ARTICLE

doi:10.1038/nature12893

Immunological and virological mechanisms of vaccine-mediated protection against SIV and HIV

Mario Roederer[1], Brandon F. Keele[2], Stephen D. Schmidt[1], Rosemarie D. Mason[1], Hugh C. Welles[1,3], Will Fischer[4], Celia Labranche[5], Kathryn E. Foulds[1], Mark K. Louder[1], Zhi-Yong Yang[1]†, John-Paul M. Todd[1], Adam P. Buzby[6], Linh V. Mach[6], Ling Shen[6], Kelly E. Seaton[7], Brandy M. Ward[5], Robert T. Bailer[1], Raphael Gottardo[8], Wenjuan Gu[9], Guido Ferrari[5], S. Munir Alam[7], Thomas N. Denny[7], David C. Montefiori[5], Georgia D. Tomaras[7], Bette T. Korber[4], Martha C. Nason[9], Robert A. Seder[1], Richard A. Koup[1], Norman L. Letvin[6]‡, Srinivas S. Rao[1], Gary J. Nabel[1]† & John R. Mascola[1]

A major challenge for the development of a highly effective AIDS vaccine is the identification of mechanisms of protective immunity. To address this question, we used a nonhuman primate challenge model with simian immunodeficiency virus (SIV). We show that antibodies to the SIV envelope are necessary and sufficient to prevent infection. Moreover, sequencing of viruses from breakthrough infections revealed selective pressure against neutralization-sensitive viruses; we identified a two-amino-acid signature that alters antigenicity and confers neutralization resistance. A similar signature confers resistance of human immunodeficiency virus (HIV)-1 to neutralization by monoclonal antibodies against variable regions 1 and 2 (V1V2), suggesting that SIV and HIV share a fundamental mechanism of immune escape from vaccine-elicited or naturally elicited antibodies. These analyses provide insight into the limited efficacy seen in HIV vaccine trials.

Among the five human efficacy trials of HIV-1 vaccines, only one has shown some success in preventing HIV infection. In the RV144 trial, a combination viral vector and protein immunization achieved a modest 31% efficacy in a cohort of low-risk adults in Thailand[1]. In-depth immunological correlates analysis suggested that specific antibody responses to the HIV-1 envelope variable regions 1 and 2 (V1V2) region correlated with protection, whereas an immunoglobin A (IgA) response showed a negative association[2,3]. Virus sequencing of the breakthrough infections in RV144 suggested a possible vaccine-mediated selection pressure against certain virus variants[4]; the mechanism of immune pressure remains elusive, but may include elicitation of antibodies targeting V1V2 of those variants[5]. In contrast, the recent HVTN 505 trial, using a DNA-prime, recombinant adenovirus type 5 (rAd5) boost, was halted for futility with no vaccine efficacy[6].

Infection of nonhuman primates with SIV represents the best available animal model for testing vaccine concepts for protecting against HIV infection, and mucosal challenge with SIV can be used to model human mucosal HIV exposure[7]. Several SIV challenge studies have shown partial protection from acquisition; in some cases, there has been an association to elicited antibodies, but a strong immunological mechanism or correlate has not been identified[8–13]. Here, we used a repetitive intrarectal challenge using a SIV_{smE660} challenge virus that was unmatched to the vaccines[14]. The E660 virus swarm is heterogeneous, comprising groups or clusters of viruses ranging from neutralization sensitive to resistant[15]. We reasoned that, in the absence of complete protection, the naturally occurring diversity of neutralization profiles would provide the most informative correlates analysis.

Our goals were to define cellular and humoral immune correlates of immunity, and to understand the mechanism leading to protection against SIV infection. Our immunogens included 'T-cell mosaics' designed to optimize coverage of epitope diversity for cellular responses[16,17]. We designed a four-arm study to define mechanisms of vaccine protection: (1) mosaic Gag; (2) mosaic heterologous envelope (Env); (3) heterologous Env based on a natural SIV_{mac239} sequence; and (4) control vaccine. Our primary questions were whether Env immunization is sufficient and/or necessary to provide protection against acquisition, whether Gag (alone) immunization provide any protection against acquisition, and finally whether the use of 'T-cell mosaic' Envs provide additional benefit over a natural Env sequence.

The number of acquisition end points in this study was similar to a large human efficacy study. We demonstrated that an Env-elicited immune response is necessary and sufficient to provide protection from acquisition. Importantly, by integrating immunological and virological analyses, we elucidated antibody-mediated mechanisms of protection and discovered a fundamental mechanism of virus escape from antibody-mediated control, shared by SIV and HIV, that has broad implications for understanding vaccine-mediated protection and potentially for vaccine design.

Vaccine immunogenicity

80 Indian origin rhesus macaques were enrolled in a DNA-prime, rAd5 boost immunization study. Animals were randomized into four groups of 20 based on *TRIM5α* (also known as *TRIM5*) alleles, gender, age and weight. All animals received three shots of DNA at 4-week intervals, followed by rAd5 at week 30[14]. The control group received vectors that contained no inserts; the second group ('mosaic Gag') received two SIV Gag mosaic immunogens[17]; the third group ('mosaic

[1]Vaccine Research Center, NIAID, NIH, Bethesda, Maryland 20892, USA. [2]SAIC-Frederick, Frederick National Laboratory, NIH, Frederick, Maryland 21702, USA. [3]George Washington University, Washington DC 20052, USA. [4]Los Alamos National Laboratories, Los Alamos, New Mexico 87545, USA. [5]Department of Surgery, Duke University, Durham, North Carolina 27710, USA. [6]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02115, USA. [7]Human Vaccine Institute, Duke University, Durham, North Carolina 27710, USA. [8]Fred Hutchison Cancer Research Center, Seattle, Washington 98109, USA. [9]Biostatistics Research Branch, NIAID, NIH, Bethesda, Maryland 20892, USA. †Present address: Sanofi-Pasteur, Cambridge, Massachusetts 02139, USA. ‡Deceased.

Env') received two SIV Env mosaic immunogens (78% and 87% sequence identity to SIV_{smE543}, a clone similar to E660[16]); and the fourth group ('mac239 Env') received an immunogen encoding SIV_{mac239} Env (83% sequence identity to E543). Envelope sequences are shown in Supplementary Table 1, and sequence distances in Supplementary Table 2.

Vaccination elicited the expected cellular (Extended Data Fig. 1) and humoral (Extended Data Fig. 2) responses. Notably, compared to mac239 Env immunization, mosaic Env induced modestly lower and qualitatively different humoral responses (Extended Data Fig. 2). Mapping of the antibody response to unglycosylated linear peptides (Extended Data Fig. 2c) revealed that mac239 Env elicited a broader response than mosaic Env. Overall, immunization elicited mild neutralization and antibody-dependent cellular cytotoxicity activity against a limited set of viral strains (Extended Data Fig. 2d–g).

SIV challenge outcome

To test vaccine efficacy against infectious challenge, we exposed animals weekly to intrarectal administration of E660 at a dose that infects ~30% of control animals per exposure[14]. Each animal was challenged up to 12 times or until it had detectable plasma viraemia. Immunization with mac239 Env provided significant protection against acquisition, whereas mosaic Env immunization did not achieve significance (Fig. 1a). There was no difference in acquisition between Gag-immunized animals and control animals. For protection against acquisition, vaccine efficacy (V_E: the reduction in the rate of infection at each challenge)[18,19] was 69% for mac239 Env (Fig. 1d).

All infected animals that received active immunization showed 0.7 to 1.1 $\log_{10}$ decrease in peak viral load (VL) on average (Fig. 1b, c and Extended Data Fig. 3b). The best control of acute VL occurred in the mosaic Env arm, whereas the mosaic Gag arm showed the best long-term control (Fig. 1d). We confirmed previous findings that animals with certain alleles of *TRIM5α* showed better innate control of infection and pathogenesis[14] (Extended Data Fig. 3d, e). Due to the stratification by *TRIM5α* alleles in our study, including this genotype as a covariate in analyses does not affect our conclusions. All three vaccine arms showed protection against loss of CD4 cells (Extended Data Fig. 3c). Thus, the mosaic Env constructs elicited effective T-cell responses that protected against pathogenic effects of infection, despite their inability to block acquisition.

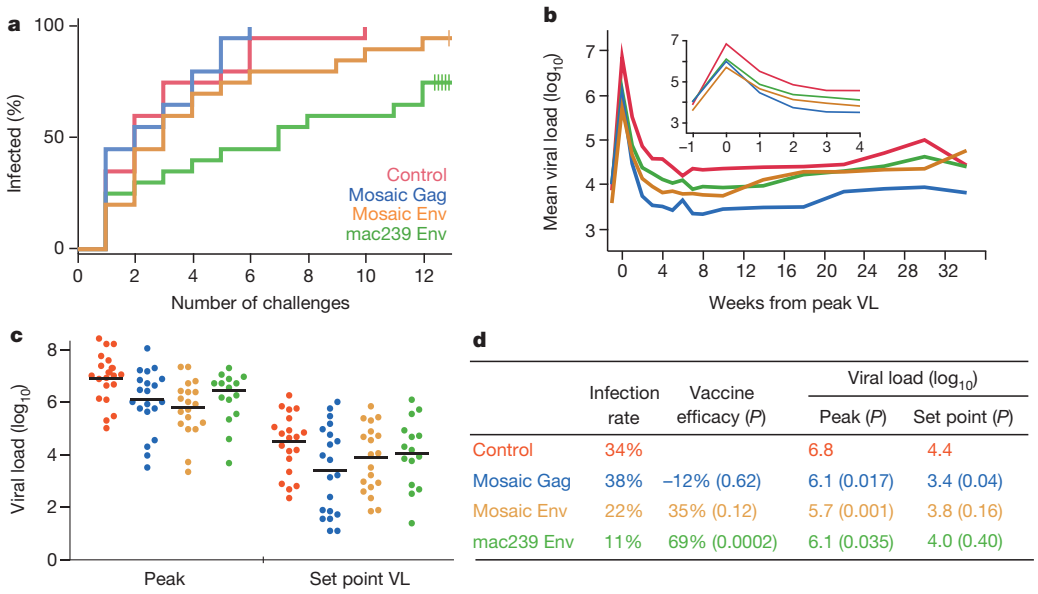
Transmitted founder analysis

Because E660 is a viral swarm with 1.8% sequence diversity, the number of transmitted founder (T/F) viruses can be determined by single-genome amplification (SGA). For every infected animal, sequencing was done on plasma from the earliest time point with detectable plasma VL, 1 week after infection: thus, the inferred sequences represent the original infecting viruses[7]. Both Env arms showed a significant decrease in T/F variants (Fig. 2a). From these data, an efficacy can be calculated by the reduction of T/F variants per challenge; theoretically, this value estimates V_E for a very low (clinically relevant) infectious dose. Immunization with mac239 Env reduced the number of T/F variants by 81%, and mosaic Env reduced T/F by 51% (Fig. 2b).

Phylogenetic analysis using all complete Env sequences did not reveal an obvious clustering of T/F variants by vaccine arm. However, a strong 'sieving' effect was discerned by examining individual amino acid variants. Over the Env coding sequence, the 133 T/F sequences showed variation at 63 sites (Supplementary Table 3); 20 positions in the cytoplasmic domain or with rare variation (<5 of 133 T/F) were excluded from further analysis. Among the remaining sites, we found significant differences in variant representation in the Env vaccinated arms compared to the control and Gag arms (Fig. 2c and Extended Data Fig. 4). The strongest effect was seen at positions 23, 45 and 47. The consensus T/F sequence at these positions (VTR) was found in a majority of T/F viruses in the control and Gag arms. In contrast, variant sequences (IAK) were significantly overrepresented in the Env-immunized arms. Thus, immunization with Env sequences induced an immune response selecting against virions with the VTR signature.

