Published in final edited form as:

AIDS. 2011 September 10; 25(14): 1711–1719. doi:10.1097/QAD.0b013e328349c693.

Association of Polymorphisms in the LEDGF/p75 Gene (*PSIP1*) with Susceptibility to HIV-1 Infection and Disease Progression

Paradise Madlala^{a,b}, Rik Gijsbers^c, Frauke Christ^c, Anneleen Hombrouck^c, Lise Werner^d, Koleka Mlisana^d, Ping An^e, Salim S. Abdool Karim^d, Cheryl A. Winkler^e, Zeger Debyser^c, and Thumbi Ndung'u^{a,d}

^aHIV Pathogenesis Programme, University of KwaZulu-Natal, Durban, South Africa

^bDepartment of Genetics, University of KwaZulu-Natal, Pietermaritzburg, South Africa

^cMolecular Medicine, Katholieke Universiteit Leuven, B-3000 Leuven, Flanders, Belgium

^dCentre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa

^eBasic Research Laboratory, Science Applications International Corporation-Frederick, National Cancer Institute-Frederick, Frederick, Maryland, United States of America

Abstract

Objective—LEDGF/p75, encoded by the *PSIP1* gene, interacts with HIV-1 integrase and targets HIV-1 integration into active genes. We investigated the influence of polymorphisms in *PSIP1* on HIV-1 acquisition and disease progression in black South Africans.

Methods—Integrase binding domain (IBD) of LEDGF/p75 was sequenced in 126 participants. Four haplotype tagging SNPs, SNP1-SNP4 and one exonic SNP, SNP5 were genotyped in 195 HIV-1 seronegative, 52 primary and 403 chronically infected individuals using TaqMan assays. LEDGF/p75 expression was quantified by real-time RT-PCR. The impact of Q472L mutation on the interaction with HIV-1 IN was measured by AlphaScreen.

Results—rs2277191A was more frequent among seropositives ($p=0.06$, Fisher's exact test), and among individuals followed longitudinally trended towards association with higher likelihood of HIV-1 acquisition (RH=2.21, $p=0.08$; Cox model) and it was also associated with rapid disease progression (RH=5.98, $p=0.04$; Cox model). rs12339417C was associated with slower decline of CD4⁺ T cell ($p=0.02$) and lower levels of LEDGF/p75 ($p<0.01$). Seroconverters had higher preinfection levels of LEDGF/p75 ($p<0.01$) but levels decreased after HIV infection ($p=0.02$).

Conclusions—Genetic variants of *PSIP1* may affect HIV-1 outcomes. Further studies are needed to confirm the effect of genetic variation of *PSIP1* on HIV-1 pathogenesis in different cohorts.

Introduction

Human immunodeficiency virus type 1 (HIV-1) requires host cell for productive infection [1]. Lens epithelium-derived growth factor p75 (LEDGF/p75) also known as PC4 or SFRS1 interacting protein 1 (PSIP1), is ubiquitously expressed in all tissues and cell types (by convention, the gene and its protein are referred to as *PSIP1* and LEDGF/p75, respectively).

Author Contributions

PM, CW and ZD and TN conceived the study and designed the experiments. PM, RG, AH and PA performed the experiments. PM, RG, FC, AH, CW, ZD and TN analyzed and interpreted the data. SAK, KM designed and oversaw the CAPRISA 002 study cohort. TN oversaw the Sinikithemba study cohort. All authors critically reviewed the manuscript.

LEDGF/p75 is a member of the hepatoma-derived growth factor (HDGF)-related protein family (HRP-family) whose members are involved in chromosomal replication, transcription and chromatin structure [2-8]. LEDGF/p75 interacts with HIV-1 integrase (IN) through specific binding of the integrase binding domain (IBD) of LEDGF/p75 to the catalytic core domain of IN to tether HIV-1 to the chromosome and target HIV-1 integration into active genes [9-16]. Disruption of the interaction between LEDGF/p75 and IN, either by IN mutations or LEDGF/p75 knockdown, inhibits HIV-1 replication [17-20], confirming the important role of host LEDGF/p75 as an HIV-1 replication cofactor.

Association studies of the influence of human genetic variation on HIV-1 replication may reveal the essential *in vivo* host factors that interact with HIV-1 and their epidemiologic importance at the population level [21]. This approach has been used mostly in studies conducted on populations from developed countries [21-24]. However, there are differences in allele frequencies among potential disease influencing gene variants between ethnic groups and geographically separated populations [25]. Therefore, host genetic studies of HIV-1 infection need to be extended to developing world populations heavily burdened with HIV/AIDS. We investigated the influence of genetic variation in *PSIP1* on HIV-1 infection and disease progression in black South Africans.

