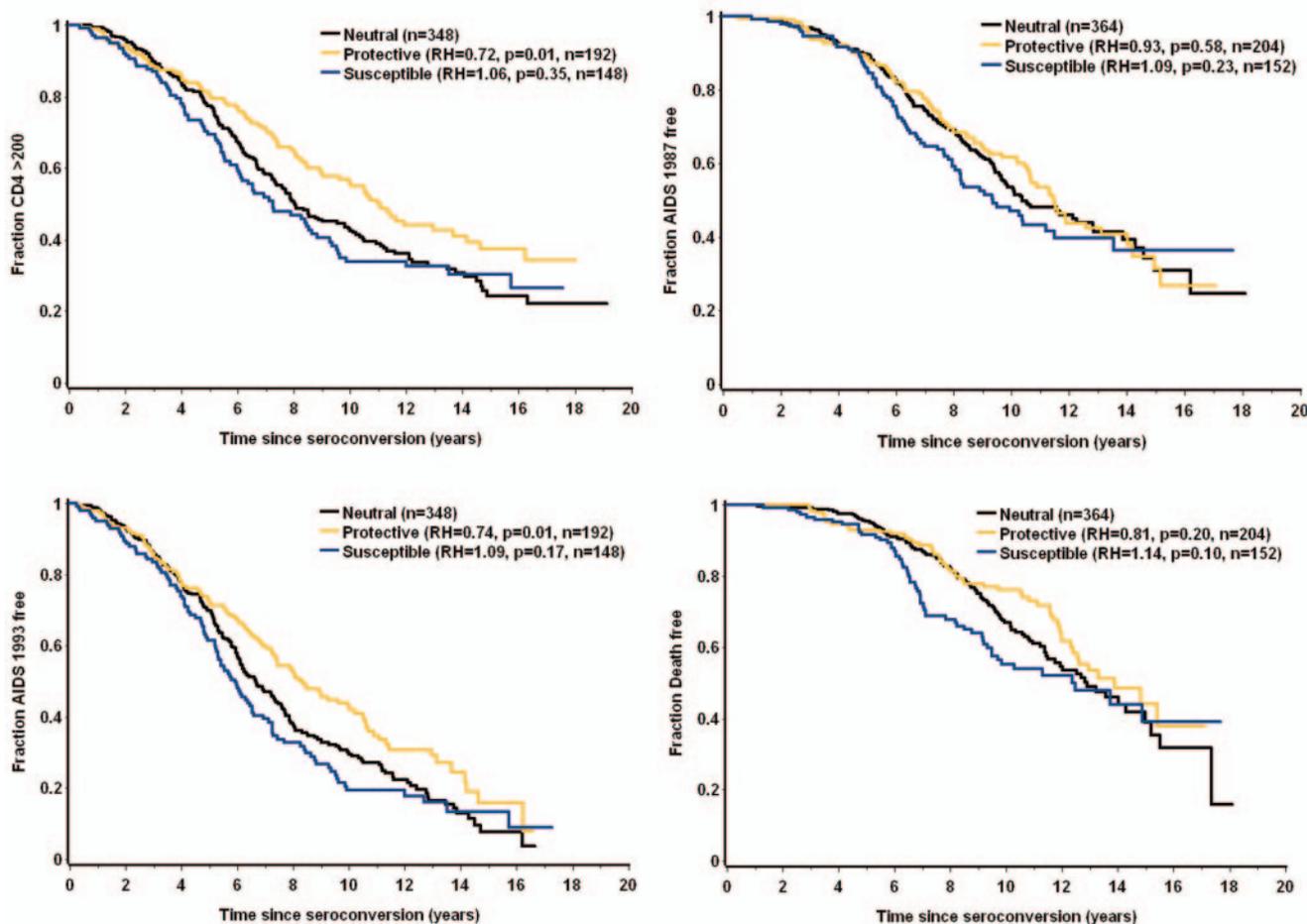

FIG. 2. Kaplan-Meier plots for the three *TSG101* haplotypic groups. Four AIDS outcomes were analyzed using the Cox proportional-hazards model.
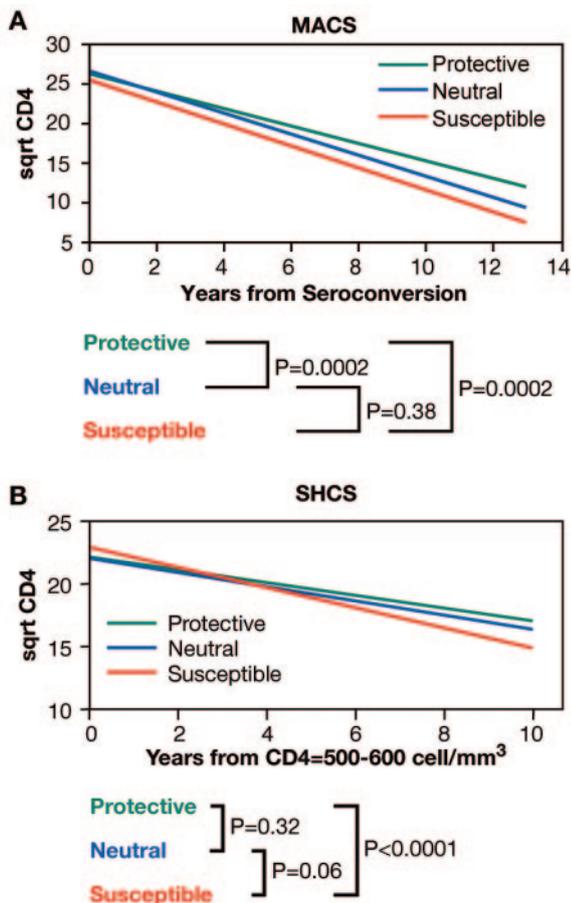
FIG. 3. Decline in CD4 T-cell count among AIDS patients with different *TSG101* haplogenotypes. The data were analyzed using a random-effects linear model (see Materials and Methods). (A) Square-root CD4 counts for 380 individuals from the MACS cohort were plotted as a function of time from seroconversion, and fitted lines for each genotype group were generated based on 4,918 measurements. Slopes for protective, neutral, and susceptible genotypic groups are as follows (95% CI in parentheses): -1.11 (-1.22, -1.00), -1.33 (-1.42, -1.25), and -1.39 (-1.50, -1.27), respectively. (B) CD4 data from the SHCS patients were plotted starting from the point at which CD4 cell counts fell in the range of 500 to 600 cells/mm^3. By chance, the susceptible haplogenotypic group had a mean CD4 count that was higher in this range than that of the protective or neutral group. The fitted lines are based on 3,551 measurements for 310 people. Slopes (95% CI) for protective, neutral, and susceptible groups are -0.52 (-0.64, -0.39), -0.57 (-0.69, -0.47), and -0.81 (-0.99, -0.64), respectively.

