The Journal of Infectious Diseases

BRIEF REPORT

# Upregulation of *PTPRC* and Interferon Response Pathways in HIV-1 Seroconverters Prior to Infection

**Yunqi Li,[1] Francois Lefebvre,[2] Edith Nakku-Joloba,[3] Allan Ronald,[4] Glenda Gray,[5] Guy de Bruyn,[5] James Kiarie,[6,7] Connie Celum,[7,8,9] Mark J. Cameron,[10,a] Jairam R. Lingappa,[7,8,11,a] and Romel D. Mackelprang[7,12,a]**

[1]Institute for Public Health Genetics, University of Washington, Seattle, Washington, USA; [2]Canadian Centre for Computational Genomics-Montréal Node, Montreal, Quebec, Canada; [3]School of Public Health, Makerere University, Kampala, Uganda; [4]Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, Canada; [5]Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa; [6]Department of Obstetrics and Gynecology, University of Nairobi, Nairobi, Kenya; [7]Department of Global Health, University of Washington, Seattle, Washington, USA; [8]Department of Medicine, University of Washington, Seattle, Washington, USA; [9]Department of Epidemiology, University of Washington, Seattle, Washington, USA; [10]Department of Population and Quantitative Health Science, Case Western Reserve University, Cleveland, Ohio, USA; [11]Department of Pediatrics, University of Washington, Seattle, Washington, USA; and [12]Department of Obstetrics and Gynecology, University of Washington, Seattle, Washington, USA

Human immunodeficiency virus 1 (HIV-1) exposed seronegative (HESN) individuals may have unique characteristics that alter susceptibility to HIV-1 infection. However, identifying truly exposed HESN is challenging. We utilized stored data and biospecimens from HIV-1 serodifferent couple cohorts, in which couples' HIV-1 exposures were quantified based on unprotected sex frequency and viral load of the partner with HIV-1. We compared peripheral blood gene expression between 15 HESN and 18 seroconverters prior to infection. We found *PTPRC* (encoding CD45 antigen) and interferon-response pathways had significantly higher expression among individuals who went on to become seropositive and thus may be a signature for increased acquisition risk.

**Keywords.** host resistance; HIV-1; acquisition; gene expression; microarray.

A well-documented feature of human immunodeficiency virus 1 (HIV-1) infection is that some individuals have high levels of exposure to HIV-1 yet remain uninfected [1, 2]. Over the last decade, efforts have been made to explain their lower susceptibility to HIV-1 to better understand correlates of protection. One challenge for such studies is the difficulty of reliably quantifying HIV-1 exposure, thereby potentially misclassifying a lack of HIV-1 exposure as reduced susceptibility to HIV-1 infection [1]. Here, we use epidemiologic data from HIV-1 serodifferent heterosexual couples (eg, one partner with HIV-1 and the other not at enrollment), including the frequency of unprotected sex and HIV-1 viral load of the partner with HIV-1, to identify persons with high levels of exposure to HIV-1 who did not acquire HIV-1 during follow-up, as previously described [1]. Based on this epidemiologic characterization, we compared peripheral blood gene expression profiles between HIV-1 seroconverters prior to HIV-infection and highly HIV-1–exposed seronegative (HESN) subjects to identify genes and pathways associated with risk of HIV-1 infection.

## METHODS

### Study Participants

This study included HIV-1 seroconverters and HESN participants from Soweto, South Africa and Kampala, Uganda who were enrolled in the Partners in Prevention HSV/HIV Transmission Study (Partners cohort) [3] or the Couples Observational Study (COS cohort) [4]. All participants were HIV-1 seronegative at enrollment, had an enrolled heterosexual partner with HIV-1, and had HIV-1 serology testing every 3 months for the duration of follow-up (24 months for Partners cohort, and 12 months for the COS cohort). Collection of whole blood for gene expression analyses every 6 months was implemented late in the Partners cohort follow-up, and throughout follow-up for the COS cohort. HIV-1 seroconverters and their partners had plasma HIV-1 RNA sequenced to confirm that virus that was transmitted was the same as the virus in samples from the partner with HIV-1 who was enrolled in the study. Estimated dates of HIV-1 infection were determined retrospectively by reverse transcriptase polymerase chain reaction (RT-PCR) of plasma collected at visits prior to seroconversion, as previously described [5]. HIV-1 seroconverters were included in the study if they had at least one whole blood RNA sample collected at a visit prior to detection of HIV-1 RNA.