Methods

Study participants

The Center for the AIDS Programme of Research in South Africa Acute Infection 002 (CAPRISA AI 002) [26, 27] and the Sinikithemba [28] cohorts were used for this study. The CAPRISA AI 002 cohort is an ongoing observational natural history study of HIV-1 subtype C infection established in Durban, KwaZulu-Natal, South Africa in 2004. HIV negative females (n=245) at high risk for HIV infection were enrolled into Phase I of the study. Participants in this cohort were screened monthly for recent HIV-1 infection by two rapid HIV-1 antibody tests (Abbott Laboratories, Tokyo, Japan) and Capillus (Trinity Biotech, Jamestown, NY, USA). HIV-1 antibody negative samples were tested for HIV-1 RNA in batches of 10 plasma samples per pool using the Ampliscreen v1.5 assay (Roche Diagnostics, Rotkreuz, Switzerland), which has a detection limit of 10 copies/ml. Samples that tested positive in pooled plasma were individually tested by quantitative RNA (Amplicor v2.0, Roche Diagnostics) and HIV enzyme immunoassay (BEP 2000; Dade Behring, Marburg, Germany) to identify HIV-1 infection. CD4⁺ T cell counts were determined by a 4-parameter FACSCalibur flow cytometer (Becton Dickinson). Participants with acute HIV-1 infection were enrolled into Phase II of the study on the basis of a reactive HIV antibody test within 3 months of previously negative results or positive HIV RNA PCR in the absence of antibodies. Date of infection was estimated by taking the midpoint between the last HIV antibody-negative result and the first HIV antibody-positive result or 14 days before the first positive HIV RNA PCR assay result for those identified as antibody negative but HIV RNA positive. An additional 34 acutely infected participants (who met the criteria for acute infection, as aforementioned) were recruited from other ongoing CAPRISA cohorts. Participants in Phase II were monitored weekly for 3 weeks, fortnightly for 2 months then monthly for 9 months and quarterly thereafter.

The Sinikithemba cohort comprises 450 antiretroviral naïve, HIV-1 subtype C chronically infected adults enrolled from McCord Hospital (Durban, South Africa) from August 2003 to 2008 and followed longitudinally [28, 29]. Sociodemographic characteristics, plasma viral load and CD4 cell count measurements were obtained at baseline. CD4 cell counts and viral loads were measured every 3 and 6 months respectively from enrollment. Viral loads were determined using the automated Cobas Amplicor HIV-1 Monitor test (version 1.5; Roche

Diagnostics). CD4⁺ T cells were enumerated using the Multitest kit (CD4/CD3/CD8/CD45) on a FACSCalibur flow cytometer (Becton Dickinson).

Ethical approval for this study was obtained from Biomedical Research Ethics Committee of the University of KwaZulu-Natal. All participants provided written informed consent.

Identification of polymorphisms in the C terminal region of *PSIP1*

Single nucleotide polymorphisms (SNPs) in the C terminal region of *PSIP1* were screened in a panel of 83 seronegative (SN) and 43 seropositive (SP) black South African women from the CAPRISA AI 002 cohort by resequencing (Table 1).

Genotyping of SNPs in *PSIP1* gene from the CAPRISA AI 002 participants

Four haplotype tagging (tag) SNPs rs2277191, rs1033056, rs12339417 and rs10283923 in this manuscript designated SNP1, SNP2, SNP3 and SNP4, respectively, available from NCBI dbSNP (<http://www.ncbi.nlm.nih.gov/SNP>) and HapMap databases (<http://www.hapmap.org>) selected by considering location, spacing, and allele frequency of at least 10% and one exonic SNP rs61744944 (Q472L) designated, SNP5 also available from the aforementioned databases (Fig. 1A, and Table 1) were genotyped in 247 patient samples (195 SNs and 52 SPs) from the CAPRISA AI 002 cohort. The tag SNPs were chosen because they are tag SNPs for *PSIP1* in the Yoruba from Nigeria (<http://www.snp.cshl.org>) and SNP5 based on the preliminary analysis of the sequencing data that suggested an association between SNP5 mutation and lower CD4 count and higher viral load ($P < 0.01$ for both, data not shown). Genotyping was performed by TaqMan SNP assay as per manufacturer's protocol (Applied Biosystems). TaqMan assays were obtained from the Assay-by-Demand service of Applied Biosystems (<http://www.appliedbiosystems.com>). Eight negative water controls were included in each run to rule out PCR contamination. Samples were genotyped in duplicate and genotypes were accepted after confirmation of no contamination or inconsistencies between duplicates.

Genotyping of SNPs in *PSIP1* gene from the Sinikithemba participants

PSIP1 SNPs associated with HIV-1 outcomes in the CAPRISA AI 002 cohort were further analyzed in the larger Sinikithemba cohort of 450 SPs, in order to confirm the results obtained from the CAPRISA AI 002. Genotyping assays were performed as described above.