Individuals from different *TSG101* haplotype groups exhibit different rates of CD4$^+$ T-cell decline. The availability of longitudinal CD4 T-cell counts in two of our AIDS cohorts, the MACS and the SHCS (http://www.shcs.ch), provided the opportunity to determine whether the influence of *TSG101* variants on AIDS progression might involve differential rates of CD4 T-cell decline. CD4 measurements over time were plotted, and the slopes of the fitted lines were compared between the three genotypic groups. For 380 MACS patients, CD4 T-cell decline was estimated from measurements obtained over a period of 13 years since seroconversion (Fig. 3A). An average of 12.9 CD4 measurements per individual, ranging from 2 to
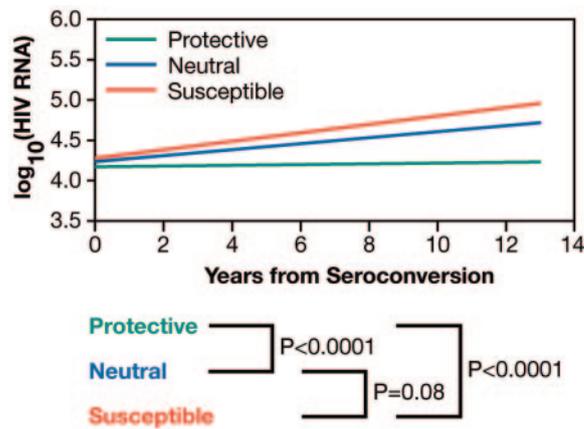

FIG. 4. Viral-load increase over time in 373 MACS patients. Fitted lines for the three *TSG101* genotypic groups were constructed based on 3,569 measurements of log$_{10}$ HIV RNA over time after seroconversion. Statistical analysis was performed using the multivariate random-effects linear model (see Materials and Methods). The best-fit lines for the protective, neutral, and susceptible haplogenotypic groups have slopes (95% CI) of 0.01 (-0.01, 0.02), 0.04 (0.03, 0.05), and 0.05 (0.04, 0.07), respectively.