For each HIV-1 seroconverter case, we identified all HESNs from the same study site and cohort, and with the same sex as the seroconverter, and selected the HESN control as the participant with the longest follow-up and highest HIV-1 exposure (based on male circumcision status of male HESNs, self-reported frequency of unprotected sex acts with their

Received 02 June 2022; editorial decision 22 December 2022; published online 13 January 2023

[a]M. J. C., J. R. L., and R. D. M contributed equally.

Correspondence: R. D. Mackelprang, PhD, Department Obstetrics and Gynecology, University of Washington, 1959 NE Pacific Street, Box 356460, Seattle, WA 98195 (romelm@uw.edu); and J. R. Lingappa, MD, PhD, Department of Global Health, University of Washington, 325 Ninth Avenue, Box 359927, Seattle, WA 98195 (lingappa@uw.edu).

**The Journal of Infectious Diseases® 2023;227:714–9**


https://doi.org/10.1093/infdis/jiac498

enrolled partner, and the plasma HIV-1 RNA level of that partner, as previously described [1]). Informed consent for research into the effect of genes on HIV infection was obtained from all participants. Study protocols describing this were approved by human subjects research committees at the University of Washington, local study sites, and affiliated institutions.

### Study Samples

Peripheral blood RNA samples (PAXgene, Preanalytic Inc) were collected every 6 months, with HIV-1 status assessed by rapid HIV enzyme immunoassay (EIA; Alere Determine, Organics Ltd) every 3 month, with positive rapid tests confirmed by HIV enzyme-linked immunosorbent assay (ELISA), and HIV-1 western blot. For HESN subjects, RNA was extracted from each PAXgene sample collected during follow-up. For HIV-1 seroconverters, RNA was generated from PAXgene samples collected up to the last visit where the participant remained HIV-1 RT-PCR negative.

### mRNA Expression Microarrays

Microarray analysis was performed on blood sample as previously described [6]. The microarray data were then preprocessed, filtered to remove probes that were less reliable and less informative, and quantile normalized to make the observations across arrays comparable. The data have also been $\log_2$ transformed to symmetrize gene expression distributions.

### Statistical Analyses

Differential gene expression analysis was conducted to identify genes associated with risk of HIV-1 acquisition. To avoid misclassifying individuals with insufficient HIV exposure as biologically resistant to HIV, we used an extreme phenotypes design comparing individuals who were highly HIV exposed but remained seronegative to individuals with lower HIV exposure who still seroconverted to HIV positive. We applied a linear model with the empirical Bayes method to improve power by shrinking estimated gene variance towards a pooled estimate through the Limma package (https://git.bioconductor.org/packages/limma), and with intersubject correlation to account for repeated samples from the same subject. Models were adjusted for age, biological sex, and study site to control confounding. Multiple testing was accounted for using the false discovery rate (FDR). Gene set enrichment analysis (GSEA) was used to identify gene sets associated with HIV-1 acquisition based on the Hallmark gene sets from the Molecular Signatures Database (MSigDB; http://www.gsea-msigdb.org/). Normalized enrichment score was generated from GSEA to quantify the distribution of gene set across a list of genes ranked by fold-change in differential gene expression analyses.

## RESULTS

We identified 25 whole-blood RNA samples from 15 HESN, and 21 samples from 18 HIV-1 seroconverters. HESN and seroconverters were similar in age, biological sex, country of residence, and study cohort (Supplementary Table 1). HESN and seroconverters were predominantly female (87% and 78%) and most were enrolled in Uganda (93% and 83%). HESN reported higher HIV-1 exposure than seroconverters. Among HESN, half reported unprotected sex with their partner with HIV-1 throughout the study, while only a third of those who went on to seroconvert reported unprotected sexual activity. The median log plasma HIV-1 RNA of the partners of HESNs was 5.3 copies/mL (interquartile range [IQR], 5.2 to 5.5 copies/mL), which was higher than the median of partners of HIV-seroconverters (4.7 copies/mL; IQR, 4.2 to 5.0 copies/mL).