Mechanism of virus selection in vaccinees

To define the mechanism of vaccine-mediated selection against viral variants, we measured the neutralization profile of all 40 Env-immunized animals against pseudo-typed viruses. CP3C-P-A8 ('CP3C' for brevity), a clone from the E660 swarm, is a neutralization-sensitive virus and has the amino acids VTR at positions 23, 45 and 47. CR54-PK-2A5 ('CR54'),



d

	Infection rate	Vaccine efficacy (*P*)	Viral load ($\log_{10}$) Peak (*P*)	Set point (*P*)
Control	34%		6.8	4.4
Mosaic Gag	38%	–12% (0.62)	6.1 (0.017)	3.4 (0.04)
Mosaic Env	22%	35% (0.12)	5.7 (0.001)	3.8 (0.16)
mac239 Env	11%	69% (0.0002)	6.1 (0.035)	4.0 (0.40)

Figure 1 | Protection against SIV challenge. **a**, The fraction infected animals in each arm following each of 12 challenges is shown. Five animals in the mac239 Env arm and one animal in the mosaic Env arm remained uninfected after 12 challenges. **b**, For each arm, the geometric mean plasma viral load (RNA copies per ml) for infected animals is shown. Each animal is synchronized to its peak VL. Inset, expanded scale for the acute phase. **c**, The peak and set point plasma viral load distributions for all infected animals. **d**, The infection rate is the fraction of infections out of the total number of exposures; vaccine efficacy was calculated as described in the methods.

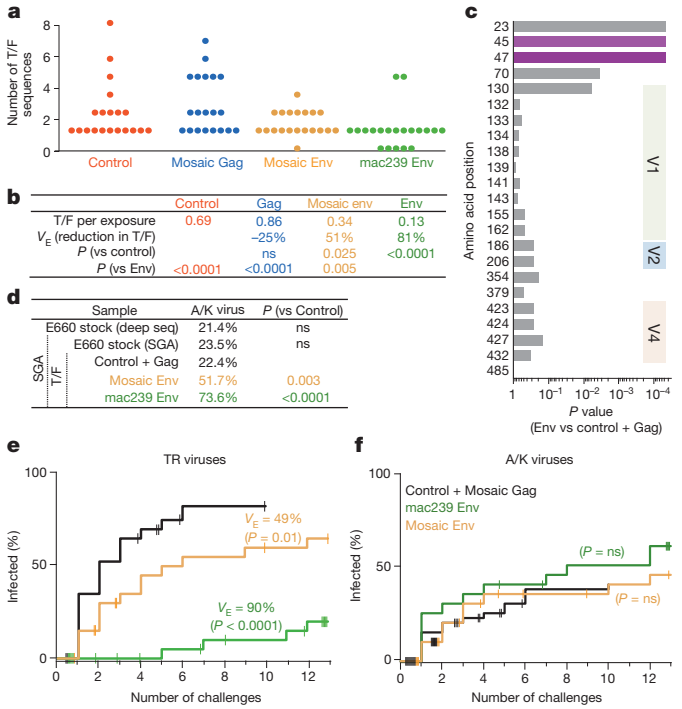


Figure 2 | Analysis of transmitted/founder (T/F) viruses. **a**, The distribution of unique T/F viruses in the first virus-positive plasma sample is shown for all 80 animals. **b**, The average number of T/F viruses per exposure event was calculated. Here, vaccine efficacy (V_E) is computed as the reduction in the number of T/F viruses (ns (not significant), $P > 0.05$). **c**, For each position in Env, the P value is shown for a permutation test comparing the fraction of viruses with the consensus amino acid in the Env T/F vs the control and Gag T/F. P values at positions 23, 45 and 47 remain significant after correction for multiple comparisons. **d**–**f**, Based on the sequence at positions 45 and 47, T/F viruses were divided into 'TR' (45T+47R) and 'A/K' (45A or 47K) viruses. **d**, Proportion of A/K viruses in the E660 challenge stock was measured by deep sequencing or by SGA, and among T/F in the immunization arms by SGA. A Fisher's exact test was performed to determine the significance of the difference in A/K viruses compared to the Control+Gag arms (ns, $P > 0.05$). **e**, **f**, Cumulative infection probabilities by TR or A/K viruses was done using a non-parametric estimate for competing risks[23]; the V_E and P values are computed using likelihoods from a modified Hudgens and Gilbert leaky vaccine model[18]. Tick marks indicate censoring of animals solely infected by the other virus type (challenges 1–12), or remaining uninfected after 12 challenges.

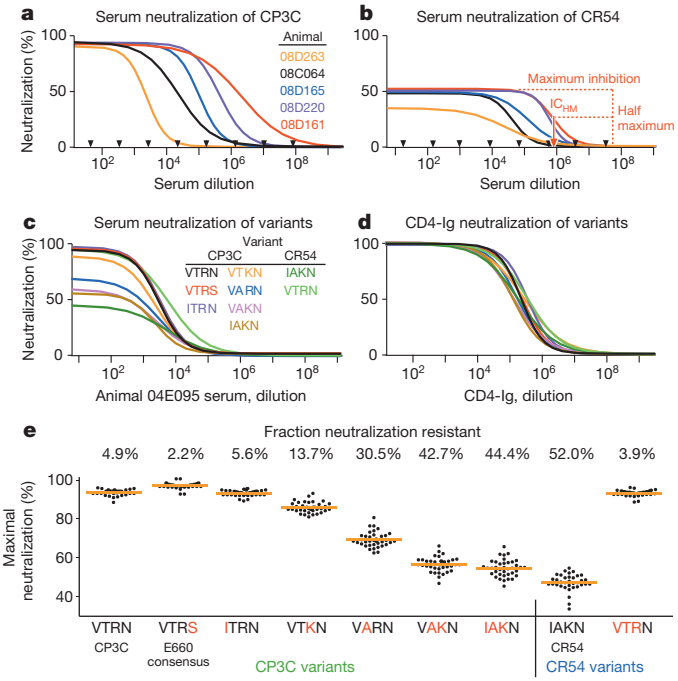


Figure 3 | Sequences accounting for neutralization resistance. **a**, **b**, Neutralization curves of CP3C, a sensitive clone from E660 (**a**), and of CR54, a resistant clone from E660 (**b**), using dilutions of sera from five Env-immunized animals (selected to show the range of potency). Black arrows indicate which dilutions were tested in duplicate; curves represent nonlinear least squares regressions of a four-parameter binding model. Nearly 100% of CP3C virions, but only 40–50% of CR54 virions, can be neutralized by immune sera. Red dashed lines show how IC_{HM} is derived for animal 08D161. **c**, **d**, Neutralization curves of 9 viral variants using serum from one animal (**c**) or CD4-Ig (**d**). The parent virus into which mutations were made is listed, along with the amino acids at positions 23, 45, 47 and 70. **e**, All variants were assayed using serial dilutions of sera from all 40 Env-immunized animals. Shown is the maximum fraction of each virus that was neutralized (determined by regression analysis). Red letters indicate amino acid substitutions compared to the parent virus. The numbers above the graphic indicate the mean resistant fraction for each virus.

another clone from the E660 swarm, is a neutralization-resistant virus with IAK at these positions. Sera from immunized animals completely neutralized CP3C, with an inhibitory concentration potency (IC_{HM}, defined as the dilution giving half-maximal inhibition) that varied 1,000-fold (Fig. 3a). In contrast, the same sera only achieved a maximum of ~50% inhibition of CR54. Importantly, this shows that CR54 is a heterogeneous population of virions despite being genetically clonal: half of the virions are easily neutralized by antisera, and half are completely resistant.

We introduced variants of four amino acids into CP3C and CR54 Env to test which might be responsible for modulating neutralization resistance. These amino acid variations did not change the potency of the antisera (IC_{HM} varied less than twofold; Fig. 3c and Extended Data Fig. 5a), and did not change the sensitivity to neutralization by CD4-Ig (Fig. 3d). However, variant sequences affected the fraction of neutralization-resistant virions.

Despite the wide breadth of epitopes targeted by vaccine-elicited antisera (Extended Data Fig. 2), there was little variation in the fraction of each virus that was neutralization resistant (Fig. 3e). Moreover, an identical neutralization-resistant fraction was observed for a panel of SIV monoclonal antibodies directed near the CD4 binding site or the V1V2 loop (Extended Data Fig. 5b, c). Therefore, even clonally derived virions are remarkably heterogeneous: a fraction are easily neutralized, and the remainder are completely resistant to antibody-based neutralization.

Generation of this resistant Env phenotype was favoured by amino acid substitutions in the C1 region. By making point mutations, we showed that the T45A and R47K mutations individually result in increased resistance. Together, changing these two amino acids converts the sensitive CP3C Env to a nearly fully resistant phenotype, and the resistant CR54 to fully sensitive. For parsimony in subsequent analyses, we divided E660 viruses into two categories: viruses with both 45T and 47R ('TR'), which are putatively neutralization sensitive, and viruses with either 45A or 47K ('A/K'), which should be generally resistant to vaccine-elicited sera.

Deep sequencing and SGA of Env genes showed that ~20% of the E660 challenge swarm were neutralization-resistant A/K viruses (Fig. 2d). The same proportion was found among infecting T/F sequences in the control and Gag arms, demonstrating that there is no innate selection for or against A/K sequences. Furthermore, A/K infections resulted in the same peak and set point plasma VL, indicating that these viruses are no more or less fit than TR viruses (Extended Data Fig. 6). However, vaccine-elicited responses strongly selected against infection by TR viruses—such that, in the mac239 Env arm, the infrequent (neutralization-resistant) A/K variants comprised nearly 75% of T/F viruses.

We next computed the V_E against A/K and TR viruses separately. The TR (sensitive) variants are highly susceptible to vaccine-mediated

control, with a V_E of 90% (Fig. 2e). In contrast, the V_E against A/K viruses did not reach significance (Fig. 2f). Thus, the heterogeneous neutralization of even clonal SIV virions, programmed by C1 amino acid variations, represents a novel mechanism of immune escape from Env-specific antibodies.

Immune correlates of risk of infection

A panoply of cellular and humoral assays quantifying vaccine-elicited responses were performed at baseline, peak post-boost, and pre-challenge time points. We found strong associations between several antibody responses and probability of infection, but no associations between T-cell responses and delayed acquisition.

Given that the E660 swarm is comprised of both neutralization-sensitive (TR) and -resistant (A/K) genotypes, it made sense to analyse correlates in two ways: first, by including all infections, irrespective of variant; and second, by separating the two types of infections. Because the vaccine is largely ineffective against A/K viruses, pooling A/K-infected with TR-infected animals may mask potential correlates.

