Potential for immune-driven viral polymorphisms to compromise antiretroviral-based pre-exposure prophylaxis for prevention of HIV-1 infection

Hiroyuki Gatanaga[1,2*], Zabrina L. Brumme[3,4], Emily Adland[5], Gustavo Reyes-Terán[6], Santiago Avila-Rios[6], Tsunefusa Hayashida[1], Takayuki Chikata[2], Art F. Y. Poon[7], Philip Goulder[5], Masafumi Takiguchi[2], Shinichi Oka[1,2], on behalf of the International HIV Adaptation Collaborative[#]

[1]AIDS Clinical Center, National Center for Global Health and Medicine, Tokyo, Japan, [2]Center for AIDS Research, Kumamoto University, Kumamoto, Japan, [3]Faculty of Health Sciences, Simon Fraser University, Burnaby, Canada, [4]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada, [5]Department of Paediatrics, University of Oxford, Oxford, UK, [6]Centre for Research in Infectious Diseases, National Institute of Respiratory Diseases, Mexico City, Mexico, [7]Department of Pathology and Laboratory Medicine, Western University, London, ON, Canada.

Category: Major Articles

Word count: abstract 244 words, text 2635 words, with 5 figures & 2 supplemental figures.

Key words: rilpivirine, E138X, HLA-B*18, escape mutation, replication fitness

Running title: E138X in HIV-1 RT and HLA-B*18.

[#]The International HIV Adaptation Collaborative:

other authors are listed in the appendix.

***Corresponding author:**

Hiroyuki Gatanaga, M.D.

AIDS Clinical Center, National Center for Global Health and Medicine

1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan

Tel: +81-3-3202-7181

Fax: +81-3-5273-6483

E-mail: higatana@acc.ncgm.go.jp

Abstract

Objective: Long-acting rilpivirine is a candidate for pre-exposure prophylaxis (PrEP) for prevention of HIV-1 infection. However, rilpivirine resistance mutations at reverse transcriptase codon 138 (RT-E138X) occur naturally in a minority of HIV-1-infected persons; in particular those expressing Human Leukocyte Antigen (HLA)-B*18 where RT-E138X arises as an immune escape mutation. We investigate the global prevalence, B*18-linkage and replicative cost of RT-E138X and its regional implications for rilpivirine PrEP.

Methods: We analysed linked RT-E138X/HLA data from 7772 patients from 16 cohorts spanning five continents and five HIV-1 subtypes, alongside unlinked global RT-E138X and HLA frequencies from public databases. E138X-containing HIV-1 variants were assessed for *in vitro* replication as a surrogate of mutation stability following transmission.

Results: RT-E138X variants, where the most common were rilpivirine resistance-associated mutations E138A/G/K, were significantly enriched in HLA-B*18-positive individuals globally ($p=3.5x10^{-20}$) and in all HIV-1 subtypes except A. RT-E138X and B*18 frequencies correlated positively in 16 cohorts with linked HIV/HLA genotypes (Spearman's $R=0.75$; $p=7.6x10^{-4}$) and in unlinked HIV/HLA

data from 43 countries (Spearman's R=0.34, p=0.02). Notably, RT-E138X frequencies approached (or exceeded) 10% in key epidemic regions (e.g. Sub-Saharan Africa, Southeastern Europe) where B*18 is more common. This, along with the observation that RT-E138X variants do not confer *in vitro* replicative costs, supports their persistence and ongoing accumulation in circulation over time.

Conclusions: Results illustrate the potential for a natural immune-driven HIV-1 polymorphism to compromise antiretroviral-based prevention, particularly in key epidemic regions. Regional RT-E138X surveillance should be undertaken before use of rilpivirine PrEP.

Introduction

Treatment of HIV-1 infection with combination antiretroviral therapy (cART) has significantly reduced HIV-related morbidity and mortality [1, 2] and can also significantly reduce the risk of onward HIV-1 transmission [3-5]. Despite ongoing efforts to expand cART access globally however, over two million new HIV-1 cases continue to be reported each year [6]. The expanded use of antiretrovirals by HIV-uninfected persons for prevention of HIV-1 infection (termed "pre-exposure prophylaxis", or PrEP) has gained traction in recent years, after seminal studies demonstrated that oral administration of the HIV-1 nucleotide reverse transcriptase inhibitor tenofovir disoproxil fumarate (TDF) combined with the nucleoside reverse transcriptase inhibitor emtricitabine can protect against HIV-1 infection in high-risk groups, provided that high rates of adherence are maintained [7-9]. Though PrEP is now approved for use in the USA, Canada and elsewhere, adherence limitations of daily dosing and the risk of transmission of HIV-1 strains resistant to PrEP components represent two major barriers to PrEP efficacy. To address the former, long-acting injectable formulations of antiretroviral agents, notably the investigational integrase inhibitor cabotegravir and the second-generation non-nucleoside reverse transcriptase inhibitor rilpivirine, are being