A total of 336 probes were differentially expressed between HESN and seroconverters, including 125 that were upregulated in HESN and 211 that were downregulated (for a complete list of differentially expressed probes, see Supplementary Table 2) based on FDR < 0.05. *PTPRC* (protein tyrosine phosphatase receptor type C; 2 probes, adjusted *P* value [adj*P*] = $6.6 \times 10^{-10}$, $\log_2$ fold-change [logFC] = 1.8; adj*P* = $5.0 \times 10^{-10}$, logFC = 1.5), *SPATA13* (spermatogenesis associated 13; adj*P* = $5.1 \times 10^{-10}$, logFC = 1.1), *RORA* (RAR related orphan receptor A; adj*P* = $1.1 \times 10^{-9}$, logFC = 0.7), and *RCSD1* (RCSD domain containing 1; adj*P* = $2.5 \times 10^{-9}$, logFC = 0.7) had the largest fold changes and smallest adjusted *P* values (Figure 1*A*). *PTPRC* encodes the CD45 antigen, a transmembrane protein on differentiated hematopoietic cells involved in immune activation, particularly in lymphocytes. The other highly ranked genes are also involved in immune response: *RORA* encodes a nuclear receptor and impacts T-helper cell differentiation to TH-17 cells [7]; *SPATA13* is a guanine nucleotide exchange factor that has been associated with diverse neurologic, pulmonary, and other outcomes; and *RCSD1* encodes a protein that regulates actin filament assembly in several lymphoid organs [8]. Additional differentially expressed genes that have been associated with HIV-1 outcomes in other studies include *NFAT5* [9] and *SH3KBP1* [10]. Sensitivity analyses adjusted for the infected-partners' HIV-1 RNA levels, however, did not change the significance or magnitude of the associations for genes described here.

Analysis of 50 gene sets previously defined for specific biological pathways (see "Methods") identified 21 that were associated with HIV-1 acquisition with FDR < 0.05 (for a complete list of enriched gene sets, see Supplementary Table 3). In particular, several immune or inflammation-related pathways had small adjusted *P* values (Figure 2A) and large normalized enrichment scores (NES), including interferon-α response (adj*P* = $1.8 \times 10^{-3}$, NES = 2.6; Figure 2B), interferon-γ response (adj*P* = $1.8 \times 10^{-3}$, NES = 2.5; Figure 2B), IL6-JAK-STAT3 signaling (adj*P* = $1.8 \times 10^{-3}$, NES = 1.9), allograft rejection