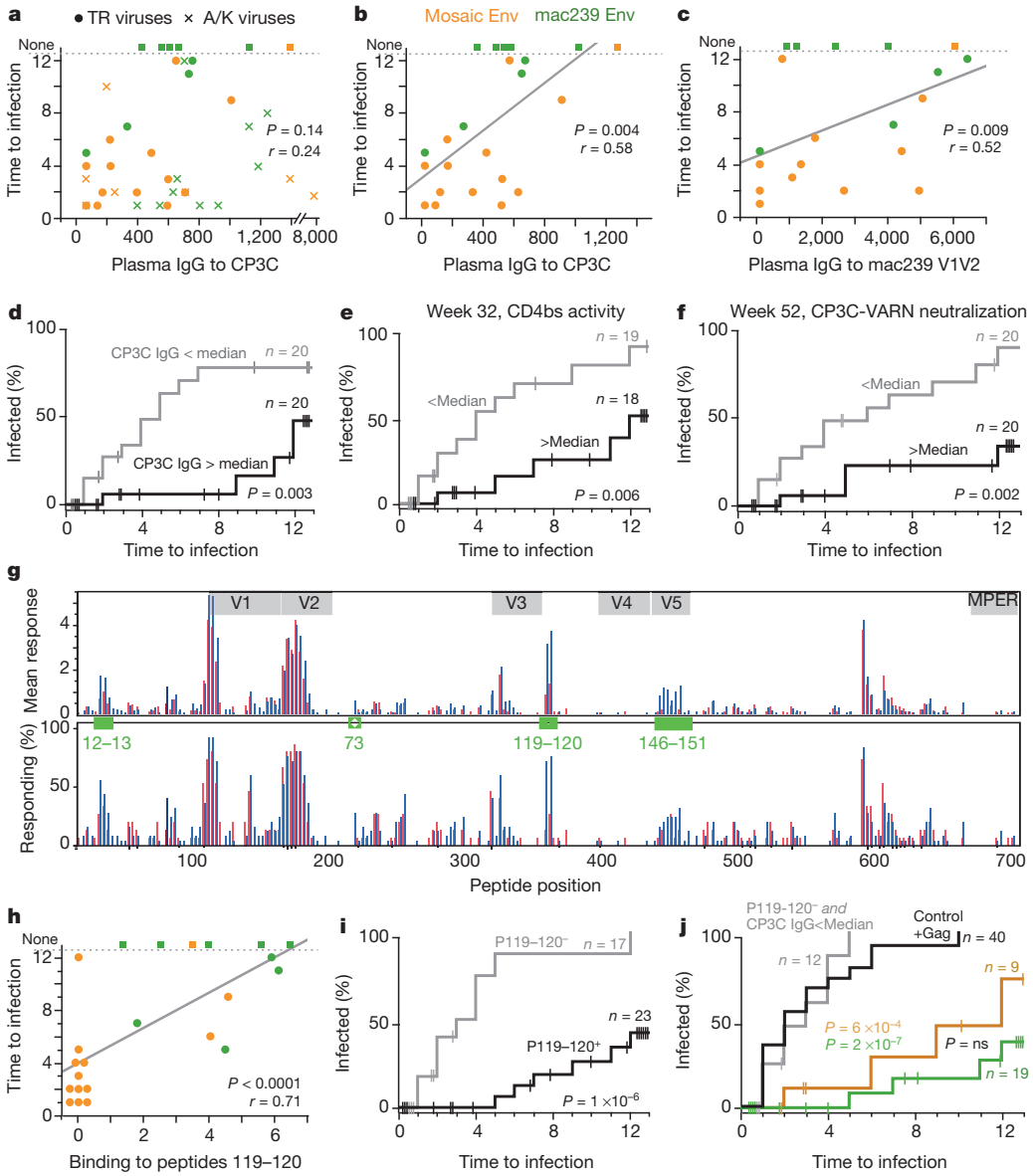


Figure 4 | Immunological correlates of risk. **a**, **b**, Week 52 plasma IgG against the CP3C envelope is graphed against time to infection (uninfected animals were assigned a value of 13). No significant correlation was found when all infection events were considered (**a**); however, by excluding animals infected solely with A/K viruses, a strong predictive relationship is seen (**b**). The line is from a linear regression for illustration; statistics are based on Spearman correlation. **c**, Week 52 plasma IgG against the mac239 V1V2 is significantly associated with protection against TR viruses, and also against all viruses (Extended Data Fig. 7b). **d**, Kaplan–Meier (KM) analysis was performed by dividing the 40 Env-immunized animals in two equal groups based on the anti-CP3C IgG responses (median = 570). Animals remaining uninfected or infected solely with A/K viruses were censored as shown by vertical lines. **e**, KM analysis comparing Env-immunized animals with higher vs lower week 32 serum activity against the CD4 binding site of envelope. **f**, KM analysis comparing Env-immunized animals with higher vs lower week 52 neutralization activity against virus pseudotyped with a CP3C Env containing a T45A mutation ('VARN'), a sequence shared by E543. **g**, The mean response (upper) and proportion of responders (lower) against each linear peptide is shown for animals grouped by time to infection: 1–3 challenges (red) vs 4 or more challenges (blue). Green boxes highlight regions potentially associated with protection identified by a Fisher's exact test; overlapping peptide numbers are in green, with sequences given in Supplementary Table 5. **h**, Average binding to the linear C3 peptides 119 and 120 (P119–120) correlates strongly with time to infection. **i**, KM analysis comparing Env-immunized animals with a positive response to C3 peptides to those with a negative response. **j**, KM analyses comparing all animals in the control and Gag arms (black), all Env-immunized animals having a CP3C IgG response below 570 and a negative C3 peptides response (grey), and animals in either Env arm having either antibody response.

The data shown in Fig. 4 illustrate these analyses. Among all 40 Env-vaccinated animals, plasma IgG binding to CP3C gp120 Env at the time of challenge did not correlate significantly with time to infection (Fig. 4a). In contrast, when we excluded animals who were infected solely with A/K viruses, we found a strong correlation with IgG binding to CP3C gp120 (Fig. 4b), but not other Envs (Extended Data Fig. 7a, b). We grouped all Env-immunized animals by those with an IgG response to CP3C above or below 570 (the median value, corresponding to an end point titre of approximately 1:1,000). Animals with the higher response had a 75% lower rate of infection by TR viruses (Fig. 4d).

Correlation with time to infection was also observed for plasma antibody avidity (Extended Data Fig. 7e), CD4-binding site activity (Fig. 4e) and neutralization of some viral strains (Fig. 4f). These data indicate that the quality of the antibody response is important. Thus, we investigated binding to specific regions within the Env.

By comparing peptide-binding data for animals grouped by time to infection (Fig. 4g), we identified four linear epitopes possibly associated with protection. There was a strong association between the breadth amongst these four epitopes and time to infection (Extended Data Fig. 8a–c). In contrast, there was no significant association with the breadth of response across all Env epitopes (Extended Data Fig. 8d). Thus, both breadth and magnitude of the response to selected epitopes are strong correlates of protection from acquisition.

The response to C3 (peptides 119+120) was the most significantly associated with protection, whether all viruses (Extended Data Fig. 8e, f) or just TR viruses were considered (Fig. 4h, i). This epitope corresponds to the α2 helix of Env and was identified as a neutralization target in HIV-1[20,21]. In a multivariable model, both IgG to CP3C ($P = 0.004$) and binding to the C3 peptides ($P = 0.02$) provided independent prediction of time to infection. We thus compared animals that had neither a response to the C3 peptides nor IgG to CP3C ($n = 12$, combining both Env arms) to animals with either response (mac239 Env: $n = 19/20$; mosaic Env: $n = 9/20$). For animals with neither antibody response, the rate of infection (12 infections in 27 exposures, 44%), and the proportion of infections with only A/K viruses (3/12, 25%) was not different from the control (unvaccinated) or Gag arms. In contrast, animals with either antibody response were primarily infected with resistant A/K viruses, and the V_E was >90% against TR viruses (Fig. 4j).

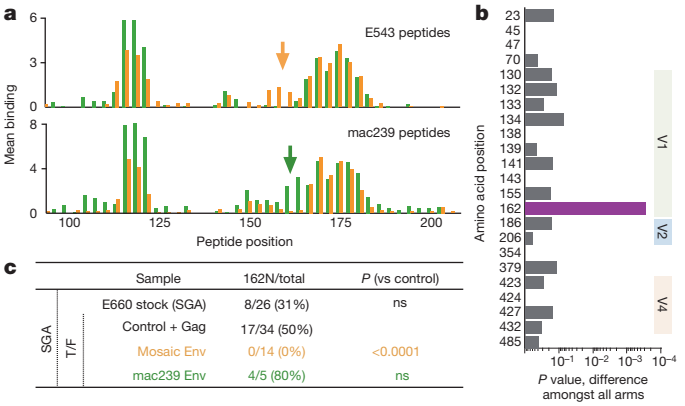


Figure 5 | Vaccine-mediated selection at V1V2. a, The binding of plasma from all 40 Env-immunized animals to linear 15-mer peptides spanning the V1V2 region of either E543 (top) or mac239 (bottom) was measured; bars represent the average binding for the 20 mosaic-immunized (orange) or the 20 mac239-immunized (green) animals. Arrows indicate an area of V1V2 showing vaccine-specific responses, encompassing amino acids 154–170. **b**, Sieving analysis was done as in Fig. 2c, but after excluding neutralization-resistant A/K viruses. The only significant association with immunization arm was at position 162. **c**, Representation of 162N (vs 162S) as determined by SGA for TR viruses in the swarm or in T/F. Note that all immunogens encode 162N, so selection is likely to be mediated against a neighbouring epitope; this epitope is found only in (one of) the mosaic immunogens, and occurs in linkage disequilibrium with 162N in the E660 swarm.

V1V2 and vaccine-specific sieving

In the human RV144 trial, antibody binding to HIV V1V2 was a primary (inverse) correlate of risk against infection. Similarly, antibody to the SIV V1V2 predicted protection against infection (Fig. 4c). The mosaic and mac239 immunogen sequences varied significantly in this region (Supplementary Table 1), and consequently elicited somewhat different antibody responses (Fig. 5a). To determine if these responses are associated with sieving, we analysed variation in T/F sequences (as in Fig. 2), after censoring vaccine-nonresponsive A/K viruses. This analysis revealed a strong selection associated with position 162 (Fig. 5b). The mosaic immunization completely selected against TR viruses with 162N (0/14 viruses, compared to 17/34 in the control; Fig. 5c). In contrast, the mac239 immunogen may have selected against the other variant (162S), although there were too few TR virus infections in this group to reach significance. Amongst A/K virus infections, there was no significant difference in representation of the 162N/S variants across vaccine arms. These data show that selection against V1V2 sequences by the SIV vaccines is limited to neutralization-sensitive viruses and, within those, selection is vaccine-sequence-specific and thus not broad.

Antibody escape mechanism in HIV

To assess whether our findings extend to HIV, we measured the inhibition of 51 distinct HIV-1 envelope pseudotyped viruses by the V1V2-specific monoclonal antibodies PG9 and PG16. As we saw for neutralization-resistant A/K SIV viruses, neutralization of some clonal HIV strains was incomplete; that is, a fraction of virions could not be neutralized (Fig. 6a). We examined the influence of sequence variation of these HIV envelopes on the fraction of neutralization-resistant virus (Fig. 6b); the most significant association was at position 47, with 47R viruses being sensitive (Fig. 6c). Sequence alignment with SIV envelope shows that position 47 in HIV is in a similar area of C1 as is position 47 in SIV (Fig. 6d); the similar signature (arginine vs lysine) indicates that a common mechanism of neutralization escape may be shared by SIV and HIV.

