ANTIRETROVIRAL THERAPY BLOCKS NATURAL SELECTION ON PROTECTIVE AND DISEASE-SUSCEPTIBLE HLA-B ALLELES IN HIV-1 INFECTION

Nicholas G Herbert[1], Gabriela Cromhout[2,3], Nomonde Bengu[2,4], Rowena Fillis[5], Samantha Kannie[6], Jeroen van Lobenstein[6], Kogielambal Chinniah[7], Constant Kapongo[4], Roopesh Bhoola[5], Malini Krishna[6], Noxolo Mchunu[2], Thilona Arumugam[8], Krista Dong[9,10,11], Sunetra Gupta[12], Christian Brander[13-15], Mary Carrington[9,16,17], Thumbi Ndung'u[2,9,18,19], Veron Ramsuran[8,20], Bridget S Penman[12†], Philip JR Goulder[1,2,9,18*†]

[1]Peter Medawar Building for Pathogen Research, Department of Paediatrics, University of Oxford, Oxford, United Kingdom
[2]HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa
[3]Dept of Paediatrics, University of KwaZulu-Natal, Paediatrics, Durban, South Africa, Paediatrics, Empangeni, South Africa
[4]Queen Nandi Regional Hospital, Empangeni, South Africa
[5]Harry Gwala Regional Hospital, Pietermaritzburg, South Africa
[6]General Justice Gizenga Mpanza Regional Hospital, Stanger, South Africa
[7]Mahatma Gandhi Memorial Hospital, Durban, South Africa
[8]School of Laboratory Medicine and Medical Sciences College of Health Sciences University of KwaZulu-Natal, Durban, South Africa
[9]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA
[10]Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA, USA
[11]Harvard Medical School, Boston, MA, USA
[12]Department of Biology, University of Oxford, Oxford, United Kingdom
[13]Irsicaixa-AIDS Research Institute, 08916 Barcelona, Spain
[14]Facultat de Medicina, Universitat de Vic-Universitat Central de Catalunya (UVic-UCC), 08500 Vic, Spain
[15]ICREA, Passeig de Lluís Companys, 23, Barcelona, Spain
[16]Basic Science Program, Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD, USA
[17]Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA
[18]Africa Health Research Institute, Durban, South Africa
[19]Division of Infection and Immunity, University College London, London, United Kingdom
[20]Centre for the AIDS Programme of Research in South Africa (CAPRISA) University of KwaZulu-Natal, Durban, South Africa

[†] These authors contributed equally

*Corresponding author
Corresponding author information:
Email: philip.goulder@paediatrics.ox.ac.uk
Address: Peter Medawar Building for Pathogen Research, 3S Parks Rd, Oxford OX1 3SY
Tel (Office): (+44) 1865 281884

Key words: Natural selection, HIV, vertical transmission, ART

Key words: Natural selection, HIV, vertical transmission, ART

Classification: Biological Sciences, Population Biology

SIGNIFICANCE STATEMENT

Despite the widely-accepted hypothesis that infectious disease drives HLA diversity, specific examples are rare. We show, in an antenatal cohort in KwaZulu-Natal, South Africa, that maternal HLA-B genotype significantly impacts HIV-1 survival and vertical transmission in the absence of antiretroviral therapy (ART). The introduction of ART in 2004, however, substantially slowed this process of natural selection. Modelling the HIV-1 epidemic in KwaZulu-Natal over a 45-year period between 1990-2035 in the absence of ART, we estimate HIV-1 driving a 38% decline in population frequency of disease-susceptible HLA-B alleles and doubling the frequency of protective HLA-B alleles. Such changes are consistent with previous work describing selective sweeps of MHC-I in HIV-1/SIVcpz-infected chimpanzees, resulting ultimately in persistence of chimpanzee MHC-I protective against AIDS. However, ART has prevented changes of this magnitude occurring over decades in humans.

ABSTRACT

MHC polymorphism is explained by natural selection driven by the MHC-dependent impact of certain infections, inflammatory conditions, autoimmune diseases and cancers. However, examples of human disease driving this process are rare. We evaluated the impact of HIV-1 in altering HLA-I frequencies in KwaZulu-Natal, South Africa, and the influence of antiretroviral therapy (ART) on this process. In a historical mother-child cohort in the pre-ART era (1998-2005), HIV-1 survival and vertical transmission were both strongly HLA-B dependent: 'disease-susceptible' HLA-B alleles (HLA-B*18/B*45:01/B*58:02) increased adult AIDS progression and vertical transmission (OR 1.6, p=0.01), whereas 'protective' HLA-B alleles (HLA-B*57/B*58:01/B*81:01) slowed AIDS progression, and decreased vertical transmission (OR 0.57, p=0.002). By contrast, in contemporary antenatal KwaZulu-Natal cohorts in the ART era (2015-2025) the impact of HLA-B on HIV-1 disease outcome and vertical transmission is dramatically reduced. Using these and reported data, we constructed a model to estimate the impact of HIV-1 on HLA-B frequencies in KwaZulu-Natal, both in the prevailing setting of ART and in a hypothetical counterfactual scenario where ART was never rolled out. Over the 45-year period 1990-2035, in the absence of ART, the proportion of the population possessing any 'protective' HLA-B allele was projected to increase from 23% to 42% (allele frequencies increasing from 0.12 to 0.24), and the proportion of the population possessing any 'disease-susceptible' HLA-B allele was projected to decrease from 28% to 18% (allele frequencies declining from 0.15 to 0.092). The introduction of ART radically slows HLA-B frequency change. These data therefore demonstrate the potential for natural selection from an infectious disease to alter human population genetics within decades, and for the successful roll-out of therapy to halt this process.

INTRODUCTION

The HLA class I region is the most polymorphic in the human genome (1). The association of particular HLA types with outcome from a range of human diseases, especially infectious and autoimmune diseases, but also other inflammatory conditions and cancers (2), is consistent with the notion that HLA-associated human disease drives HLA polymorphism. However, specific examples of selection pressure mediated by human pathogens operating to alter gene frequencies are rare.

The HIV-1 epidemic provides a unique opportunity to investigate the potential impact of a human infection in driving altered HLA frequencies, since HLA-I associations with differential HIV-1 disease progression are well-documented (3–5), and with just over 80 million people infected with HIV-1 since the initial cases were described in 1981 (6), the numbers affected could be sufficient to observe population-level impact on HLA-I frequencies. In particular, this applies to the countries in sub-Saharan Africa that have been the worst-affected, accounting for two-thirds of people living with HIV-1 (PLWH). South Africa alone accounts for >20% of PLWH worldwide, and, within this country, KwaZulu-Natal has been the most severely affected province, with antenatal seroprevalence rates maintained at a relatively stable level of ~40% since antiretroviral therapy roll-out 20 years ago (7). For these reasons, the studies described below focus on cohorts we have studied both pre-ART (1998-2005) and subsequent to ART roll-out (2015-2025) in KwaZulu-Natal, South Africa.

Studies in KwaZulu-Natal have shown that, in ART-naïve infection, the HLA-I molecules most strongly associated with low plasma viral loads, high absolute CD4 counts and therefore slow progression to AIDS are HLA-B*57, HLA-B*58:01 and HLA-B*81:01 (8, 9). The HLA-I molecules associated with high plasma viral loads, low absolute CD4 counts and rapid progression to AIDS are HLA-B*18, HLA-B*45:01 and HLA-B*58:02. The mechanism underlying these HLA-I associations with HIV-1 disease outcome has been related to the ability of HLA class I molecules such as HLA-B*57 to present multiple epitopes within the abundant and highly conserved Gag

capsid protein that enable HIV-specific CD8+ T-cells to recognise and kill virus-infected target cells effectively (10–17). At the same time, the selection of viral escape mutations is severely limited by the cost to viral replicative capacity ('fitness') of variation within the highly conserved regions of the HIV-1 proteome (12, 18–23). By contrast, 'disease-susceptible' HLA such as HLA-B*58:02 present no Gag-specific CD8+ T-cell epitopes (10, 24) and therefore exert little or no selection pressure on this critical region of the virus (25).

Thus, time to AIDS and death in the pre-ART era was strongly associated with the specific HLA-I molecules expressed. In addition, the risk of adult HIV-1 transmission is strongly related to viral load (26–28), and in children born to mothers LWH in the pre-ART era, the risk of vertical transmission was also linked with maternal plasma viral load (29). We therefore hypothesised that, in the pre-ART era, mothers expressing 'disease-susceptible' HLA would have higher plasma viral loads and therefore would be more likely to transmit vertically than mothers expressing 'protective' HLA-I. In the pre-ART era, the prognosis for children following vertical transmission was especially poor, with a 60% 2-year mortality (30). The prognosis for even HIV-exposed, uninfected children whose mothers had AIDS was also poor (30). Thus, we hypothesised that the ability of mothers expressing 'disease-susceptible' HLA-I to have children who themselves would survive to reproductive age would be substantially reduced in the pre-ART era, both because of reduced survival of the mothers and of their children.

To address these hypotheses, we investigated two antenatal cohorts in KwaZulu-Natal, one prior to the ART era (31) (1998-2005), and one subsequent to it (32, 33) (2015-2025). To simulate the impact of HIV-1 on frequencies of 'disease-susceptible' and 'protective' HLA-I in this population, we developed an evolutionary-epidemiological model of HLA-B frequencies based on KwaZulu-Natal. We fitted this model to observations from both the pre-ART and post-ART era, to estimate the changes in HLA-B frequencies that could have occurred in KZN throughout this period. We also conducted a hypothetical counterfactual simulation where ART was never rolled out.

RESULTS

Natural selection on protective and disease-susceptible *HLA-B* alleles in the pre-ART era

Antiretroviral therapy (ART) was introduced in South Africa in 2004, but access was limited initially, with less than 1% of the 5.3m people living with HIV-1 in South Africa in 2004 receiving treatment in that year (34, 35). We studied antenatal cohorts (8, 31) in KwaZulu-Natal, South Africa, in the pre-ART era (1998-2005), at a time when HIV-1 prevalence was increasing from 35% to plateau at a relatively stable ~40% (Fig 1A). Mothers testing seropositive for HIV-1 infection were enrolled during pregnancy and, as expected (29), plasma viral loads (pVL) were higher in vertically transmitting mothers compared to non-transmitting mothers (87,400 median versus 20,100, $p<0.0001$, Fig 1B). We had previously shown in the Zulu-Xhosa population in KwaZulu-Natal (8, 9) that HLA-B*57, HLA-B*58:01 and HLA-B*81:01 were 'protective' against HIV-1 disease progression, being associated with low pVL and high absolute CD4 counts, and that HLA-B*18, HLA-B*45:01 and HLA-B*58:02 are 'disease-susceptible', being associated with high pVL and low absolute CD4 counts. Not unexpectedly, therefore, these same associations were observed in the cohort of mothers enrolled antenatally here (Fig 1CD, FG). Furthermore, as we had hypothesised, the frequencies of disease-susceptible HLA-I were significantly higher in the transmitting mothers and their infants than in the non-transmitting mothers ($p=0.01$ (odds ratio 1.6) and $p=0.006$ (odds ratio 1.7)), respectively; these were also higher than in a cohort of HIV-uninfected adults in KwaZulu-Natal ($p=0.003$ and $p=0.003$, respectively, Fig 2A). Similarly, the frequencies of protective HLA-I were significantly lower in the transmitting mothers and their infants than in the non-transmitting mothers ($p=0.002$ (odds ratio 0.47) and $p=0.004$ (odds ratio 0.5), Fig 2B) and somewhat lower than in the control cohort of uninfected adults, although this latter difference did not reach statistical significance. Thus, in ART-naïve infection, 'protective' HLA-B are associated with slow HIV-1 disease progression to AIDS and death, whereas 'disease-susceptible' HLA-B alleles are associated with rapid progression to AIDS and death in the mothers, as well as an increased risk of vertical transmission and subsequently reduced infant survival. This process of natural selection, if continued without intervention, would therefore tend to increase the frequencies of 'protective' HLA-B and reduce the population frequencies of the disease-susceptible HLA-B alleles.