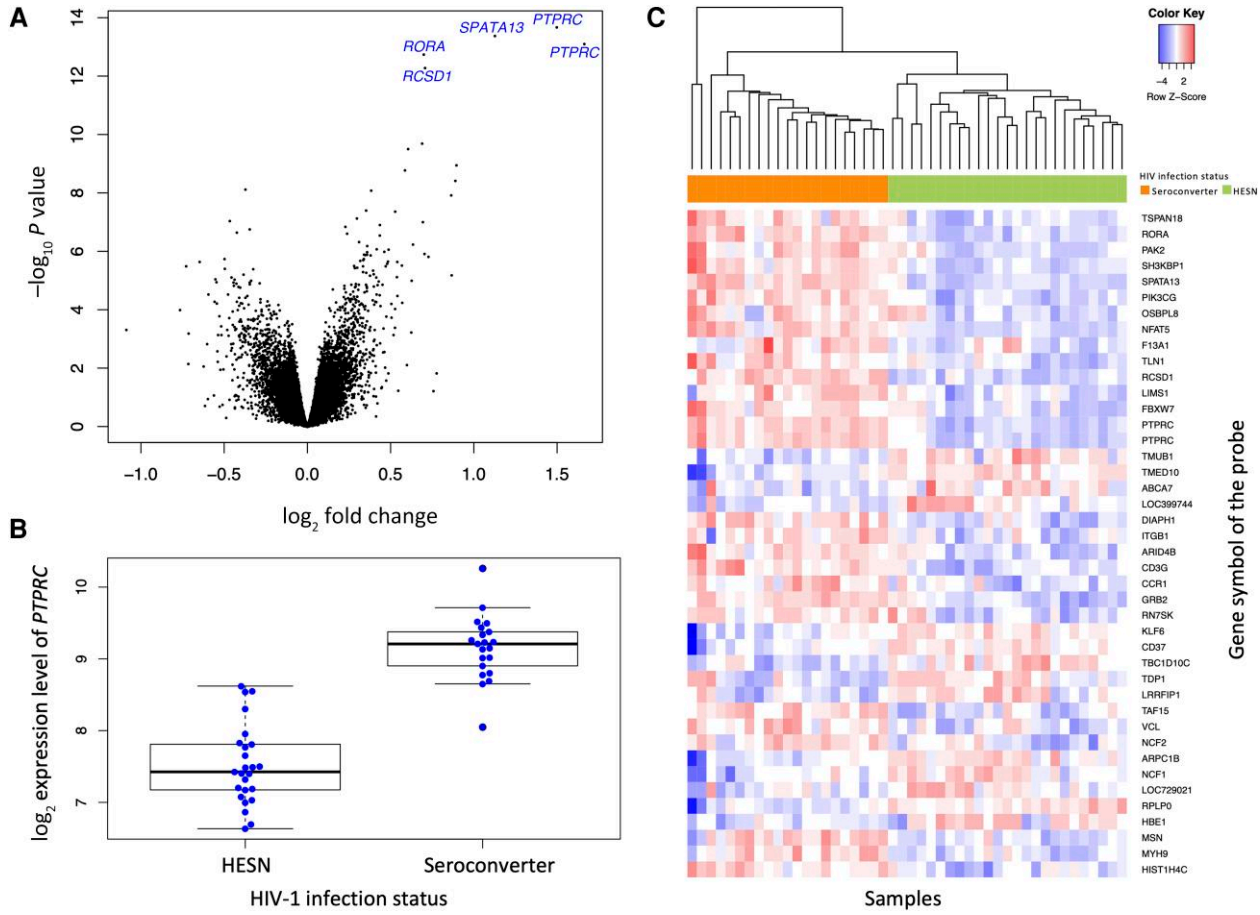


**Figure 1.** Differential gene expression analysis reveals unique biomarkers of seroconverter status. *A*, Volcano plot indicating the $\log_2$ fold-changes and *P* values for all tested probes. The top 5 signals with greater fold changes and small *P* values are annotated with the human gene names in italics. *B*, Boxplot (Bold line-median, box-interquartile range, top and bottom hortizontal lines-maximum and minimum values excluding outliers) with bee swarm plot superimposed for *PTPRC* gene expression between HIV-1 seroconverter and HIV-1 exposed seronegative (HESN) participants. Higher expression of *PTPRC* gene is observed in seroconverter compared to HESN. *C*, Heatmap containing the differentially expressed genes with an absolute $\log_2$ fold-change >0.5 and adjusted *P* value less than 0.05 under false discovery rate correction.

(adj$P = 1.8 \times 10^{-3}$, NES = 1.9), E2F targets (adj$P = 1.83 \times 10^{-3}$, NES = 1.8), and complement (adj$P = 1.83 \times 10^{-3}$, NES = 1.7).

Given their large fold change, small *P* values, and coregulation, we were particularly interested in the 2 *PTPRC* probes most strongly associated with HIV-1 acquisition. CD45, the protein product of *PTPRC*, has 8 isoforms: alternative splicing of exons 4, 5, and 6 differentiates these isoforms (exons 2 and 3 are included in all isoforms). The full microarray dataset included 4 *PTPRC* probes annealing to distinct genic regions with two probes (IDs, 2600408 and 6180288) associated with significant fold-change in our analysis, 1 (ID, 6650193) with a marginally significant fold change, and 1 (ID, 870095) with a nonsignificant fold change. We aligned the 4 probe sequences to the *PTPRC* sequence through the Basic Local Alignment Search Tool (BLAST), and found that the high fold-change probes, 2600408 and 6180288, mapped to exon 2 and exon 3, respectively (Supplementary Figure 1); while the nonsignificant probe, 870095, mapped to intron 3; and the marginally significant probe, 6650193, mapped to exon 6. These probes do not differentiate distinct CD45 isoforms.

