**Extended 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 IFNγ, 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 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; MFI AUC, MFI 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 percent 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.

**Extended Figure 3. 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[24]. 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.

**Extended 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 acids 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 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 2-fold – i.e., C1 sequence variations do not impact $IC_{HM}$ 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 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 (i.e., with multiple T/F per animal). Bars indicate the interquartile range of values. The peak and setpoint 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 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 Figure 8. Immunological correlates of risk – Breadth of binding to linear peptides. a,** A multivariate 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.