considered for use as PrEP [10-12]. However, whereas HIV-1 variants resistant to TDF are still relatively uncommon in most areas [13, 14], and integrase-resistant variants rarer still [15], rilpivirine-resistant variants, in particular those with mutations at the 138th position of HIV-1 reverse transcriptase (RT-E138X), occur naturally at a significant frequency in some areas [14, 16-18]. Indeed, RT-E138X was previously shown to be significantly enriched among HIV-1 subtype B infected individuals carrying the Human Leukocyte Antigen (HLA) class I-B*18 allele [19, 20]; RT-E138X variants were later confirmed to be HLA-B*18-restricted CD8+ cytotoxic T lymphocyte (CTL) escape mutations occurring at the second (HLA/peptide anchor) position of a B*18-restricted CTL epitope spanning HIV-1 reverse transcriptase codons 137-144 [21].

As rilpivirine PrEP will presumably fail to protect against infection by RT-E138X-containing strains, establishing their regional prevalence is paramount. However, no previous study has characterized RT-E138X prevalence and its association with HLA-B*18 frequency across HIV-1 subtypes and global populations. In light of recent studies demonstrating ongoing adaptation of HIV-1 to HLA class I [22], we combine *in vitro* experiments with analyses of host and viral genetic data to test the hypothesis that differential global RT-E138X

frequencies reflect the B*18-mediated selection and subsequent stable transmission of these variants in human populations, and interpret our observations in context of the regional implications for the use of rilpivirine as PrEP.

105

Methods

Cohorts and data sources

RT-E138X and HLA-B*18 frequencies in global cohorts (linked analysis)

Plasma HIV-1 RNA-derived reverse transcriptase sequences linked to high-resolution HLA-B typing data from 7647 chronically HIV-infected, ART-naïve individuals from 15 established contemporary cohorts in Canada [20], USA [20], Mexico [23, 24], Belize [25], Guatemala [26], Honduras [27], Nicaragua [28], Panama [29], UK [22], Uganda [30], Botswana [22], South Africa [22], Japan [21], Vietnam [31], and Australia [20], and one published study including 125 HLA-B locus-typed, chronically HIV-infected, ART-naïve individuals from France [18] (retrieved via a PubMed search of English-language articles published between 2010-2015 using keywords '"HIV" and "resistance" that yielded 6803 articles, of which only this one featured linked RT-E138X and B*18 frequency data) were analyzed with respect to RT-E138X and HLA-B*18 carriage. In total, the analyzed data were derived from 7772 HIV-infected ART-naïve individuals and these data comprised HIV-1 group M subtypes A (n=233), B (n=5641), C (n=1063), D (n=186), CRF01_AE (n=438) and others/recombinants (n=86) (subtype data were not available for the French study). Study subjects from all of our 15 cohorts gave

written informed consent for their participation and/or specimens were anonymized by IRB-approved procedures. This study was approved by all relevant institutional review boards.

RT-E138X and HLA-B*18 frequencies in global databases (unlinked analysis)

Unlinked RT-E138X and HLA-B*18 frequencies were additionally retrieved from the Stanford University HIV Drug Resistance [32] and Allele Frequencies [33] databases respectively using custom Python web scripts [34]. Briefly, HLA frequencies at allele-level resolution were retrieved for all published studies of N=100 or greater, yielding 27480449 HLA alleles from 85 countries. Similarly, a total of 44934 HIV-1 sequences from unique therapy-naive subjects were retrieved from a total of 132 countries (mean 123 sequences/country, interquartile range [IQR]=19-279). Pairwise alignment of the translated sequences against the HIV-1 subtype B reference strain HXB2 was performed using *HyPhy* as described in [34]. Analyses of RT-E138X and HLA-B*18 frequencies were limited to countries with a minimum of 100 HIV-1 sequences; 43 countries were thus represented in both RT-E138X and HLA data sets. Here, analysis was performed using “traditional” allele frequencies, where the denominator is the number of alleles in the diploid human genome (2N).