Weakened natural selection on *HLA-B* alleles in the ART era

By 2023, an estimated 5.9m (77%) of the 7.7m PLWH in South Africa were receiving ART (36). Long-term survival is now the expected outcome in PLWH receiving ART. Similarly, ART has dramatically reduced HIV-1 transmission rates, since HIV-1 transmission is effectively completely prevented when viraemia is suppressed by ART (37). Accordingly, the intra-uterine vertical transmission rate is 0.4% in the current ART era (32), compared to the vertical transmission rate for *in utero* infection of 7.5% in the pre-ART era (31, 38). Most of the vertical transmissions in the current era arise when, for social reasons, the mothers have not been able to access ART, or have not been aware of their HIV-1 status, either having become infected themselves during pregnancy or testing seropositive for the first time during the pregnancy (32). Thus, we hypothesised that, in the current ART era, both long-term survival in PLWH and vertical transmission are no longer related to HLA-I-mediated immune control of HIV-1 viraemia, but are now more strongly affected by HLA-independent factors that determine whether ART is accessed.

To test this hypothesis with respect to vertical transmission, we studied a second mother-child cohort enrolled in KwaZulu-Natal during the ART era (2015-2025) (32). We compared the frequencies of protective and disease-susceptible HLA-B molecules in transmitter mother-child pairs with those of a contemporaneous HIV-uninfected cohort of age-matched women in KwaZulu-Natal (39, 40) and observed that there are no significant differences between the frequencies of either of these HLA-B alleles between HIV-uninfected women and transmitter mothers and their infants (Fig 2C-D). The numbers of study participants studied in the ART era were sufficient to provide >98% power (<2% type II error rate) to detect a significant difference between the groups had the difference between transmitter mothers and HIV-ve adults in the ART era been the same as that which we observed in the pre-ART era. Comparing HLA-B frequencies in the transmitter pairs in the pre-ART era versus the transmitter pairs in current ART era, as hypothesised, we observed that the frequencies of ‘disease-susceptible’ HLA-B alleles were significantly lower in the ART era compared to the pre-ART era, whereas these do not differ between the respective HIV-uninfected cohorts (Fig 2E, Table S1). As anticipated, ‘protective’

HLA-B alleles were somewhat increased in frequency in the transmitter mother-child pairs in the ART era compared to the pre-ART era, albeit not significantly so (Fig 2F).

These analyses were also undertaken individually for the three 'disease-susceptible' and for the three 'protective' HLA-B alleles (Fig 3A-F). The most prevalent of these six alleles is the disease-susceptible HLA-B*58:02. This allele was enriched in the transmitter mother-child pairs in the pre-ART era (phenotypic frequency 33%) but not so in the current ART era (phenotypic frequency 22%) (Fig 3C).

With respect to the relationship between HLA-B expression and survival in adult LWH in the current ART era, as described above, in the antenatal clinic in KZN we studied, 90% of mothers LWH have undetectable plasma viral loads at delivery and 95.7% had plasma viral loads of <1000 c/ml (Fig S1A). These findings are in close agreement with published data for all antenatal clinics in KZN, reporting >80% of mothers with undetectable plasma viral loads, and 94.1% of mothers with plasma viral loads <1000 c/ml (7). In the 0.4% of mothers LWH in KZN who transmit the virus in the ART era, there is no significant effect of HLA-B on plasma viral load. However, it is noteworthy that, even though absolute CD4 counts are relatively high among transmitter mothers in the ART era (mean 487 versus 320 cells/ul in the pre-ART era), an HLA-B effect on absolute CD4 counts and CD4% remains (Fig S1CD). Thus, although ART is weakening natural selection on HLA-B alleles, the fact that in adults ART is not initiated immediately following acquisition of HIV, and in a subset of people, illustrated by transmitter mothers, ART adherence is suboptimal, demonstrates that the selection pressure imposed by HIV-1 on HLA-B has not completely halted even in the ART era.

ART weakens selection of disease-susceptible, high-expressing *HLA-A* alleles during vertical transmission

Having observed evidence of the HIV-1 epidemic driving changes in HLA-B frequencies, we next considered the possibility that the same selection pressures might be operating on the HLA-A locus. Although differences between HLA-A molecules have substantially less impact on HIV-1

pVL in ART-naïve individuals than differences in HLA-B molecules (8, 9), high HLA-A expressing alleles are associated with high pVL and low absolute CD4 counts in ART-naïve adult and paediatric infection, mediated through inhibition of NKG2A+ NK cell responses by high-expressing HLA-A alleles (41, 42). In the pre-ART and ART era cohorts studied here, surprisingly, we observed no difference overall in HLA-A expression levels between the different groups (Fig S2A). However, individuals with disease-susceptible, high-expressing HLA-A alleles (z-score >0.5) were significantly increased in frequency amongst the transmitter mother-child pairs in the pre-ART era compared to the ART era ($p<0.0001$, Fig S2B). Similarly, children with protective, low-expressing HLA-A alleles (z-score <-0.5) were somewhat more frequent among the transmitter mother-child pairs in the ART era compared to the pre-ART era ($p=0.007$, Fig S2C). In analyses of individual HLA-A alleles, the lowest-expressed and therefore most protective HLA-A allele is HLA-A*74:01, and children expressing HLA-A*74:01 were present at a higher frequency in the ART era versus the pre-ART era ($p=0.005$ pre-Bonferroni-Holm correction for multiple tests; $p=0.085$ post-correction, Fig S3D). Taking into account the fact that HLA-A*74:01 is in linkage disequilibrium with the strongly protective HLA-B*57:03, we reanalysed these data removing individuals expressing HLA-B*57:03. There remained a somewhat increased frequency of the protective HLA-A*74:01 in the ART era children versus the pre-ART era children ($p=0.017$, uncorrected for multiple tests).

Simulating the KZN HIV-1 epidemic: ART radically slows allele frequency changes

To evaluate the impact of ART on HLA-B population genetics in the KwaZulu-Natal HIV-1 epidemic, we developed an evolutionary-epidemiological model of HIV-1 transmission and treatment (Fig 4) in a population containing 3 possible types of HLA-B allele: protective, disease-susceptible and indifferent. We began our simulation in the year 1990, using allele frequencies inferred from the HIV-negative HLA-typed dataset from the year 2000, since this was the earliest time point for which we had genetic data. We assumed that genotype frequencies were at Hardy Weinberg equilibrium, and started our simulation with an HIV-1 prevalence of 0.016 (the proportion of HIV-infected women in KZN reported in the National HIV Surveillance Programme 1990 antenatal survey (44)). We assumed, based on behavioural observations (48), that the

parameter determining the horizontal transmissibility of HIV-1 could take 3 different values in 3 different time periods (1990-1998; 1999-2008 and 2009 onwards), and that ART treatment rates, when non-zero, took 2 different values (2004-2016 and 2017 onwards) (see Supporting Information). We obtained suitable values of these transmission and treatment parameters by fitting the model to antenatal HIV survey data reported for KZN province (7, 43–45), and to ART treatment data from Conan *et al* (46) and the Sentinel Survey (7), using Bayesian MCMC (full details in Supporting Information). The datapoints used for fitting are visualised in figure 4A.

Amornkul *et al* (47) report times to AIDS for HIV-1 C (the clade of virus that predominates in Southern Africa) and also reports the impact of having protective or disease susceptible HLA-B alleles on times to AIDS in sub Saharan African populations. Our pre-ART data allows us to estimate the odds ratios for disease susceptible or protected HLA-B genotypes vertically transmitting HIV relative to genotypes containing neither such allele (Table S2). We brought together this information to generate three different HLA-B evolutionary scenarios: (i) a best estimate scenario, in which HLA-B specific times to AIDS and risks of vertical transmission are taken from the relevant hazard ratio or odds ratio calculated directly from the aforementioned data sources; (ii) a conservative scenario, in which the impact of HLA-B on times to AIDS and risks of vertical transmission are the minimal values implied by the data (e.g. the lower bound of the 95% confidence interval of a hazard ratio or odds ratio if the ratio is greater than 1 or the upper bound of that interval if the ratio is less than 1), and (iii) an extreme scenario, in which the impact of HLA-B on times to AIDS and risks of vertical transmission are the maximum values implied by the data (e.g. the upper bound of the 95% confidence interval of a hazard ratio or odds ratio if the ratio is greater than 1 or the lower bound of that interval if the ratio is less than 1). Full details of these scenarios are provided in the Supporting Information and Table S3.

Despite the devastating impact of the HIV-1 epidemic on KZN throughout the 1990s and early 2000s, frequencies of protective and disease-susceptible alleles only shifted from 0.12 to 0.14 (protective) and 0.15 to 0.14 (susceptible) between 1990 and 2004 in our best-estimate HLA-B properties scenario (Fig 4). This equates to 28% of the population possessing any susceptible

allele in 1990 decreasing to 25% of the population possessing any susceptible allele in 2004 (and 23% of the population possessing a protective allele in 1990 increasing to 27% of the population possessing a protective allele in 2004). Within 15 years of ART becoming available, allele frequencies start to stabilise, with the protective allele at a frequency of 0.17 and the susceptible allele at a frequency of 0.12 in 2019 in our best estimate scenario. These frequencies only shift slightly more by 2035 (0.18 protective; 0.12 susceptible), resulting in the 2035 population having 22% with any susceptible allele and 32% with any protective allele. The extreme HLA properties scenario and the conservative HLA properties scenario result in greater and smaller overall allele frequency changes respectively but show the same stabilising effect of ART (Fig 4A).

In the absence of ART, in our best estimate scenario, the susceptible allele is at a frequency of 0.092 in 2035 (a decline of 38% compared to its initial allele frequency in 1990), and the protective allele has reached an allele frequency of 0.24 (double its allele frequency in 1990). This equates to just 18% of the population possessing any susceptible allele and 42% of the population possessing any protective allele in 2035, had there been no introduction of ART.