Discussion

Immune correlate studies that interrogate both virus sequences and immune responses can provide key insights on mechanisms of protection from HIV-1 acquisition. Using a nonhuman primate model

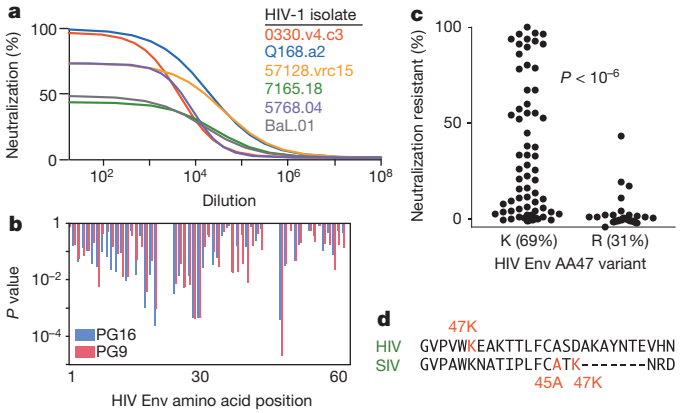


Figure 6 | C1 Sequences and HIV Env neutralization. a, Neutralization profiles of 51 different HIV-1 strains by the V1V2 antibodies PG9 and PG16 were determined. Example curves of PG16 on six viruses are shown. As for SIV A/K viruses, a variable fraction of each clonal virus is completely neutralization resistant; the remainder is highly sensitive. **b**, The influence of variants at each position in envelope on the fraction of neutralization resistant virus is shown as a *P* value from a Fisher's test; shown is the C1 region. **c**, The most significant association for all positions in Env was amino acid 47. The distribution of the fraction of neutralization-resistant virus is shown for the two variants, 47K and 47R. **d**, Alignment of SIV and HIV Env proteins in the middle of the C1 region, highlighting the positions of the neutralization signatures.

with a number of acquisition end points similar to large human efficacy studies, we demonstrated that an Env-elicited immune response is necessary and sufficient to provide protection from acquisition. We identified antibody-based correlates including responses to several epitopes. In our study, SIV Env T-cell mosaic immunogens elicited more effective T-cell responses, but less effective antibody responses. With respect to the virus, we identified a strong sieving effect of Env immunization, selecting for minor variants in the challenge swarm. And finally, we identify a sequence signature in the SIV Env, possibly shared by HIV, that programs the neutralization phenotype of the viruses through a mechanism affecting the entire antigenic surface of the protein.

Among our three vaccine groups, there was no association between protection from infection and protection from pathogenesis (for example, VL control). This suggests that humoral responses that effectively block acquisition are not necessarily correlated with cellular responses that control pathogenesis. Furthermore, we show that the Env-induced CTL suppressed acute viraemia better than Gag CTL, but suppressed chronic viraemia less effectively (Fig. 1b, d). Our data also show that vaccination resulted in reduced T/F viruses in breakthrough infections. This suggests that the primary mechanism of protection is by lowering the effective infectious dose, that is, *in vivo* neutralization.

Analysis of the sequences of breakthrough viruses revealed an amino acid signature, in the C1 region of Env, of viruses more likely to escape this neutralization. By creating point mutations that interconverted the neutralization profile of well-characterized viral envelopes, we defined a minimal two-amino-acid signature at positions 45 and 47 (TR vs A/K). Importantly, introduction of the A/K signature resulted in a fraction of clonally derived Env proteins having a 'global' antigenic change. This was manifested as resistance to polyclonal sera from dozens of animals, as well as resistance to monoclonal antibodies directed to the CD4 binding site or the V1V2 loops. Thus, the mechanism of resistance probably includes post-transcriptional modification, such as alternative glycosylation or folding, capable of masking the majority of epitopes on the viral Env.

We identified a hierarchy within this neutralization escape mechanism. This phenotype can occur for only a specific domain of the Env, such as for V1V2-directed antibodies against SIV (Extended Data Fig. 5b) and HIV (Fig. 6). This probably occurs through alternative glycosylation pathways restricted to this site. Resistance can also be global, affecting virtually any epitope, as we show for the SIV envelope (Extended Data Fig. 5a). A hierarchy was observed *in vivo*, in that sieving at the V1V2 domain was only observed in viruses lacking the global resistance phenotype (that is, in TR but not A/K viruses).

The observation that C1 amino acid variations can lead to alternative Env structures is consistent with data from ref. 22, where a single amino acid substitution was found to confer co-receptor dual tropism on mac239. Notably, the mutation responsible for the altered structure at the distant V3 loop was 47E—that is, within the signature we identified as conferring altered antigenicity upon SIV Env.

It is notable that all viruses in the mac251 swarm contain the resistant A/K signature. This may account for the weak correlation with vaccine-induced antibody in previous studies[11,14]. It is likely that the resistant Env form can be neutralized by antibodies targeting 'sites of vulnerability' (that is, rare epitopes conserved across all structures); for SIV, as it is in HIV[21], one of these may be the α2 helix. Antibody responses to this peptide were not only highly correlated with protection against TR viruses (Fig. 4g), but also showed a trend for protection against infection with the A/K viruses ($P = 0.07$). Likewise, the CD4-Ig molecule fully neutralized A/K viruses, suggesting that an appropriately targeted antibody to the CD4 binding site could have a similar effect.

By restricting our correlates analysis to exclude infections resulting from neutralization-resistant viruses (which are insensitive to the vaccine responses), we identified several strong correlates of risk of infection. All of these correlates derive from antibody measures, and include the magnitude of binding, the avidity of binding, and the breadth to selected epitopes of the SIV envelope. The importance of taking into account the virology is underscored by our analysis of the mosaic arm: despite this arm not achieving statistically significant protection overall (Fig. 1), we could identify active immune mechanisms (Fig. 4) as well as identify a mosaic immunogen-specific sieving effect in V1V2 (Fig. 5).

Our study provides insight into the possible reasons for the failure of HVTN505 and the limited protection in RV144[1]. Vaccination using our specific SIV Env expression vectors generated an antibody response ineffective against specific variants and protected against the subset of neutralization-sensitive viral variants (Fig. 2e). On the basis of data here, we propose that HVTN505 failed owing to an inability to elicit antisera that completely neutralized circulating HIV-1 strains, which are primarily neutralization-resistant. In contrast, the moderate success of RV144 suggests that antibodies were elicited that could neutralize some viruses circulating in that cohort; these sensitive viruses were susceptible to the vaccine-matched V1V2, leading to sieving. In any case, it will be critical to apply integrated analyses to HIV vaccine trials similar to what we did for this SIV study: that is, to clone and determine the neutralization profile of T/F viruses in the placebo arms (defining resistance of the circulating strains) and the active arms (to determine if the vaccine selectively blocked a subset of viruses), to optimally assess factors associated with vaccine-mediated protection.

Deciding which vaccine products to advance into large, expensive efficacy trials is difficult and complex. It is reasonable to postulate that any highly efficacious candidate will need to elicit antibodies targeting universal sites of vulnerability (that is, epitopes shared by the heterogeneous forms of even clonal virions), or to separately elicit antibodies targeting each structural form. Thus, understanding the biophysical basis for this viral heterogeneity will be crucial for designing vaccines capable of completely blocking HIV.

In conclusion, we identified a sequence signature of the SIV Env that distinguishes broadly neutralization-resistant viruses. By taking this signature of T/F viruses from breakthrough infections into account, we found several strong correlates of risk against infection, all based on antigen-specific antibody measurements—even for the mosaic vaccine arm that did not, upon initial analysis, reach statistically significant protection. We found that this signature, although probably not unique, is shared by SIV and HIV, and may underlie a fundamental mechanism of immune escape in both vaccinated and naturally infected subjects. Finally, our combined virological and immunological analyses provide insight into the biology of vaccine-mediated control, and lay a foundation for analysis and advancement of future HIV vaccines.

METHODS SUMMARY

Animals were handled in accordance with the standards of the American Association for the Accreditation of Laboratory Animal Care (AAALAC) and meet NIH standards as set forth in the Guidelines for Care and Use of Laboratory Animals. The animal protocol, VRC 10-332, was approved by the Vaccine Research Center IACUC. Functionality of all immunogens (mac239 and mosaic, Env and Gag) was confirmed by multiple assays. Animals were randomized into four groups of 20 based on *TRIM5α* alleles, gender, age and weight. Animals were challenged weekly with a dose of SIV_{smE660} previously shown to infect unvaccinated animals approximately 30% per exposure, as described[14]. Weekly challenges were initiated at week 53 (6 months after rAd5 boost), and were halted when an animal became PCR positive for viral RNA, or after 12 exposures. All immunological and virological assays performed for correlation analyses were qualified or validated, and performed by investigators blind as to group assignment and challenge outcome.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 September; accepted 21 November 2013.
Published online 18 December 2013.

1. Rerks-Ngarm, S. *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N. Engl. J. Med.* **361,** 2209–2220 (2009).

2. Haynes, B. F. *et al.* Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N. Engl. J. Med.* **366,** 1275–1286 (2012).
3. Tomaras, G. D. *et al.* Vaccine-induced plasma IgA specific for the C1 region of the HIV-1 envelope blocks binding and effector function of IgG. *Proc. Natl Acad. Sci. USA* **110,** 9019–9024 (2013).
4. Rolland, M. *et al.* Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. *Nature* **490,** 417–420 (2012).
5. Liao, H. X. *et al.* Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. *Immunity* **38,** 176–186 (2013).
6. Hammer, S. M. *et al.* Efficacy Trial of a DNA/rAd5 HIV-1 Preventive Vaccine. *N. Engl. J. Med.* **369,** 2083–2092 (2013).
7. Keele, B. F. *et al.* Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. *J. Exp. Med.* **206,** 1117–1134 (2009).
8. Hidajat, R. *et al.* Correlation of vaccine-elicited systemic and mucosal nonneutralizing antibody activities with reduced acute viremia following intrarectal simian immunodeficiency virus SIVmac251 challenge of rhesus macaques. *J. Virol.* **83,** 791–801 (2009).
9. Lai, L. *et al.* Prevention of infection by a granulocyte-macrophage colony-stimulating factor co-expressing DNA/modified vaccinia Ankara simian immunodeficiency virus vaccine. *J. Infect. Dis.* **204,** 164–173 (2011).
10. Schell, J. B. *et al.* Significant protection against high-dose simian immunodeficiency virus challenge conferred by a new prime-boost vaccine regimen. *J. Virol.* **85,** 5764–5772 (2011).
11. Barouch, D. H. *et al.* Vaccine protection against acquisition of neutralization-resistant SIV challenges in rhesus monkeys. *Nature* **482,** 89–93 (2012).
12. Flatz, L. *et al.* Gene-based vaccination with a mismatched envelope protects against simian immunodeficiency virus infection in nonhuman primates. *J. Virol.* **86,** 7760–7770 (2012).
13. Lai, L. *et al.* SIVmac239 MVA vaccine with and without a DNA prime, similar prevention of infection by a repeated dose SIVsmE660 challenge despite different immune responses. *Vaccine* **30,** 1737–1745 (2012).
14. Letvin, N. L. *et al.* Immune and genetic correlates of vaccine protection against mucosal infection by SIV in monkeys. *Sci. Transl. Med.* **3,** 81ra36 (2011).
15. Lopker, M. *et al.* Heterogeneity in neutralization sensitivities of viruses comprising the simian immunodeficiency virus SIVsmE660 isolate and vaccine challenge stock. *J. Virol.* **87,** 5477–5492 (2013).
16. Fischer, W. *et al.* Distinct evolutionary pressures underlie diversity in simian immunodeficiency virus and human immunodeficiency virus lineages. *J. Virol.* **86,** 13217–13231 (2012).
17. Fischer, W. *et al.* Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nature Med.* **13,** 100–106 (2007).
18. Hudgens, M. G. & Gilbert, P. B. Assessing vaccine effects in repeated low-dose challenge experiments. *Biometrics* **65,** 1223–1232 (2009).
19. Hudgens, M. G. *et al.* Power to detect the effects of HIV vaccination in repeated low-dose challenge experiments. *J. Infect. Dis.* **200,** 609–613 (2009).
20. Gray, E. S. *et al.* Isolation of a monoclonal antibody that targets the alpha-2 helix of gp120 and represents the initial autologous neutralizing-antibody response in an HIV-1 subtype C-infected individual. *J. Virol.* **85,** 7719–7729 (2011).
21. Moore, P. L. *et al.* Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection. *PLoS Pathog.* **5,** e1000598 (2009).
22. Del Prete, G. Q. *et al.* Derivation and characterization of a simian immunodeficiency virus SIVmac239 variant with tropism for CXCR4. *J. Virol.* **83,** 9911–9922 (2009).
23. Gray, R. T. A class of *K*-sample tests for comparing the cumulative incidence of a competing risk. *Ann. Stat.* **16,** 1141–1154 (1988).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the following individuals: W. Shi, L. Wu, S.-Y. Ko, L. Wang and W.-P. Kong for immunogen construction; M. M. Donaldson, S.-F. Kao, D. Quinn, J. Owuor, K. Denison, H. Balachandran, C. Luedemann, W. T. Williams, G. Overman, A. Deal, C. Brinkley and L. Racz for technical assistance with immunology assays; A. Ault for assistance managing NHP studies; S. O'Connor for deep-sequencing data; M. Seaman for providing plasmids encoding E660 envelopes; F. McCutchan, J. Overbaugh and J. Kim for HIV-1 strains; and P. Gilbert for advice with using the Aalen and Johansen model. This work was supported by the Intramural Research Program of the Vaccine Research Center, NIAID, NIH; by NIH contracts HHSN261200800001E (B.F.K., W.G.) and HHSN27201100016C (D.C.M.); by NIH grant AI100645 (B.T.K., W.F.); and by the Bill and Melinda Gates Foundation grant OPP1032317.

