Published in final edited form as:

HIV Clin Trials. 2013 ; 14(6): 292–302. doi:10.1310/hct1406-292.

Genetic Predictors of Cervical Dysplasia in African American HIV-Infected Women: ACTG DACS 268

Michelle S. Cespedes¹, Sarah L. Kerns², Robert S. Holzman¹, Paul J. McLaren^{3,4}, Harry Ostrer⁵, and Judith A. Aberg¹

¹New York University School of Medicine, New York

²Mount Sinai School of Medicine, New York

³Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

⁴Broad Institute of MIT and Harvard, Cambridge, Massachusetts

⁵Albert Einstein College of Medicine, Bronx, New York

Abstract

Objective—To examine genome-wide associations in HIV-infected women with a history of cervical dysplasia compared with HIV-infected women with no history of abnormal Papanicolaou (Pap) tests.

Design—Case-control study using data from women analyzed for the HIV Controllers Study and enrolled in HIV treatment-naïve studies in the AIDS Clinical Trials Group (ACTG).

Methods—Genotyping utilized Illumina HumanHap 650 Y or 1MDuo platforms. After quality control and principal component analysis, ~610,000 significant single nucleotide polymorphisms (SNPs) were tested for association. Threshold for significance was $P < 5 \times 10^{-8}$ for genome-wide associations.

Results—No significant genomic association was observed between women with low-grade dysplasia and controls. The genome-wide association study (GWAS) analysis between women with high-grade dysplasia or invasive cervical cancer and normal controls identified significant SNPs. In the analyses limited to African American women, 11 SNPs were significantly associated with the development of high-grade dysplasia or cancer after correcting for multiple comparisons. The model using significant SNPs alone had improved accuracy in predicting high-grade dysplasia

© 2013 Thomas Land Publishers, Inc.

Address for correspondence: Michelle Cespedes, MD, New York University School of Medicine, 550 First Ave, BCD Room 558, New York, NY 10016; phone: 212-263-6565; fax: 212-263-8264; Michelle.Cespedes@nyumc.org.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

Conflict of interest: The authors report no conflicts of interest.

Previous publication: Presented in part at the 19th Conference on Retroviruses and Opportunistic Infections; March 5–8, 2012; Seattle, Washington.

Additional contributions: The authors would like to thank Susan Koletar and Evelyn Zheng for their contributions as the ALLRT liaisons and Janet Anderson for her contributions to the study concept.

in African American women compared to the use of clinical data (area under the receiver operating characteristic curve for genetic and clinical model = 0.9 and 0.747, respectively).

Conclusions—These preliminary data serve as proof of concept that there may be a genetic predisposition to developing high-grade cervical dysplasia in African American HIV-infected women. Given the small sample size, the results need to be validated in a separate cohort.

Keywords

cervical dysplasia; GWAS; risk markers

HIV and disease caused by human papillomavirus (HPV) disproportionately affect women of color worldwide. In the United States, African American women represent 64% of all HIV-infected women despite only representing 13% of the population. Worldwide, cervical cancer is one of the leading causes of death in women, second only to breast cancer.¹ African American women have the highest death rate from cervical cancer at 4.3 per 100,000 compared to a rate of 2.1 per 100,000 for White women.² HIV-infected women have higher prevalence and persistence of HPV, the causative agent of cervical cancer.^{3,4}

Genetic polymorphisms in HLA type and cell cycle arrest genes have been implicated in cervical cancer risk, but results have been inconsistent in different cohorts mainly of European descent.^{5–10} There is a growing recognition that genetic risk determinants from a particular racial or ethnic group may not be generalizable to another racial subset.^{11–13} Characterization of the role of genetic determinants on the development of cervical cancer may increase understanding of the pathophysiology of the disease, improve early detection and treatment, and decrease mortality.

Genome-wide association studies (GWASs) have the advantage of using new technologies to rapidly analyze hundreds of thousands of single nucleotide polymorphisms (SNPs) and relate them to clinical traits and conditions. Recent GWASs have identified genetic polymorphisms associated with the risk of HIV progression, a consistent confounder for cervical cancer among women with HIV infection.^{14–19} Few of these studies have included HIV-infected women or a meaningful number of minority subjects.

The aim of this study was to identify SNPs and by extension genes in predominately minority HIV-infected women that may be associated with the development of cervical cancer. It addresses a health disparity in oncogenomics research. As the global HIV epidemic continues to disproportionately affect women of color, special insight about the care and prevention of morbidity in this vulnerable population is warranted.

METHODS

Study Population and Case Definition

This was a nested case-control study of the HIV Controller's Study that used a cohort of HIV-infected women who enrolled in parent AIDS Clinical Trials Group (ACTG) treatment-naïve randomized control trials. HIV-infected women from 4 ACTG trials of antiretroviral-naïve individuals (protocols ACTG384, A5095, A5142, and A5202) were included.^{20–23}

Consents for each study were approved by the institutional review board at each of the participating ACTG study sites. All women included in our study consented under protocol A5128 for use of stored specimens for genetic testing. Women who had genomic data generated from the HIV Controller's Study and had gynecologic data collected from the ACTG longitudinal study A5001 (ALLRT) were included in our analysis. The study was approved by the National Institute of Allergy and Infectious Diseases (NIAID)– sponsored ACTG Scientific Agenda Steering Committee and the New York University School of Medicine Institutional Review Board.