Viral replication assessment

In order to delineate the effect of RT-E138X on viral replication capacity, recombinant HIV-1s were produced in the subtype B NL4-3 reference backbone (HIV-1_{NL4-3}) [21] and single-strain HIV-1 replication assays were performed as described in [35]. Briefly, MT-2 cells ($1x10^5$) were exposed to a standardized viral inoculum (500 blue-cell-forming units [BFU] in MAGIC-5 cells) for 2 hours, washed twice with phosphate-buffered saline (PBS), and cultured in 1 ml of complete medium. The culture supernatants were harvested every other day, and p24 Gag amounts were determined using a chemiluminescence enzyme immunoassay (Fuji-Rebio, Tokyo, Japan). Replication assays were performed in triplicate and repeated three times using independently generated virus preparations. To further delineate the effect of RT-E138X, competitive HIV-1 replication assays were also performed [35]. Freshly prepared H9 cells ($3x10^5$) were exposed to mixtures of paired virus preparations at various ratios for 2 hours, washed twice with PBS, and cultured. On day 1, one-third of the infected H9 cells were harvested and washed twice with PBS, and proviral DNA were sequenced (denoted passage 0). The viral culture which best approximated a 50:50 mixture

on day 1 was further propagated. Culture supernatants were transferred weekly to new uninfected H9 cells for 18 additional passages. DNA extracted from cells harvested at the end of each passage was subjected to PCR amplification of the HIV-1 RT region followed by direct DNA sequencing; the relative proportions of each viral population were estimated by their relative chromatogram peak heights.

Statistical analysis

The association between RT-E138X and B*18 carriage in all study cohorts, overall and stratified by HIV-1 subtype, was assessed using Fisher's exact test. Sequences containing amino acid mixtures at RT codon 138 were counted as variants. The correlation between RT-E138X prevalence and HLA-B*18 frequency was evaluated using Spearman’s rank correlation. Differences in p24 Gag production by wild-type versus RT-E138X-containing HIV-1 strains in the viral replication assay was assessed using Student's t-test. All tests of significance were two-sided with α defined as $p<0.05$. All statistical analyses were performed with SPSS version 17.0.

Results

Analysis of linked viral/HLA genotypes from 7772 antiretroviral naive HIV-infected patients from 16 cohorts spanning five continents revealed a highly significant overall enrichment of RT-E138X among HLA-B*18-positive compared to HLA-B*18-negative individuals, with 11.9% and 2.1% of B*18-positive and B*18-negative individuals respectively harbouring RT-E138X ($p=3.5x10^{-20}$; Figure 1A). The most common variants at this position, comprising 97.6% of all those observed, were E138A, followed by E138G and E138K (Figure 1B, C), all major rilpivirine resistance-associated mutations [36]. On average therefore, 11.6% of HLA-B*18-positive patients worldwide naturally harbour rilpivirine-resistant HIV-1 variants.

Stratified by cohort, E138X was enriched in B*18-expressing individuals in all countries except Belize (where no E138X variants were observed), Vietnam and the UK (where overall E138X frequencies were <1%); in the remaining 13 countries, E138X frequencies ranged from 5.6-21% in B*18-expressing persons compared to only 0.37-10% in non-B*18-expressing persons (Figure 1D). Moreover, the enrichment of RT-E138X variants in B*18-expressing persons was statistically significant in 7 of these 13 countries (Japan, Mexico, Guatemala,

South Africa, USA, Uganda, and Canada). These observations confirm reproducible selection of RT-E138X by HLA-B*18 in human populations globally.

Consistent with recent reports of population-level HIV-1 adaptation to HLA class I alleles in human populations [22], E138X prevalence correlated significantly with HLA-B*18 frequency in the 16 antiretroviral-naive study cohorts. This remained true regardless of whether E138X frequencies were computed in the overall population (Spearman's R=0.75; $p=7.6x10^{-4}$; Figure 2A) or only among the B*18-negative population subset (Spearman's R=0.66, p=0.0055; Figure 2B). The positive relationship between RT-E138X and B*18 population frequencies was further corroborated by an analysis of unlinked RT-E138X and HLA-B*18 data from 43 countries (see methods) (Spearman's R=0.34; p=0.02, Figure 3). Overall, these data strongly suggest that RT-E138X variants are accumulating in regions where the restricting HLA-B*18 allele is more common. Strikingly, the areas where population RT-E138X frequencies are most elevated (e.g. approaching or exceeding 10% in some countries) tend to be "key" epidemic regions in terms of high HIV prevalence (e.g. Sub Saharan Africa) or concentrated HIV-1 epidemics (e.g. Southeastern Europe).

