was selected for each MacaM protein model, yielding the set of orthologous *C. atys* protein predictions most similar to the *M. mulatta* protein models. The spliced CDS sequence for each Caty_1.0 transcript prediction was extracted with gffread (utility from cufflinks v.2.1.1). Caty_1.0 transcript prediction CDS sequences were screened against the *de novo* RNA-seq assembly transcript models by alignment with BLAT (v.34) and an aligment score was calculated as the number of matching bases minus the number of CDS sequence bases missing in alignment gaps normalized by the CDS sequence length.

This score penalizes bases missing from the CDS sequence without penalizing extra sequence that may have been added to the RNA-seq transcript model during the assembly process. Only predicted CDS sequences that had a score >0.99 were retained as supported by RNA-seq data. The MacaM best match selected Caty1.0 protein models were then cross-referenced with the RNA-seq supported Caty_1.0 transcript models to eliminate protein models without RNA-seq evidence. The protein alignments to MacaM for these models were then re-examined to find genes for which the alignment identity was less than 97%, where there were gaps in the alignments or the alignment was not the full length of the protein model. These two species share a common ancestor about 10–11 million years ago, and therefore the expectation is that most proteins will be >97% identical. This was confirmed by using a maximum likelihood amino acid model (WAG amino acid matrix) to estimate sequence distances between the *C. atys* and *M. mulatta* orthologues (Extended Data Fig. 1). Proteins of interest for differential response to lentivirus infection may be more divergent than expected on average. These represent potentially divergent genes and were further screened against the Gene Ontology (GO) term 'immune response'. This list of divergent immune genes was then further curated by manual inspection of multiple alignments of cDNA transcript and genomic sequences of *C. atys* (Caty_1.0), *M. mulatta* (MacaM) and human (GRCh38.p7). Multiple alignment analysis was performed using Multalin (http://multalin.toulouse.inra.fr/). *TLR4* and *ICAM2* sequence alignments were generated using Jalview.

Gene family evolution methods. In order to identify rapidly evolving gene families along the *C. atys* lineage, we obtained peptides from human, chimpanzee, orangutan, gibbon, macaque, baboon, vervet, marmoset and mouse from ENSEMBL 83[32]. The *C. atys* peptides were obtained from NCBI[33]. To ensure that each gene was counted only once, we used only the longest isoform of each protein in each species. We then performed an all-versus-all BLAST search on these filtered sequences[34]. The resulting *e* values from the search were used as the main clustering criterion for the MCL program to group peptides into gene families[35]. This resulted in 14,889 clusters. We then removed all clusters only present in a single species, resulting in 10,967 gene families. We also obtained an ultrametric tree from a previous study and added sooty mangabey based on its divergence time from baboon (TimeTree)[36,37].

With the gene family data and ultrametric phylogeny as input, we estimated gene gain and loss rates (λ) with CAFE v.3.0[38]. This version of CAFE is able to estimate the amount of assembly and annotation error (ε) present in the input data using a distribution across the observed gene family counts and a pseudo-likelihood search. CAFE is then able to correct for this error and obtain a more accurate estimate of λ. We find an ε of about 0.04, which implies that 4% of gene families have observed counts that are not equal to their true counts. After correcting for this error rate, we find $\lambda = 0.0020$. These values for ε and λ are on par with those previously reported for mammalian datasets[38,39] (Extended Data Table 3b). Using the estimated λ value, CAFE infers ancestral gene counts and calculates *P* values across the tree for each family and lineage to assess the significance of any gene family changes along a given branch. CAFE uses Monte Carlo re-sampling to assess if a given family is rapidly evolving. For those families found to be rapidly evolving ($P < 0.01$), it then calculates *P* values for each lineage within the family using the Viterbi method. Those lineages with low *P* values ($P < 0.01$) are said to be rapidly evolving.

We observed 1,561 rapidly evolving families across the 10 species of mammals sampled here. Extended Data Table 3c summarizes the gene family changes for all 10 species. Humans have the highest average expansion rate across all families at 0.20 whereas gibbons have the lowest at −0.09, meaning that they have the most gene family contractions. *C. atys* has undergone 535 gene family expansions of which 96 are rapid expansions and 340 gene family contractions of which 48 are rapid contractions.

Genetic distance between *C. atys* and *M. mulatta* orthologues. The amino acid sequences of 9,257 *C. atys* proteins with RNA-seq support (Fig. 1) were aligned to *M. mulatta* orthologues as described above. We then used the codeml package from PAML (v.4.9a) on each of these alignments with the WAG amino acid rate matrix to calculate maximum likelihood genetic distances between the two sequences[40]. A histogram was generated from these distances with R (Extended Data Fig. 1a).