Figure 4B displays the allele frequency changes that occur by 2035 in the absence of ART, with four different values for the baseline probability of vertical transmission. The baseline probability of vertical transmission is the probability of vertical transmission of HIV from infected mothers who are homozygous for indifferent HLA-B alleles. The protective or disease susceptibility effects of other genotypes are included in the model by changing their risks of vertical transmission relative to this baseline. When the baseline probability is 0%, the only mechanism driving HLA-B frequency changes is the impact of HIV on lifespan. As the baseline probability of vertical transmission is increased, we can see that vertical transmission does have an effect on allele frequencies, but that this effect is small compared to that driven by the impact of HIV on lifespan.

In our best-estimate and extreme HLA-B evolutionary scenarios, protective and disease susceptible HLA-B genotypes affect both times to AIDS and vertical transmission of HIV, and the greater the level of vertical transmission, the greater the increase in protective allele (or decrease

in susceptible allele) frequencies by 2035 (Fig 4B). In our conservative scenario, HLA-B has a relatively small effect on times to AIDS and does not affect the probability of vertical transmission (see Supporting Information and tables S2 and S3). Figure 4B thus illustrates the counter intuitive fact that if HLA-B genotype affects time to AIDS but not vertical transmission, higher rates of vertical transmission of HIV reduce the rate of allele frequency change. If vertical transmission is experienced to the same extent by all HLA-B genotypes, the relative advantage of having a longer time to AIDS is lessened.

DISCUSSION

We have shown here, within a maternal cohort in KwaZulu-Natal, South Africa in the pre-ART era (1998-2005), HIV-1 survival and vertical transmission are both HLA-B dependent. AIDS-free survival and non-transmission are favoured by the expression of 'protective' HLA-B (HLA-B*57/58:01/81:01) whereas rapid progression to AIDS and vertical transmission are favoured by the expression of 'disease-susceptible' HLA-B (HLA-B*18/45:01/58:02). Using data from this cohort and other published studies to describe this process of natural selection, we developed a model to estimate the impact of HIV-1 on HLA-B frequencies over time in KZN. Our best estimate of the interaction between HIV and HLA-B suggests that, in the absence of ART, HIV selection could have doubled the frequencies of protective HLA-B alleles and reduced the frequencies of disease-susceptible HLA-B allele frequencies by 38% over a 45-year period. Even more extreme allele frequency changes are within the bounds of possibility (Figure 4A). Evaluation of a KwaZulu-Natal cohort in the ART era (2015-2025) showed that the HLA-B-dependent effects on HIV-1 survival and vertical transmission are substantially weakened by ART, and our model suggests this weakening is sufficient to radically slow HLA-B allele frequency change.

These studies are consistent with previous work on chimpanzees, proposing that, 2-3 million years ago, a selective sweep of chimpanzee MHC-I occurred as a result of widespread HIV-1/SIVcpz infection, removing animals expressing MHC-I associated with HIV-1/SIVcpz-associated disease (49). The chimpanzees surviving today have a relatively small repertoire of MHC-I molecules that are, however, associated with resistance to HIV-1/SIVcpz AIDS (50–52). Bonobos

appear to have an even more diminished MHC class I repertoire compared to chimpanzees, which suggests that the selective sweep may have predated the speciation of common chimpanzees and bonobos (53). It is striking that the HIV-1-specific epitopes targeted by CD8+ T-cells restricted by these protective chimpanzee MHC-I molecules, (Patr-03:01, Patr-B*01:01/03:01/05:01) are virtually the identical Gag-specific epitopes that are presented by protective HLA-B (HLA-B*27/57/58:01) molecules in humans (5, 51). Similarly, in rhesus macaques (*Macaca mulatta*) and pig-tailed macaques (*Macaca nemstrina*), that are not naturally infected with HIV-1 or SIV, immune control of experimental SIV infection is observed in association with protective MHC-I molecules (Mamu-A*01, Mamu-B*08, Mamu-B*17, Mamu-A1*065:01 (90-120-Ia), Mane-A*10) that target virtually the identical epitopes presented by protective HLA in humans, HLA-B*27/57/58:01/81:01 (5, 54–58). These findings support the hypothesis that immune control of immunodeficiency virus infection is mediated by MHC-I molecules that can present Gag epitopes for recognition by virus-specific CD8+ T-cells. These epitopes are highly abundant and conserved, with the result that virus-infected target cells are recognised and killed early in the viral life cycle (14), and immune escape is limited by purifying selection against low fitness variants (22, 23).

The impact of HIV-1 described here in driving increases in protective HLA-B allele frequencies and decreases in disease-susceptible HLA-B allele frequencies over the course of the epidemic in the absence of ART is also broadly consistent with previous studies that have estimated changes in gene frequencies as a result of HIV-1 -mediated natural selection (59, 60). The analysis by Cromer *et al*, based on a hypothetical population where the HIV prevalence plateaus at 30%, focused on the disease-susceptible HLA-B allele combination, HLA-B*35-Px/B53, in people living with HIV-1 B clade infection, the clade of virus that predominates in the Western world. In this study, the authors estimate a 50% decrease in frequency of this ‘frail’ HLA-B allele combination over 50 years. These estimates are similar to the 38% reduction over 45 years of the disease-susceptible HLA-B allele combination HLA-B*18/45:01/58:02 calculated here for the KwaZulu-Natal population affected by C clade HIV-1.

What is distinct about the current study is, first, that it is focused on a population in Sub-Saharan Africa, where two-thirds of the global HIV-1 pandemic is centred; second, the HIV-1 seroprevalence figures here of 30-40% corresponds to the specific situation for antenatal populations in KZN (Fig 1A); likewise, survival times in people living with C clade infections are derived from data in people living in Sub-Saharan Africa, and these differ from the survival figures in the PLWH in North America (the average time to AIDS when infected with clade C is 4.5 years (47), as opposed to the average time to AIDS of ~10 years seen in North America (61)); third, we have taken into account the impact of HLA on vertical transmission and child survival, again using actual data; fourth, we have included analysis of 'protective' HLA-B, where the increase over time (doubling over 45yrs in our best estimate scenario) is faster than the decrease in disease-susceptible HLA-B; and, finally, we have incorporated the impact of ART roll-out in the current analysis.

The current study is also broadly consistent with a study of the impact of CCR5 variants (59), the coreceptor that facilitates entry of HIV-1 into target CD4+ T-cells via binding to CCR5 in addition to CD4. That HLA has strongest genetic effect on HIV-1 outcome is abundantly clear from several genome-wide association studies (3, 62–65). The HIV-resistant CCR5 variants most prevalent in Africa populations affect HIV-1 survival by approximately 2yrs (66), and these effects are therefore weaker than those of protective HLA-B (see Table S3). Correspondingly, the estimates are that the CCR5 variants affecting HIV-1 survival would alter more slowly than HLA-B alleles affecting HIV-1 survival, with HIV-resistant CCR5 variants increasing by 33% over 100 years and HIV-susceptible variants decreasing by 50% over 100 years.

Although the major HLA effects on HIV-1 outcome are through HLA-B (3, 62–65), differences in HLA-A genes also affect HIV-1 disease outcome, high-expressed HLA-A being associated with high viraemia and rapid progression to AIDS (41, 42). We did not observe evidence of an impact of HLA-A on AIDS progression and HIV-1 survival in the adult cohorts we studied here. However, we did observe a striking significant enrichment of individuals with high expressing HLA-A alleles

(z-score >0.5) in the transmitter mother-child pairs during the pre-ART era compared to non-transmitters and seronegative individuals ($p<0.0001$). This enrichment had disappeared in the ART era transmitter pairs ($p<0.0001$, comparing high expressing HLA-A frequencies of transmitter pairs in the pre-ART versus ART era). These data are therefore consistent with the HLA-B data in showing that disease-susceptible HLA-I alleles were enriched in transmitter pairs in the pre-ART era, and that following the advent of ART, they are no longer significantly different from uninfected cohorts. Overall, however, the part played by HLA-A in affecting survival of children born to mothers living with HIV-1 is quite small relative to the impact of HLA-B on adult survival, and, given that any HLA-A effect on adult survival is only evident after removal of the HLA-B effect, it is unlikely that significant changes in HLA-A frequencies would result over the course of an ART-free HIV-1 epidemic.

It is important to highlight limitations of this study. One limitation is the size of our cohorts, and the time points at which populations were sampled. We do not show a change in frequency of protective and disease-susceptible HLA-I resulting from the differential impact of HIV-1 on survival and reproductive rate described above. The reason is, first, that HLA data from the same Zulu ethnic group being studied here are not available pre-1990. Second, HLA typing at this time depended on serological methods (67) and HLA types were 2-digit only. As a result, detailed, large cohort (>1000 individuals) studies during this period were unable to distinguish, for example, between HLA-B*58:01 (protective) and HLA-B*58:02 (disease-susceptible), and HLA-B*81 did not feature at all, being difficult to distinguish from HLA-B*07 or HLA-B*42. Up to 10% of individuals were not successfully typed owing to limitations in the class I antisera available. Third, although HIV-1 seroprevalence in antenatal mothers increases exponentially during the period 1990-2000, >80% of antenatal mothers were seronegative for most of this decade, and at the population level HIV-1 seroprevalence would be lower still, especially among males who account for <40% of infections (68). Thus, changes in HLA frequencies resulting from the differential impact of HIV on disease outcomes would take some time to become evident at the population level. This is reflected in the data generated by the model here and is consistent also with the model of Cromer et al (60).

We do observe the phenotypic frequency of protective HLA-B alleles increasing marginally from 22.7% in the uninfected pre-ART cohort from the year 2000 to 23.6% in the uninfected ART era cohort from 2015-2025 (Fig 2F); but this difference is not statistically significant. Nevertheless, the magnitude of the change is consistent with the behaviour of the model (specifically, between 2000 and 2020 in the model, the proportion of the population with any protective allele at all changes from 25% to 30%). We do not observe an equivalent change in the phenotypic frequency of disease-susceptible HLA-B alleles between the two uninfected cohorts (Fig 2F).

It is important also to note that, whereas in the analyses of HLA frequencies in the pre-ART era, a cohort of non-transmitter mothers were available to compare against the transmitter mother-child pairs, such a cohort was not available in the ART era. In the pre-ART era HIV-uninfected adults served as an additional comparator group, and while the HLA frequencies in this group did not differ significantly from those in the non-transmitter mothers, the disease-susceptible HLA-B frequency in HIV-uninfected adults was significantly lower than in the transmitter mother-child pairs. In the ART era, ART coverage is high and, specifically, ~95% of antenatal mothers have low or undetectable viral loads (<1000 c/ml; (7) & our data, Fig S1). With these levels of viral suppression, it is not surprising that birth transmission rates in the ART era are <1% (32). Non-transmitter mothers in the ART era therefore represent >99% of mothers LWH. Given this strong evidence of very effective control of HIV among non-transmitter mothers in the ART era, it is reasonable to argue that HIV-uninfected adults in the ART era represent an appropriate contrast transmitting mother-child pairs. Of note, the risk of adults (or mothers) themselves becoming HIV infected is independent of their HLA-I type. This has been shown both in GWAS studies (69, 70) and in studies that have focused specifically on HLA-I susceptibility to infection (71, 72).