While relative enrichment of E138A in HIV-1 subtype C compared to subtype B has previously been reported [37], RT-E138X distribution is incompletely characterized in other major HIV-1 subtypes though the data in Figure 3 suggests that it arises in all major HIV-1 subtypes and circulating recombinant forms (CRFs) globally. We therefore stratified RT-E138X prevalence by subtype in our 15 cohorts with linked HIV/HLA data. RT-E138X was most common in subtype C (6.7%), followed by D (4.3%), A (3.9%), B (1.9%) and CRF01_AE (0.68%) (all comparisons except CRF01_AE significant at $p<0.05$ with subtype B used as the reference group). These frequencies corroborated those from 60518 unique HIV-infected reverse transcriptase sequences from reverse transcriptase inhibitor-naïve patients retrieved from the Stanford HIV drug resistance database (Spearman's $R=0.90$; $p=0.037$ and Supplemental Figure 1), indicating that our cohorts are not unrepresentative of the pandemic. Also consistent with our overall analyses (Figure 1B, C), rilpivirine resistance-associated mutations E138A followed by E138G and E138K accounted for >95% of RT codon 138 variants across HIV-1 subtypes A-G and CRFs 01_AE and 02_AG in the Stanford database (Supplemental Figure 1). These data further

confirm that rilpivirine-resistance mutations arise naturally in all major HIV-1 subtypes and CRFs, and that their frequencies tend to be higher in non-B subtypes.

Visualization of RT-E138X distribution confirmed substantial frequency differences globally, with regions hardest hit by the epidemic, notably Sub-Saharan African nations where HIV-1 subtype C predominates, harbouring the highest RT-E138X burdens (Supplemental Figure 2). RT-E138X frequency in published HIV-1 reverse transcriptase sequences from Zimbabwe, for example, exceeds 15%. RT-E138X prevalence was also elevated in certain Southeastern European nations dominated by HIV-1 subtype F epidemics where B*18 frequency is also elevated (e.g. Romania [38]) (Figure 3 and Supplemental Figure 2). The relationship between HLA-B*18 carriage and RT-E138X frequency was also heretofore incompletely uncharacterized for non-B HIV-1 subtypes: but stratification of our 15 cohorts by HIV-1 subtype revealed statistically significant associations between E138X and HLA-B*18 carriage in HIV-1 subtypes B, C, and D, and a marginally significant association in CRF01_AE (Figure 4). Together, these data indicate that regional RT-E138X prevalence depends on both HIV-1 subtype as well as population HLA-B*18 frequency.

HIV-1 immune escape mutations conferring little or no cost to viral fitness are the most likely to accumulate in circulation, as there is no pressure to revert to consensus following transmission to an individual lacking the restricting HLA allele. To assess the replicative cost of RT-E138X (as a surrogate of its potential to persist upon transmission) we constructed recombinant HIV-1 variants harbouring E138A, E138G, and E138K and assessed their *in vitro* replicative competence compared to wild-type HIV-1 (E138). E138X variants exhibited no significant differences in replicative kinetics with respect to wild-type HIV-1 in a 10-day monoculture assay (Figure 5A) nor in a competition assay spanning 18 passages of one week each (Figures 5B, C, D). The lack of *in vitro* replicative costs, combined with the significant positive correlation between RT-E138X and HLA-B*18 frequencies in all regions (Figures 2, 3) and across most major HIV-1 subtypes (Figure 4) strongly suggest that, once selected in HLA-B*18-positive individuals, E138X mutations are stably maintained even after transmission to HLA-B*18-negative individuals.

Discussion

Regardless of whether antiretrovirals are used for HIV-1 treatment or prevention, the presence of drug resistance mutations will undermine their efficacy. RT-E138K is the most commonly identified mutation upon virologic failure of rilpivirine-containing ART [39], conferring 2 to 3-fold decreased susceptibility to this drug when present alone [36]; E138A and E138G confer the same level of rilpivirine resistance as E138K [21]. Using large global datasets of linked HIV/HLA genotypes, we demonstrate that E138X variants (most commonly E138A, E138G, or E138K) naturally occur in persons expressing HLA-B*18 allele in the majority of global regions and in the majority of major HIV-1 group M subtypes and CRFs. Notably, regional RT-E138X frequency correlated positively and significantly with HLA-B*18 frequency in over 40 countries - with subtype C epidemics in Sub-Saharan Africa (e.g. Zimbabwe) and subtype F epidemics in South-Eastern Europe (e.g. Romania) exhibiting the highest E138X frequencies globally (>15% and nearly 10%, respectively). This observation is unlikely to be explained by confounding by regional rilpivirine use (as use of this drug is negligible in resource-limited settings) or by cross-resistance to other nonnucleoside reverse transcriptase inhibitors (as E138X specifically confers

resistance to rilpivirine [21, 36]). Instead, our observations, combined with the lack of *in vitro* replicative costs of E138X (Figure 5) are consistent with the B*18-mediated selection and subsequent stable transmission persistence of RT-E138X variants upon transmission, leading them to accumulate in circulating HIV-1 sequences to a degree that mirrors regional HLA-B*18 frequencies [22].