Author Contributions M.R., R.A.S., R.A.K., G.J.N., N.L.L., S.S.R. and J.R.M. designed and supervised the study. W.F., Z.-Y.Y., B.T.K. and G.J.N. designed and manufactured immunogens. J.-P.M.T., N.L.L. and S.S.R. supervised nonhuman primate procedures. M.R., B.F.K., K.E.F., A.P.B., L.V.M., K.E.S., B.M.W., R.T.B., R.G., G.F., S.M.A., T.N.D., D.C.M., G.D.T., R.A.K. and J.R.M. supervised assays and performed primary data analysis. D.C.M. provided HIV-1 strains. S.D.S., R.D.M., H.C.W., C.L., M.K.L., L.V.M., L.S., K.E.S. and B.M.W. performed assays. M.R., W.G. and M.C.N. aggregated data and performed statistical analysis. M.R. and J.R.M. wrote the manuscript.

Author Information T/F sequences are deposited in GenBank under accession numbers KF602252–KF603880. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.R. (Roederer@NIH.gov).

METHODS

Animals. Animals were handled in accordance with the standards of the American Association for the Accreditation of Laboratory Animal Care (AAALAC) and meet NIH standards as set forth in the Guidelines for Care and Use of Laboratory Animals. The animal protocol, VRC 10-332, was approved by the Vaccine Research Center IACUC. All animals were Indian-origin rhesus macaques, male or female, approximately 3–5 years of age. Animals selected for the study were negative for MHC class I alleles Mamu-A01, -B08, and -B17[24]. Animals were typed by PCR for *TRIM5α* alleles, and categorized as having 0, 1 or 2 restrictive alleles[25]. 80 animals were randomized into four arms based on the following criteria: *TRIM5α* allele category, gender, age and weight. Blood was collected at regular intervals (weekly or biweekly). Peripheral blood mononuclear cells (PBMC) were prepared; a small number were reserved for phenotyping for absolute cell counts, and the remainder were viably cryopreserved (in fetal bovine serum containing 10% DMSO and stored in liquid phase nitrogen until analysis). Plasma was frozen at −80 °C for virological and serological analysis. Sample size ($n = 20$ per arm) was chosen to have an 80% probability to detect a vaccine efficacy of 50%[18].

Immunization. The design of the mosaic immunogens has been previously described[16]. Briefly, an input data set was assembled to include all available SIV_{sm} complete genome sequences that were either directly isolated from sooty mangabeys or (in a small number of cases) had been minimally passaged in tissue culture. So that viral challenges could reasonably approximate real-world post-vaccination exposure to unknown virus strains, we specifically excluded from the mosaic sequence input all sequences from the mac251 lineage (including mac239) as well as isolates of (and closely related sequences to) smE660. Mosaic sequence cocktails were generated sequentially, so that a single-sequence mosaic was generated first, and a second sequence was subsequently generated as the best complement to the first[26,27]. Coverage values of potential T-cell epitopes (as amino acid 9-mers) have been published[16]. Mosaic coding sequences were introduced into the same DNA and rAd5 backbones as the mac239 Env. All rAd5 vectors were produced by GenVec. The mosaic Env were given as gp160 in both the DNA prime and the rAd5 boost. The mac239 natural Env immunogens (DNA and rAd5) are identical to what was previously used[14], and was given as a gp140 in the DNA prime, and a gp145 in the rAd5 boost.

Functionality of all immunogens (mac239 and mosaic, Env and Gag) was confirmed by multiple assays. Expression of Env and Gag from DNA and rAd5 vectors in 293 cells in tissue culture was assessed by western blot analysis, and was found to be comparable. Immunogenicity of each vector/insert combination was confirmed in mouse studies.

Primates were immunized as previously described[14]: a total of four times; 4 mg of DNA was given intramuscularly at weeks 0, 4, and 8 and 10^{10} particles of rAd5 was given intramuscularly at week 28. The two Gag mosaic immunogens were mixed before administration, as were the two Env mosaic immunogens.

SIV challenge. Animals were challenged weekly with a dose of smE660 previously shown to infect unvaccinated animals approximately 30% per exposure, as previously described[14]. Weekly challenges were initiated at week 53 (6 months after rAd5 boost), and were halted when an animal became PCR positive for viral RNA, or after 12 exposures. There was no statistically significant change in the rate of infection within any group over time, indicating that infection was stochastic and there was no selection for innately resistant animals.

Assays. All immunological and virological assays performed for correlation analyses were performed by investigators blind as to group assignment and challenge outcome.

Virology. To quantify SIV viral load, viral RNA from plasma was isolated using a Qiagen QIAsymphony Virus/Bacteria Midi kit on Qiagen's automated sample preparation platform, the QIAsymphony SP. Viral RNA from 500 µl of plasma was eluted into 60 µl of elution buffer. All subsequent reactions were setup using Qiagen's automated PCR setup platform, the QIAgility. 25 µl of viral RNA was annealed to a target specific reverse primer 5′-CACTAGGTGTCTCTGCACTAT CTGTTTTG-3′ then reverse transcribed using SuperScript III RT (Invitrogen) and PCR nucleotides (Roche) along with RNase Out (Invitrogen) using an optimized version of the manufacturer's protocol. Resulting cDNA was treated with RNase H (Applied Biosystems) according to manufacturer's protocol. 10 µl of cDNA was then used to setup a real-time PCR using Gene Expression Mastermix (Applied Biosystems) along with target specific labelled probe 5′-/56-FAM/CT TCCTCAGTGTGTTTCACTTTCTCTTCTGCG/3BHQ_1/-3′ and forward 5′-GTC TGCGTCATCTGGTGCATTC-3′ and reverse primers 5′-CACTAGGTGTCTC TGCACTATCTGTTTTG-3′ (custom synthesis by Integrated DNA Technologies). Real-time PCR was performed on an Applied Biosystems StepOne Plus platform using the standard curve protocol. The RNA standard was transcribed from the pSP72 vector containing the first 731 bp of the SIV_{mac239} or SIV_{smE660} *gag* gene using the MEGAscript T7 kit (Ambion Inc.), quantitated by optical density (OD), and serially diluted to generate a standard curve. The quality of the RNA standard was assessed using an Agilent Bioanalyzer with RNA Nano 6000 chips (Agilent Inc.). The sensitivity of this assay has been shown to be 250 copies per ml.

The number of transmitted/founder (T/F) variants was determined by single genome amplification (SGA) of the full-length envelope gene as previously described[7]. The number of sequences analysed per animal was 21.2 ± 4.8 (mean ± s.d.), with a range of 10–38. There was no difference in number of sequences analysed by group.

All 1,629 sequences are deposited in GenBank under accession numbers KF602252–KF603880.

SIV envelope constructs. Sequences of the CP3C-P-A8 envelope (referred to in this paper as 'CP3C' for brevity) and CR54-PK-2A5 ('CR54') are shown in Supplementary Table 1. These sequences were used to produce protein for binding assays as well as pseudotyped viruses. Mutations were designed into each virus to create individual amino acid variants as listed in Fig. 3; the relevant portion of the envelope, encompassing the C1 region, is shown aligned in Supplementary Table 4. SIV Env mutant plasmids were generated by site-directed mutagenesis by GeneImmune Biotechnology.

Immunology. Intracellular cytokine staining for antigen-specific responses was performed using a qualified assay as described[28]. Cells were stimulated with overlapping 15mer peptides from Gag or Env from mac239 or smE543 (a clone similar to smE660). Data are shown for stimulations with E543 peptides. For breadth analysis, IFN-γ ELIspots were performed as described[29], using pools of 10 peptides from each protein.

Raw peptide microarray data (PepStar) were pre-processed and normalized as previously described[30]. Responses to peptides from mac239 or E543 were measured; data are shown for E543 only, except in Fig. 5a where data from both sets are shown. For each peptide, the mean binding from 10 control animals was subtracted from the value for each vaccinated animal. The distribution of resulting values was used to define a cut-off value of 1.2 for positivity: a large fraction of peptide responses constituted a near-normal distribution centred on 0 (after background subtraction); the 10th percentile of this distribution was −1.2; thus, +1.2 is an estimate of the 90th percentile of a completely negative response. For breadth analysis, positive responses to partially overlapping peptides were considered to comprise a single epitope.

SIV-specific humoral IgG and IgA levels were evaluated by a standardized antibody binding multiplex assay as previously described[31,32]. IgA levels were low and are shown only as MFI for the lowest dilution tested. IgG levels are shown as MFI AUC (area under the curve) computed over the dilutions in the linear range of the assay. Avidity was quantified by surface plasmon resonance (SPR) as previously described[2,33].

Viral neutralization assays. Neutralization was evaluated using three distinct assays. (1) Plasma neutralization of viral replication in PBMC was performed as previously described[14]. (2) Env-pseudovirus neutralization was measured using single-round-of-infection SIV Env-pseudoviruses with TZM-bl target cells stably expressing high levels of CD4 and the co-receptors CCR5 and CXCR4[34,35]. Tat-regulated luciferase gene expression was quantified to determine the reduction in virus infection. Neutralization curves were fit by nonlinear least squares regression, and the 50% inhibitory concentrations (IC_{HM}) was computed as the antibody concentration required to achieve 50% of maximal inhibition. (3) Replication competent SIV was used to infect TZM-bl cells as above, with cloned or uncloned swarm SIVs essentially as described[36]. Briefly, neutralization assays were performed with serial dilutions of heat-inactivated (56 °C, 1 h) samples. Diluted samples were pre-incubated with virus (~150,000 relative light unit equivalents) for 1 h at 37 °C before addition of cells. Following 48 h incubation, cells were lysed and luciferase activity determined using a microtitre plate luminometer and BriteLite Plus Reagent (Perkin Elmer). Neutralization titres are the reciprocal sample dilution or concentration (for sCD4) at which relative luminescence units (RLU) were reduced by 50% compared to RLU in virus control wells after subtraction of background RLU in cell control wells.