LEDGF/p75 mRNA expression analysis

Peripheral blood mononuclear cells (PBMCs) were isolated from 57 SNs and 38 SPs from the CAPRISA AI 002 cohort. RNA was extracted from 2×10^6 PBMCs immediately after thawing using Trizol LS reagent (Invitrogen, Carlsbad, USA) according to the manufacturer's protocol and the integrity of RNA was confirmed using 3-(N-morpholino) propanesulfonic acid (MOPS) gels. RNA was reverse transcribed using the iScript complementary DNA (cDNA) synthesis kit (Bio-Rad). LEDGF/p75 mRNA expression was quantified by real-time PCR using SYBR Green Chemistry on a Roche Lightcycler version 1.5 (Roche Diagnostics) (Table 1).

Expression and purification of recombinant proteins

His₆-tagged HIV-1 integrase, 3xflag-tagged LEDGF/p75, MBP-JPO2 and MBP-pogZ were purified for AlphaScreen applications as described previously [30, 31].

AlphaScreen

The AlphaScreen assay is a technique used to measure protein-protein affinity interactions. To measure the influence of SNP5T (Q472L) on the binding affinity of LEDGF/p75 we performed AlphaScreen assay according to the manufacturer's protocol (Perkin Elmer, Benelux). Reactions were performed in 25 µl volume in 384-well Optiwell™ microtiter plates (Perkin Elmer). The reaction buffer contained 25 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM MgCl₂, 0.01% (v/v) Tween-20 and 0.1% (w/v) bovine serum albumin. 300nM His₆-tagged IN was incubated with 100 nM flag-LEDGF/p75 for an hour at 4°C. Subsequently 5 µl of Ni-chelate-coated acceptor beads and 5 µl anti-flag donor beads were added to a final concentration of 20 µg/ml of both beads. Proteins and beads were incubated for 1 hour at 30°C. Exposure of the reaction to direct light was avoided and the emission of light from the acceptor beads was measured in the EnVision plate reader (Perkin Elmer, Benelux) and analyzed using the EnVision manager software. Assays with JPO2 or pogZ, respectively, were essentially performed as described previously [30, 31].

Complemented cell lines

Complemented HeLaP4/CCR5 knockdown cells (A3 clone) were generated and grown as described earlier [32]. Briefly, the Q472L mutation was introduced in pLNC_LEDGF BC-Ires-Bsd and MLV-based vectors were generated [32]. Following transduction, cells were selected with 3 µg/ml blasticidin (Invitrogen, Merelbeke, Belgium). Protein expression was verified by Western blot analysis and immunocytochemistry (data not shown).

Virus strains

The molecular clone pNL4.3 was obtained through the NIH AIDS Research and Reference Reagent Program. Virus stock (HIV_{NL4.3}) was produced as described earlier [32].

HIV-1 breakthrough assay

Cells were seeded at 30,000 cells per well in a 6-well dish and infected as described earlier with minor modifications [20]. Briefly, cells were infected with 56,000 pg p24 HIV_{NL4.3} in a total volume of 2 ml (MOI 0.01). 24 hrs later, cells were washed twice with 1xPBS prior to addition of 4 ml of fresh medium. HIV replication was monitored by sampling the culture medium for p24 ELISA (HIV-1 p24 ELISA kit, Perkin Elmer).

Quantitative PCR

Integrated proviral copies were quantified by real-time quantitative PCR (Q-PCR) on genomic DNA using the iQ5 Multicolor RT PCR detection system (BioRad, Nazareth, Belgium). In order to allow quantification of integrated proviral copies in HIV-1_{NL4.3} infected cells, cells were subcultured at day 6 and grown under AZT/ritonavir treatment for 10 days, 0.5µM and 1.5µM, respectively, i.e. 25-fold IC₅₀ as determined in MT4/MTT assay [33] to eliminate all non-integrated viral DNA. Genomic DNA was extracted using the GenElute mammalian genomic DNA miniprep kit (Sigma, Bornem, Belgium); for each reaction 100 ng was used for Q-PCR. Integrated copies for HIV-1_{NL4.3} were quantified using a Gag-derived primer-probe set. Each 25 µl reaction contained 12.5 µl 2x iQ Supermix (Biorad), 40 nM primer and 40 nM probe. RNaseP was used as house-keeping gene control (TaqMan RNaseP control reagent, Applied Biosystems, the Netherlands). All samples were run in quadruplet and subjected to 3 min at 95°C, 50 cycles of 95°C for 10s and 55°C for 30 s. Data were analyzed with iQ5 Optical System software (BioRad, Nazareth, Belgium).

Statistical analysis

The difference in allele frequency distribution between the HIV-1-positive and HIV-1-negative group was determined by Fisher's exact test (FET) for each SNP to test the null

hypothesis that allele frequencies were the same in the two groups. The effect of each SNP on HIV-1 viral load and CD4⁺ T cell count was determined using a Generalized Estimating Equation (GEE) model [34] taking into account longitudinal measures for each participant. Viral loads were log-transformed and the square root of CD4⁺ T cell count was used to normalize their measurements.