It is also notable that the regions where long-acting PrEP has the potential to make the greatest impact on reducing HIV-1 incidence are also the areas where RT-E138X prevalence is highest. HIV-1 subtype C, by virtue of its dominance in Sub-Saharan Africa and other regions bearing a disproportionate HIV-1 infection burden, is the most prevalent subtype globally. It also features the highest natural burden of RT-E138X (6.4% in nearly 10000 unique treatment-naive sequences examined, with frequencies exceeding 15% in published HIV-1 sequences from Zimbabwe). Long-acting rilpivirine is currently being evaluated as PrEP in some areas including sub-Saharan Africa [11, 12, 40, 41], but data from this current and previous studies [37] strongly suggest that rilpivirine will fail to protect against infection by the substantial minority of RT-E138X-containing strains circulating in these areas.

Critically, whereas HIV-1 genotypic resistance testing can be used to identify transmitted or *de novo* selected resistance mutations in HIV-infected persons (thus avoiding the prescription of antiretrovirals for which the autologous virus harbors resistance), no analogous personalized screening can be performed in the context of HIV-1 prevention. Instead, regional Molecular Epidemiologic surveillance of circulating HIV-1 variants represents the only means to inform the regional selection of antiretrovirals used as PrEP (as well as to inform the regional selection of first-line antiretroviral regimens in settings where drug resistance genotyping is not routinely available [24]). Given that rilpivirine PrEP will presumably fail to protect against infection by RT-E138X containing strains, and that these strains occur at appreciable frequencies in areas where HIV-1 prevention efforts are most needed, we strongly advocate that regional Molecular Epidemiologic surveillance of HIV-1 reverse transcriptase sequence variation be undertaken prior to the selection of antiretrovirals used as HIV-1 prevention, and that rilpivirine not be used as single-agent PrEP in areas with elevated E138X frequencies.

Appendix

Other authors of the International HIV Adaptation Collaborative include: Humberto Valenzuela-Ponce, Maribel Soto-Nava, Claudia García-Morales (Centre for Research in Infectious Diseases, National Institute of Respiratory Diseases, Mexico City, Mexico), Carlos R. Mejía-Villatoro (Infectious Diseases Clinic, Roosevelt Hospital, Guatemala City, Guatemala), Giang Van Tran, Kinh Van Nguyen (National Hospital of Tropical Diseases, Dong Da District, Hanoi, Vietnam), Rita I. Meza (HIV National Laboratory, Honduran Ministry of Health, Tegucigalpa, Honduras), Elsa Y. Palou (Hospital Escuela Universitario, Tegucigalpa, Honduras), Ivette Lorenzana (Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras), Juan M. Pascale (Gorgas Memorial Institute for Health Studies, Panama City, Panama), Guillermo Porras-Cortés (Hospital Metropolitano Vivian Pellas, Managua, Nicaragua), Marvin Manzanero (Ministry of Health, Belmopan, Belize), Guinevere Q. Lee (British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada), Jeffrey N. Martin (University of California San Francisco, San Francisco, CA, USA), David Bangsberg (Center for Global Health, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA), Mary N. Carrington (Cancer and Inflammation Program,

Leidos Biomedical Research, Inc., Frederick National Laboratory, Frederick, USA), Mina John, Simon Mallal (Murdoch University, Perth, Australia), Roger Shapiro (Department of Immunology and Infectious Diseases, Harvard T H Chan School of Public Health, Boston, MA, USA), John Frater (Nuffield Department of Medicine, Medical Sciences Division, University of Oxford, Oxford, UK).

Acknowledgements

The authors would like to thank Dr. Richard Harrigan for data access.

Contribution

HG designed the study, did the literature search, analyzed and interpreted data, and drafted the manuscript. ZLB collected, analyzed and interpreted data, oversaw analysis, and edited the manuscript. EA, GRT, SAR, HVP, MSN, CGM, CRMV, TC, GVT, KVN, RIM, EYP, IL, JMP, GPC, MM, and GQL collected and analyzed data. TH performed viral experiments, statistical analysis, and produced figures. JNM, DB, MNC, MJ, SM, RS, JF, PG, and MT collected data and oversaw analysis. AFYP retrieved and analyzed data from global databases. SO oversaw analysis and contributed to manuscript.