A further point to note in relation to the finding of no significant difference in disease-susceptible HLA-B frequencies in the ART era between HIV-uninfected adults and transmitter mother-child pairs is the possibility that sample sizes were insufficient to detect the difference of the magnitude that was observed in the pre-ART era. The sample sizes of n=524 in the HIV-ve adults and n=317

488 in the transmitter mothers in the ART era provide >98% power (<2% type II error rate) to detect a
489 significant difference of this magnitude between these two groups in the ART era.
490
491 Whilst the analyses presented here indicate that the process of natural selection driven by HIV in
492 the pre-ART era has been substantially slowed by ART, it is not possible to say that this process
493 has been halted altogether. The reason for this is that ART coverage is not 100% from the time of
494 acquisition of infection, and, as highlighted in the paper, disease progression to AIDS in Sub-
495 Saharan Africa for people who express disease-susceptible HLA-B may only take 3-4 years (see
496 supporting information and Table S3).This is well illustrated by the analysis of the 0.4% of mothers
497 LWH who are transmitters: although viral load is rapidly suppressed by ART, the differential
498 impact of HIV on CD4 counts according to HLA-B type shows that this process of natural
499 selection is not completely halted (Fig S1). Indeed, were ART to be interrupted for any sustained
500 period at the population level for whatever reason, it is clear that the HLA-dependent effects
501 observed in the pre-ART era would be activated very rapidly once again.
502
503 An important limitation of our modelling analysis is the relative lack of information about how
504 different HLA-B alleles combine to impact HIV-1 progression. The classification of HLA-B alleles
505 as being relatively protective or risky for HIV-1 is well-supported by multiple lines of evidence, but
506 the existence and completeness of any dominance effects is less clear. Furthermore, relatively
507 few data are available for HIV-1 subtype C, the type which predominates in KZN. We used the
508 average time-to-AIDS data reported by Amornkul *et al* (47) for subtype C in our model. We used
509 Amornkul's hazard ratios for the presence/absence of B*57 and B*45 to adjust this time to AIDS
510 up or down for "protective" and "risky" genotypes, since these are the only available estimates of
511 the impact of HLA-B on time to AIDS that are specific to sub Saharan African populations.
512 However, Amornkul *et al*'s HLA-B results are based on their entire dataset (which is not
513 exclusively subtype C). We had to treat results for the presence/absence of B*57 and B*45 as
514 indicative of all protective and disease susceptible allele containing genotypes for the purposes of
515 our model, which was an extrapolation. We further assumed that a compound heterozygote for a
516 protective and a disease susceptible HLA-B allele would experience the population average time

517 to AIDS (i.e. we assumed the protective and risky effects cancelled each other out) – see Table
518 S3. Such heterozygotes are relatively rare in the model, but it would still have been far better if
519 estimates of the relative risk of HIV-1 C progression for all possible genotypic combinations were
520 available to parametrise the model.
521
522 We made the simplifying assumption that the onset of AIDS meant an individual no longer
523 contributed to HIV-1 transmission and would no longer successfully reproduce. Fertility rates and
524 rates of sexual contact are much lower after the onset of AIDS in women (73), but it would have
525 been more realistic to model a continuous decline according to WHO disease stages, especially if
526 it were possible to stratify the rate of progression between these stages by HLA genotype.
527
528 In addition to the HIV/AIDS specific limitations detailed above, our model makes multiple
529 simplifying assumptions in representing a population within an evolutionary-epidemiological
530 framework (which attempts to account for rapid infectious disease dynamics and long-term
531 population genetic trends within the same model). We assumed a constant birth rate over time,
532 and assumed there was no difference in the fertility of women LWH versus women without HIV.
533 We modelled only the female population, implicitly assuming that the impact of HIV on women
534 and their HLA frequencies would be mirrored in the male population. In fact, HIV rates in KZN
535 show sex specific patterns, with incidence rates peaking at lower ages in women than in men,
536 and with shifts in this pattern over time (74); to account for this would require extensive data not
537 only on those age specific patterns but also on how sexual contact between different age groups
538 has changed, or not, over time. Despite these challenges, it would be extremely interesting to
539 incorporate greater population stratification in future models and explore which mechanisms
540 accelerate or decelerate HLA allele frequency changes.
541
542 As noted in the Introduction, it is widely accepted that MHC/HLA polymorphism is driven by
543 natural selection from infectious diseases. Superficially, our results appear to contradict this, since
544 the model would indicate that selection from HIV-1 reduces HLA diversity (indeed, the long-term
545 outcome of our model is that disadvantageous alleles are driven to extinction within a few

hundred years in the extreme and best-estimate scenarios, and are at extremely low levels within 600 years even in the conservative scenario, Fig S3). However, we do not predict this would be the long-term outcome in any population, even in the absence of ART. Two processes not captured by our model are (i) co-evolution, whereby the pathogen evolves alongside the host, as previously described for HIV in relation to protective HLA-B (75); and (ii) selection from multiple antigenically diverse pathogens. We have shown the impact of a single pathogen over a short time. Over evolutionary time, processes similar to those shown here would play out repeatedly, driven by different pathogens and favouring different MHC/HLA types. Even if lost from a sub-population, MHC/HLA types can be re-introduced via migration or even introgression between species. Selection from a HIV/SIV like pathogen may have skewed MHC frequencies in the ancestors of chimpanzees and bonobos as proposed by de Groot *et al* (49, 53), but apes, like humans, also have to contend with completely different pathogens such as *Laverania* malaria parasites, whose evolutionary impact on bonobo MHC frequencies has also recently been demonstrated (76). It would not be realistic to expect any primate to have solely adapted to HIV/SIV.

In summary, here we demonstrate that HLA-B genotype impacts not only HIV-1 survival rates but also the vertical transmission of HIV-1, in the absence of ART. We also demonstrate the potential for an infectious disease to rapidly alter HLA frequencies. The HIV-1 pandemic in KZN, if left untreated, could have doubled protective HLA-B allele frequencies and decreased disease susceptible HLA-B alleles by 38% within 45 years. The introduction of ART, together with high levels of adherence and successful suppression of viraemia, is substantially weakening this process of natural selection.

MATERIALS AND METHODS

Study populations

The non-transmitting mothers and transmitter pairs from the pre-ART era were part of a historical cohort of early-treated infants living with HIV-1 followed in 2002–2005 in Durban, KwaZulu-Natal, South Africa (PEHSS cohort, Paediatric Early HAART and Structured Treatment Interruption Study) (31). Antenatal mothers were recruited from October 2002, and paediatric study subjects were enrolled between July 2003 and September 2005. At this time, single dose nevirapine (sd-NVP) was the only intervention available for prevention of mother-to-child transmission. Mothers were given sd-NVP (200 mg) at the onset of labour and infants were given sd-NVP (2mg/kg) within 72h of birth. Plasma viral loads were determined from 662 mothers and 71 infants using Roche Amplicor assay version 1.5 (Roche Molecular Systems, Branchburg, New Jersey, USA). The HIV-uninfected control cohort from the pre-ART era was a cohort of 110 blood donor adults enrolled by the Natal Blood Transfusion Service, Pinetown, Durban, KwaZulu-Natal, South Africa in the year 2000. The PEHSS cohort was approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal, Durban and the Institutional Review Board of the Massachusetts General Hospital, Boston, Massachusetts, USA.

The uninfected cohort from the ART era (2015-2025) comprised 492 females from the Females Rising through Education, Support and Health (FRESH) study (39) and 32 uninfected mothers from the Ucwaningo Lwamawele cohort (meaning ‘Learning from Twins’) study, both of which are ongoing studies being conducted in Durban, KwaZulu-Natal. The ART era transmitter pairs are part of the Ucwaningo Lwabantwana (meaning ‘Learning from Children’) cohort which is a cohort of >300 in utero-infected children enrolled and followed in KwaZulu-Natal, South Africa from 2015-2025 (32). At birth 92% of mothers are receiving combination antiretroviral therapy. The Ucwaningo Lwabantwana cohort is an ongoing study which has been approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (BF450/14) and the Oxfordshire Research Ethics Committee (06/Q1604/12). Written informed consent for the infant’s and mother’s participation in the study was obtained from the mother or the infant’s legal guardian.

The FRESH study was approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (BE699/18), where all study participants provided written informed consent.

All individuals from the cohorts included in this study were of African descent, from KwaZulu-Natal, South Africa.

HLA typing

Samples from the PEHSS cohort (non-transmitting mothers and transmitter pairs from 2002-2005) were HLA typed using a targeted next-generation sequencing (NGS) method. Briefly, locus-specific primers are used to amplify HLA-A and B (exons 1 to 4) and C (exons 1 to 5) genes with Fluidigm Access Array (Fluidigm Singapore PTE Ltd, Singapore). The Fluidigm polymerase chain reaction (PCR) amplicons are pooled and subjected to sequencing on an Illumina MiSeq sequencer (Illumina, San Diego, CA 92122 USA). HLA alleles and genotypes are called using the Omixon HLA Explore (version 2.0.0) software (Omixon, Budapest, Hungary). For the HIV-negative KZN adults from 2000, genomic DNA samples were initially typed to an oligo-allelic level using Dynal RELITM reverse Sequencing Specific Oligonucleotide (SSO) kits for the HLA-A, -B, -C loci (Dynal Biotech). Refining the genotype to the allele level was performed using the Dynal Biotech sequence-specific priming kits in conjunction with the previous SSO type. Where alleles were still not defined to the allele level, bespoke sequence-specific priming primer mixes were utilised. All HLA class I alleles in the IMGTallele release 2.4.0 were considered for typing (77). For samples from the ART era transmitter pairs (2015-2025), FRESH cohort and Ucwaningo Lwamawele cohort, HLA typing was done using in house locus-specific primers which were used to amplify HLA-A/B and C (exons 2 and 3) genes. The products were purified and sequenced using in-house designed primers on an AB3730 DNA analyser. Traces were analysed using bespoke software and BioEdit.

Statistics

For scatterplots, median values and IQRs are indicated. Comparisons were performed using Fisher's exact tests for categorical variables and Mann–Whitney U test for continuous variables. In Figure 1C-G, those with both a 'protective' and a 'disease-susceptible' HLA-B allele were excluded from the analyses. In the analyses shown in Fig 2-3, all individuals were included. P-value less than 0.05 was considered significant. More information regarding statistical tests for Figure 1 and 2 can be found in Table S1. All calculations and graphs performed using GraphPad Prism v10.2.3 (GraphPad Software) except for the evolutionary-epidemiological modelling (see below).