Cases were divided into 2 groups: subjects with documented dysplasia of any grade and subjects with high-grade cervical dysplasia or invasive cervical cancer (HG). Subjects classified as HG were defined as having a tissue diagnosis by colposcopy or cone biopsy of cervical intraepithelial neoplasia (CIN) II, CIN III, or invasive cervical cancer on pathology. The control group was comprised of HIV-infected women with no history of cervical dysplasia at study entry and no incident abnormal Pap of any grade during follow-up. Subjects with atypical squamous cells of unknown significance (ASCUS) without documentation of dysplasia were not included in the analysis. Demographic, clinical, and comorbidity data were extracted from encounter forms that were collected during the longitudinal follow-up period.

Genotyping and Quality Control

DNA was isolated from lymphocytes through Ficoll separation. Cases and controls were genotyped using the Illumina HumanHap 650Y (Illumina, San Diego, CA) chip for the early samples and the Human-1M-duo chip for the majority (~75%) of the samples. An equal proportion of cases were genotyped on the Human-1M-duo chip compared to controls.

Quality control filtering and association analysis were carried out using PLINK (pngu.mgh.harvard.edu/~Purcell/plink/) whole genome association analysis toolset.²⁴ SNPs were excluded from the analysis if they had no genotype for more than 5% of individuals, were not in Hardy-Weinberg equilibrium among control subjects (by use of threshold $P < .00001$), or had a minor allele frequency of less than 5%. The minor allele was defined as the allele with the lower frequency among the total sample analyzed. After removing samples with poor genotyping call rate (<95%), we selected a set of common, unlinked SNPs with very low missingness (<0.1%) for principal components analysis (PCA) with EIGENSTRAT (<http://www.hsph.harvard.edu/faculty/alkes-price/software/>) in order to check ancestry.

Data Analysis

Quality-control filtering, identity by state clustering, and association analysis were carried out by using the PLINK genetic analysis software. PLINK provided population stratification detection, gene-based tests of association, and multimarker predictors. The main analyses were performed by using 2×2 tables that compared the frequencies of each of the 2 alleles (A and a) among case and control subjects; a 2-sided chi-square test was used to assess the statistical significance of the difference in genotype frequency. We considered SNP associations with $P < 5 \times 10^{-8}$ from the genome-wide scan as statistically significant.

Population structure was assessed by PCA, which included a randomly selected subset of 100,000 SNPs for all study subjects as well as 90 individuals of each of 3 reference populations for the International HapMap Project (European Americans from Utah, Yorubans from Nigeria, and a combined population from Tokyo and Beijing).²⁵ PCA information from the same 90 African and 90 European HapMap individuals was included to assist in estimating ancestry proportions.

General logistic regression modeling was performed by use of R, version 2.13.2. Models examined were those suggested by univariate analysis and by prior knowledge regarding causes of abnormalities on Pap tests. When automated model selection was performed, we used R's step Akaike Information Criterion (AIC) function to perform a backwards elimination of variables to arrive at the equation that minimized the AIC.

RESULTS

Patient Characteristics

Three hundred fifty-one women from the HIV Controller's Study who had gynecologic data recorded as part of ACTG ALLRT were reviewed. Seventy-six women were excluded due to the absence of the cervix without documentation of the indication for the hysterectomy, an active gynecologic infection that may have interfered with interpretation of their Pap test and no subsequent recorded gynecologic data, or inadequate data. Fifty-one women had abnormal Pap tests with ASCUS or evidence of inflammation and were excluded in this analysis. Ninety-one women had cervical dysplasia of any grade. Eighteen (20%) of these 91 remaining women had high-grade dys-plasia (CIN II, CIN III) or cervical cancer. The case participants were compared to 133 controls who had normal Pap tests. Demographic information is summarized in Table 1.

No differences were observed for nadir CD4+ T-cell count or baseline HIV-1 RNA viral load at study entry between participants in the case and control groups. Case participants with any grade of dysplasia were more likely to report a history of genital HSV and vaginal candidiasis when compared to controls. The subset of women with HG had significantly higher rates of substance abuse and trichomonas infection compared to controls.

Genome-wide Association Test

We performed Bonferonni corrected allelic association tests to compare the allele frequencies for 613,335 SNPs between case and control subjects. No significant genome-wide associations were obtained when women with no history of dysplasia were compared to women with any level of dysplasia including CIN I. When women with HG were compared to those with normal Pap tests, 3 SNPs located on chromosome 2 were significantly associated with the development of HG. Two were located in PRKCE, the protein kinase C epsilon gene, and the third in HECW2, the gene that encodes E3 ubiquitin protein ligase that stabilizes tp73 ($P = 1.78 \times 10^{-8}$, 4.43×10^{-8} , and 4.43×10^{-8} , respectively) (Figure 1).