CD4 binding inhibition by sera was determined as described[37] with the following modifications. Plate-bound CP3C Env was incubated with or without a 1:5 dilution of plasma at 37 °C for 1 h. After washing, wells were incubated with 50 µg ml^{-1} CD4-Ig–Biotin at 37 °C for 1 h. Plates were washed to remove excess CD4-Ig–Biotin and incubated with streptavidin horseradish peroxidase at 37 °C for 1 h. Inhibition was calculated as the fraction of the signal in wells with plasma to those without.

Statistics. The analyses presented here used a variety of techniques. Comparisons of continuous end points between groups were based on *t*-tests and analysis of variance, log-transformed when appropriate (for example, viral load). Comparisons of groups with respect to number of challenges until infection used the discrete time survival model assuming a leaky vaccine effect[18]. A comparison of the goodness-of-fit of possible models showed that the likelihood of the leaky model performed better than the all-or-none model (and the null hypothesis), and performed similarly to a

model that allowed both types of effects[18]. For the cumulative incidence of A/K vs TR viruses, we used nonparametric estimates that allowed for competing risks[23]. V_E against each virus type was computed by modifying the Hudgens and Gilbert leaky vaccine model[18] to account for two infection types, with P values computed from likelihood ratio tests. We did not formally test for heterogeneity of per-exposure probability. However, two pieces of evidence support lack of heterogeneity. First, the leaky model is a better fit than the all-or-none model (which should be sensitive to one type of heterogeneity). Second, we evaluated the risk of infection as a function of challenge number, and found no statistically significant change over time in any group.

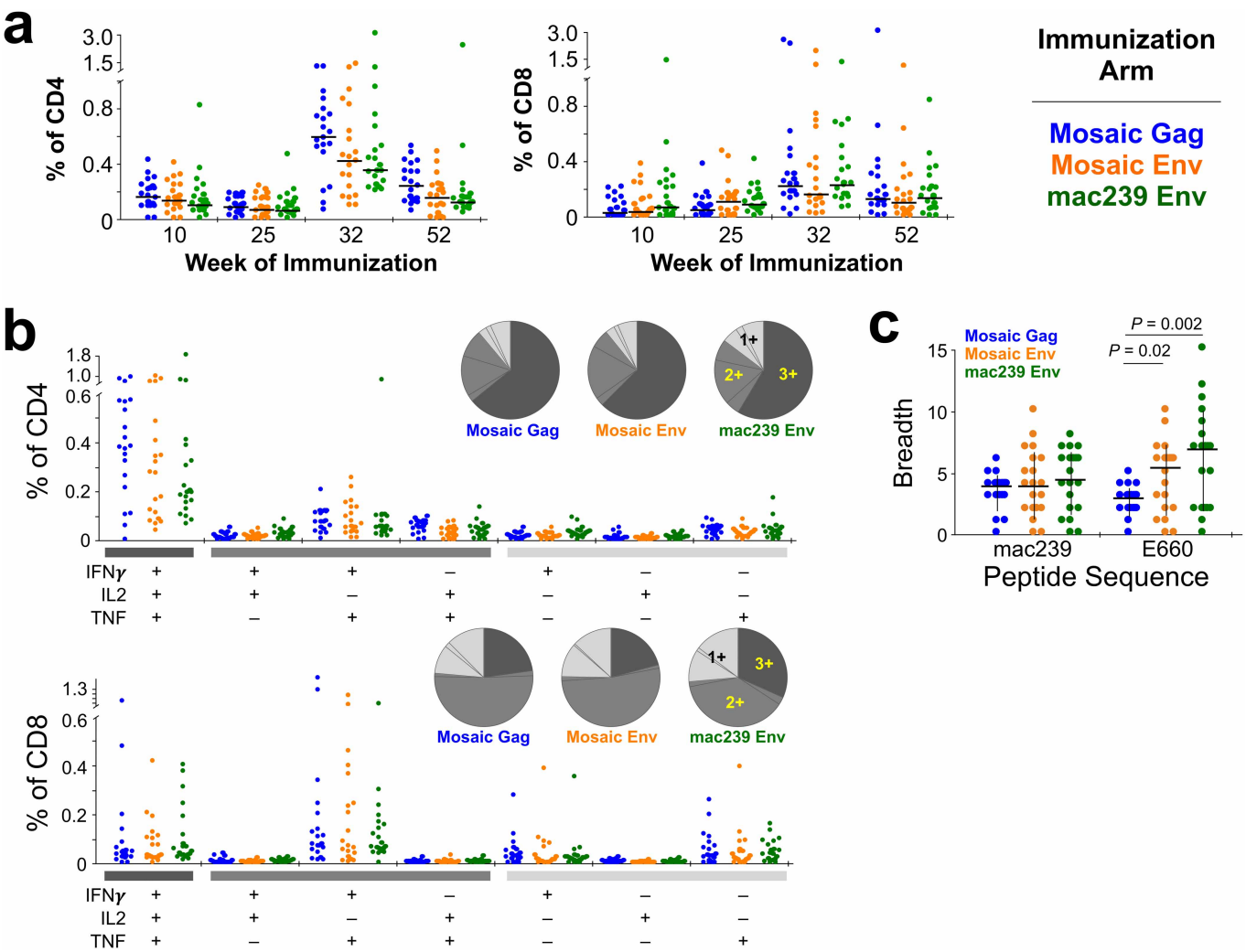
Associations between immunological measurements and number of challenges were based on similar models, with continuous predictors dichotomized at the median. Immunological predictors that showed some association using this method were investigated further using Cox Proportional Hazard models, both univariate and multivariate. All correlations were based on Spearman's rho (non-parametric), to handle variables with censored readings below or above assay limits, as well as to include uninfected animals in correlations based on number of infections. P values reported are not corrected for the number of comparisons between multiple immunological assays and outcome, except as noted.

In this paper we identified several potential correlates of risk of infection[38,39] that deserve further investigation and confirmation. Since these are not measureable in unvaccinated animals, the correlations with time to infection we report might be related to either differential effects of vaccines or to unspecified differences between the immune systems of the animals.

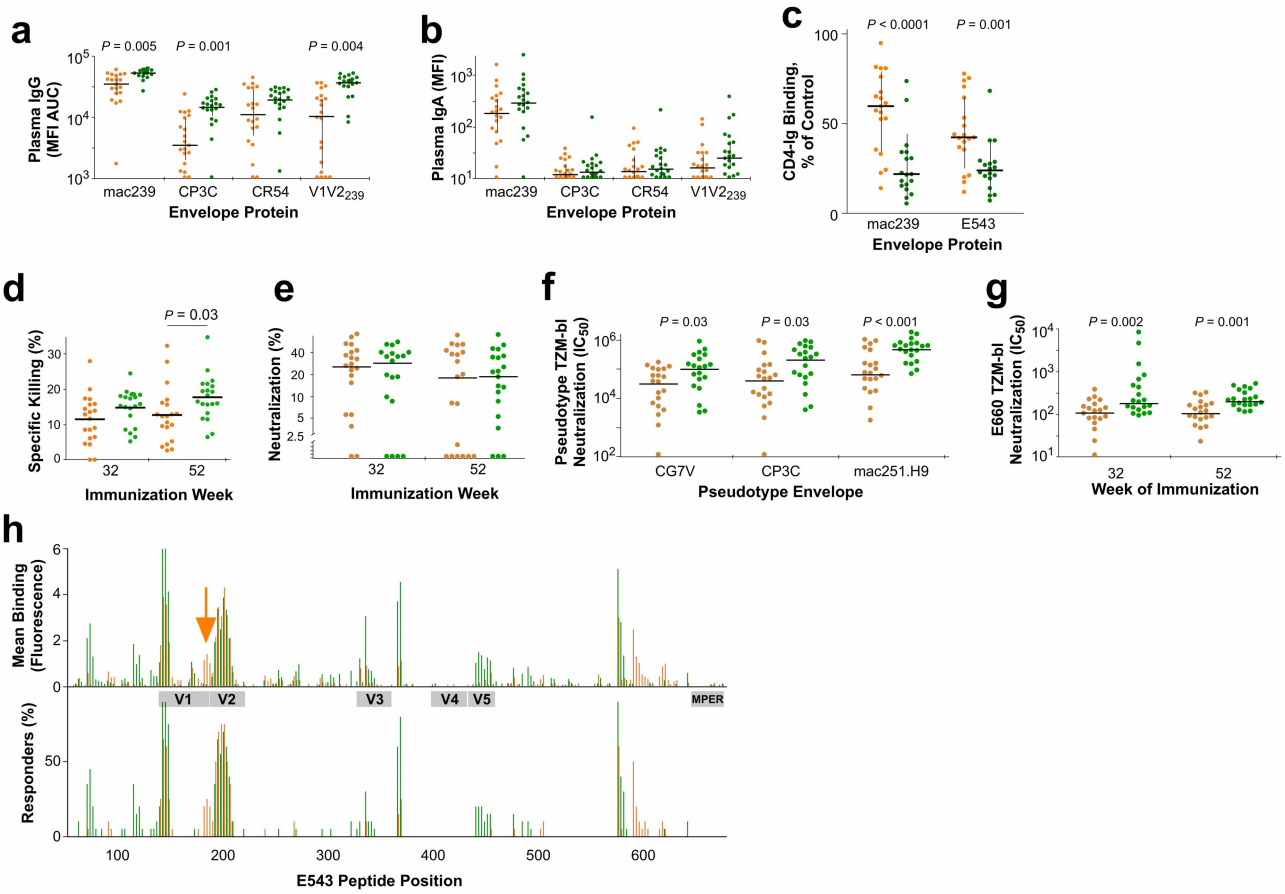
For sieve analysis, only Env positions with at least five variants amongst all sequences were considered. The distribution of variants at those positions passing this minimum threshold was compared between groups using a permutation-based version of the Fisher's exact test, using 10,000 permutations. P values reported in Fig. 2c and Extended Data Fig. 4 are not corrected for multiple comparisons, but positions 23, 45 and 47 (in both analyses) remain significant ($P \leq 0.002$) after such correction. Similarly, position 162 in Fig. 5b is significant after correction for multiple comparisons.

As expected from our stratification, *TRIM5α* alleles were found to have no effect on the conclusions of vaccine effects on protection; as follows: analysis of the discrete time-to-infection model using only *TRIM5α*-resistant animals did not change V_E. Cox proportional hazard modelling of time-to-infection by group did not change when *TRIM5α* was included as a covariate. Similarly, the importance of virus type infection (A/K vs TR) was unaffected by inclusion of *TRIM5α*. Finally, immunological correlates analyses (prediction of time-to-infection by antibody measures) did not change when *TRIM5α* was added as a covariate. The number of animals in each arm (out of 20) with homozygous *TRIM5α*-sensitive alleles was 8 (control), 9 (mosaic Env), 7 (mosaic Gag) and 8 (mosaic Gag).