Modelling

We used a system of linked ordinary differential equations to model a population in which individuals belonged to one of 6 different HLA-B genotypes (homozygous for a disease-susceptible HLA-B; homozygous for an indifferent HLA-B; homozygous for a protective HLA-B; heterozygous for disease-susceptible and indifferent; heterozygous for protective and indifferent, or heterozygous for protective and disease-susceptible). As well as being classified by genotype, individuals were either susceptible to HIV-1, infected with HIV-1 or infected with HIV-1 and treated, and the rate of becoming infected with HIV-1 depended upon the current prevalence of untreated HIV-1 in the population. Untreated HIV-1 infection led to AIDS at genotype-specific rates, based on observations of sub-Saharan African populations (47) – see Supporting Information for a full description. Vertical transmission of HIV-1 occurred at maternal-genotype-specific rates, based on the data in this manuscript, and we assumed that children who were infected vertically would not survive to reproductive age. Since any estimate of the impact of HLA-B on HIV progression or on vertical transmission is inherently uncertain, we devised three scenarios: a best estimate scenario, in which HLA-B genotypes were assigned times to AIDS and probabilities of vertically transmitting HIV based on reported hazard ratios and odds ratios (see Supporting Information); a conservative scenario, in which HLA-B genotypes were assigned the least possible impact on times to AIDS and probabilities

650 of vertically transmitting HIV implied by the data (using the upper bound of the confidence interval
651 for a hazard ratio or odds ratio if that ratio is below 1, and the lower bound if that ratio is above 1),
652 and an extreme scenario, in which HLA-B genotypes were assigned the greatest possible impact
653 on times to AIDS and probabilities of vertically transmitting HIV implied by the data (the lower bound
654 of a confidence interval for a hazard ratio or odds ratio if that ratio is below 1, and the upper bound
655 if that ratio is greater than 1). A full explanation of these scenarios, and the values used, are
656 provided in the Supporting Information.

657 We began our simulations in 1990, with realistic frequencies of the protective, disease-susceptible
658 and indifferent alleles (based on a HIV-negative HLA typed dataset from the year 2000), genotypes
659 at Hardy Weinberg equilibrium, and with HIV-1 prevalence at 0.016, (the proportion of HIV-infected
660 women in KZN reported in the National HIV Surveillance Programme 1990 antenatal survey (44)).
661 We intended the model to apply to antenatal populations, hence fitted it to antenatal survey data as
662 well as data on ART treatment levels specific to women in KZN (7)(see Supporting Information and
663 Table S6).

664 We chose a set of realistic demographic parameters (see Table S3) and set the baseline probability
665 of vertical transmission of HIV-1 at 35%, based on vertical transmission probabilities estimated by
666 de Cock *et al* (78) (whose maximum estimate is 45%, if mothers breastfeed for up to 2 years). We
667 fitted horizontal HIV-1 transmission parameters and ART treatment rate parameters using Bayesian
668 MCMC with a Binomial likelihood function, since the HIV prevalence data and ART treatment data
669 consist of proportions (see Supporting Information for full details of the fitting process). We fitted
670 different transmission parameters to the time periods 1990-1998, 1999-2008 and 2009-onwards to
671 account for known behavioural changes in South African populations, and different treatment rate
672 parameters for 2004-2016 and 2017- onwards to account for the introduction of Universal Test and
673 Treat in 2016. We performed this fitting process separately for each of our best estimate,
674 conservative and extreme HLA-B property scenarios, since different times to AIDS for a proportion
675 of the population require different transmission parameters to best fit the available data.

To generate counterfactual (no ART) simulations, we took the transmission parameter fitted for 1999-2008 for the relevant scenario and allowed that level of transmission to continue indefinitely.

To investigate the potential impact of differing levels of vertical transmission of HIV-1, we varied the baseline probability of vertical transmission between 0 and 45% (figure 4B), keeping all other aspects of each scenario the same (i.e. the “best estimate” scenario still used the relevant “best estimate” parameters from tables S3 and S5, likewise the other HLA-B scenarios).

The full equations for the model are given in the Supporting Information, and the parameters of the model are explained in detail in Tables S3-5. The model was implemented in Matlab, using the ode45 solver.

ACKNOWLEDGEMENTS AND FUNDING SOURCES

This work received funding from the Wellcome Trust (PG WTIA Grant WT104748MA) and the National Institutes of Health (AI133673 and U01AI168655 to P.G.; 1UM1AI164566 (the Pediatric Adolescent Virus Elimination (PAVE) Martin Delaney Collaboratory Grant) to P.G. This study was also supported by the EPIICAL project (https://www.epiical.org) to P.G., supported by PENTA-ID foundation (https://penta-id.org/) and funded through an independent grant by ViiV Healthcare UK. The ‘Ucwaningo Lwamawele’ study was also supported by the grant number U01AI168655 to P.G. The funders played no role in either the design or execution of the study. This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under Contract No. 75N91019D00024. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This Research was supported in part by the Intramural Research Program of the NIH, Frederick National Lab, Center for Cancer Research. The authors would like to

acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work. https://doi.org/10.5281/zenodo.22558.

REFERENCES

1. J. Robinson, *et al.*, The IMGT/HLA database. *Nucleic Acids Res* **39**, D1171–D1176 (2011).
2. J. Trowsdale, J. C. Knight, Major Histocompatibility Complex Genomics and Human Disease. *Annu Rev Genomics Hum Genet* **14**, 301–323 (2013).
3. M. Carrington, B. D. Walker, Immunogenetics of Spontaneous Control of HIV. *Annu Rev Med* **63**, 131–145 (2012).
4. P. J. R. Goulder, D. I. Watkins, Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* **8**, 619–630 (2008).
5. P. J. R. Goulder, B. D. Walker, HIV and HLA Class I: An Evolving Relationship. *Immunity* **37**, 426–440 (2012).
6. UNAIDS, Global HIV & AIDS statistics — Fact sheet. *https://www.unaids.org/en/resources/fact-sheet* (2024).
7. Tendesayi Kufa-Chakezha, Nosipho Shangase, Carl Lombard, Samuel Manda, Adrian Puren, "The 2022 Antenatal HIV Sentinel Survey" (2024).
8. P. Kiepiela, *et al.*, Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* **432**, 769–775 (2004).
9. A. Leslie, *et al.*, Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. *J Virol* **84**, 9879–9888 (2010).
10. P. Kiepiela, *et al.*, CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med* **13**, 46–53 (2007).
11. H. Streeck, *et al.*, Recognition of a Defined Region within p24 Gag by CD8 $^{+}$ T Cells during Primary Human Immunodeficiency Virus Type 1 Infection in Individuals Expressing Protective HLA Class I Alleles. *J Virol* **81**, 7725–7731 (2007).
12. A. Schneidewind, *et al.*, Escape from the Dominant HLA-B27-Restricted Cytotoxic T-Lymphocyte Response in Gag Is Associated with a Dramatic Reduction in Human Immunodeficiency Virus Type 1 Replication. *J Virol* **81**, 12382–12393 (2007).
13. B. Julg, *et al.*, Enhanced Anti-HIV Functional Activity Associated with Gag-Specific CD8 T-Cell Responses. *J Virol* **84**, 5540–5549 (2010).
14. J. B. Sacha, *et al.*, Gag-Specific CD8+ T Lymphocytes Recognize Infected Cells before AIDS-Virus Integration and Viral Protein Expression. *The Journal of Immunology* **178**, 2746–2754 (2007).
15. J. B. Sacha, *et al.*, Differential Antigen Presentation Kinetics of CD8 $^{+}$ T-Cell Epitopes Derived from the Same Viral Protein. *J Virol* **82**, 9293–9298 (2008).
16. J. B. Sacha, *et al.*, Gag- and Nef-specific CD4 $^{+}$ T cells recognize and inhibit SIV replication in infected macrophages early after infection. *Proceedings of the National Academy of Sciences* **106**, 9791–9796 (2009).
17. M. Kawada, *et al.*, Gag-Specific Cytotoxic T-Lymphocyte-Based Control of Primary Simian Immunodeficiency Virus Replication in a Vaccine Trial. *J Virol* **82**, 10199–10206 (2008).

18. H. Crawford, *et al.*, Compensatory Mutation Partially Restores Fitness and Delays Reversion of Escape Mutation within the Immunodominant HLA-B*5703-Restricted Gag Epitope in Chronic Human Immunodeficiency Virus Type 1 Infection. *J Virol* **81**, 8346–8351 (2007).
19. J. Martinez-Picado, *et al.*, Fitness Cost of Escape Mutations in p24 Gag in Association with Control of Human Immunodeficiency Virus Type 1. *J Virol* **80**, 3617–3623 (2006).
20. T. Miura, *et al.*, HLA-B57/B*5801 Human Immunodeficiency Virus Type 1 Elite Controllers Select for Rare Gag Variants Associated with Reduced Viral Replication Capacity and Strong Cytotoxic T-Lymphotye Recognition. *J Virol* **83**, 2743–2755 (2009).
21. A. Schneidewind, *et al.*, Structural and Functional Constraints Limit Options for Cytotoxic T-Lymphocyte Escape in the Immunodominant HLA-B27-Restricted Epitope in Human Immunodeficiency Virus Type 1 Capsid. *J Virol* **82**, 5594–5605 (2008).
22. T. C. Friedrich, *et al.*, Reversion of CTL escape–variant immunodeficiency viruses in vivo. *Nat Med* **10**, 275–281 (2004).
23. A. J. Leslie, *et al.*, HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med* **10**, 282–289 (2004).
24. K. C. Ngumbela, *et al.*, Targeting of a CD8 T Cell Env Epitope Presented by HLA-B*5802 Is Associated with Markers of HIV Disease Progression and Lack of Selection Pressure. *AIDS Res Hum Retroviruses* **24**, 72–82 (2008).
25. P. C. Matthews, *et al.*, Central Role of Reverting Mutations in HLA Associations with Human Immunodeficiency Virus Set Point. *J Virol* **82**, 8548–8559 (2008).
26. L. N. Broyles, R. Luo, D. Boeras, L. Vojnov, The risk of sexual transmission of HIV in individuals with low-level HIV viraemia: a systematic review. *The Lancet* **402**, 464–471 (2023).
27. T. C. Quinn, *et al.*, Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1. *New England Journal of Medicine* **342**, 921–929 (2000).
28. M. S. Cohen, *et al.*, Antiretroviral treatment of HIV-1 prevents transmission of HIV-1: where do we go from here? *The Lancet* **382**, 1515–1524 (2013).
29. P. M. Garcia, *et al.*, Maternal Levels of Plasma Human Immunodeficiency Virus Type 1 RNA and the Risk of Perinatal Transmission. *New England Journal of Medicine* **341**, 394–402 (1999).
30. M.-L. Newell, *et al.*, Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. *The Lancet* **364**, 1236–1243 (2004).
31. W. Mphatswe, *et al.*, High frequency of rapid immunological progression in African infants infected in the era of perinatal HIV prophylaxis. *AIDS* **21**, 1253–1261 (2007).
32. J. R. Millar, *et al.*, HIGH-FREQUENCY failure of combination antiretroviral therapy in paediatric HIV infection is associated with unmet maternal needs causing maternal NON-ADHERENCE. *EClinicalMedicine* **22**, 100344 (2020).
33. N. Bengu, *et al.*, Sustained aviremia despite anti-retroviral therapy non-adherence in male children after in utero HIV transmission. *Nat Med* **30**, 2796–2804 (2024).
34. R. Steinbrook, The AIDS Epidemic in 2004. *New England Journal of Medicine* **351**, 115–117 (2004).
35. A. Larsen, *et al.*, Timing and Predictors of Initiation on Antiretroviral Therapy Among Newly-Diagnosed HIV-Infected Persons in South Africa. *AIDS Behav* **23**, 375–385 (2019).
36. UNAIDS, South Africa UNAIDS Factsheet 2023. *https://www.unaids.org/en/regionscountries/countries/southafrica* (2025).
37. M. S. Cohen, *et al.*, Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *New England Journal of Medicine* **375**, 830–839 (2016).
38. P. J. Goulder, S. R. Lewin, E. M. Leitman, Paediatric HIV infection: the potential for cure. *Nat Rev Immunol* **16**, 259–271 (2016).