32, was considered in the analysis. The SHCS cohort primarily consisted of seroprevalent individuals, so in this case, CD4 T-cell decline was measured over a period of 10 years starting at the point where CD4 counts fell in the range of 500 to 600 cells/mm^3 (Fig. 3B). For 310 SHCS patients, an average of 11.5 CD4 T-cell count data points (ranging from 2 to 52) per individual were available for analysis. In both cohorts, the effects of the three genotypic groups on CD4 T-cell decline corresponded to the effects these genotypes had on AIDS progression, where the decline was progressively steeper in the following order: A/C-C/C (susceptible group) > A/A-B/C (neutral group) > A/B-B/B (protective group) (Fig. 3). Differences in CD4 T-cell decline between the protective and susceptible groups were highly significant in both cohorts (P = <0.0001 to 0.0002).

Effects of the *TSG101* haplotypes on viral load increase over time. Given the central role of TSG101 in HIV-1 budding and the known correlation between the viral load and the rate of AIDS progression (23), we tested whether the effect of *TSG101* variation on CD4 T-cell decline and AIDS progression was also reflected in viral load changes over time. For this analysis, the increase in viral load over time among 373 MACS patients was measured, stratified by the three *TSG101* genotypic groups (Fig. 4). An average of 9.6 measurements (ranging from 2 to 28 measurements) of viral load per patient over a period of 13 years since seroconversion were available for the analysis. No significant difference in HIV-1 RNA levels between the three genotypic groups was observed at the time of seroconversion. Strikingly, however, highly significant differences in log$_{10}$ HIV-1 RNA slopes were observed between the protective and susceptible groups (P < 0.0001), strongly suggesting that variation in the *TSG101* gene affects the HIV-1 viral load, potentially through the differential efficiency of *TSG101* variants to mediate viral budding.

TSG101 genotypes and HIV-1 infection. The *TSG101* genotypic groups were also tested for potential effects on HIV-1

TABLE 3. Frequencies of *TSG101* genotypic groups among
HIV-1-positive vs. HIV-1-negative Caucasian individuals

Genotypic group	% SC[a]/HREU[b] (OR, 95% CI)	% SC/SN[c] (OR, 95% CI)	% SC/OSN[d]/HREU (P value for trend[e])
Neutral	51/51 (1.0, 0.7–1.5)	51/48 (1.1, 0.8–1.5)	51/45/51 (0.76)
Protective	28/36 (0.7, 0.5–1.0)	28/36 (0.7, 0.5–0.9)	28/36/36 (0.03)
Susceptible	21/13 (1.7, 1.0–2.9)	21/16 (1.5, 1.0–2.2)	21/19/13 (0.04)

[a] SC, seroconverters (n = 728).
[b] HREU, high-risk exposed uninfected people (n = 134).
[c] SN, all seronegatives combined (n = 219).
[d] OSN, seronegatives other than HREU (n = 85).
[e] Mantel-Haentzel chi-square test.

infection by comparing the distributions of these genotypes in HIV⁺ patients with those in seronegative or high-risk exposed uninfected individuals. Although only marginally significant values (P = 0.02 to 0.08) were determined (Table 3), the effects of these genotypes on HIV-1 infection each paralleled their respective effects on the viral load, on CD4 T-cell decline, and on the rate of progression to AIDS.