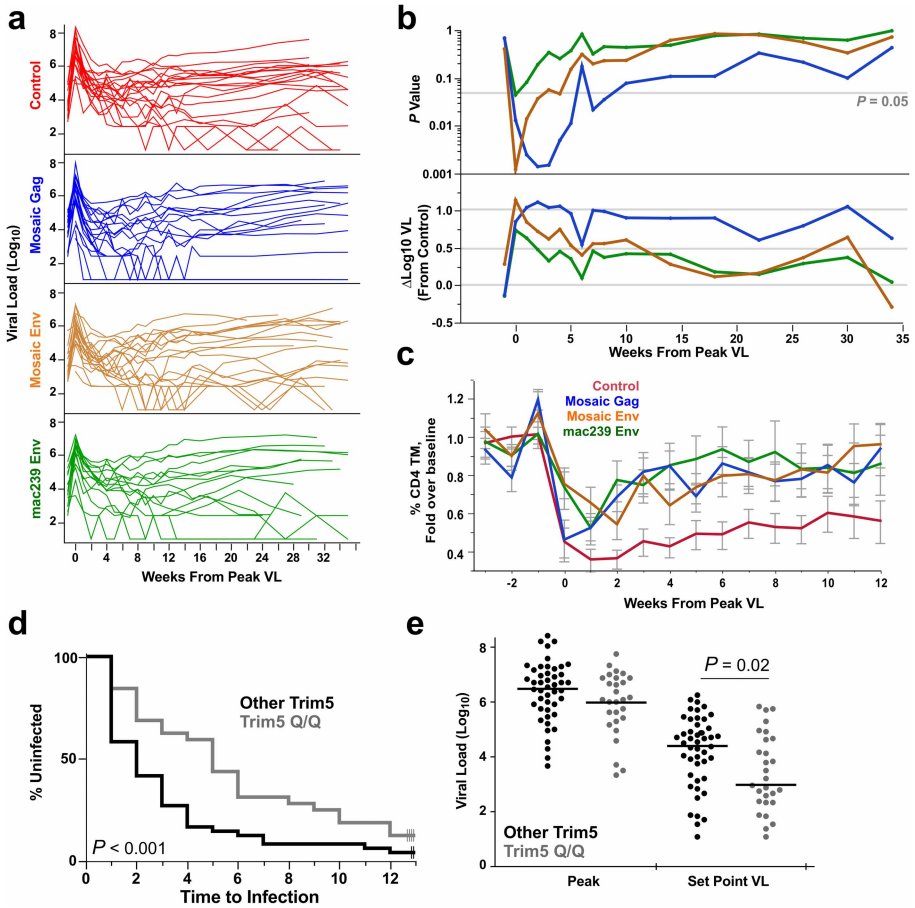
24. Lim, S. Y. *et al.* Contributions of *Mamu-A*01* status and *TRIM5* allele expression, but not *CCL3L* copy number variation, to the control of SIVmac251 replication in Indian-origin rhesus monkeys. *PLoS Genet.* **6,** e1000997 (2010).
25. Lim, S. Y. *et al.* TRIM5alpha modulates immunodeficiency virus control in rhesus monkeys. *PLoS Pathog.* **6,** e1000738 (2010).
26. Fenimore, P. W. *et al.* Designing and testing broadly-protective filoviral vaccines optimized for cytotoxic T-lymphocyte epitope coverage. *PLoS ONE* **7,** e44769 (2012).
27. Yusim, K. *et al.* Genotype 1 and global hepatitis C T-cell vaccines designed to optimize coverage of genetic diversity. *J. Gen. Virol.* **91,** 1194–1206 (2010).
28. Donaldson, M. M. *et al.* Optimization and qualification of an 8-color intracellular cytokine staining assay for quantifying T cell responses in rhesus macaques for pre-clinical vaccine studies. *J. Immunol. Methods* **386,** 10–21 (2012).
29. Santra, S. *et al.* Heterologous prime/boost immunizations of rhesus monkeys using chimpanzee adenovirus vectors. *Vaccine* **27,** 5837–5845 (2009).
30. Imholte, G. C. *et al.* A computational framework for the analysis of peptide microarray antibody binding data with application to HIV vaccine profiling. *J. Immunol. Methods* **395,** 1–13 (2013).
31. Bolton, D. L. *et al.* Comparison of systemic and mucosal vaccination: impact on intravenous and rectal SIV challenge. *Mucosal Immunol.* **5,** 41–52 (2012).
32. Tomaras, G. D. *et al.* Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. *J. Virol.* **82,** 12449–12463 (2008).
33. Alam, S. M. *et al.* Human immunodeficiency virus type 1 gp41 antibodies that mask membrane proximal region epitopes: antibody binding kinetics, induction, and potential for regulation in acute infection. *J. Virol.* **82,** 115–125 (2008).
34. Shu, Y. *et al.* Efficient protein boosting after plasmid DNA or recombinant adenovirus immunization with HIV-1 vaccine constructs. *Vaccine* **25,** 1398–1408 (2007).
35. Wu, L. *et al.* Enhanced exposure of the CD4-binding site to neutralizing antibodies by structural design of a membrane-anchored human immunodeficiency virus type 1 gp120 domain. *J. Virol.* **83,** 5077–5086 (2009).
36. Li, M. *et al.* Human immunodeficiency virus type 1 *env* clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79,** 10108–10125 (2005).
37. Zhao, J. *et al.* Preclinical studies of human immunodeficiency virus/AIDS vaccines: inverse correlation between avidity of anti-Env antibodies and peak postchallenge viremia. *J. Virol.* **83,** 4102–4111 (2009).
38. Qin, L., Gilbert, P. B., Corey, L., McElrath, M. J. & Self, S. G. A framework for assessing immunological correlates of protection in vaccine trials. *J. Infect. Dis.* **196,** 1304–1312 (2007).
39. Plotkin, S. A. & Gilbert, P. B. Nomenclature for immune correlates of protection after vaccination. *Clin. Infect. Dis.* **54,** 1615–1617 (2012).



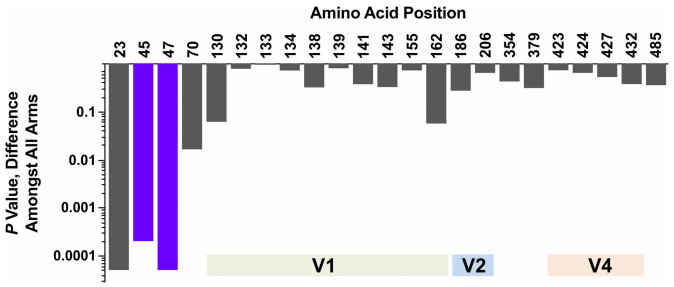
Extended Data Figure 1 | Cellular immunogenicity of vaccines. Gag- or Env-specific CD4 and CD8 T-cell responses measured by intracellular cytokine stimulation. Total T-cell responses were similar in all three active arms. **a**, Induction of T-cell responses are shown as the fraction of CD4 or CD8 memory T cells producing IFN-γ, IL2 or TNF in response to stimulation with overlapping peptides matched to the E660 challenge strain. Time points include peak post-DNA prime (week 10), pre-boost (week 25), peak post-rAd5 boost (week 32) and pre-challenge (week 52). **b**, The quality of the week 32 T-cell response is shown by the fraction of CD4 or CD8 cells responding to overlapping peptide pools matched to the E660 challenge swarm. There was no difference in the quality between any of the groups at any time point. **c**, Mosaic vaccination did not significantly improve the breadth of the T-cell response. Responses to pools of 10 overlapping peptides corresponding to mac239 Env or Gag, or the smE543 Env or Gag were tested for responses measured by ELISpot for the week 32 samples. Graphed is the number of positive pools (out of 23 for Env, and 13 for Gag) for each animal by group.



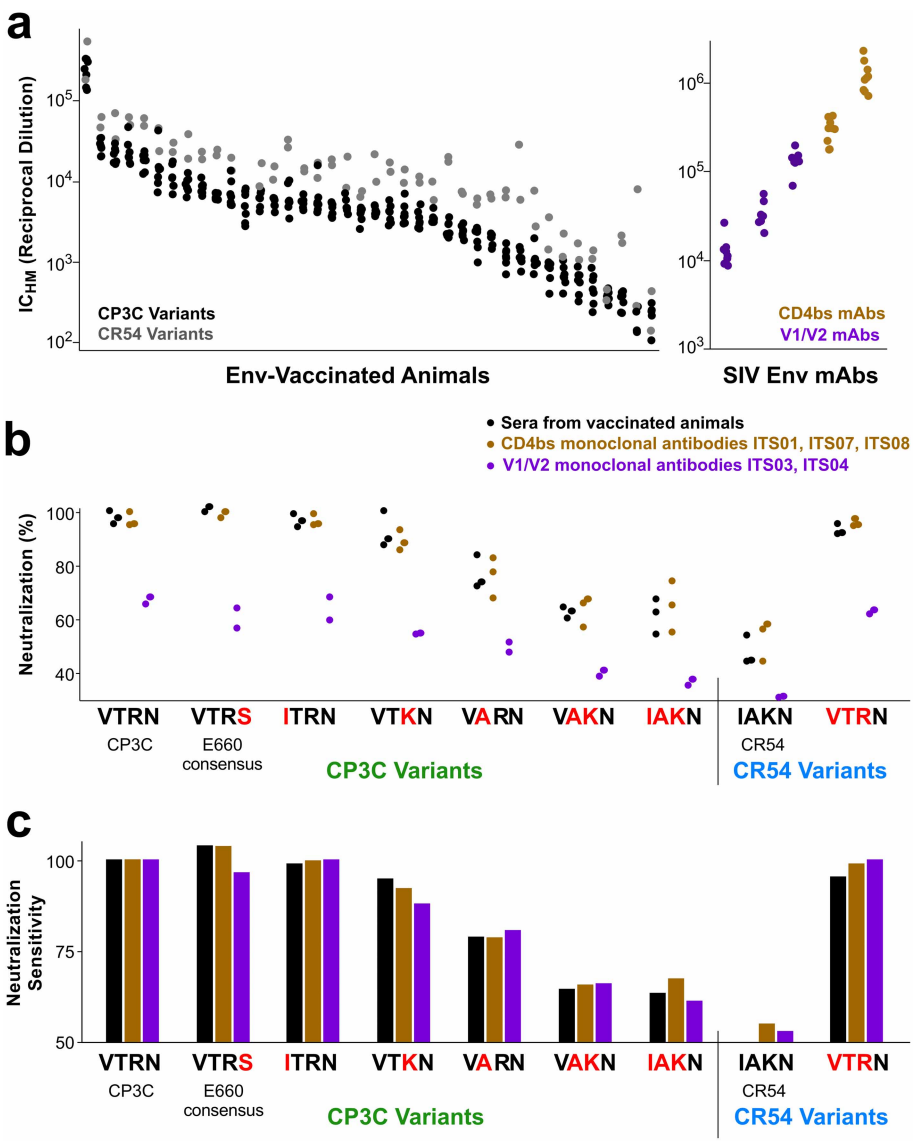
Extended Data Figure 2 | Humoral immunogenicity of vaccines. Mosaic immunization induced mildly lower humoral responses that were qualitatively different. **a**, **b**, Plasma IgG (**a**) or IgA (**b**) responses at week 32 were quantified against SIV envelope proteins derived from mac239, E660-CP3C, E660-CR54, or a mac239 V1V2 polypeptide expressed on a J08 scaffold. MFI, mean fluorescence intensity using a bead-based Luminex platform. AUC, area under the curve. **c**, CD4-binding site activity was measured by the ability of sera to cross-block CD4-Ig binding to mac239 or smE543 envelopes. **d**, Antibody-dependent cellular cytotoxicity mediated killing of SIV-infected target PBMC, shown as per cent specific killing. **e**, PBMC neutralization assay showing no substantial difference between immunization arms or time since vaccination. **f**, Neutralization by week 32 plasma was measured against three envelope-pseudotyped viruses. **g**, Neutralization of the E660 challenge stock using the TZM-bl indicator cell line. **h**, Week 32 plasma antibody binding to overlapping peptides spanning the SIV E543 envelope was quantified for the two envelope immunization arms. The mean response for all 20 animals in each arm (top) or the fraction of animals responding (bottom) is shown for peptides from the extracellular portion of the envelope. The arrow indicates an area near the V1V2 junction targeted by the mosaic but not the mac239 immunogen; several other areas, including C1, V3, C3 and V5, were better targeted by the mac239 immunogen.