39. K. L. Dong, *et al.*, Detection and treatment of Fiebig stage I HIV-1 infection in young at-risk women in South Africa: a prospective cohort study. *Lancet HIV* **5**, e35–e44 (2018).
40. T. Ndung'u, K. L. Dong, D. S. Kwon, B. D. Walker, A FRESH approach: Combining basic science and social good. *Sci Immunol* **3** (2018).
41. V. A. Vieira, *et al.*, An HLA-I signature favouring KIR-educated Natural Killer cells mediates immune control of HIV in children and contrasts with the HLA-B-restricted CD8+ T-cell-mediated immune control in adults. *PLoS Pathog* **17**, e1010090 (2021).
42. V. Ramsuran, *et al.*, Elevated *HLA-A* expression impairs HIV control through inhibition of NKG2A-expressing cells. *Science (1979)* **359**, 86–90 (2018).
43. H. G. Küstner, J. P. Swanevelder, A. van Middelkoop, National HIV surveillance in South Africa--1993-1995. *South African Medical Journal* **88**, 1316–20 (1998).
44. H. G. Küstner, J. P. Swanevelder, A. Van Middelkoop, National HIV surveillance--South Africa, 1990-1992. *South African Medical Journal* **84**, 195–200 (1994).
45. "NATIONAL HIV AND SYPHILIS SERO-PREVALENCE SURVEY OF WOMEN ATTENDING PUBLIC ANTENATAL CLINICS IN SOUTH AFRICA" (2001).
46. N. Conan, *et al.*, Increase in HIV viral suppression in KwaZulu-Natal, South Africa: Community-based cross sectional surveys 2018 and 2013. What remains to be done? *PLoS One* **17**, e0265488 (2022).
47. P. N. Amornkul, *et al.*, Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa. *AIDS* **27**, 2775–2786 (2013).
48. L. F. Johnson, T. B. Hallett, T. M. Rehle, R. E. Dorrington, The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. *J R Soc Interface* **9**, 1544–1554 (2012).
49. N. G. de Groot, *et al.*, Evidence for an ancient selective sweep in the MHC class I gene repertoire of chimpanzees. *Proceedings of the National Academy of Sciences* **99**, 11748–11753 (2002).
50. N. G. de Groot, R. E. Bontrop, The HIV-1 pandemic: does the selective sweep in chimpanzees mirror humankind's future? *Retrovirology* **10**, 53 (2013).
51. N. G. de Groot, *et al.*, AIDS-protective HLA-B*27/B*57 and chimpanzee MHC class I molecules target analogous conserved areas of HIV-1/SIV $_{cpz}$. *Proceedings of the National Academy of Sciences* **107**, 15175–15180 (2010).
52. S. Balla-Jhagjhoorsingh, Differential cytotoxic T-lymphocyte (CTL) responses in HIV-1 immunised sibling chimpanzees with shared MHC haplotypes. *Immunol Lett* **66**, 61–67 (1999).
53. N. G. de Groot, *et al.*, Limited MHC class I intron 2 repertoire variation in bonobos. *Immunogenetics* **69**, 677–688 (2017).
54. T. M. Allen, *et al.*, CD8 $^{+}$ Lymphocytes from Simian Immunodeficiency Virus-Infected Rhesus Macaques Recognize 14 Different Epitopes Bound by the Major Histocompatibility Complex Class I Molecule Mamu-A*01: Implications for Vaccine Design and Testing. *J Virol* **75**, 738–749 (2001).
55. N. J. Maness, *et al.*, Comprehensive Immunological Evaluation Reveals Surprisingly Few Differences between Elite Controller and Progressor *Mamu-B*17* -Positive Simian Immunodeficiency Virus-Infected Rhesus Macaques. *J Virol* **82**, 5245–5254 (2008).
56. B. R. Mothé, *et al.*, Dominance of CD8 Responses Specific for Epitopes Bound by a Single Major Histocompatibility Complex Class I Molecule during the Acute Phase of Viral Infection. *J Virol* **76**, 875–884 (2002).
57. M. Z. Smith, *et al.*, Analysis of Pigtail Macaque Major Histocompatibility Complex Class I Molecules Presenting Immunodominant Simian Immunodeficiency Virus Epitopes. *J Virol* **79**, 684–695 (2005).

58. T. Nomura, T. Matano, Association of MHC-I genotypes with disease progression in HIV/SIV infections. *Front Microbiol* **3** (2012).
59. P. Schliekelman, C. Garner, M. Slatkin, Natural selection and resistance to HIV. *Nature* **411**, 545–546 (2001).
60. D. Cromer, S. M. Wolinsky, A. R. McLean, How fast could HIV change gene frequencies in the human population? *Proceedings of the Royal Society B: Biological Sciences* **277**, 1981–1989 (2010).
61. P. Bacchetti, A. R. Moss, Incubation period of AIDS in San Francisco. *Nature* **338**, 251–253 (1989).
62. C. Vergara, *et al.*, Multiancestry sex-stratified genomic associations with HIV viral load and controller status from the ICGH. *JCI Insight* **8** (2023).
63. The International HIV Controllers Study, The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. *Science (1979)* **330**, 1551–1557 (2010).
64. P. J. McLaren, *et al.*, Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proceedings of the National Academy of Sciences* **112**, 14658–14663 (2015).
65. P. J. McLaren, *et al.*, Association Study of Common Genetic Variants and HIV-1 Acquisition in 6,300 Infected Cases and 7,200 Controls. *PLoS Pathog* **9**, e1003515 (2013).
66. E. Gonzalez, *et al.*, Race-specific HIV-1 disease-modifying effects associated with *CCR5* haplotypes. *Proceedings of the National Academy of Sciences* **96**, 12004–12009 (1999).
67. P. I. TERASAKI, J. D. McCLELLAND, Microdroplet Assay of Human Serum Cytotoxins. *Nature* **204**, 998–1000 (1964).
68. Q. A. Karim, *et al.*, Stabilizing HIV prevalence masks high HIV incidence rates amongst rural and urban women in KwaZulu-Natal, South Africa. *Int J Epidemiol* **40**, 922–930 (2011).
69. A. K. Shevchenko, *et al.*, Genome-wide association study reveals genetic variants associated with HIV-1C infection in a Botswana study population. *Proceedings of the National Academy of Sciences* **118** (2021).
70. J. Lane, *et al.*, A genome-wide association study of resistance to HIV infection in highly exposed uninfected individuals with hemophilia A. *Hum Mol Genet* **22**, 1903–1910 (2013).
71. N. Vince, *et al.*, HLA Class I and KIR Genes Do Not Protect Against HIV Type 1 Infection in Highly Exposed Uninfected Individuals With Hemophilia A. *J Infect Dis* **210**, 1047–1051 (2014).
72. X. Gao, *et al.*, HLA-B alleles associate consistently with HIV heterosexual transmission, viral load, and progression to AIDS, but not susceptibility to infection. *AIDS* **24**, 1835–1840 (2010).
73. A. Ross, *et al.*, HIV-1 disease progression and fertility. *AIDS* **18**, 799–804 (2004).
74. A. Akullian, *et al.*, Large age shifts in HIV-1 incidence patterns in KwaZulu-Natal, South Africa. *Proceedings of the National Academy of Sciences* **118** (2021).
75. R. Payne, *et al.*, Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. *Proceedings of the National Academy of Sciences* **111** (2014).
76. E. E. Wroblewski, *et al.*, Malaria-driven adaptation of MHC class I in wild bonobo populations. *Nat Commun* **14**, 1033 (2023).
77. M. Bunce, "PCR-Sequence-Specific Primer Typing of HLA Class I and Class II Alleles" in *MHC Protocols*, (Humana Press), pp. 143–172.
78. K. M. De Cock, *et al.*, Prevention of Mother-to-Child HIV Transmission in Resource-Poor Countries. *JAMA* **283**, 1175 (2000).

887

888 **FIGURE LEGENDS**

889 **Figure 1. The HIV-1 epidemic curve among antenatal woman (1990-2022) and pVL and CD4**
890 **counts from pre-ART mothers (1998-2005).** A. HIV-1 epidemic curve among antenatal woman in
891 KwaZulu Natal (1990-2022). This graph was adapted from the Antenatal HIV Sentinel Survey (7).
892 These data points were used to fit our model in Fig 4. B. Viral loads at pregnancy of non-transmitting
893 and transmitting mothers. C. Viral loads at pregnancy compared between mothers who had one
894 'protective' (HLA-B*57/58:01/81:01) or one 'disease-susceptible' (HLA-B*18/45:01/58:02) HLA-B
895 allele. D. Viral loads at pregnancy compared between mothers with either one 'protective or
896 'disease-susceptible' HLA-B alleles in non-transmitting and transmitting mothers. E. CD4 counts at
897 pregnancy of non-transmitting and transmitting mothers. F. CD4 counts at pregnancy compared
898 between mothers with either one 'protective' or 'disease-susceptible' HLA-B alleles. G. CD4 counts
899 at pregnancy compared between mothers with either one 'protective or 'disease-susceptible' HLA-
900 B alleles in non-transmitting and transmitting mothers. One-tailed Mann-Whitney tests were used
901 to determine the p values shown. All p values <0.025 shown remained significant after Bonferroni-
902 Holm correction for multiple tests (Table S1). Those mothers with both protective and disease-
903 susceptible HLA-B alleles were excluded. Numbers correspond to median values and error bars
904 indicate IQR.