***TLR4* gene tree.** *TLR4* nucleotide sequences for 17 primate species were obtained from the NCBI GenBank resource (human: NM_138554.4; rhesus macaque: XM_015116960.1; sooty mangabey: manually curated XM_012091593.1; bonobo: NM_001279223.1; Nancy Ma's night monkey: XM_012472756.2; drill: XM_011973281.1; colobus monkey: XM_011950060.1; crab-eating macaque: NM_001319615.1; squirrel monkey: XM_003925187.2; baboon: XM_003911309.4; pig-tailed macaque: NM_001305889.1; marmoset: XM_017975811.1; gorilla: XM_004048514.2; chimpanzee: NM_001144863.1; orangutan: AB445642.1; African green monkey: XM_007968248.1; gibbon: XM_003264057.3). These sequences were aligned with PASTA2 and we then constructed a maximum likelihood gene tree with RAxML3, performing 100 bootstrap replicates[41,42] (Extended Data Fig. 7). Finding low bootstrap support amongst nodes ancestral to sooty mangabey, drill and baboon, we counted the number of sites that were discordant with respect to the gene tree topology. That is, the number of sites in which baboon and *C. atys* share the same state and *C. atys* and drill share a different state with an outgroup species (one of the two other Old World monkeys).

Sample collection and processing. Peripheral blood samples from SIV-negative rhesus macaques and SIV-negative sooty mangabeys were collected by venipuncture according to standard procedures at the Yerkes National Primate Research Center of Emory University and in accordance with US National Institutes of Health guidelines. Human blood samples were obtained from healthy donors at the Yerkes National Primate Research Center in accordance with Institutional Review Board protocol IRB0004582 and all relevant ethical regulations. Informed consent was obtained from all blood donors. Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood using Ficoll density-gradient centrifugation.

***In vitro* TLR-ligand stimulation assay.** The assay used in this study is a modified version of the procedure previously described[43]. Ultrapure LPS (*Escherichia coli* 0111:B4) and monophosphoryl lipid-A (*Salmonella minnesota*) were purchased from Invivogen. Whole blood collected in EDTA vacutainers was diluted 1:4 with RPMI 1640 medium and 195 μl aliquots were transferred to 96-well, round-bottom micro-titre plates. Agonists were diluted in RPMI 1640 and 5 μl were applied to the wells at the following final concentrations: LPS, 1,000–10 ng ml^{-1}; lipid-A, 10–1 μg ml^{-1}. Suspensions were then mixed by pipet and incubated at 37 ° C, 5% CO_2 for 4 h). After incubation, plates were centrifuged at 700 r.p.m. for 10 min, and 120 μl of cell-free supernatant was removed and stored at −80 °C until the assay was carried out. Each TLR ligand at a given concentration was performed in triplicate for each animal.

Cytokine bead array (CBA). Samples were obtained from sooty mangabeys and rhesus macaques housed at the YNPRC. Sooty mangabeys were naturally infected at the YNPRC and rhesus macaques had been infected previously with SIV_{smm} as previously described[19]. Supernatant levels of TNF and IL-6 were measured using the human inflammation CBA kit (BD Biosciences Immunocytometry Systems) according to the manufacturer's instructions, with the modification that the sample volumes for supernatant, antibody-coupled bead mix and PE-conjugated detection antibody solution were all reduced to 25 μl instead of 50 μl[44]. After incubation, samples were washed with 2% paraformaldehyde in PBS, resuspended in 150 μl PBS, and analysed using a FACSCalibur flow cytometer (BD Biosciences Immunocytometry Systems). The average of triplicate cytokine measurements was used as the representative value for individual animals, and variations in cytokine levels between species groups were tested for statistical significance using unpaired *t*-tests in Prism 6.0. To quantify the level of *TLR4* mRNA, and to perform linear regression of TLR-signalling molecules with TNF and IL6 cytokine levels, in the LPS-stimulated blood samples in the longitudinal SIV_{smm}-infected samples, we used microarray expression data from matched whole-blood samples; these data are available from the NCBI Geo database (accession GSE16147).

Plasma viral load measurement. Quantification of SIV_{smm} plasma viral RNA levels were quantified using qPCR as described previously[45,46].

RNA-seq analysis of LPS-stimulated monocytes. RNA-seq analysis was conducted at the Yerkes Nonhuman Primate Genomics Core Laboratory (http://www.yerkes.emory.edu/nhp_genomics_core/). $CD14^+$ monocytes were isolated from Ficoll-isolated PBMCs using CD14 MicroBeads according to the manufacturer's instructions (Miltenyi Biotec). Subsequently, 0.4×10^6 cells were stimulated for 6 h with 10 ng ml^{-1} LPS and then immediately lysed in 350 μl RLT buffer (Qiagen). RNA was purified using Micro RNEasy columns (Qiagen) and RNA quality was assessed using Agilent Bioanalyzer. Then, 10 ng of total RNA was used as input for mRNA amplification using 5′ template-switch PCR with the Clontech SMART-Seq v.4 Ultra Low Input RNA kit, according to the manufacturer's instructions. Amplified mRNA was fragmented and appended with dual indexed barcodes using Illumina NexteraXT DNA Library Prep kits. Libraries were validated by capillary electrophoresis on an Agilent 4200 TapeStation, pooled and sequenced on an Illumina HiSeq 3000 using (100 bp paired-end reads) at an average read depth of 18 million. RNA-seq data were analysed by alignment and annotation to either