## DISCUSSION

Our study found a signature of differentially expressed genes in whole blood that differentiated HESN with exposure to a partner with HIV and seroconverters, who were sampled prior to infection. Most notably, individuals who went on to acquire HIV exhibited higher expression of *PTPRC* (Figure 1B) prior to HIV-1 infection. We also found strong associations with gene sets involved in interferon-α and -γ responses.

*PTPRC* has previously been shown to be important for T-cell activation and HIV-1 infection. Two variants have been identified in *PTPRC* exon 4: the C77G mutation was observed in UK residents and is known to cause abnormal splicing with altered expression of CD45RA in activated or memory T cells [11], and the A54G missense mutation was observed in Ugandans [12],

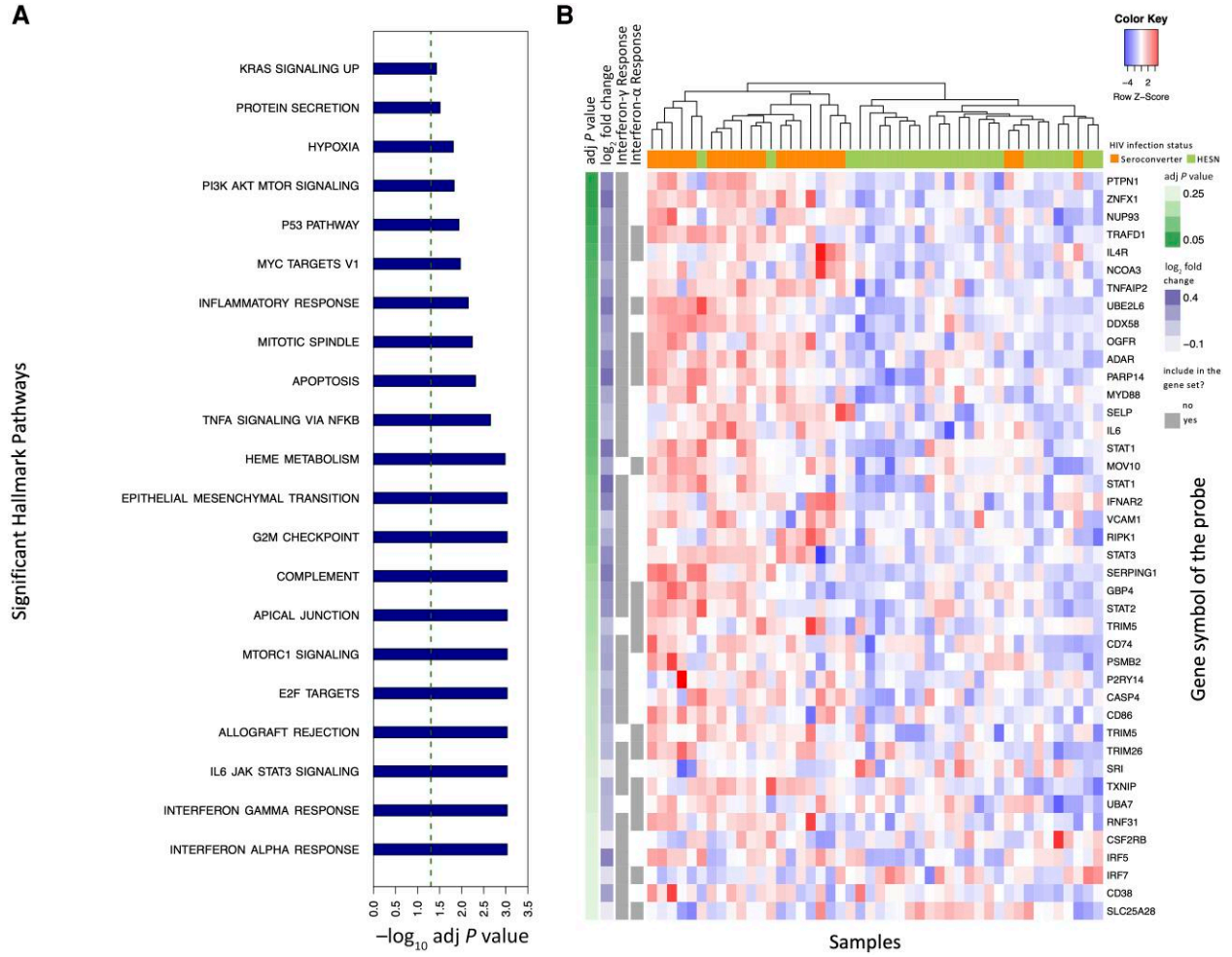


**Figure 2.** Gene set enrichment analysis-based pathway analysis of seroconverter differentially expressed genes. *A,* Bar plot indicating the normalized enrichment score of significant Hallmark pathways with adjusted *P* values less than .05. Positive values indicate an enrichment in seroconverters, and negative value indicate enrichment in seronegative participants. Several immune or inflammation-related pathways are identified with high scores, including interferon responses and JAK-STAT signaling. *B,* Heatmap containing leading edges of the interferon-α and -γ pathways with adjusted *P* value (in differential gene expression analysis) less than .25. Cell color represents the gene expression level—higher expression in red and lower expression in blue. Samples are annotated with HIV infection status, with green indicating HESNs and orange indicating seroconverters. Genes are annotated with adjusted *P* values and $\log_2$ fold-changes from differential gene expression analysis and whether they are included in the gene sets. Seroconverters show a higher expression of leading-edge genes in interferon response pathways.

both of which been associated with increased acquisition of HIV-1. Data from humanized mice showed that the frequency of human CD45-positive cells is significantly associated with increased HIV-1 susceptibility [13]. Our data now add to this context, providing a direct link between elevated CD45 expression and increased risk of HIV-1 acquisition.