905

906 **Figure 2. Disease-associated HLA-B phenotypic frequencies in mother-child pairs in**
907 **KwaZulu-Natal, South Africa from the pre-ART era (1998-2005) and ART era (2015-2025).** A.
908 Phenotypic frequencies of disease-susceptible HLA-B alleles from the pre-ART era. B. Phenotypic
909 frequencies of protective HLA-B alleles from the pre-ART era. C. Phenotypic frequencies of
910 disease-susceptible HLA-B alleles from the ART era. D. Phenotypic frequencies of protective HLA-
911 B alleles from the ART era. E. Phenotypic frequencies of disease-susceptible HLA-B alleles from
912 the pre- and ART era. F. Phenotypic frequencies of protective HLA-B alleles from the pre- and ART

era. One-sided Fisher's exact tests were used to determine the p values shown. All p values <0.025 shown remained significant after Bonferroni-Holm correction for multiple tests (Table S1) . Dark grey; HIV negative adults from pre-ART era. Red; non-transmitting mothers from pre-ART era. Dark blue; transmitting mothers from pre-ART era. Light blue; children from pre-ART era. Light grey; HIV negative mothers from ART era. Dark green; transmitting mothers from ART era. Light green; children from ART era.

Figure 3. Phenotypic frequencies of individual protective and disease-susceptible HLA-B alleles across cohorts. A-C. Phenotypic frequencies of disease-susceptible HLA-B alleles; HLA-B*18, 45:01 and 58:02, respectively. D-F. Phenotypic frequencies of protective HLA-B alleles; HLA-B*57, 58:01 and 81:01, respectively. The p values shown are from Fisher's exact tests uncorrected for multiple comparisons.

Figure 4. Simulating a KZN-like HIV-1 epidemic and its consequences for HLA-B allele frequencies. A. A time series of our evolutionary epidemiological model from 1990 to 2035. The upper row of panel A displays HIV-1 prevalence in women, for our best estimate HLA-B properties scenario (see figure S6 for the equivalent figures for the conservative and extreme scenarios). The lower row displays frequencies of HLA-B alleles deemed disease-susceptible (blue) or protective (red). The allele trajectories associated with our best estimate, conservative and extreme HLA-B scenarios are indicated by different line styles, with the best estimate scenario a solid line, the conservative scenario a dashed line and the extreme scenario a dotted-dashed line. In the panels on the left, we have fitted the model to antenatal survey data on HIV-1 prevalence (squares) and ART uptake (triangles). We assumed, based on behavioural observations, that the parameter determining the transmissibility of HIV-1 could take 3 different values in 3 different time periods (1990-1998; 1999-2008 and 2009 onwards), and that treatment rates could take 2 different values (2004-2016 and 2017 onwards). Each of these transmission rates and treatment rates were fitted

to the data. In addition to fitting the model to data between 1990 and 2022 we extrapolate to 2035. Note that the trajectory of this extrapolation is a continuation of the trend seen in the final few years of data, and any interpretation of the extrapolation depends on the strength of our belief in this trend. Panels on the right illustrate a counterfactual simulation. Both left- and right-hand panels are identical up until the year 2004. In the counterfactual simulation, ART is not rolled out, and we take the transmission parameter that has been fitted for the period 1999-2008 and allow transmission at that rate to continue in the absence of ART. The consequences for HIV-1 transmission and HLA-B allele frequencies are shown. In all of panel A, vertical transmission of HIV-1 occurs at a probability of 35% for mothers who are homozygous for the indifferent HLA-B allele (and is increased or decreased from this baseline depending on maternal genotype, see Supporting Information for further details). A vertical transmission probability of 35% is within the range of probabilities of vertical transmission for HIV-1 positive breastfeeding mothers estimated by de Cock *et al* (78) (whose maximum estimate is 45% in mothers breastfeeding for 2 years). In panel B. we explore how changing the baseline level of vertical transmission impacts HLA-B allele frequencies, in the absence of ART, to understand the potential changes in HLA-B allele frequencies that have been averted by ART. Each bar chart displays allele frequency changes (from the starting conditions in 1990) that we predict could have occurred by 2035, for the indicated HLA-B properties scenario and for the probability of vertical transmission in homozygotes for the indifferent allele given on the x axis.

Fig 1

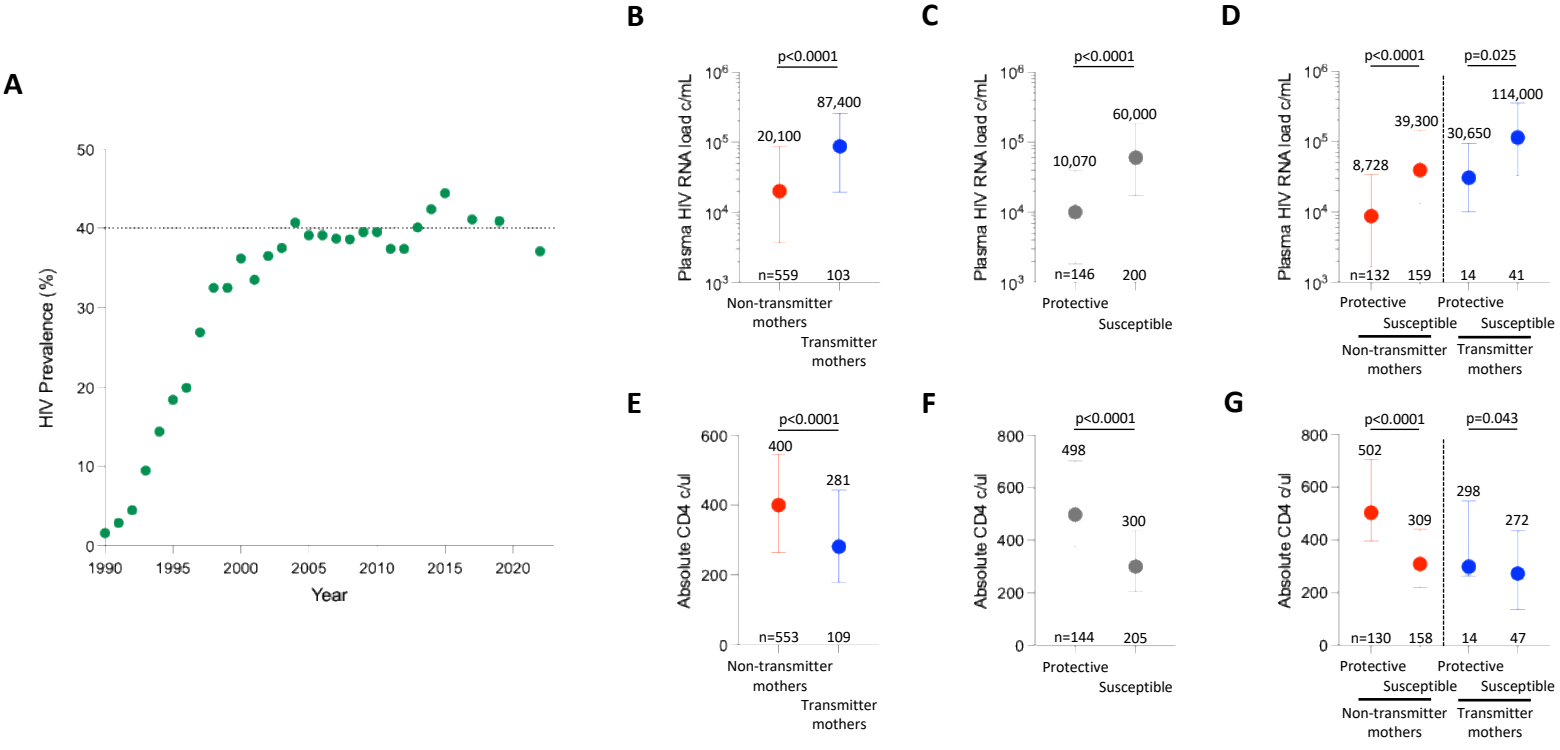

A
HIV Prevalence (%)
Year
B
Plasma HIV RNA load c/mL
p<0.0001
20,100
87,400
n=559
103
Non-transmitter mothers
Transmitter mothers
C
Plasma HIV RNA load c/mL
p<0.0001
10,070
60,000
n=146
200
Protective
Susceptible
D
Plasma HIV RNA load c/mL
p<0.0001
p=0.025
8,728
39,300
30,650
114,000
n=132
159
14
41
Protective
Susceptible
Non-transmitter mothers
Transmitter mothers
E
Absolute CD4 c/ul
p<0.0001
400
281
n=553
109
Non-transmitter mothers
Transmitter mothers
F
Absolute CD4 c/ul
p<0.0001
498
300
n=144
205
Protective
Susceptible
G
Absolute CD4 c/ul
p<0.0001
p=0.043
502
309
298
272
n=130
158
14
47
Protective
Susceptible
Non-transmitter mothers
Transmitter mothers