Kaplan-Meier survival statistics and the Cox proportional hazards model (Cox model) were used to assess the effect of each SNP on time to HIV-1 infection after enrollment and on the rate of progression to AIDS defined as CD4⁺ T cell decline to less than 350 cells/ μ l (CD4<350). Decline in CD4 levels was determined and compared for the group with one or two copies of the minor allele to a reference group with two copies of the major frequency allele (dominant genetic model), for each SNP. The significance of genotypic associations and relative hazard (RH) was determined by unadjusted Cox regression analysis for the dominant genetic model.

LEDGF/p75 mRNA expression levels were compared between SNs and SPs by performing dot plot graphical representation, nonparametric statistical analysis, and correlation (Pearson). Values were expressed as median values. Differences between the 2 groups were evaluated using Dunn's multiple comparison test, whereas the Mann-Whitney *U* test was used for any 2-group comparisons. The software used for the analysis was SAS version 9.1.3 (SAS Institute Inc., Cary, NC).

Results

Study design and selection of patient samples

195 SNs plus 52 SPs from the CAPRISA AI 002 cohort and 403 SPs from the Sinikithemba cohort were genotyped. 34 samples from the CAPRISA AI cohort and 48 samples from the Sinikithemba cohort were excluded from all analysis due to sample unavailability or poor quality genotype data.

Identification of Q472L as an exonic polymorphism in the C-terminus of LEDGF/p75

The *PSIP1* gene is 46.9 kb long and consists of fifteen exons (Fig. 1A). We sequenced part of the C-terminal region to screen for SNPs in 83 SNs and 43 SPs from the CAPRISA AI 002 cohort. 14 previously described SNPs (dbSNP, www.ncbi.nlm.nih.gov) were discovered (data not shown). Only 1 SNP (rs61744944) was in the exon region and nonsynonymous (Q472L).

Effect of *PSIP1* SNPs on Susceptibility to HIV-1 Infection in the CAPRISA AI cohort

The CAPRISA AI 002 cohort comprises high risk individuals who were initially identified as SN and then followed longitudinally. This study design allowed us to test whether genetic variation in *PSIP1* was associated with susceptibility to HIV-1 infection. The minor allele frequencies (MAF) of SNPs and haplotype frequencies studied are shown (Fig. 1A). Kaplan-Meier survival analysis of time to HIV-1 acquisition suggested a trend of association between SNP1A and HIV-1 acquisition (RH=2.21, 95% CI=0.92–5.28; uncorrected *P*=0.08, Cox model) (Table 2).

Association of *PSIP1* SNPs with CD4⁺ T cell counts and viral load in the CAPRISA AI 002 cohort

The approximate time of infection was known for SPs in the CAPRISA AI 002 cohort [26]. Since viral loads and CD4⁺ T cell counts fluctuate significantly during acute HIV-1 infection, we analyzed the data in two intervals post infection, 0-3 months (acute) and 3-12

months (early chronic) to detect possible differences between genotypes during these phases of infection.

A dominant model analysis using GEE showed significant associations between SNP4G and lower CD4⁺ T cell count ($P<0.01$, uncorrected) (Fig. 1B) and between SNP5T and lower viral load ($p=0.02$, uncorrected) during acute phase of infection (Fig. 1C). SNP1A was associated with lower CD4⁺ T cell count ($P<0.01$, uncorrected) and SNP3C with higher CD4⁺ T cell count ($P<0.01$, uncorrected) during early chronic phase of infection (Fig. 1D). The minor allele of all SNPs, except for SNP1 were associated with lower viral loads ($P<0.05$, uncorrected) during early chronic phase of infection (Fig. 1E).

Effect of *PSIP1* SNPs and haplotypes on CD4⁺ T cell decline in the CAPRISA AI 002 cohort

We tested association of genetic variation with the rate of CD4 decline to CD4<350 in the CAPRISA cohort. SNP1A was associated with rapid CD4⁺ T cell decline ($P=0.04$ uncorrected, Cox model) (Table 3). Haplotype 4 (HAP4)—the only haplotype carrying SNP1A—was also trending towards association with rapid CD4⁺ T cell decline (RH = 2.46, 95% CI =0.87-6.93, $P=0.09$) (Table 3). However, the effect of SNP3C on CD4 decline could not be assessed in the CAPRISA AI 002 cohort as there were only three SNP3C participants whose CD4 count had dropped to 350 cells/ μ l and below, at the time of this analysis.

Effect of *PSIP1* SNPs on CD4⁺ T cell decline in the Sinikithemba cohort

Interpretation of the data from the CAPRISA AI 002 cohort was complicated by small sample size of the cohort and fluctuations in CD4⁺ T counts and viral loads that characterize acute and early chronic phases of HIV-1 infection. To further elucidate the role of *PSIP1* genetic variants in HIV-1 pathogenesis, we extended our analysis to the larger Sinikithemba cohort from the same geographical area. SNP2 and SNP4 were excluded from this analysis because they were not associated with CD4⁺ T cell count during early chronic phase of infection in the CAPRISA AI 002 cohort. Since the Sinikithemba cohort is a seroprevalent cohort with unknown dates of infection, we assessed the influence of each of SNPs on disease progression by comparing the slopes of CD4⁺ T cell decline levels using the dominant model. SNP3C was associated with delayed disease progression as reflected by a slower rate of CD4⁺ T cells decline to CD4<350 ($P=0.02$ uncorrected, Cox model) (Fig. 1G).