DISCUSSION

The consistent protective (haplotype B) and susceptible (haplotype C) influences of *TSG101* haplotypes on multiple outcomes after HIV-1 exposure, including longitudinal viral load levels, CD4⁺ T-cell decline, and subsequent disease progression, as well as a moderate effect on HIV-1 infection, support a physiological role for genetic variation near/within the *TSG101* gene in HIV-1 pathogenesis. Whether the (−183, +181) haplotypic variants have a direct effect on these outcomes remains in question. We did not find any nonsynonymous polymorphisms or 3′ untranslated region SNPs linked to the −183, +181 variants that could explain the observed effects of these variants on HIV-1 disease. Further, the (−183, +181) haplogenotypes did not show any significant difference in the levels of *TSG101* mRNA transcription in peripheral blood lymphocytes or purified CD4 cells as tested by real-time PCR, and no differences were observed in promoter activity using reporter constructs (data not shown). Likewise, no alternatively spliced forms of *TSG101* mRNA associating with the genotypes were observed, negating splice variation as the underlying mechanism of the associations described. Although we were not able to detect an influence of the *TSG101* haplotypes on gene transcription or mRNA splicing, we cannot rule out the possibility that the *TSG101* SNPs may affect transcription/splicing under specific physiological conditions.

The protective and susceptible effects of *TSG101* haplotypes on AIDS progression are not constant over time, as indicated by the Kaplan-Meier curves (Fig. 2). The relative shapes of the curves suggest that *TSG101* variation has little or no effect during the early (~4 years) and late (>12 years) stages of infection, but rather, only at an intermediate time period. Other host genetic, viral, or environmental factors that affect disease progression during the early and late stages of infection may override effects of *TSG101*. Alternatively, the TSG101 interaction with HIV and/or consequences of this interaction could be different during these two extreme time intervals compared with the intermediate period. However, the effects of *TSG101* haplotypes occurring during the intermediate time

interval were strong enough to be evident when the entire patient cohort was used in the categorical analysis, as well as the longitudinal analyses of CD4 T-cell decline and viral-load increase.

Recently, the −183 variant, corresponding to haplotype C, was shown to associate with lower virus production ex vivo, a paradoxical finding given its association with faster CD4 T-cell decline (3) and susceptibility to AIDS reported here. Although this finding appears to contradict the genetic epidemiological findings presented here, the ex vivo assay may not be physiologically relevant to the described effect of *TSG101* on AIDS progression given its time dependence as discussed above.

The HapMap genotype data (http://www.hapmap.org) suggest that *TSG101* is located in a region of strong LD: the corresponding haplotype block defined according to Gabriel et al. (11) spans 118 kb. This block includes the entire *TSG101* gene; its partial paralogue, *UEV-3* (15); and two additional gene fragments (the lactate dehydrogenase A-like 6A gene [GenBank accession no. NM144972] and a computationally predicted gene that may encode a protein similar to the mitochondrial carrier homolog 1 [XM497268]). Therefore, the (−183, +181) haplogenotypes may mark the true disease variant(s) through LD. In support of this, no effect on disease progression was detected in a smaller sample of African-American seroconverters (although longitudinal CD4 and viral-load data from these individuals were not available for analysis; also, the susceptible haplotype group was observed at a frequency of only 4%). However, given the requirement of TSG101 for HIV-1 budding in vitro, it seems likely that the effect described herein is due to variation in *TSG101* and not to polymorphism in a neighboring locus. If so, this study represents the first genetic epidemiological evidence to support previous in vitro studies indicating a primary role for TSG101 in HIV-1 pathogenesis.

REFERENCES

1. **Babst, M., G. Odorizzi, E. J. Estepa, and S. D. Emr.** 2000. Mammalian tumor susceptibility gene 101 (TSG101) and the yeast homologue, Vps23p, both function in late endosomal trafficking. Traffic **1:**248–258.
2. **Bishop, N., A. Horman, and P. Woodman.** 2002. Mammalian class E vps proteins recognize ubiquitin and act in the removal of endosomal protein-ubiquitin conjugates. J. Cell Biol. **157:**91–101.
3. **Bleiber, G., M. May, R. Martinez, P. Meylan, J. Ott, J. S. Beckmann, and A. Telenti.** 2005. Use of a combined ex vivo/in vivo population approach for screening of human genes involved in the human immunodeficiency virus type 1 life cycle for variants influencing disease progression. J. Virol. **79:**12674–12680.
4. **Bouamr, F., J. A. Melillo, M. Q. Wang, K. Nagashima, M. de Los Santos, A. Rein, and S. P. Goff.** 2003. PPPYVEPTAP motif is the late domain of human T-cell leukemia virus type 1 Gag and mediates its functional interaction with cellular proteins Nedd4 and Tsg101. J. Virol. **77:**11882–11895.
5. **Burgdorf, S., P. Leister, and K. H. Scheidtmann.** 2004. TSG101 interacts with apoptosis-antagonizing transcription factor and enhances androgen re-

ceptor-mediated transcription by promoting its monoubiquitination. J. Biol. Chem. **279:**17524–17534.