Fig 2

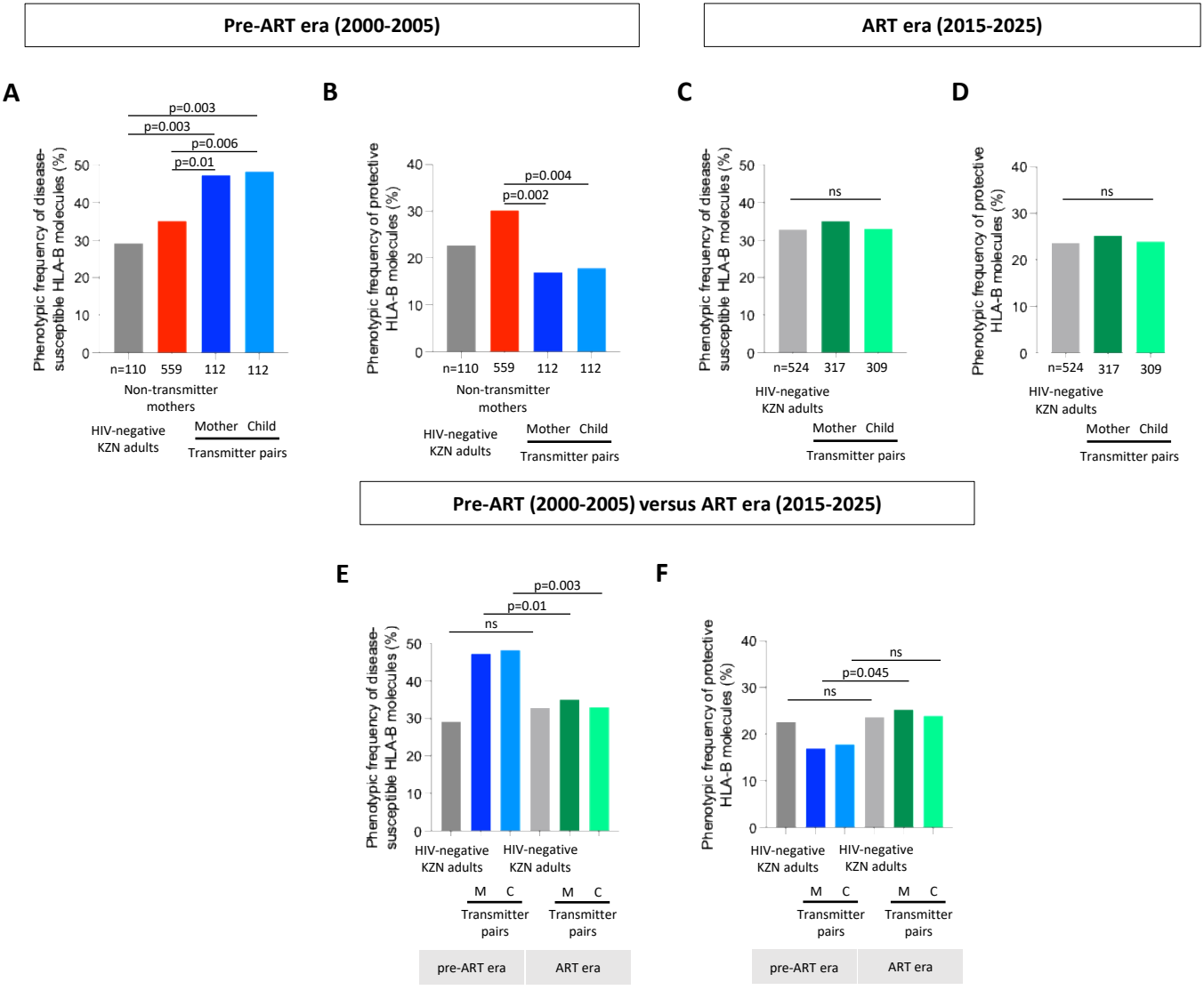

Pre-ART era (2000-2005)
ART era (2015-2025)
A
B
C
D
Phenotypic frequency of disease-susceptible HLA-B molecules (%)
Phenotypic frequency of protective HLA-B molecules (%)
p=0.003
p=0.003
p=0.006
p=0.01
p=0.004
p=0.002
ns
n=110 559 112 112
n=524 317 309
Non-transmitter mothers
HIV-negative KZN adults
Mother Child
Transmitter pairs
Pre-ART (2000-2005) versus ART era (2015-2025)
E
F
p=0.003
p=0.01
ns
p=0.045
M C
Transmitter pairs
pre-ART era
ART era

Fig 3

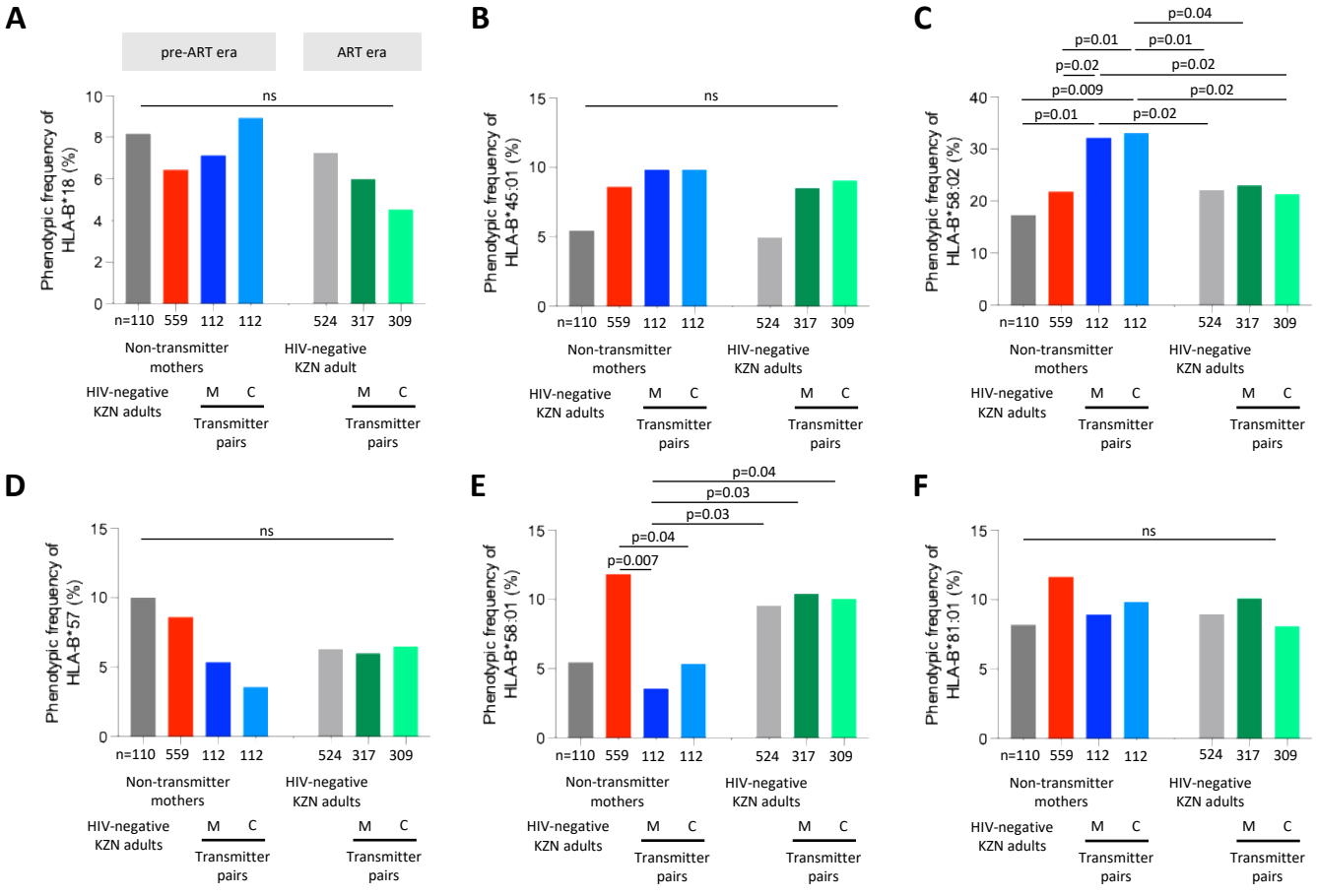
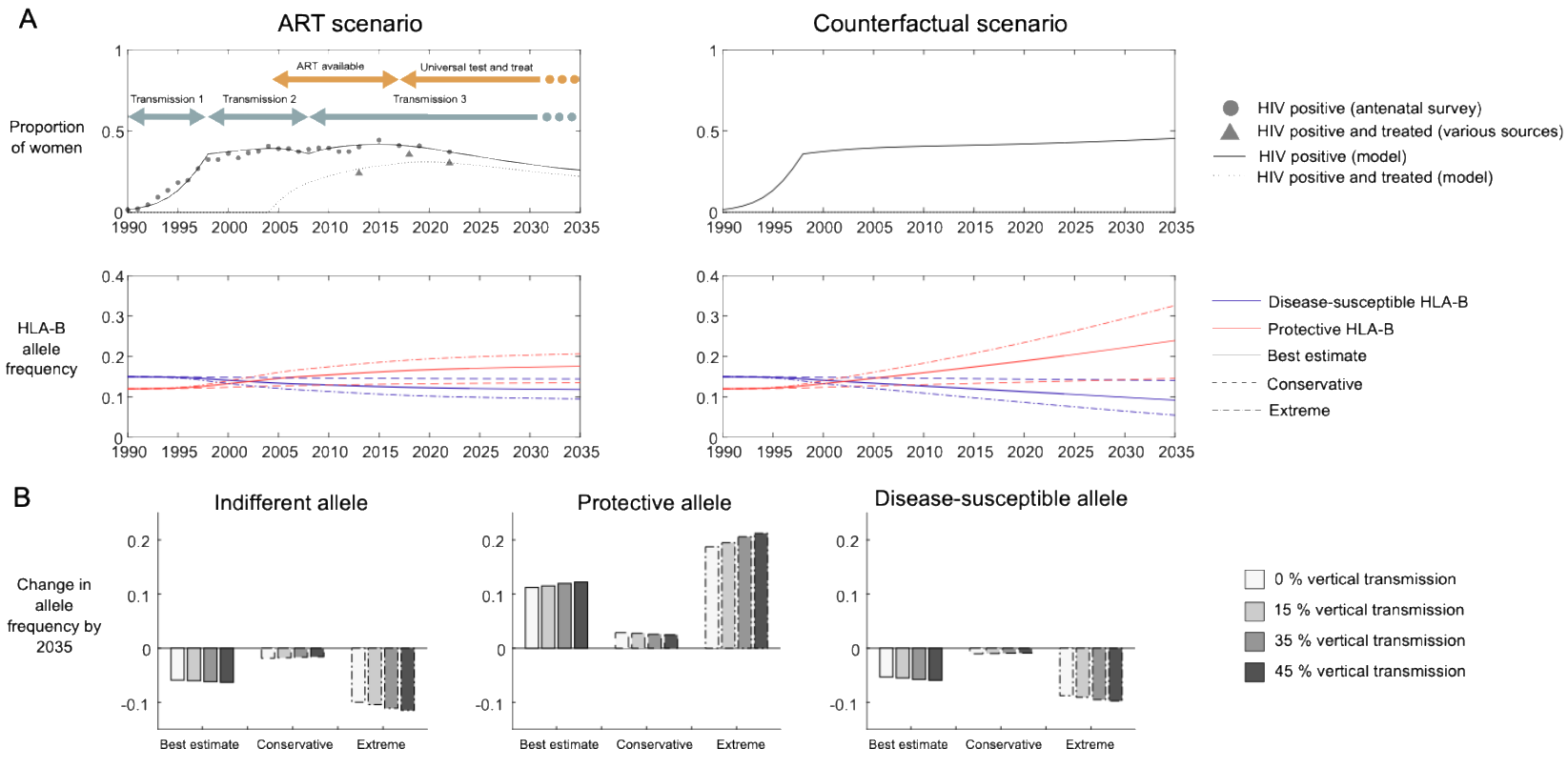

A
pre-ART era
ART era
ns
Phenotypic frequency of HLA-B*18 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adult
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs
B
ns
Phenotypic frequency of HLA-B*45:01 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adults
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs
C
p=0.04
p=0.01
p=0.01
p=0.02
p=0.02
p=0.009
p=0.02
p=0.01
p=0.02
Phenotypic frequency of HLA-B*58:02 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adults
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs
D
ns
Phenotypic frequency of HLA-B*57 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adults
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs
E
p=0.04
p=0.03
p=0.03
p=0.04
p=0.007
Phenotypic frequency of HLA-B*58:01 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adults
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs
F
ns
Phenotypic frequency of HLA-B*81:01 (%)
n=110 559 112 112 524 317 309
Non-transmitter mothers
HIV-negative KZN adults
HIV-negative KZN adults
M C
Transmitter pairs
M C
Transmitter pairs

Fig 4


A
ART scenario
Counterfactual scenario
Proportion of women
ART available
Universal test and treat
Transmission 1
Transmission 2
Transmission 3
HIV positive (antenatal survey)
HIV positive and treated (various sources)
HIV positive (model)
HIV positive and treated (model)
HLA-B allele frequency
Disease-susceptible HLA-B
Protective HLA-B
Best estimate
Conservative
Extreme
B
Indifferent allele
Protective allele
Disease-susceptible allele
Change in allele frequency by 2035
Best estimate
Conservative
Extreme
0 % vertical transmission
15 % vertical transmission
35 % vertical transmission
45 % vertical transmission