Funding

This work was funded by the Japan Agency for Medical Research and Development (HG and MT). The International HIV Adaptation Collaborative (IHAC) is supported by a grant from the Canadian Institutes for Health Research (CIHR) (PJT-148621, to ZLB, MJ and SM). The UARTO cohort is supported by grants from the National Institutes of Health P30 AI027763, R01 MH054907 and UM1 CA181255 to DB and JNM. ZLB is supported by a Scholar Award from the Michael Smith Foundation for Health Research. JF is supported by the Medical Research Council. AFYP was supported by a New Investigator Award from CIHR. PG is supported by the NIH RO1 AI46995 and Wellcome Trust WT104748MA.

Potential conflicts of interest.

HG has received honoraria from ViiV Healthcare, Co., MSD K.K., Janssen Pharmaceutical K.K., and Torii Pharmaceutical, Co. SO has received honoraria and research grants from MSD K.K., AbbVie GK, Janssen Pharmaceutical K.K., Pfizer, Co., ViiV Healthcare, and Roche Diagnostics K.K., and has received honoraria from Astellas Pharmaceutical K.K., Bristol-Myers K.K., Daiichisankyo,

Co., Dainippon Sumitomo Pharma, Co., GlaxoSmithKline, K.K., Taisho Toyama Pharmaceutical, Co., and Torii Pharmaceutical, Co. All other authors declare no conflict of interest.

370

References

1. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med **1998;** 338:853-60.
2. Walensky RP, Paltiel AD, Losina E, Mercincavage LM, Schaclman BR, Sax PE, et al. The survival benefits of AIDS treatment in the United States. J Infect Dis **2006;** 194:11-9.
3. Montaner JS, Hogg R, Wood E, Kerr T, Tyndall M, Levy AR, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet **2006;** 368:531-36.
4. Grinsztejn B, Hosseinipour MC, Ribaudo HJ, Swindells S, Eron J, Chen YQ, et al. Effects of early versus delayed initiation of antiretroviral treatment on clinical outcomes of HIV-1 infection: results from the phase 3 HPTN 052 randomised controlled trial. Lancet Infect Dis **2014;** 14:281-90.
5. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med **2016;** 375:830-9.

6. UNAIDS. Global HIV Statistics. Fact Sheet November 2016. http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf Accessed 18 February 2017.
7. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med **2010;** 363:2587-99.
8. Molina JM, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, et al. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. N Engl J M **2015;** 373:2237-46.
9. McCormack S, Dunn DT, Desai M, Dolling DI, Gafos R, Gilson R, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. Lancet **2016;** 387:53-60.
10. Andrews CD, Heneine W. Cabotegravir long-acting for HIV-1 prevention. Curr Opin HIV AIDS **2015;** 10:258-63.
11. Jackson A, McGowan I. Long-acting rilpivirine for HIV prevention. Curr Opin HIV AIDS **2015;** 61:1336-41.

12. McGowan I, Dezzutti CS, Siegel A, Engstrom J, Nikiforov A, Duffill K, et al. Long-acting rilpivirine as potential pre-exposure prophylaxis for HIV prevention (the MWRI-01 study): an open-label, phase 1, compartmental, pharmacokinetic and pharmacodynamic assessment. Lancet HIV **2016;** 3:e569-78.
13. Chan PA, Huang A, Kantor R. Low prevalence of transmitted K65R and other tenofovir resistance mutations across different HIV-1 subtypes: implications for pre-exposure prophylaxis. J Int AIDS Soc **2012;** 15:17701.
14. Rhee SY, Blanco JL, Jordan MR, Taylor J, Lemey P, Varghese V, et al. Geographic and temporal trends in the molecular epidemiology and genetic mechanisms of transmitted HIV-1 drug resistance: an individual-patient- and sequence-level meta-analysis. PLoS Med **2015;** 12:e1001810.
15. Stekler JD, McKernan J, Milne R, Tapia KA, Mykhalchenko K, Holte S, et al. Lack of resistance to integrase inhibitors among antiretroviral-naïve subjects with primary HIV-1 infection, 2007-2013. Antivir Ther **2015;** 20:77-80.
16. Alvarez M, Monge S, Chueca N, Guillot V, Viciana P, Anta L, et al. Transmitted drug resistance to rilpivirine in newly diagnosed antiretroviral naïve adults. Clin Microbiol Infect **2015;** 21:104.e1-5.

17. Lambert-Niclot S, Charpentier C, Storto A, Fofana DB, Soulié C, Fourati S, et al. Prevalence of pre-existing resistance-associated mutations to rilpivirine, emtricitabine and tenofovir in antiretroviral-naïve patients infected with B and non-B subtype HIV-1 viruses. J Antimicrob Chemother **2013;** 68:1237-42.

18. Jeulin H, Foissac M, Boyer L, Aqrinier N, Perrier P, Kennel A, et al. Real-life rilpivirine resistance to rilpivirine and potential emergence of an E138A-positive HIV strain in north-eastern France. J Antimicrob Chemother **2014;** 69:3095-102.

19. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. PLoS One **2009;** 4:e6687.

20. Carlson JM, Brumme CJ, Martin E, Lisgarten J, Brockman MA, Le AQ, et al. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. J Virol **2012;** 86:13202-16.

21. Gatanaga H, Murakoshi H, Hachiya A, Hayashida T, Chikata T, Ode H, et al. Naturally selected rilpivirine-resistant HIV-1 variants by host cellular immunity. Clin Infect Dis **2013;** 57:1051-55.

22. Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M, et al. Adaptation of HIV-1 to human leukocyte antigen class I. Nature **2009;** 458:641-45.

23. Ávila-Ríos S, García-Morales C, Garrido-Rodríguez D, Ormsby CE, Hemández-Juan R, Andrade-Villanueva J, et al. National prevalence and trends of HIV transmitted drug resistance in Mexico. PLoS One **2011;** 6:e27812.

24. Ávila-Ríos S, García-Morales C, Matías-Florentino M, Romero-Mora KA, Tapia-Trejo D, Quiroz-Morales VS, et al. Pretreatment HIV-drug resistance in Mexico and its impact on the effectiveness of first-line antiretroviral therapy: a nationally representative 2015 WHO survey. Lancet HIV **2016;** 3:e579-91.

25. Ávila-Ríos S, García-Morales C, Tapia-Trejo D, Mejía-Villatoro C, Pascale JM, Porras-Cortes G, et al. HIV drug resistance in antiretroviral treatment-naïve individuals in the Mesoamerican region, 2010-2014. XXIV International HIV Drug Resistance Workshop, 2015, Seattle, WA, USA.

26. Ávila-Ríos S, García-Morales C, Garrido-Rodríguez D, Tapia-Trejo D, Girón-Callejas AC, Mendizábal-Burastero R, et al. HIV-1 drug resistance surveillance in antiretroviral treatment-naïve individuals from a reference hospital in Guatemala. AIDS Res Hum Retroviruses **2015;** 31:401-11.

27. Ávila-Ríos S, García-Morales C, Tapia-Trejo D, Meza RI, Muñez SM, Parham L, et al. HIV-1 drug resistance surveillance in Honduras after a decade of widespread antiretroviral therapy. PLoS One **2015;** 10:e0142604.

28. Ávila-Ríos S, García-Morales C, Matías-Florentino M, et al. HIV drug resistance in antiretroviral treatment-naïve individuals in the largest public hospital in Nicaragua, 2011-2015. PLoS One **2016;** 11:e0164156.

29. Mendoza Y, Castillo Mewa J, Matínez AA, Zaldívar Y, Sosa N, Arteaga G, et al. HIV-1 antiretroviral drug resistance mutations in treatment naïve and experienced Panamanian subjects: impact on national use of EFV-based schemes. PLoS One **2016;** 11:e0154317.

30. Lee GQ, Bangsberg DR, Muzoora C, Boum Y, Oyugi JH, Emenyonu N, et al. Prevalence and virologic consequences of transmitted HIV-1 drug resistance in Uganda. AIDS Res Hum Retroviruses **2014;** 30:896-906

31. Van Tran G, Chikata T, Carlson JM, Murakoshi H, Nguyen DH, Tamura Y, et al. A strong association of human leukocyte antigen-associated Pol and Gag mutations with clinical parameters in HIV-1 subtype A/E infection. AIDS **2016;** 30:681-89.

32. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res **2003;** 31:298-303.

33. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res **2015;** 43:D784-88.

34. Karakas A, Brumme ZL, Poon AF. Global database-driven assessment of HIV-1 adaptation to the immune repertoires of human populations. J Virol **2015;** 89:10693-95.

35. Gatanaga H, Hachiya A, Kimura S, Oka S. Mutations other than 103N in human immunodeficiency virus type 1 reverse transcriptase (RT) emerge from K103R polymorphism under non-nucleoside RT inhibitor pressure. Virology **2006;** 344:354-62.

36. Wensing AM, Calvez V, Günthard HF, Johnson VA, Paredes R, Pillay D, et al. 2017 update of the drug resistance mutations in HIV-1. Top Antivir Med **2017;** 24:132-41.

37. Sluis-Cremer N, Jordan MR, Huber K, Wallis CL, Bertagnolio S, Mellors JW, et al. E138A in HIV-1 reverse transcriptase is more common in subtype C than B: implications for rilpivirine use in resource-limited settings. Antiviral Res **2014;** 107:31-34.