The analysis of women from all races who had HG compared with their 3 closest genetically matched controls determined via PCA failed to yield significant results. To analyze the

contribution of racial diversity driving the genome-wide results seen in the HG analysis, we repeated the analysis using only case and control participants of self-reported African American descent. In this analysis, 11 SNPs were significantly associated with the development of HG (Table 2). The most significant SNP, rs10736156, lies within the coding region of GBF1, an important host factor for viral replication, and is associated with a 9-fold increase in risk of developing high-grade dysplasia ($P = 1.98 \times 10^{-12}$; odds ratio [OR], 9.29; 95% CI, 2.2–40.2).

To further remove variation related to ethnic or environmental factors, we compared African American cases to their 3 closest genetically matched controls as determined by PCA. In this analysis, rs2253178 associated with GOLIM4 on chromosome 3 reached the threshold for genome-wide significance. In the analyses limited to African Americans, the associated genes of the 2 SNPs with the highest genome-wide significance, GBF1 and GOLIM4, are both localized to the Golgi apparatus.

Logistic regression that incorporated the first 3 principal components were performed for all permutations of the case-control analyses including women with high-grade disease compared to women with normal Paps regardless of race and the analyses limited to African American women. No statistically significant polymorphisms were identified on testing for dominant or recessive models or with the use principal component covariates on logistic regression.

African Ancestry Association with HG

The mismatch of ancestry in case and control participants is a common confounder in GWAS analyses. We estimated the proportion of African ancestry in African American subjects to describe the potential role of ancestry in our genome-wide association analytical results. We performed PCA in these subjects and reference groups of European, African, and Asian individuals genotyped by the International HapMap Project. We found no statistically significant difference in the proportion of African ancestry between African American case and control participants, suggesting that no underlying genetic difference exists between the groups that can be attributed to race (Figure 2).

Classification of Cervical Dysplasia Subjects Based on Clinical and Genetic Factors

For illustrative purposes, we investigated the utility of clinical and genetic factors for differentiation of subjects who developed HG and controls, using logistic regression modeling. These analyses were limited to African American subjects. Clinical covariates were variables that have been associated with an increased risk of malignancy for other cancers and all of the gynecologic variables available. Table 3 shows univariate odds ratios for each variable and multivariate odds ratios after adjustment for all other listed variables. In the model using clinical covariates, only a history of trichomonas was significant on univariate analysis but the significance dissipated on multivariate analysis. Using all the clinical variables listed in Table 3, we performed a backwards elimination using the lowest AIC as the criterion for best equation (stepAIC function in R). The logistic regression model that includes hepatitis C status and trichomonas history alone was superior to other possible combinations on multivariate analysis (P s = .027 and .077, respectively).

Next, we developed a logistic regression model using genetic SNPs derived from the GWAS analysis limited to African American women. We included only SNPs with significant Bonferroni *P* values and excluded other variables as none were associated with high-grade dysplasia on the clinical multivariate model. The final model using the “best fit” determined by AIC included 9 SNPs and yielded improbable odds ratios on multivariate analysis due to the small sample size of the study (Table 3). Among the SNPs included in the genetic model, a higher percentage of cases were homozygous or heterozygous for more than one risk allele compared to control subjects (90% vs 9%). None of the 72 African American control subjects were homozygous for any high-risk allele, and only one case subject had the absence of any high-risk allele for any SNP.

To compare the utility of the clinical and the genetic information for predicting the presence of high-grade dysplasia among the African American women in the study, we examined the area under the receiver operating characteristic (ROC) curve associated with the best subset of clinical or genetic information. Using the AIC to select the best subset of clinical information, the final equation included smoking status and a history of the following: chlamydia or gonorrhea infection, hepatitis C infection, syphilis, trichomonas, and vaginal candidiasis. For this equation, the area under the ROC curve was 0.747. Using the AIC, the equation with the best subset of genetic information included the presence of any minor allele to rs7819278, rs4814615, or rs12193019. rs12193019 was the fourth highest ranking SNP in the analysis of women of all races; its associated gene, RIMS1, encodes a RAS gene superfamily member. For this equation, the area under the ROC curve was 0.9.

DISCUSSION

To our knowledge, this is the first study to use a genome-wide approach to attempt to identify genetic variations in HIV-infected women for the development of high-grade cervical dysplasia or cancer. Despite a marked decrease in the prevalence of all other AIDS-defining cancers since the introduction of effective antiretroviral treatments for HIV infection, there has been little change in the prevalence of HPV-related disease in HIV-infected patients.²⁶

The natural history of cervical cancer is complex, and HIV infection and immune response to treatment are likely confounders.²⁷ It has been well documented that CIN I lesions can spontaneously regress in the absence of HIV therapy. In HIV-infected women, antiretroviral treatment modifies the duration of the presence of cervical HPV DNA but has little effect on regression of HPV-associated disease.^{28–30} In our analysis, no associations were detected comparing women with mild dysplasia to either high-grade cases or to controls with normal Pap tests. This suggests that the women who developed more advanced disease are genetically different than those who have evidence of HPV disease yet do not progress in disease stage.

