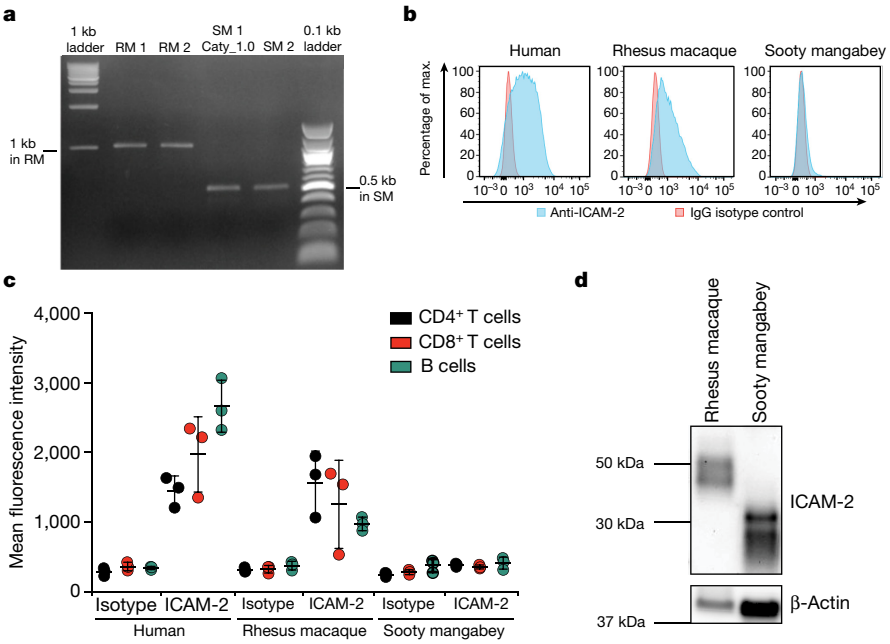

Figure 2 | Genomic deletion in Ca*ICAM2* results in a truncated and dysfunctional protein. a, PCR to confirm a putative 0.5-kb deletion in the Ca*ICAM2*. **b**, ICAM-2 surface expression of primary $CD4^+$ cells by flow cytometry. $n = 3$; representative plots for **c**. **c**, ICAM-2 surface expression in B cells, $CD4^+$ and $CD8^+$ T cells from human, rhesus macaques and sooty mangabeys. $n = 3$ biologically independent samples for each species. **d**, ICAM-2-specific western blot using peripheral blood mononuclear cells from *M. mulatta* and *C. atys*. $n = 3$ *M. mulatta*; $n = 2$ *C. atys*; one representative biological sample per species is shown. For gel source data, see Supplementary Figs 1, 3.

molecular weight form of ICAM-2 could be detected intracellularly by western blot in *C. atys* cells (Fig. 2d), thus demonstrating the presence of the predicted truncated ICAM-2 protein. Overall, these data indicate that the presence of a species-specific gene sequence difference in Ca*ICAM2* results in the abrogation of surface expression of this protein in *C. atys*. Further studies are needed to elucidate potential links between this truncated form of ICAM-2 and the remarkable immunological features of SIV infection in this species.

TLR-4 is a pattern recognition receptor that senses lipopolysaccharides (LPS) on gram-negative bacteria and initiates pro-inflammatory cytokine induction, maturation and activation in macrophages, dendritic cells and other immune cells. During pathogenic HIV or SIV infections, exacerbated TLR-4 stimulation and concomitant pro-inflammatory signalling elicited by microbial translocation is considered a primary mechanism that underlies HIV-induced chronic immune activation[10,11]. Here, we found that the TLR-4 protein sequences of *M. mulatta* and *C. atys* are markedly different at the C terminus (Extended Data Fig. 5a). We confirmed the underlying difference in the *TLR4* nucleotide sequence by Sanger sequencing (Extended Data Fig. 5b, c). We next analysed the genomic DNA sequence of *TLR4* in 10 additional sooty mangabeys and found that the observed DNA sequence difference was present in all individuals (Extended Data Fig. 6a). Alignment of TLR-4 protein sequences from different primate species revealed that the 17-amino-acid longer C-terminal sequence is only found in natural SIV hosts, such as African green monkey, drill and colobus monkey (Fig. 3a), whereas non-natural hosts, including *M. mulatta* and baboons show expression of the short TLR-4 C-terminal sequence.

The divergence of TLR-4 amino acid sequences amongst Old World primates shows an interesting pattern of molecular evolution. First, the genomic sequence encoding the *TLR4* C terminus is defined by a 1-bp deletion causing a frame shift in all Old World monkeys, both natural and non-natural hosts, including colobine and cercopithecine lineages, but it is not found in either hominoids (apes and humans) or platyrrhines (New World monkeys) (Extended Data Fig. 6b). This suggests that this mutation occurred after the hominoid–Old World monkey divergence approximately 25 million years ago[12]. Second, there is a G-to-A nucleotide substitution in the non-natural host Old World monkeys (baboons and macaques) that creates a truncated protein in these species[8] (Extended Data Fig. 6b). Although a naive analysis of this pattern would suggest two independent mutational changes in *TLR4*, the short internal branch of the species tree implies that incomplete lineage sorting of an ancestral polymorphism could also generate this pattern[13] (Fig. 3b). To test this hypothesis, we examined the *TLR4* gene tree among 17 primate species. While generally supporting the relationships among these species (Fig. 3b), the analysis also found a number of nucleotide positions—spaced throughout the gene—consistent with incomplete lineage sorting between *C. atys*, baboon and *M. mulatta* (Extended Data Fig. 7). The incomplete lineage sorting hypothesis is also more likely, given that balancing selection is often found to be acting on immune-related genes. Therefore, even though baboons are believed to be more closely related to sooty mangabeys and drills than to rhesus macaques, the phylogeny of Old World monkeys is compatible with the possibility of a single G-to-A mutation creating the truncated form of the protein in the common ancestor of baboons, rhesus macaques and sooty mangabeys[12,14] (Fig. 3b).

We next investigated potential differences in TLR-4 function between *M. mulatta* and *C. atys*. Our previous work has shown that macrophages from *C. atys* exhibit higher expression of tetherin, APOBEC and TRIM5α in response to LPS compared to *M. mulatta*[15]. This is consistent with the relative resistance of *C. atys* macrophages to *in vivo* SIV infection after experimental $CD4^+$ T cell depletion compared to SIV-infected *M. mulatta* macrophages[16]. Here we analysed cytokine gene expression and protein production after LPS stimulation, and found reduced mRNA expression and secretion of TNF (also known as TNF-α) and IL-6 in cells from *C. atys* compared to *M. mulatta* (Fig. 3c, d). Because some commercial LPS preparations contain lipoprotein contaminants that can induce TLR-2 signalling, we confirmed the TLR-4 specificity of the reduced LPS response using the selective TLR-4 agonist[17] lipid-A (Extended Data Fig. 8a, b). Next, we found that the species-specific differences between *C. atys* and *M. mulatta* in LPS-induced TNF and IL-6 production were maintained in acute and chronic infection (Fig. 3e and Extended Data Fig. 8c). Additionally, we did not observe any difference in the mRNA levels