LEDGF/p75 mRNA expression levels in the CAPRISA AI 002 cohort

Quantitative RT-PCR on LEDGF/p75 mRNA levels was performed in 57 SNs and 38 SPs. PBMCs obtained from SPs had reduced LEDGF/p75 expression compared to PBMCs obtained from SN subjects ($P=0.02$) (Fig. 2A). For 13 of SPs, samples pre- (baseline) and post-HIV-1 infection were available for analysis. Baseline PBMCs from seroconverters had higher LEDGF/p75 expression compared to PBMC from non-seroconverters ($P<0.01$) (Fig. 2B). SNP3C was significantly associated with reduced levels of LEDGF/p75 ($P<0.01$) (Fig. 2C) in both the SN and SP groups and slower rate of CD4⁺ T cell decline (fig. 1G) in the SP group.

Functional analysis of Q472L LEDGF/p75

To determine the impact of the Q472L mutation on the interaction with IN- since this exonic mutation was located adjacent to the integrase binding domain (amino acids 347-429)- we measured LEDGF/p75-IN affinity in an AlphaScreen assay and evaluated the effect of the Q472L mutation on the interaction with JPO2 and pogZ, two cellular binding partners of LEDGF/p75 [30, 31]. Potent RNAi-mediated knockdown of LEDGF/p75 also severely hampers HIV replication affecting the integration step of the provirus [20, 35], a phenotype

that is readily rescued upon re-introduction of RNAi resistant LEDGF/p75 (LEDGF BC) [32]. In a similar setting, we complemented potent LEDGF/p75 knockdown cells (KD) with the Q472L variant (LEDGF BC Q472L). Q472L did not alter the binding affinity of LEDGF/p75 for IN, JPO2 and pogZ and complementation of LEDGF/p75-depleted cells with mutant LEDGF/p75 (Q472L) rescued HIV-1 replication to near wild-type levels (data not shown).

Discussion

LEDGF/p75 promotes HIV-1 replication *in vitro* [10-12, 14, 16, 35, 36]. In this study, we investigated the association between genetic variation in *PSIP1* and susceptibility to HIV-1 infection and disease progression in two South African cohorts. We found that SNP1A was associated with higher likelihood of HIV-1 acquisition, lower CD4⁺ T cell counts during the early chronic phase of infection and rapid CD4⁺ T cell depletion in CAPRISA 002 acute infection cohort. SNP3C was associated with higher CD4⁺ T cell count and lower viral load during the early chronic phase of infection in the CAPRISA AI 002 cohort.

The Sinikithemba cohort comprises chronically HIV-1 infected individuals with unknown date of infection and therefore we could examine the effect of SNPs on trajectory of CD4⁺ T cell decline over 6 years of follow-up. SNP3C was significantly associated with delayed disease progression as was reflected by a slower rate of CD4⁺ T cell decline in the Sinikithemba cohort. The protective effect of SNP3C was consistent between the two cohorts suggesting that SNP3C may be associated with reduced HIV-1 replication.

Next, we investigated the association between *PSIP1* genetic variation and expression levels of LEDGF/p75 and found that SNP3C was significantly associated with lower levels of LEDGF/p75 which corroborates the plausible role of SNP3C in reducing HIV-1 replication. These findings suggest that SNP3C reduces the mRNA levels of LEDGF/p75 thereby inhibiting HIV-1 replication, a finding that is consistent with *in vitro* knockdown studies [20, 37]. Pre-infection PBMCs from seroconverters had significantly higher LEDGF/p75 mRNA levels compared to nonseroconverters. HIV-1 infection reduced LEDGF/p75 expression to lower levels compared to uninfected individuals. These findings suggested that high levels of LEDGF/p75 may increase the likelihood of HIV-1 acquisition and the rate of disease progression confirming the that LEDGF/p75 is an important host factor for productive HIV-1 infection [11]. On the other hand, SNP1 associated with differential susceptibility and early HIV-1 infection outcomes in the CAPRISA acute infection cohort but not in the Sinikithemba chronic infection cohort. This SNP was not associated with differential LEDGF/p75 mRNA expression levels and future studies will need to address other possible contributing mechanisms.

Lastly, we analyzed the impact of the Q472L mutation on LEDGF/p75-IN interaction and plausible effects on cellular binding factors of LEDGF/p75, JPO2 and pogZ. Q472L did not alter the binding affinity of LEDGF/p75 for IN, JPO2 and pogZ and complementation of LEDGF/p75-depleted cells with mutant LEDGF/p75 (Q472L) rescued HIV-1 replication to near wild-type levels. It is not inconceivable that this mutation may affect other LEDGF/p75 functions such as integration site targeting.