Based on the function of the identified genes, we attempted to determine the biological plausibility of their possible roles in the development of cervical dysplasia. Two SNPs identified in an analysis that included women of all races localized to PRKCE, a member of a family of kinases that serve as major receptors of phorbol esters, a class of tumor

promoters.^{31,32} PRKCE is an oncoprotein that disrupts the reactivation of tumor suppressor pRB and promotes transcriptional elongation of the c-myc oncogene. The same analysis identified HECW2 as significantly associated. HECW2 enhances the activation of transcription by tp53, the gene that encodes p53.³³ P53 regulates the cell cycle and functions as a tumor suppressor. Interestingly, pRB and p53 are the targets of the HPV early transcription proteins E6 and E7, which are responsible for HPV's oncogenic properties.^{34–37} The E6 protein inhibits p53 function, thereby promoting loss of cell cycle control and inhibiting apoptosis. HPV E7 binds to and induces the destabilization and degradation of pRB accelerating DNA synthesis and cell cycle progression.

GBF1 and GOLIM4 had the greatest significance after Bonferroni correction in the genome-wide analyses that were limited to all African Americans and African American cases compared to their 3 closest genetically matched controls, respectively. GBF1, a guanine nucleotide exchange factor that regulates the recruitment of proteins by mediating ADP ribosylation factor, is crucial for hepatitis C, coxsackievirus B3, and poliovirus replication.^{38–40} GOLIM4, another resident golgi gene, processes proteins synthesized in the rough endoplasmic reticulum and assists in the transport of protein cargo through the Golgi apparatus.⁴¹ Two additional genes identified in the analysis of African American subjects have been implicated in cancer pathology. SH3PXD2A is involved in the invasiveness of some cancer cells, and ATF3 is involved in the cellular stress response and induced by a variety of signals, including many of those encountered by cancer cells.^{42–44}

The evolutionary history of African Americans is complex, and genetic profiles that include them can be analyzed to determine the genetic contributions that can be attributed to their different racial heritages. Admixture mapping in both anthropological and genetic epidemiologic studies is a technique to map susceptibility alleles in complex genetic diseases associated with continental ancestry.^{45,46} The central attraction of admixture mapping is that it is expected that admixture-generated linkage disequilibrium can be easily detected via high resolution marker sets, because gene flow has occurred recently in modern populations (eg, in African and Hispanic Americans in the past 20 generations) Using these data, researchers can determine whether a candidate gene that confers an increased risk is predominately inherited from the patient's African genetic makeup. Knowledge of the proportion of racial ancestry could affect predictive models or treatment algorithms, especially in resource poor settings where the minority population and the burden of cervical cancer are greater. Though no significant difference on racial ancestry was detected in our cohort, this may be an artifact of sample size. Eleven of the 12 significant SNPs in the analyses limited to African Americans exhibited mean allele frequencies that are at least 4-fold higher in the reference HapMap European population. This suggests that these polymorphisms may remain significant in a large cohort that includes European women.

Major limitations of this study were the limited clinical covariates and the small sample size. Data on the age of diagnosis of dysplasia, age of sexual debut, number of lifetime sexual partners, evidence of exposure to HPV, and HPV cervical DNA type were not compiled by the ACTG treatment-naïve studies or the ALLRT. Of the clinical covariates obtained, only infection with *Trichomonas vaginalis* and a history of substance abuse were significantly different between HG cases and controls. *Trichomonas* infection has been associated with

the development of cervical cancer in multiple studies.^{47,48} Evidence exists that trichomonas disrupts the integrity of the tight junctions between cervical epithelial cells, thus facilitating the initial infection of columnar cells by oncogenic HPV.^{49,50}

We attempted to correct for population structure given the racial diversity of the larger cohort by limiting the majority of the analyses to African American subjects. This method further decreased the small sample size and decreased the probability that we had enough power to detect true associations. The improbable odds ratios obtained on multivariate logistic regression for individual SNPs are a direct result of the small numbers of individuals who had risk alleles that were included in the model. Association studies that have several thousand subjects are more likely to yield true reproducible findings with even small odds ratios. Despite this, studies with more modest sample sizes can detect strong associations. The small sample size in our study makes it is unlikely that our results are reproducible, but the study demonstrates the proof of concept that there may be a predisposition for the development of high-grade cervical dysplasia. The HIV Epidemiology Research Study and Women's Interagency HIV Study, 2 large multisite, observational prospective cohort studies of HIV-infected women and uninfected women reporting HIV risk behavior, can serve as validation cohorts for the credibility of the results. Both studies collected and stored cervicovaginal and serum samples biannually for several years and have over 6,000 women enrolled. It will be interesting to explore whether these observations remain in an analysis of a larger cohort and to further investigate the interaction of genetic and clinical variables on the development of HG in this high-risk population.