38. Mehta SR, Wertheim JO, Delport W, Ene L, Tardei G, Duiculescu D, et al. Using phylogeography to characterize the origins of the HIV-1 subtype F epidemic in Romania. Infect Genet Evol **2011;** 11:975-79,

39. Rimsky L, Vingerhoets J, Van Eygen V, Eron J, Clotet B, Hoogstoel A, et al. Genotypic and phenotypic characterization of HIV-1 isolates obtained from patients on rilpivirine therapy experiencing virologic failure in the phase 3 ECHO and THRIVE studies: 48-week analysis. J Acquir Immune Defic Syndr **2012;** 59:39-46.

40. Glaubius RL, Hood G, Penrose KJ, Parikh UM, Mellors JW, Bendavid E, et al. Cost-effectiveness of injectable preexposure prophylaxis for HIV prevention in South Africa. Clin Infect Dis **2016;** 63:539-47.

Figure legends

Figure 1. RT-E138X mutations are significantly enriched in HLA-B*18-positive individuals globally **A.** Prevalence of RT-E138X mutations in HLA-B*18-positive (black bar) and B*18-negative (white bar) individuals among N=7772 patients from 16 global cohorts. Ns, percentages of E138X mutations, and the Fisher's exact test p-value are shown. **B-C,** Variations of E138X mutations in N=50 HLA-B*18-positive patients (**B**) and N=157 HLA-B*18-negative patients (**C**), shown as percentages. Black bars indicate E138X mutations listed as primary rilpivirine resistance mutations [36]; white bars indicate other E138X mutations. **D,** Prevalence of RT-E138X mutations in HLA-B*18-positive (black bars) and -negative (white bars) patients, stratified by cohort. Percentages of E138X mutations, numbers of study subjects, odds ratios and p values (calculated for cohorts with a minimum of N=10 HLA-B*18-positive patients) are shown. Odds ratios were not calculated for p-values >0.05.

Figure 2. RT-E138X frequency generally mirrors regional HLA-B*18 frequency. **A,** Correlation between RT- E138X prevalence in the overall population and HLA-

B*18 frequency in the 16 study cohorts. **B,** Correlation between E138X prevalence in HLA-B*18-negative patients and HLA-B*18 frequency in the 16 study cohorts. Error bars represent 95% confidence limits, obtained using a binomial error distribution. Spearman's R and p values are shown.

Figure 3. Correlation between E138X prevalence and HLA-B*18 frequency in 43 countries for which unlinked HIV-1 and HLA frequencies were retrieved from public databases. Colors denote the predominant HIV-1 subtype in that country (defined as that observed in >75% of sequences in the Los Alamos HIV database; https://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp); countries where the predominant subtype was <75% are additionally denoted with a black filled semicircle. Cameroon's complex mixed-subtype epidemic is indicated with a full black circle.

Figure 4. Prevalence of E138X mutations in HLA-B*18-positive and -negative patients, stratified by HIV-1 subtype. This analysis includes N=7647 patients from Belize, Vietnam, Japan, Mexico, Guatemala, Honduras, Nicaragua, UK, South Africa, USA, Australia, Uganda, Canada, Panama, and Botswana. "Other" HIV-1

subtypes include subtype F, G, and recombinants other than CRF_01AE. Percentages of E138X mutations, numbers of study subjects, p-values (calculated for subtypes with a minimum of N=10 HLA-B*18-positive patients) and odds ratios (calculated for all associations with $p<0.05$) are indicated.

Figure 5. Effects of E138X mutations on HIV-1 replication fitness. **A,** Replication kinetics of wild-type (E138) and E138X (E138A, E138G, E138K)-containing HIV-1. MT-2 cells were exposed to each virus preparation and p24 concentration in culture supernatant was measured every other days. **B-D,** Competitive HIV-1 replication assay. H9 cells were exposed to mixtures of wild-type (E138) and each of E138X (E138A, E138G, E138K) variant virus preparations. Every week, the supernatant of the virus culture was passaged to new uninfected H9 cells. The cells harvested at the end of each passage were subjected to direct DNA sequencing of HIV-1 reverse transcriptase coding region. The change in the viral population was determined by relative chromatogram peak heights.

Supplemental Figure legends

Supplemental Figure 1. Variant frequency at HIV-1 RT codon 138 in N=60518 RT inhibitor-naive patients, stratified by HIV-1 subtype. Data retrieved from Stanford University HIV Drug Resistance Database on November 3, 2016. Numbers of subjects and E138X frequency are also shown.

Supplemental Figure 2. World map of E138X frequency. Data from 85 countries with a minimum of 50 published sequences in the Stanford University HIV Drug Resistance database [32] are included, along with data from Guatemala, Belize, Nicaragua and Panama from the cohorts studied herein.