We found that genes stimulated by interferon-α and -γ are upregulated in samples from seroconverters that were obtained prior to seroconversion, suggesting elevated interferon responses may increase susceptibility to HIV-1 infection. Interferon-α and -γ are cytokines elicited in response to infections and environmental stressors, and directly stimulate antiviral responses and immunoregulatory functions such as immune cell recruitment. Although interferon can contribute to controlling HIV-1 infection, it may also facilitate infection by increasing target cell activation. This may explain why we observed elevated interferon responses in seroconverters prior to infection, and is consistent with the hypothesis that low baseline immune activation, or immune quiescence, is associated with protection from HIV acquisition [14].

Previous data also suggested that interferon production may be partly regulated by CD45, the protein product of *PTPRC*, which itself is an interferon-stimulated gene. CD45 was reported to be necessary for type 1 interferon production in dendritic cells [15]. Together, elevated CD45 and interferon responses may indicate a state of increased immune activation with increased target cells available for HIV infection.

Notably, variation in the *PTPRC* gene was not associated with risk of HIV-1 infection in our whole-genome sequencing study [2], suggesting that this differentially expressed gene signature may represent an epigenetic pathway. Our analysis was not powered to detect associations between *PTPRC* expression and host genetic variants [2]. We also were not able to perform qPCR or quantify PTPRC protein levels due to lack of remaining samples from the relevant participant visits, thus limiting our ability to use these assays as additional confirmation of our results. Another limitation in our study is that difference in partner HIV RNA levels in 2 groups could have confounded the association between *PTPRC* and HIV status if higher HIV RNA levels were associated with lower *PTPRC* expression, which is unlikely. We performed sensitivity analyses in which models were also adjusted for partner HIV RNA levels, which should remove any residual confounding. For *PTPRC*, adjustment did not change the significance or size of the associations. Our extreme phenotype design was used to increase the likelihood of identifying biological factors associated with risk of HIV infection. The possibility that this could also identify a factor associated with risky behaviors is highly unlikely because we include both African heterosexual men and women in this analysis—groups with very different levels of power to choose their sexual prevention behaviors. Other limitations include possible social desirability bias that result in inaccurate response to the sexual behavior questions, leading to less accurate HIV exposure quantification. Thus, further evaluation of the relationship between protein-level changes, such as cell surface CD45 expression, and genetic variation in *PTPRC* expression is warranted for better understanding of the correlates of HIV-1 susceptibility.

## Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

## Notes

***Acknowledgments.*** The study team thank the participants and staff in the Partners HSV/HIV Transmission Study, the Couples Observational Study. We also thank Xuanlin Hou from University of Washington International Clinical Research Center data team for data management assistance, and Michael Cartwright from Case Western Reserve University and Li Pan from Vaccine and Gene Therapy Institute of Florida.

***Financial support.*** This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH) (grant numbers 5R21AI087427-02 and 5R01AI134293-03).

***Potential conflicts of interest.*** All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

Presented in part: Cold Spring Harbor Asia Conference on Infectious Disease Genomics, Suzhou, China, 7–11 May 2012.

## References

1. Mackelprang RD, Baeten JM, Donnell D, et al. Quantifying ongoing HIV-1 exposure in HIV-1-serodiscordant couples to identify individuals with potential host resistance to HIV-1. J Infect Dis **2012**; 206:1299–308.
2. Mackelprang RD, Bamshad MJ, Chong JX, et al. Whole genome sequencing of extreme phenotypes identifies variants in CD101 and UBE2V1 associated with increased risk of sexually acquired HIV-1. PLoS Pathog **2017**; 13:e1006703.
3. Celum C, Wald A, Lingappa JR, et al. Acyclovir and transmission of HIV-1 from persons infected with HIV-1 and HSV-2. N Engl J Med **2010**; 362:427–39.
4. Lingappa JR, Petrovski S, Kahle E, et al. Genomewide association study for determinants of HIV-1 acquisition and viral set point in HIV-1 serodiscordant couples with quantified virus exposure. PLoS One **2011**; 6:e28632.
5. Lingappa JR, Thomas KK, Hughes JP, et al. Partner characteristics predicting HIV-1 set point in sexually acquired HIV-1 among African seroconverters. AIDS Res Hum Retroviruses **2013**; 29:164–71.
6. Caskey M, Lefebvre F, Filali-Mouhim A, et al. Synthetic double-stranded RNA induces innate immune responses similar to a live viral vaccine in humans. J Exp Med **2011**; 208:2357–66.
7. Wang R, Campbell S, Amir M, et al. Genetic and pharmacological inhibition of the nuclear receptor RORalpha regulates TH17 driven inflammatory disorders. Nat Commun **2021**; 12:76.
8. Eyers CE, McNeill H, Knebel A, et al. The phosphorylation of CapZ-interacting protein (CapZIP) by stress-activated protein kinases triggers its dissociation from CapZ. Biochem J **2005**; 389:127–35.
9. Ranjbar S, Tsytsykova AV, Lee SK, et al. NFAT5 regulates HIV-1 in primary monocytes via a highly conserved long terminal repeat site. PLoS Pathog **2006**; 2:e130.
10. Landi A, Vermeire J, Iannucci V, et al. Genome-wide shRNA screening identifies host factors involved in early endocytic events for HIV-1-induced CD4 down-regulation. Retrovirology **2014**; 11:118.
11. Tchilian EZ, Wallace DL, Dawes R, et al. A point mutation in CD45 may be associated with an increased risk of HIV-1 infection. AIDS **2001**; 15:1892–4.
12. Stanton T, Boxall S, Bennett A, et al. CD45 Variant alleles: possibly increased frequency of a novel exon 4

CD45 polymorphism in HIV seropositive Ugandans. Immunogenetics **2004**; 56:107–10.
13. Nguyen PV, Wessels JM, Mueller K, et al. Frequency of human $CD45^{+}$ target cells is a key determinant of intravaginal HIV-1 infection in humanized mice. Sci Rep **2017**; 7: 15263.
14. Card CM, Ball TB, Fowke KR. Immune quiescence: a model of protection against HIV infection. Retrovirology **2013**; 10:141.
15. Montoya M, Dawes R, Reid D, et al. CD45 is required for type I IFN production by dendritic cells. Eur J Immunol **2006**; 36:2150–8.