6. **Carstens, M. J., A. Krempler, A. A. Triplett, M. Van Lohuizen, and K. U. Wagner.** 2004. Cell cycle arrest and cell death are controlled by p53-dependent and p53-independent mechanisms in Tsg101-deficient cells. J. Biol. Chem. **279:**35984–35994.

7. **Cox, D. R., and D. Oakes.** 1984. Analysis of survival data. Chapman and Hall, New York, N.Y.

8. **Cullen, M., J. Noble, H. Erlich, K. Thorpe, S. Beck, W. Klitz, J. Trowsdale, and M. Carrington.** 1997. Characterization of recombination in the HLA class II region. Am. J. Hum. Genet. **60:**397–407.

9. **Demirov, D. G., and E. O. Freed.** 2004. Retrovirus budding. Virus Res. **106:**87–102.

10. **Demirov, D. G., A. Ono, J. M. Orenstein, and E. O. Freed.** 2002. Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. Proc. Natl. Acad. Sci. USA **99:**955–960.

11. **Gabriel, S. B., S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, and D. Altshuler.** 2002. The structure of haplotype blocks in the human genome. Science **296:**2225–2229.

12. **Garrus, J. E., U. K. von Schwedler, O. W. Pornillos, S. G. Morham, K. H. Zavitz, H. E. Wang, D. A. Wettstein, K. M. Stray, M. Cote, R. L. Rich, D. G. Myszka, and W. I. Sundquist.** 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell **107:**55–65.

13. **Hittelman, A. B., D. Burakov, J. A. Iniguez-Lluhi, L. P. Freedman, and M. J. Garabedian.** 1999. Differential regulation of glucocorticoid receptor transcriptional activation via AF-1-associated proteins. EMBO J. **18:**5380–5388.

14. **Katzmann, D. J., M. Babst, and S. D. Emr.** 2001. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell **106:**145–155.

15. **Kloor, M., P. Bork, A. Duwe, R. Klaes, M. von Knebel Doeberitz, and R. Ridder.** 2002. Identification and characterization of UEV3, a human cDNA with similarities to inactive E2 ubiquitin-conjugating enzymes. Biochim. Biophys. Acta **1579:**219–224.

16. **Krempler, A., M. D. Henry, A. A. Triplett, and K. U. Wagner.** 2002. Targeted deletion of the Tsg101 gene results in cell cycle arrest at G₁/S and p53-independent cell death. J. Biol. Chem. **277:**43216–43223.

17. **Li, L., and S. N. Cohen.** 1996. Tsg101: a novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells. Cell **85:**319–329.

18. **Li, L., J. Liao, J. Ruland, T. W. Mak, and S. N. Cohen.** 2001. A TSG101/MDM2 regulatory loop modulates MDM2 degradation and MDM2/p53 feedback control. Proc. Natl. Acad. Sci. USA **98:**1619–1624.

19. **Long, J. C., R. C. Williams, and M. Urbanek.** 1995. An E-M algorithm and testing strategy for multiple-locus haplotypes. Am. J. Hum. Genet. **56:**799–810.

20. **Lu, Q., L. W. Hope, M. Brasch, C. Reinhard, and S. N. Cohen.** 2003. TSG101 interaction with HRS mediates endosomal trafficking and receptor downregulation. Proc. Natl. Acad. Sci. USA **100:**7626–7631.

21. **Martin-Serrano, J., T. Zang, and P. D. Bieniasz.** 2001. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat. Med. **7:**1313–1319.

22. **McIver, B., S. K. Grebe, L. Wang, I. D. Hay, A. Yokomizo, W. Liu, J. R. Goellner, C. S. Grant, D. I. Smith, and N. L. Eberhardt.** 2000. FHIT and TSG101 in thyroid tumours: aberrant transcripts reflect rare abnormal RNA processing events of uncertain pathogenetic or clinical significance. Clin. Endocrinol. **52:**749–757.

23. **Mellors, J. W., C. R. Rinaldo, Jr., P. Gupta, R. M. White, J. A. Todd, and L. A. Kingsley.** 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science **272:**1167–1170.