Acknowledgments

Michelle Cespedes participated in study conception and design, acquisition and analysis of data, and drafting of the manuscript. Sarah Kerns was involved in data analysis. Robert Hozman was involved in data analysis and manuscript preparation. Paul McLaren and Harry Ostrer were involved in study conception and design as well as optimization of data analysis. Judith Aberg participated in the study conception and design as well as manuscript preparation. All authors read and revised the manuscript.

Financial support/disclosures: This study was supported in part by NIAID AI069532, NIH 1UL1RR029893 from the National Center for Research Resources, the Grunebaum AIDS Scholarship Award, and NIAID U01 AI068636 through the ACTG MHIMP and ACTG ALLRT database.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010; 127:2893–2917. [PubMed: 21351269]
2. Howlader, N.; Noone, AM.; Krapcho, M.; Neyman, N.; Aminou, R., et al. SEER Cancer Statistics Review, 1975–2009 (Vintage 2009 Populations). Bethesda, MD: National Cancer Institute; 2012.
3. Palefsky J. Cervical human papillomavirus infection and cervical intraepithelial neoplasia in women positive for human immunodeficiency virus in the era of highly active antiretroviral therapy. *Curr Opin Oncol*. 2003; 15:382–388. [PubMed: 12960521]
4. Frisch M, Biggar RJ, Goedert JL. Human papillomavirus-associated cancers in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome. *J Natl Cancer Inst*. 2000; 92:1500–1510. [PubMed: 10995805]
5. Madeleine MM, Johnson LG, Smith AG, et al. Comprehensive analysis of HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 loci and squamous cell cervical cancer risk. *Cancer Res*. 2008; 68:3532–3539. [PubMed: 18451182]

6. Deshpande A, Wheeler CM, Hunt WC, et al. Variation in HLA class I antigen-processing genes and susceptibility to human papillomavirus type 16-associated cervical cancer. *J Infect Dis.* 2008; 197:371–381. [PubMed: 18248301]
7. Arias-Pulido H, Joste N, Wheeler CM. Loss of heterozygosity on chromosome 6 in HPV-16 positive cervical carcinomas carrying the DRB1*1501-DQB1*0602 haplotype. *Genes Chromosomes Cancer.* 2004; 40:277–284. [PubMed: 15188450]
8. Maciag PC, Schlecht NF, Souza PS, Rohan TE, Franco EL, Villa LL. Polymorphisms of the human leukocyte antigen DRB1 and DQB1 genes and the natural history of human papillomavirus infection [published online ahead of print July 3, 2002]. *J Infect Dis.* 2002; 186:164–172. [PubMed: 12134251]
9. Goncalves M, Soares E, Fernandes AP, et al. Langerhans' cell count and HLA class II profile in cervical intraepithelial neoplasia in the presence or absence of HIV infection. *Eur J Obstet Gynecol Reprod Biol.* 2004; 114:221–227. [PubMed: 15140519]
10. de Souza P, Sichero L, Maciag PC. HPV variants and HLA polymorphisms: The role of variability on the risk of cervical cancer. *Future Oncol.* 2009; 5:359–370. [PubMed: 19374542]
11. Kerns SL, Ostrer H, Stock R, et al. Genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with the development of erectile dysfunction in African-American men after radiotherapy for prostate cancer. *Int J Radiation Oncol Biol Physics.* 2010; 78:1292–1300.
12. Tzur S, Rosset S, Shemer R, et al. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Human Genet.* 2010; 128:345–350. [PubMed: 20635188]
13. Freedman ML, Haiman CA, Patterson N, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci USA.* 2006; 103:14068–14073. [PubMed: 16945910]
14. Pereyra F, Jia X, McLaren PJ, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science.* 2010; 330:1551–1557. [PubMed: 21051598]
15. Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science.* 2007; 317:944–947. [PubMed: 17641165]
16. Herbeck JT, Gottlieb GS, Winkler CA, et al. Multistage genomewide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS. *J Infect Dis.* 2010; 201:618–626. [PubMed: 20064070]
17. Limou S, Coulonges C, Herbeck JT, et al. Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term nonprogression to AIDS. *J Infect Dis.* 2010; 202:908–915. [PubMed: 20704485]
18. Siddiqui RA, Sauermann U, Altmuller J, et al. X chromosomal variation is associated with slow progression to AIDS in HIV-1-infected women. *Am J Human Genet.* 2009; 85:228–239. [PubMed: 19679225]
19. Aouizerat BE, Pearce CL, Miaskowski C. The search for host genetic factors of HIV/AIDS pathogenesis in the post-genome era: Progress to date and new avenues for discovery. *Curr HIV/AIDS Rep.* 2011; 8:38–44. [PubMed: 21221856]
20. Daar E, Tierney C, Fischl MA, et al. Atazanavir plus ritonavir or efavirenz as part of a 3-drug regimen for initial treatment of HIV-1: A randomized trial. *Ann Intern Med.* 2011; 154:445–151. [PubMed: 21320923]
21. Gandhi RT, Spritzler J, Chan E, et al. Effect of baseline-and treatment-related factors on immunologic recovery after initiation of antiretroviral therapy in HIV-1 Y positive subjects: Results from ACTG 384. *J Acquir Immune Defic Syndr.* 2006; 42:427–434.
22. Gulick R, Ribaudo H, Shikuma C, et al. Three- vs four-drug antiretroviral regimens for the initial treatment of HIV-1 infection. *JAMA.* 2006; 296:769–781. [PubMed: 16905783]
23. Haubrich RH, Riddler SA, DiRienzo AG, et al. Metabolic outcomes in a randomized trial of nucleoside, nonnucleoside and protease inhibitor-sparing regimens for initial HIV treatment. *AIDS.* 2009; 23:1109–1118. [PubMed: 19417580]
24. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Human Genet.* 2007; 81:559–575. [PubMed: 17701901]