Conclusion

These findings demonstrate that genetic variation in *PSIP1* may influence susceptibility to HIV-1 infection and disease progression, which provides *in vivo* evidence that LEDGF/p75 is an important host cofactor for HIV-1 replication. However, due to small sample size and

heterogeneous nature of our cohorts, our findings should be interpreted with caution and will need to be replicated in additional studies.

Acknowledgments

We thank Taryn Page, Yuchun Zhou, Beth Binns-Roemer, Sofie Vets and Nam Joo Van der Veken for excellent technical support. We acknowledge the participants and their clinicians who participated in the CAPRISA AI 002 and Sinikithemba studies. The CAPRISA 002 study was supported by the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), U.S. Department of Health and Human Services (grant U19 AI 51794). This study was funded by the seventh framework programme (FP7) of the European commission (THINC, HEALTH-F3-2008-201032). Additional funding was provided by the South African Department of Science and Technology/National Research Foundation Research Chairs Initiative. This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. Greene WC, Debyser Z, Ikeda Y, et al. Novel targets for HIV therapy. *Antiviral Res.* 2008; 80:251–65. [PubMed: 18789977]
2. Dietz F, Franken S, Yoshida K, Nakamura H, Kappler J, Gieselmann V. The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. *Biochem J.* 2002; 366:491–500. [PubMed: 12006088]
3. Ganapathy V, Daniels T, Casiano CA. LEDGF/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. *Autoimmun Rev.* 2003; 2:290–7. [PubMed: 12965181]
4. Ge H, Si Y, Wolffe AP. A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. *Mol Cell.* 1998; 2:751–9. [PubMed: 9885563]
5. Sewram S, Singh R, Kormuth E, et al. Human TRIM5alpha expression levels and reduced susceptibility to HIV-1 infection. *J Infect Dis.* 2009; 199:1657–63. [PubMed: 19388851]
6. Sharma P, Singh DP, Fatma N, Chylack LT Jr, Shinohara T. Activation of LEDGF gene by thermal- and oxidative-stresses. *Biochem Biophys Res Commun.* 2000; 276:1320–4. [PubMed: 11027629]
7. Shinohara T, Singh DP, Fatma N. LEDGF, a survival factor, activates stress-related genes. *Prog Retin Eye Res.* 2002; 21:341–58. [PubMed: 12052388]
8. Singh DP, Fatma N, Kimura A, Chylack LT Jr, Shinohara T. LEDGF Binds to Heat Shock and Stress-Related Element to Activate the Expression of Stress-Related Genes. *Biochemical and Biophysical Research Communications.* 2001; 283:943–955. [PubMed: 11350077]
9. Llano M, Vanegas M, Fregoso O, et al. LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J Virol.* 2004; 78:9524–37. [PubMed: 15308744]
10. Maertens G, Cherepanov P, Pluymers W, et al. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J Biol Chem.* 2003; 278:33528–39. [PubMed: 12796494]
11. De Rijck J, Vandekerckhove L, Gijsbers R, et al. Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J Virol.* 2006; 80:11498–509. [PubMed: 16987986]
12. Hombrouck A, De Rijck J, Hendrix J, et al. Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog.* 2007; 3:e47. [PubMed: 17397262]
13. MacNeil A, Sankale JL, Meloni ST, Sarr AD, Mboup S, Kanki P. Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. *J Virol.* 2006; 80:7316–21. [PubMed: 16840312]
14. Marshall HM, Ronen K, Berry C, et al. Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One.* 2007; 2:e1340. [PubMed: 18092005]
15. Cherepanov P. LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.* 2007; 35:113–24. [PubMed: 17158150]