24. **Muromoto, R., K. Sugiyama, T. Yamamoto, K. Oritani, K. Shimoda, and T. Matsuda.** 2004. Physical and functional interactions between Daxx and TSG101. Biochem. Biophys. Res. Commun. **316:**827–833.

25. **Myers, E. L., and J. F. Allen.** 2002. Tsg101, an inactive homologue of ubiquitin ligase e2, interacts specifically with human immunodeficiency virus type 2 gag polyprotein and results in increased levels of ubiquitinated gag. J. Virol. **76:**11226–11235.

26. **O'Brien, S. J., and G. W. Nelson.** 2004. Human genes that limit AIDS. Nat. Genet. **36:**565–574.

27. **Oh, H., C. Mammucari, A. Nenci, S. Cabodi, S. N. Cohen, and G. P. Dotto.** 2002. Negative regulation of cell growth and differentiation by TSG101 through association with p21(Cip1/WAF1). Proc. Natl. Acad. Sci. USA **99:**5430–5435.

28. **Pornillos, O., S. L. Alam, R. L. Rich, D. G. Myszka, D. R. Davis, and W. I. Sundquist.** 2002. Structure and functional interactions of the Tsg101 UEV domain. EMBO J. **21:**2397–2406.

29. **Ruland, J., C. Sirard, A. Elia, D. MacPherson, A. Wakeham, L. Li, J. L. de la Pompa, S. N. Cohen, and T. W. Mak.** 2001. p53 accumulation, defective cell proliferation, and early embryonic lethality in mice lacking tsg101. Proc. Natl. Acad. Sci. USA **98:**1859–1864.

30. **Strack, B., A. Calistri, S. Craig, E. Popova, and H. G. Gottlinger.** 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell **114:**689–699.

31. **Stuchell, M. D., J. E. Garrus, B. Muller, K. M. Stray, S. Ghaffarian, R. McKinnon, H. G. Krausslich, S. G. Morham, and W. I. Sundquist.** 2004. The human endosomal sorting complex required for transport (ESCRT-I) and its role in HIV-1 budding. J. Biol. Chem. **279:**36059–36071.

32. **Sun, Z., J. Pan, W. X. Hope, S. N. Cohen, and S. P. Balk.** 1999. Tumor susceptibility gene 101 protein represses androgen receptor transactivation and interacts with p300. Cancer **86:**689–696.

33. **Sundquist, W. I., H. L. Schubert, B. N. Kelly, G. C. Hill, J. M. Holton, and C. P. Hill.** 2004. Ubiquitin recognition by the human TSG101 protein. Mol. Cell **13:**783–789.

34. **Trink, B., S. I. Pai, S. L. Spunt, V. Raman, P. Cairns, J. Jen, E. Gabrielson, S. Sukumar, and D. Sidransky.** 1998. Absence of TSG101 transcript abnormalities in human cancers. Oncogene **16:**2815–2818.

35. **VerPlank, L., F. Bouamr, T. J. LaGrassa, B. Agresta, A. Kikonyogo, J. Leis, and C. A. Carter.** 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). Proc. Natl. Acad. Sci. USA **98:**7724–7729.

36. **von Schwedler, U. K., M. Stuchell, B. Muller, D. M. Ward, H. Y. Chung, E. Morita, H. E. Wang, T. Davis, G. P. He, D. M. Cimbora, A. Scott, H. G. Krausslich, J. Kaplan, S. G. Morham, and W. I. Sundquist.** 2003. The protein network of HIV budding. Cell **114:**701–713.

37. **Wirblich, C., B. Bhattacharya, and P. Roy.** 2006. Nonstructural protein 3 of bluetongue virus assists virus release by recruiting ESCRT-I protein Tsg101. J. Virol. **80:**460–473.

38. **Xie, W., L. Li, and S. N. Cohen.** 1998. Cell cycle-dependent subcellular localization of the TSG101 protein and mitotic and nuclear abnormalities associated with TSG101 deficiency. Proc. Natl. Acad. Sci. USA **95:**1595–1600.

39. **Zhong, Q., Y. Chen, D. Jones, and W. H. Lee.** 1998. Perturbation of TSG101 protein affects cell cycle progression. Cancer Res. **58:**2699–2702.