25. International HapMap Consortium. The International HapMap Project. *Nature*. 2003; 426:789–796. [PubMed: 14685227]
26. De Vuyst H, Lillo F, Broutet N, Smith JS. HIV, human papillomavirus, and cervical neoplasia and cancer in the era of highly active antiretroviral therapy. *Eur J Cancer Prev*. 2008; 17:545–554. [PubMed: 18941376]
27. Ho G, Einstein M, Romney S, Kadish A, et al. Risk factors for persistent cervical intraepithelial neoplasia grades 1 and 2: Managed by watchful waiting. *J Low Genit Tract Dis*. 2011; 15:268–275. [PubMed: 21811178]
28. Paramsothy P, Jamieson D, Heilig CM, et al. The effect of highly active antiretroviral therapy on human papilloma-virus clearance and cervical cytology. *Obstet Gynecol*. 2009; 113:26–31. [PubMed: 19104356]
29. Fife KH, Wu JW, Squires KE, Watts DH, Andersen JW, Brown DR. Prevalence and persistence of cervical human papillomavirus infection in HIV-positive women initiating highly active antiretroviral therapy. *J Acquir Immune Defic Syndr*. 2009; 51:274–282. [PubMed: 19387354]
30. Minkoff H, Zhong Y, Burk RD, et al. Influence of adherent and effective antiretroviral therapy use on human papillomavirus infection and squamous intraepithelial lesions in human immunodeficiency virus-positive women. *J Infect Dis*. 2010; 201:681–690. [PubMed: 20105077]
31. McJilton MA, Van Sikes C, Wescott GG, et al. Protein kinase C epsilon interacts with Bax and promotes survival of human prostate cancer cells. *Oncogene*. 2003; 22:7958–7968. [PubMed: 12970744]
32. Wheeler DNK, Oberley T, Verma A. Protein kinase C epsilon is linked to 12-O-tetradecanoylphorbol-13-acetate-induced tumor necrosis factor-alpha ectodomain shedding and the development of metastatic squamous cell carcinoma in protein kinase C epsilon transgenic mice. *Cancer Res*. 2003; 63:6547–6555. [PubMed: 14559850]
33. Miyazaki K, Ozaki T, Kato C, et al. A novel HECT-type E3 ubiquitin ligase, NEDL2, stabilizes p73 and enhances its transcriptional activity. *Biochem Biophys Res Commun*. 2003; 308:106–113. [PubMed: 12890487]
34. Mantovani F, Banks L. The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene*. 2001; 20:7874–7887. [PubMed: 11753670]
35. Li T, Zhao L, Liu Z, Han Y, Fan D. Regulation of apoptosis by the papillomavirus E6 oncogene. *World J Gastroenterol*. 2005; 11:931–937.
36. Munger K, Basile J, Duensing S, et al. Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene*. 2001; 20:7888. [PubMed: 11753671]
37. Furumoto H, Irahara M. Human papilloma virus (HPV) and cervical cancer. *J Med Invest*. 2002; 49:124. [PubMed: 12323001]
38. Belov GA, Kovtunovych G, Jackson CL, Ehrenfeld E. Poliovirus replication requires the N-terminus but not the catalytic Sec7 domain of ArfGEF GBF1. *Cellular Microbiol*. 2010; 12:1463–1479.
39. Goueslain L, Alsaleh K, Horellou P, et al. Identification of GBF1 as a cellular factor required for hepatitis C virus RNA replication. *J Virol*. 2010; 84:773–787. [PubMed: 19906930]
40. Lanke KH, van der Schaar HM, Belov GA, et al. GBF1, a guanine nucleotide exchange factor for Arf, is crucial for coxsackievirus B3 RNA replication. *J Virol*. 2009; 83:11940–11949. [PubMed: 19740986]
41. Natarajan R, Linstedt AD. A cycling cis-Golgi protein mediates endosome-to-Golgi traffic. *Molec Biol Cell*. 2004; 15:4798–4806. [PubMed: 15331763]
42. Baganim Y, Madar S, Rais Y, et al. Transcriptional activity of ATF3 in the stromal compartment of tumors promotes cancer progression. *Carcinogenesis*. 2011; 32:1749–1757. [PubMed: 21900211]
43. Seals DF, Azucena EF Jr, Pass I, et al. The adaptor protein Tks5/Fish is required for podosome formation and function, and for the protease-driven invasion of cancer cells. *Cancer Cell*. 2005; 7:155–165. [PubMed: 15710328]
44. Yin X, Dewille JW, Hai T. A potential dichotomous role of ATF3, an adaptive-response gene, in cancer development. *Oncogene*. 2008; 27:2118–2127. [PubMed: 17952119]

45. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155:945–959. [PubMed: 10835412]
46. Montana G, Pritchard JK. Statistical tests for admixture mapping with case-control and cases-only data. *Am J Human Genet*. 2004; 75:771–789. [PubMed: 15386213]
47. Gram I, Macaluso M, Churchill J, Stalsberg H. *Trichomonas vaginalis* (TV) and human papillomavirus (HPV) infection and the incidence of cervical intraepithelial neoplasia (CIN) grade III. *Cancer Causes Control*. 1993; 3:231–236. [PubMed: 1319218]
48. Kharsany A, Hoosen A, Moodley J, Bagaratee J, Gouws E. The association between sexually transmitted pathogens and cervical intra-epithelial neoplasia in a developing community. *Genitourin Med*. 1993; 69:357–336. [PubMed: 8244352]
49. Gilbert R, Elia G, Beach D, Klaessig S, Singh B. Cyto-pathogenic effect of *trichomonas vaginalis* on human vaginal epithelial cells cultured in vitro. *Infect Immun*. 2000; 68:4200–4206. [PubMed: 10858237]
50. da Costa R, de Souza W, Benchimol M, et al. *Trichomonas vaginalis* perturbs the junctional complex in epithelial cells. *Cell Res*. 2005; 15:704–716. [PubMed: 16212877]

Figure 1.

Plot of $-\log_{10} P$ values from allelic chi-square tests for association of each of the 613,335 single nucleotide polymorphisms (SNPs) with the development of high-grade cervical dysplasia in HIV-infected (A) women of all races, (B) African American women only, and (C) African American cases compared to their 3 closest genetic matches determined by principal component analysis. Line represents the cutoff for genome-wide significance. SNPs with the highest significance are highlighted with a halo.

Figure 2.

Ancestry distributions using principal component analysis. **(A)** Subjects of all races with any grade of dysplasia (including CIN I) and normal controls; **(B)** African American (AA) high-grade dysplasia cases vs African American controls; and **(C)** proportional ancestry for AA high-grade cases vs AA normal controls. Included for reference are 90 HapMap individuals from the respective reference populations: Asian, African, and European.

Table 1

Demographics

	HG (<i>n</i> = 18)	Any dysplasia (<i>n</i> = 91)	Normal Pap (<i>n</i> = 133)
Mean (<i>SD</i>) age, ^a years	40.1 (8.9)	38 (9.0)	40.5 (10.1)
Race			
Black	10 (56)	48 (53)	72 (54)
White	5 (28)	25 (27)	33 (25)
Hispanic	3 (17)	17 (19)	27 (20)
Other	0	1 (1)	1 (1)
Mean (<i>SD</i>) nadir CD4	156 (150.0)	191 (161.8)	207 (156.7)
Mean (<i>SD</i>) baseline CD4	188 (157.2)	210 (169.7)	238 (188.5)
Median baseline RNA (IQR)	36,604 (14,006–130,424)	50,924 (16,038–160,363)	56,292 (18,154–146,957)
Smoking	10 (56)	57 (63)	63 (47)
Substance abuse	5 (28)*	9 (10)	10 (8)
Psychiatric disorder	8 (44)	32 (35)	39 (29)
Medical conditions			
Anemia	4 (22)	14 (15)	20 (15)
Asthma	3 (17)	15 (17)	11 (8)
Cardiovascular disease	2 (11)	9 (10)	8 (6)
Diabetes	1 (6)	6 (7)	10 (8)
Hepatitis C	3 (17)	12 (13)	9 (7)
Hypertension	3 (17)	24 (26)	33 (25)
Malignancy (non-HIV associated)	0	5 (5)	5 (4)
HIV-associated infections			
Candidiasis, oropharyngeal	7 (39)	34 (37)	39 (29)
Pneumonia, bacterial	3 (17)	11 (12)	23 (17)
Pneumonia, <i>Pneumocystis jiroveci</i>	1 (6)	10 (11)	14 (11)
Varicella zoster virus	0	14 (15)	15 (11)
Atypical OI	0	8 (9)	14 (11)
Gynecological infections			
Candidiasis, vaginal	3 (17)	23 (25)*	19 (14)
Chlamydia/gonorrhea	3 (17)	12 (13)	8 (6)
Genital warts	1 (6)	7 (8)	8 (6)
Herpes simplex virus	2 (11)	24 (26)*	18 (14)
Syphilis	2 (11)	9 (10)	6 (5)
Trichomonas	5 (28)*	16 (18)	15 (11)

Note: Values given as n (%), unless otherwise indicated. HG = cervical intraepithelial neoplasia (CIN) II, CIN III, or invasive cervical cancer; OI = opportunistic infection.