16. Cherepanov P, Maertens G, Proost P, et al. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J Biol Chem.* 2003; 278:372–81. [PubMed: 12407101]
17. Busschots K, Voet A, De Maeyer M, et al. Identification of the LEDGF/p75 Binding Site in HIV-1 Integrase. *J Mol Biol.* 2007; 365:1480–1492. [PubMed: 17137594]
18. Daelemans D, Lu R, De Clercq E, Engelman A. Characterization of a replication-competent, integrase-defective human immunodeficiency virus (HIV)/simian virus 40 chimera as a powerful tool for the discovery and validation of HIV integrase inhibitors. *J Virol.* 2007
19. Emiliani S, Mousnier A, Busschots K, et al. Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J Biol Chem.* 2005; 280:25517–23. [PubMed: 15855167]
20. Vandekerckhove L, Christ F, Van Maele B, et al. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J Virol.* 2006; 80:1886–96. [PubMed: 16439544]
21. An P, Wang LH, Hutcheson-Dilks H, et al. Regulatory polymorphisms in the cyclophilin A gene, PPIA, accelerate progression to AIDS. *PLoS Pathog.* 2007; 3:e88. [PubMed: 17590083]
22. Agrawal L, Lu X, Qingwen J, et al. Role for CCR5Delta32 protein in resistance to R5, R5X4, and X4 human immunodeficiency virus type 1 in primary CD4+ cells. *J Virol.* 2004; 78:2277–87. [PubMed: 14963124]
23. An P, Bleiber G, Duggal P, et al. APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol.* 2004; 78:11070–6. [PubMed: 15452227]
24. An P, Duggal P, Wang LH, et al. Polymorphisms of CUL5 are associated with CD4+ T cell loss in HIV-1 infected individuals. *PLoS Genet.* 2007; 3:e19. [PubMed: 17257057]
25. Winkler C, An P, O'Brien SJ. Patterns of ethnic diversity among the genes that influence AIDS. *Hum Mol Genet.* 2004; 13 Spec No 1:R9-19.
26. van Loggerenberg F, Mlisana K, Williamson C, et al. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. *PLoS One.* 2008; 3:e1954. [PubMed: 18414658]
27. Reddy K, Winkler CA, Werner L, Mlisana K, Karim SS, Ndung'u T. APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. *AIDS.* 2009
28. Kiepiela P, Ngumbela K, Thobakgale C, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med.* 2007; 13:46–53. [PubMed: 17173051]
29. Brumme Z, Wang B, Nair K, et al. Impact of select immunologic and virologic biomarkers on CD4 cell count decrease in patients with chronic HIV-1 subtype C infection: results from Sinikithemba Cohort, Durban, South Africa. *Clin Infect Dis.* 2009; 49:956–64. [PubMed: 19663693]
30. Bartholomeeusen K, Christ F, Hendrix J, et al. Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ. *J Biol Chem.* 2009; 284:11467–77. [PubMed: 19244240]
31. Bartholomeeusen K, De Rijck J, Busschots K, et al. Differential interaction of HIV-1 integrase and JPO2 with the C terminus of LEDGF/p75. *J Mol Biol.* 2007; 372:407–21. [PubMed: 17669426]
32. Gijssbers R, Ronen K, Vets S, et al. LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Mol Ther.* 18:552–60. [PubMed: 20195265]
33. Pauwels R, Andries K, Debyser Z, et al. New tetrahydroimidazo[4,5,1-jk][1,4]-benzodiazepin-2(1H)-one and -thione derivatives are potent inhibitors of human immunodeficiency virus type 1 replication and are synergistic with 2',3'-dideoxynucleoside analogs. *Antimicrob Agents Chemother.* 1994; 38:2863–70. [PubMed: 7535037]
34. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986; 42:121–30. [PubMed: 3719049]
35. Llano M, Saenz DT, Meehan A, et al. An essential role for LEDGF/p75 in HIV integration. *Science.* 2006; 314:461–4. [PubMed: 16959972]
36. Busschots K, Vercammen J, Emiliani S, et al. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J Biol Chem.* 2005; 280:17841–7. [PubMed: 15749713]

37. Brass AL, Dykxhoorn DM, Benita Y, et al. Identification of host proteins required for HIV infection through a functional genomic screen. *Science*. 2008; 319:921–6. [PubMed: 18187620]

Fig.1. Locations of *PSIP1* single nucleotide polymorphisms (SNPs) on chromosome 9p22 and SNPs influence on disease outcome

A. *PSIP1* gene: coding exons are marked with solid black boxes. Minor allele frequencies (MAF) of the SNPs are shown for the CAPRISA and Sinikithemba cohorts. **B and C**, association of individual SNPs with differential CD4⁺ T cell counts and viral loads, respectively, during the acute infection phase (≤ 3 months post infection) of the CAPRISA 002 study cohort. **D and E**, association during the early chronic phase of infection (3-12 months) of the CAPRISA 002 cohort. Rates of CD4 cell count decline stratified by genotypes for **F**- SNP1, **G**- SNP3, and **H**- SNP5 for the Sinikithemba seroprevalent cohort. Individuals homozygous for the major allele reference genotype are indicated in black, heterozygous and homozygous for the minor allele are shown in red. **MAF**-minor allele frequencies for CAPRISA and Sinikithemba cohorts **ND**-not done.

*- An asterisk represents p-values that remain significant after correction for multiple comparison (Bonferroni), with a single asterisk representing $P < 0.05$ and double asterisks $P < 0.01$.

Fig. 2. Expression of LEDGF/p75 mRNA in peripheral blood mononuclear cells (PBMCs) represented as the normalized ratio of LEDGF/p75 to GAPDH

A, Expression of LEDGF/p75 mRNA in PBMCs obtained from HIV-1 positive versus negative participants. **B**, Expression of LEDGF/p75 mRNA in PBMCs obtained seroconverters versus nonseroconverters. **C**, Expression of LEDGF/p75 mRNA as modulated by different genotypes of SNP3.

TT-wild type genotype of SNP 3

CT-heterozygous genotype of SNP 3

--the horizontal line in the middle of the points denotes the median value.