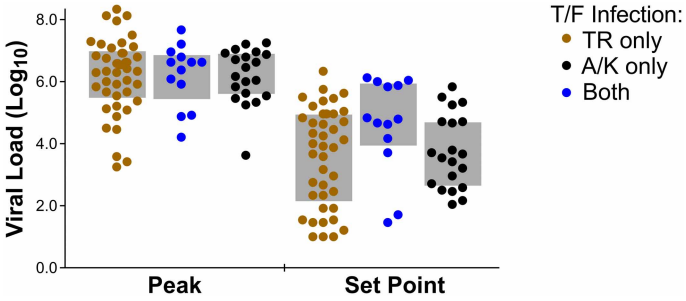
Extended Data Figure 3 | Viral pathogenesis and influence of *TRIM5α* alleles. **a**, Viral load (VL) was measured weekly until 12 weeks post peak and then monthly thereafter. Curves are shown for all 74 infected animals and are synchronized by the peak VL. **b**, For each time point, the distribution of VL in each immunization arm was compared to the control arm. The mean difference (lower) and significance of the difference (Student's *t*-test; upper) is graphed. **c**, The loss of CD4 cells following mucosal challenge is much more temperate than following intravenous challenge[31]. The most consistent measurable loss was for CD4 transitional memory cells ($CD45RA^{-}CCR7^{-}CD28^{+}$); the change in the frequency of these cells relative to the pre-infection average is shown. Other CD4 subsets showed less dramatic depletion. **d**, **e**, All 80 animals were grouped according to predicted resistance based on *TRIM5α* allelism (resistant: $TRIM5\alpha^{Q/Q}$; sensitive: all other combinations). A significant effect of genetics on acquisition (**d**) and pathogenesis (**e**) was observed. Animals were randomized equally into the four immunization arms based on *TRIM5α* genotype (for all homozygous and heterozygous genotypes).



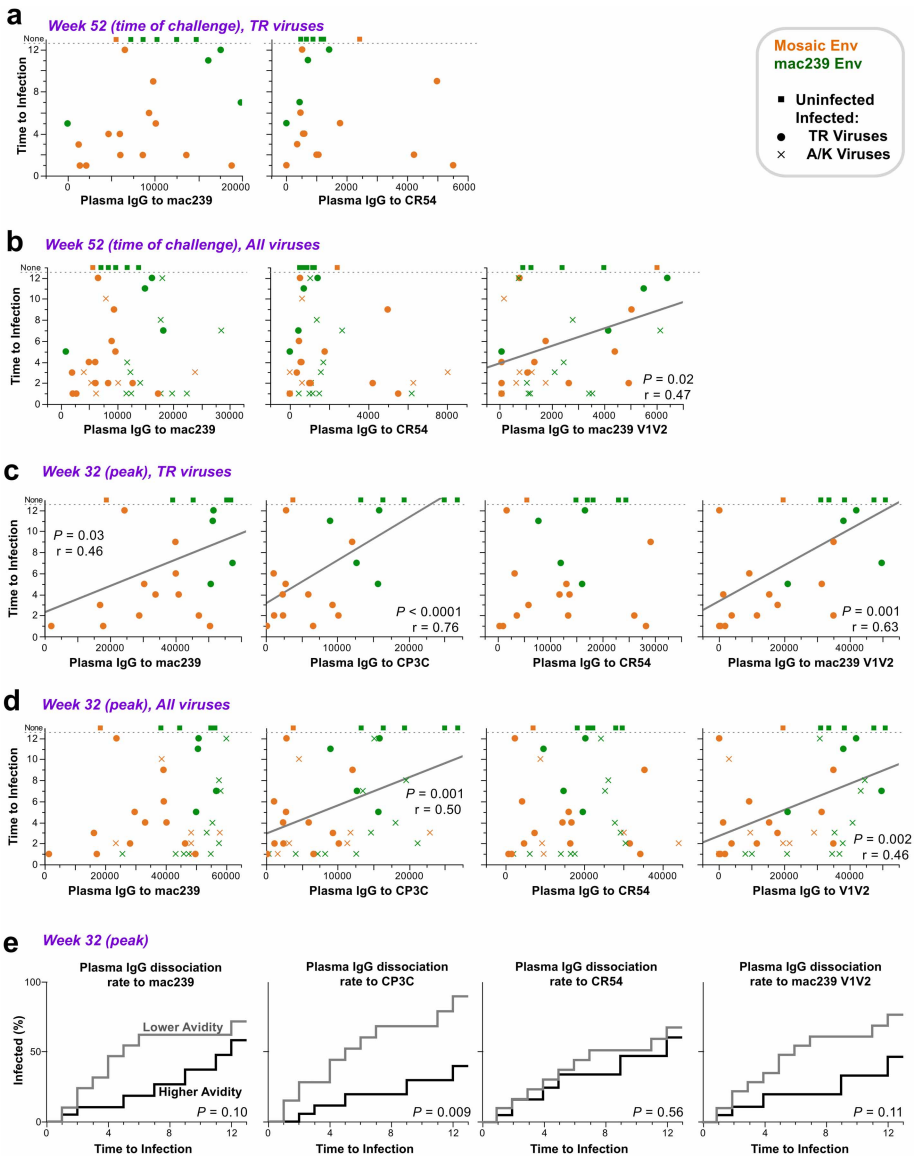
Extended Data Figure 4 | Transmitted/founder (T/F) selection in any vaccine arm. The number of T/F viruses with a variant from consensus was compared across all four arms for amino acid positions showing heterogeneity. A permutation test was used to compute the significance of a difference across all groups. The *P* values for positions 23, 45 and 47 remain significant after correction for multiple comparisons.



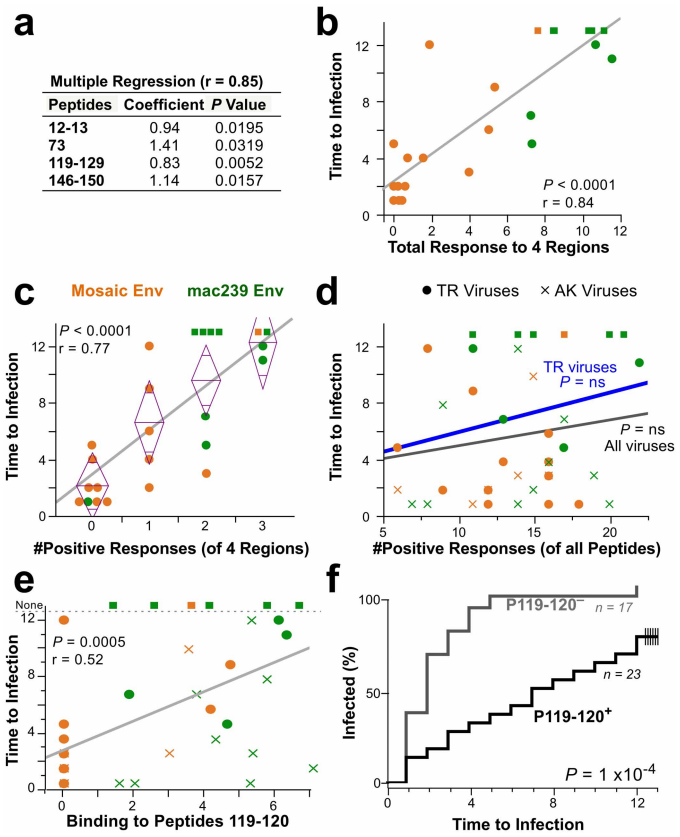
Extended Data Figure 5 | Neutralization sensitivity of variant envelopes. Nine envelope variants (Fig. 3) were evaluated for neutralization sensitivity by antisera from vaccinated animals (black or grey) and monoclonal antibodies to the CD4 binding site (brown) or the V1V2 loops (purple). **a**, The IC_{HM} (reciprocal concentration of antisera resulting in 50% of maximum neutralization) for all neutralization experiments is summarized by animal (left) or monoclonal antibody (right) for the seven CP3C variants and the two CR54 variants. The range of IC_{HM} across the viruses was less than twofold; that is, C1 sequence variations do not affect IC_{HM} but only the fraction of neutralization-resistant virions within each virus preparation (Fig. 3e). **b**, In a separate experiment, sera from three vaccinated animals and five monoclonal antibodies were compared. Note that the V1V2 antibodies only neutralize ~60% of the sensitive CP3C strain. **c**, Relative neutralization sensitivity was calculated by normalizing neutralization of each class of antibodies to 100% for CP3C.



Extended Data Figure 6 | Pathogenesis of TR and A/K viruses. Animals were divided into groups based on whether they were infected solely with TR viruses, A/K viruses, or both (that is, with multiple T/F per animal). Bars indicate the interquartile range of values. The peak and set point viral load did not differ according to which type of virus infected and replicated in the animal. In addition, no significant differences were observed when these data were split by vaccine arm.



Extended Data Figure 7 | Immunological correlates of risk, plasma IgG. **a**, **b**, Week 52 plasma IgG against the CP3C envelope is graphed against time to infection (uninfected animals were assigned a value of 13). Data are shown excluding A/K virus infections (**a**) or for all infections (**b**). Significant correlations are indicated by a linear least-squares regression line; statistics are nonparametric Spearman's tests. **c**, **d**, Similar analyses using week 32 (peak) plasma, for all TR infections (**c**), or all viral infections (**d**). **e**, Avidity to SIV envelopes was measured by Biacore; for each KM analysis, animals were divided in two equal groups based on having lower than median disassociation rate (high avidity) vs higher (low avidity), for TR infections.



Extended Data Figure 8 | Immunological correlates of risk, breadth of binding to linear peptides. **a**, A multivariable regression of time to infection vs responses to each of the four regions shown in Fig. 4g was performed. All four regions provided independent predictive power. **b**, The binding activity to all four regions was summed; the total response to these four epitopes showed a high correlation with time to infection. **c**, The number of the four regions with positive responses within each animal was computed (no animal responded to all four). The line indicates a linear regression; statistics are based on a nonparametric Spearman's test. **d**, The number of epitopes with positive responses across the entire envelope was computed for each animal. No correlation with protection (for all viruses or for only TR viruses) was seen with overall breadth. ns, $P > 0.05$. **e**, Average binding to the linear C3 peptides 119 and 120 correlates with time to infection for all animals, irrespective of virus. **f**, KM analysis comparing Env-immunized animals with a positive response to C3 peptides to those with a negative response.