^a Age at the time of entry to antiretroviral treatment-naïve parent study.

* *P* < .05 compared to control.

Table 2

Single nucleotide polymorphisms showing strongest association (Bonferroni $P < .05$) with the development of high-grade dysplasia in HIV-infected women

CHR	SNP	Associated gene	Minor allele	MAF cases	MAF controls	P^*	Bonferroni P^*	MAF CEU	MAF YRI
10	rs10736156	GBF1	C	0.643	0.049	1.98E-12	1.21E-06	0.500	0
3	rs9846423	DRD3	A	0.389	0.014	1.48E-10	9.08E-05	0.058	0
7	rs13245967	ELMO1	A	0.450	0.042	2.92E-09	0.002	0.146	0.035
4	rs12642599	SORCS2	T	0.333	0.014	3.69E-09	0.002	0.044	0.08
10	rs3824781	SH3PXD2A	T	0.350	0.021	8.18E-09	0.005	0.095	0.054
2 ^a	rs4952803	PRKCE	C	0.118	0	1.78E-08	0.011	0.500	0
13	rs9524198	GPC6	G	0.550	0.083	1.78E-08	0.011	0.288	0
16	rs10492892	CENPN	G	0.300	0.014	2.61E-08	0.016	0.226	0
2 ^a	rs12329384	HECW2	T	0.111	0	4.43E-08	0.027	0	0.058
2 ^a	rs4952805	PRKCE	G	0.111	0	4.43E-08	0.027	0.500	0
3 ^b	rs2253178	GOLIM4	A	0.600	0.050	4.83E-08	0.029	0.500	0
2	rs1362486	LOC644838	C	0.350	0.023	4.86E-08	0.030	0.167	0
1	rs11571537	ATF3	C	0.200	0	5.53E-08	0.034	0.075	0.017
19	rs1003467	ZNF536	C	0.200	0	5.53E-08	0.034	0.035	0
10	rs1537745	APBB1IP	A	0.250	0.007	5.79E-08	0.036	0.040	0.013

Note: Bonferroni correction is based on 613,335 single nucleotide polymorphisms (SNPs) obtained after frequency and genotype pruning. Reported P values are the genomic inflation factor corrected values with the genomic inflation factor based on chi-square equal to 1. CEU = HapMap genetic reference Caucasian population; CHR = chromosome; MA = minor allele; MAF = minor allele frequency; YRI = HapMap reference African American population.

^a All high grades vs normals of all races.

^b African American high grade vs genetically matched controls.

* P values expressed as $nE(-x) = n \times 10^{-x}$.

Table 3

Logistic regression models for high-grade dysplasia in HIV-infected African American women

Model using clinical variables	Univariate		Multivariable ^a	
	OR	95% CI	OR	95% CI
Smoking status	0.88	0.23–3.4	0.35	0.04–2.94
Asthma	3.35	0.56–20.19	1.57	0.09–25.95
Malignancy, non-OI	0	0–Inf	0	0–Inf
Chlamydia /gonorrhea	3.35	0.56–20.19	3.48	0.18–67.22
Diabetes	1.22	0.13–11.35	1.65	0.07–38.03
Genital warts	1.89	0.19–18.82	1.04	0.06–18.96
Hepatitis C	3.35	0.56–20.19	10.02	0.71–145.64
HSV, genital	1.55	0.29–8.38	1.10	0.1–12.58
Neuropsychiatric history	1.95	0.44–8.54	0.70	0.09–5.24
Substance abuse	2	0.36–11.11	2.11	0.1–45.98
Syphilis	3.35	0.56–20.19	4.36	0.35–53.68
Trichomonas	4.13	0.99–17.28	4.55	0.42–49.65
Candidiasis, vaginal	0.89	0.1–7.96	0.37	0.01–10.46
Nadir CD4 < 200	1	0.27–3.75	3.14	0.39–24.97
Pretreatment HIV RNA <100, 000	1.54	0.3–7.87	1.88	0.26–13.71
Model using genetic variables	OR	95% CI	OR	95% CI
rs2253178	1.01	0.26–3.9	Inf ^b	0–Inf
rs7819278	2.66	0.59–12.01	Inf	0–Inf
rs6034039	1.68	0.44–6.45	0	0–Inf
rs4814615	1.87	0.49–7.22	0	0–Inf
rs9846423	35	5.37–228.02	Inf	0–Inf
rs12642599	35	5.37–228.02	Inf	0–Inf
rs10736156	9.29	2.15–40.16	0	0–Inf
rs3824781	34.5	6.22–191.45	Inf	0–Inf
rs10492892	35	5.37–228.02	Inf	0–Inf

Note: HSV = herpes simplex virus; Inf = infinity; OI = opportunistic infection; OR = odds ratio.

^a Adjusting for all other clinical or genetic covariates listed.

^b Greater than 10²⁰.