Table 1

PCR, sequencing and real time RT-PCR primers and cycling conditions used in this study

	Gene Name, Accession and SNP number	Primer /probe name and rs numbers	PCR Primer Sequence 5'→3' TaqMan Assay ID# or Assay sequences	Cycling conditions: denaturation, annealing and extension
PCR	<i>PSIP1</i> AF 199339	LEDGF1 Seq5	F: TGG GCT CAA AGC ATTA ATC C R: TCT GTG GCG TAT ACA CAG TG	(95 °C, 10 mins), 35 cycles of (94 °C, 30 s; 60°C, 30 s; 72°C, 45 s) and (72 °C, 10 mins)
Sequencing		Seq1 Seq6 Seq2 Seq5 S9 R1	F: GCC AGA TATGAT TTA ATC TAG R: GTA GAC TTT TCC ATG ATT CCT GAC F: GCC TGT ATA TAG AAA TAC TGG R: TCT GTG GCG TAT ACA CAG TG F: CTT CAA AGG ATA CAT GC R: GTA GAC TTT TCC ATG ATT CCT GAC	(96 °C, 3 mins), 25 cycles of (96 °C, 15 s; 50°C, 15s; 60 °C, 2 s), (72 °C, 5 mins)
TaqMan Assay	SNP1 SNP2 SNP3 SNP4 SNP5	rs2277191 rs10283923 rs12339417 rs1033056 PSIP1_Q472L	C_15883595_10 C_29529242_10 C_31936110_10 C_2757693_20 PSIP1_Q472L_s AAAACCAAAGATCAAGGGAAGAAA Psp1_q472l_a TGTGAAATTGTTGGCTTTTACCA	
Real Time –RT PCR		PTZ1 PTZ2	F: GTC AAC CCC ACC GTG TTC TTC R: TTT CTG CTG TCT TTG GGA CCT TG	(95 °C, 6 s), (60 °C, 6 s) and (72 °C, 10s)
	<i>GAPDH</i> NM_002046	GAPDH1 GAPDH2	F: AAG GTC GGA GTC AAC GGA TT R: CTC CTG GAA GAT GGT GAT GG	(95 °C, 6 s), (60 °C, 6 s) and (72 °C, 10s)

2 F: → forward primer

3 R: → reverse primer

4 mins → minutes

5 s → seconds

Table 2

Association of *PSIP1* SNPs with HIV-1 acquisition in the CAPRISA AI 002 cohort, dominant model

SNP Risk group	n	Genotype			22 and 12 versus 11 (reference group)*	
		Number (Frequency)			RH (95% CI)	P - value
SNP1 G→A		GG	AG	AA		
	SN	183	158 (0.86)	22 (0.12)	3 (0.02)	
	SP	52	39 (0.75)	12 (0.23)	1 (0.01)	2.21 (0.92 – 5.28)
SNP2 G→C		GG	CG	CC		
	SN	181	66 (0.36)	77 (0.43)	38 (0.21)	
	SP	52	20 (0.38)	23 (0.44)	9 (0.17)	0.99 (0.44 – 2.24)
SNP3 T→C		TT	CT	CC		
	SN	172	86 (0.49)	64 (0.36)	22 (0.13)	
	SP	50	29 (0.58)	19 (0.38)	2 (0.04)	0.60 (0.26 – 1.37)
SNP4 A→G		AA	AG	GG		
	SN	172	108 (0.63)	53 (0.31)	11 (0.06)	
	SP	48	31 (0.65)	16 (0.33)	1 (0.02)	0.85 (0.36 – 2.01)
SNP5 A→T		AA	AT	TT		
	SN	184	162 (0.88)	21 (0.11)	1 (0.005)	
	SP	51	41 (0.80)	10 (0.20)	0 (0.00)	1.90 (0.71 – 5.09)

Kaplan-Meier survival statistics and the Cox proportional hazards model (Cox model) were used to assess the effect of each SNP on time to HIV-1 acquisition after enrollment. *P*-values uncorrected for multiple comparisons are shown.

Table 3

Effects of *PSIP1* SNPs on disease progression in the CAPRISA AI 002 cohort, dominant model

Endpoint	SNP Number	SNPs on AIDS Progression		Haplotype Number	Haplotypes on AIDS Progression	
		RH	95% CI		RH	95% CI
CD4 < 350	SNP1	5.98	2.27 – 5.82	0.04	HAP1	1.84 0.55–6.12 0.77
CD4 < 350	SNP2	2.68	0.99 – 7.24	0.70	HAP2	0.84 0.19–3.77 0.21
CD4 < 350	SNP3	1.25	0.44 – 3.56	0.29	HAP3	0.00 0.00–0.00 0.99
CD4 < 350	SNP4	2.05	0.70 – 6.02	0.90	HAP4	5.41 1.92–5.24 0.09
CD5 < 350	SNP5	1.74	0.50 – 6.05	0.71	HAP5	1.58 0.36–7.49 0.66

Cox proportional hazards model (Cox model) was used to calculate the rate of CD4 decline to less than 350 cells/ml.

P-values uncorrected for multiple comparisons are shown.